Undergraduate Texts in Mathematics

UTM

T. Kyle Petersen

Inquiry-Based Enumerative Combinatorics

One, Two, Skip a Few… Ninety-Nine, One Hundred

Springer

Undergraduate Texts in Mathematics

Undergraduate Texts in Mathematics

Undergraduate Texts in Mathematics are generally aimed at third- and fourth-year undergraduate mathematics students at North American universities. These texts strive to provide students and teachers with new perspectives and novel approaches. The books include motivation that guides the reader to an appreciation of interrelations among different aspects of the subject. They feature examples that illustrate key concepts as well as exercises that strengthen understanding.

More information about this series at http://www.springer.com/series/666

T. Kyle Petersen

Inquiry-Based Enumerative Combinatorics

One, Two, Skip a Few… Ninety-Nine, One Hundred

Springer

T. Kyle Petersen
Department of Mathematical Sciences
DePaul University
Chicago, IL, USA

ISSN 0172-6056 ISSN 2197-5604 (electronic)
Undergraduate Texts in Mathematics
ISBN 978-3-030-18307-3 ISBN 978-3-030-18308-0 (eBook)
https://doi.org/10.1007/978-3-030-18308-0

Mathematics Subject Classification (2010): 05A05, 05A10, 05A15, 05A19, 11B37, 11B39, 11B65, 11B73, 60C05

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Do you remember your first day of your junior year of high school? I do. I was in Doc Drennan's classroom at Norman High School in Norman, Oklahoma, and he asked whether any of us could recite the quadratic formula. A dozen hands shot up. Then he asked whether any of us could explain why it was true. All the raised hands were lowered.

By the end of that day in class, however, we had on the board two complete explanations of the quadratic formula. What was interesting about this was that Doc Drennan never left his seat (which, incidentally, was at the back of the classroom). How did he do it? He asked us questions. He sent various students to the board to grab the chalk and explain their ideas. He had us discuss our ideas with each other. By the end of that year, we had learned most of single-variable calculus but also some computer programming, some number theory, even a unit on fractals. I'm sure he got out of his chair at some point, but my enduring memory of Doc Drennan is of him sitting at the back of the room and asking questions, usually some variation on "And what do you think?"

Though the label wasn't in common use at the time, today we call this teaching style "Inquiry-Based Learning," or IBL. For most of the 20th century the dominant pedagogical style in college math classes was the lecture. Early in the 21st century this is probably still true, but the tide seems to be shifting. Several landmark education studies have recognized that nearly any "active learning" teaching style is superior to lecture in terms of student outcomes. (Here "active" means that the students are engaged in classroom activities that promote higher order thinking.) In 2016, the Conference Board of the Mathematical Sciences, an umbrella organization consisting of seventeen professional societies (including the AMS, ASA, MAA, NCTM, SIAM, and SOA), put out a policy statement on active learning that calls on institutions of higher education to "ensure that effective active learning is incorporated into postsecondary mathematics classrooms." As IBL is certainly an active learning style, the time is ripe for IBL to become more widely practiced in the mathematics curriculum.

The goal in an IBL course is to get the students to be as active as possible in finding their own answers to the big questions of the subject. This can make using a typical textbook problematic for an IBL classroom, since the author has already given a clear explanation for the main ideas. What if we take that same textbook, but rip out the polished examples and proofs of theorems? This is exactly what Paul Halmos did in his Linear Algebra courses. (Halmos was a practitioner of IBL before it was known as IBL.) He called these "problems courses" because what students received on day one was a stack of problems to be solved and theorems to be proved. These problems constituted the content of the course.

This is what most IBL instructors do. If you want to teach course X, you think about the big ideas from that subject, then work backward from the big ideas to a sequence of bite-sized problems that lead to the big ones. Maybe your course has a hundred problems.[1] The thought goes that if you can solve these one hundred problems then you are well on your way to mastering subject X.

Combinatorics is a very broad subject, so the difficulty in writing about the subject is not what to include, but rather what to exclude. Which hundred problems should we choose? I wanted a small, relatively self-contained book with a narrow focus for a one-semester course. I chose enumerative combinatorics with an emphasis on generating function techniques.[2] Working backward from some big goal results (to do with Eulerian numbers and Narayana numbers), I ended up with Chapters 0–9 of the book. Later, I added further topics on Refined Enumeration (Chapter 10), Probability (Chapter 11), Partitions (Chapter 12), and some connections to Number Theory (Chapter 13).

Over the years, I tinkered with the problem sequence, adding and removing particular topics that seemed more or less important over time. Mostly I added more "bridge problems" between the big ones to make sure every student in the class had an access point to the material. I have learned from experience that rarely is there such a thing as a "too easy" problem. But if there is a low floor to the difficulty of some problems, I hope that with most of them there is a high ceiling as well.

One beautiful thing about the inquiry-based approach is that students will often dream up interesting variations on the themes in the book. With this in mind, I have added a collection of essays and reading suggestions after each

[1] In fact, the first few years I taught the Combinatorics course at DePaul I had *exactly* one hundred problems: hence the subtitle "One, Two, Skip a Few... Ninety-Nine, One Hundred."

[2] In some ways I hope this book can act as an inquiry-based version of the first two chapters of Herb Wilf's wonderful book "generatingfunctionology."

chapter. These interludes between chapters represent mini-lectures that I often give in my courses. They sometimes serve as a way to bridge the gaps between the topics of one chapter and the next. Other times they point students in directions of generalization and further investigation.

Let me finish with some hints about pace for the instructor using the book for the first time. In my experience, it takes a typical group of students about ten weeks to present one hundred problems. While progress is not constant, I think ten student presentations per week is a good rule of thumb. In my class, we go over the material in Chapter 0 together on day one—I bring actual wooden Tower of Hanoi puzzles for the students to play with. Then the students jump into the problems. How far we ultimately get varies, but as long as the spirit of inquiry is at the forefront of the course, I am happy. The journey matters to me more than the destination.

That said, there is some room for student and instructor choice. The problems and chapters are meant to be done in the order listed, but from a logical point of view, Chapters 12 and 13 could be done any time after Chapter 6, and Chapter 11 has only a very mild dependence on results from Chapters 8 and 9. I think Chapter 0 is important to discuss with students early on to set the tone for the course. And while I consider Chapters 1–4 to be essential material (First Principles, Permutations, Combinations, and the Binomial Theorem), some students will have seen this material in earlier courses. Depending on the backgrounds of the students, you could begin as late as Chapter 4 or Chapter 5, with only a quick review of earlier facts.

Acknowledgements. I owe thanks to many people for helping me in the course of my mathematical career. I have had many mentors, official and otherwise. For their guidance, encouragement, and support over the years I thank (in alphabetical order): Sara Billey, Ruth Charney, Ira Gessel, Don Knuth, Jim Propp, Vic Reiner, Bruce Sagan, Richard Stanley, John Stembridge, and Peter Winkler. My various research collaborators and co-authors also deserve credit. In a sometimes indirect way, all of these people helped to shape my outlook on mathematics by asking (and helping me to ask myself) good questions.

I also want to thank those in the IBL community who helped me get my early training in IBL teaching techniques. Over the years I had support first from Harry Lucas, Jr. and the Educational Advancement Foundation, and later from Stan Yoshinobu and the Academy of Inquiry Based Learning. Dana Ernst has also been a good friend and colleague over the years.

Several colleagues at DePaul provided feedback on drafts of the book, including Andy Carroll, Chris Drupieski, Karl Liechty, and Bridget Tenner.

My family deserves thanks for putting up with me as I finished the writing of this book. I appreciate your patience!

Lastly, I want to thank Doc Drennan for asking me good questions back in my junior year of high school.

Chicago, USA T. Kyle Petersen
May 2018 DePaul University

Contents

Image Credits

Page 1: "Tower of Hanoi" by Evanherk, accessed at https://commons.wikimedia.org/wiki/File:Tower_of_Hanoi.jpeg / CC BY-SA

Page 14: "John Taine" from Wonder Stories, Dec. 1931, accessed at https://commons.wikimedia.org/wiki/File:John_Taine_WS3112.jpg / CC BY

Page 14: "Eugene Charles Catalan" by Emile Delperee, accessed at https://commons.wikimedia.org/wiki/File:Eugene_charles_catalan.jpg / CC BY

Page 14: "Portrait of Leonhard Euler" by Jakob Emanuel, accessed at https://commons.wikimedia.org/wiki/File:Leonhard_Euler.jpg / CC BY

Page 14: Fibonacci from *I benefattori dell'umanita*, accessed at https://commons.wikimedia.org/wiki/File:Fibonacci.jpg / CC BY

Page 14: Francois Edouard Anatole Lucas photograph, accessed at https://commons.wikimedia.org/wiki/File:Elucas_1.png / CC BY

Page 14: Percy MacMahon photograph, accessed at https://commons.wikimedia.org/wiki/File:Percy_MacMahon.jpg / CC BY

Page 14: Blaise Pascal portrait, accessed at https://commons.wikimedia.org/wiki/File:Blaise_pascal.jpg / CC BY

Page 14: Photograph of Neil Sloane provided by Neil Sloane, reproduced with permission

Page 98: Photograph of Neil Sloane's notebook, accessed at https://oeis.org/A000435/a000435_1.pdf, reproduced with permission from Neil Sloane

Page 99: Screenshot of On-Line Encyclopedia of Integer Sequences, accessed at http://oeis.org/, reproduced with permission from Neil Sloane

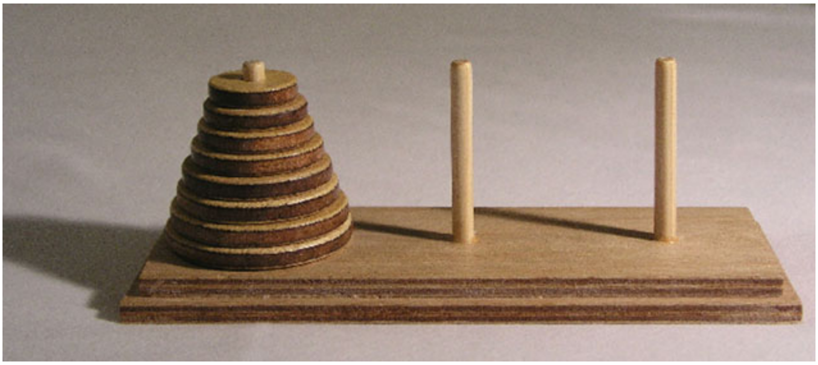
All other images are the author's own.

Chapter 0
Introduction to this book

"We are continually faced with a series of great opportunities brilliantly disguised as insoluble problems."
–John W. Gardner

An introduction to combinatorial problem solving

The Tower of Hanoi is a famous puzzle invented by Edouard Lucas in 1883. There are three pegs and eight disks of all different sizes. Initially the disks are stacked on one peg in decreasing size from the bottom up.



The Tower of Hanoi puzzle.

The object of the game is to move all eight disks onto another peg by moving only one disk at a time. However, a larger disk cannot lie on top of a smaller disk at any time.

Q: How many moves are required to solve the Tower of Hanoi puzzle?

Think about this question for yourself before turning the page.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_0

One way to approach a puzzle like this is to first generalize the question. For example, why should there be eight disks? What about n disks? If we suppose the puzzle can have any number of disks then we can scale the problem up and down at will. This leads to a good first step for any problem.

DATA COLLECTION: CONSIDER SMALL CASES FIRST.

We can see right away that it's easy to solve the puzzle if there are only one or two disks, and a moment's thought leads us to solve the three disk puzzle. We see that it takes us 1, 3, and 7 moves, respectively, to solve these cases.

We now come to an easily overlooked part of problem-solving.

NOTATION: NAME AND CONQUER.

As we are collecting our data, it is useful to give a name to the quantities we see. We will let T_n denote the minimum number of moves required to move n disks from one peg to another. ("T" for Tower.) So far, we have $T_1 = 1$, $T_2 = 3$, and $T_3 = 7$. There is an even smaller case we haven't considered yet: $T_0 = 0$, since it takes no moves to move no disks!

It often helps to make a table like the one below.

n	T_n
0	0
1	1
2	3
3	7

Now that we've handled some small cases, let's think big. How would we deal with a tower where n is large? You may have already noticed that a winning strategy for three disks is to move the top two disks onto a different peg first, then to move the big guy, then to move the smaller two back on top of the big one. In general, we can move the smallest $n-1$ disks onto a different peg (in T_{n-1} steps), move the biggest disk (in one move), then move the smaller disks back on top (in T_{n-1} more steps). With this strategy, it takes $2T_{n-1}+1$ steps to move n disks. We don't know if this is the best possible strategy, so it only gives an upper bound on the fastest way to move n disks:

$$T_n \le 2T_{n-1} + 1, \qquad \text{for } n \ge 1.$$

We'd like to say that this is an equality. To do this, we need to show there isn't some other, faster way to move the disks.

So, is there a faster way to do it? No. In order to move the biggest disk to a new peg, *all* the smaller disks must be on one peg, and by definition this *requires* T_{n-1} moves. We now need one move to get the big guy onto his new peg. To move the smaller disks back on top of the biggest one, we need, again by definition, at least T_{n-1} more steps. Thus, we have

$$2T_{n-1} + 1 \leq T_n, \qquad \text{for } n \geq 1.$$

The only way for both of our inequalities to be true is if we in fact have equality. To summarize,

$$\begin{aligned} T_0 &= 0, \\ T_n &= 2T_{n-1} + 1, \qquad \text{for } n \geq 1. \end{aligned}$$

Such a set of equalities is called a RECURRENCE RELATION: a way of getting new values in our sequence of numbers, given knowledge of some previous terms. We will see more of these later in the book.

The first nice thing about a recurrence is that it allows us to quickly generate terms in the sequence:

$$\begin{aligned} T_0 &= 0, \\ T_1 &= 2 \cdot 0 + 1 = 1, \\ T_2 &= 2 \cdot 1 + 1 = 3, \\ T_3 &= 2 \cdot 3 + 1 = 7, \\ T_4 &= 2 \cdot 7 + 1 = 15, \\ T_5 &= 2 \cdot 15 + 1 = 31, \\ T_6 &= 2 \cdot 31 + 1 = 63, \text{ and so on.} \end{aligned}$$

An obvious drawback of this approach is that if we want, say, the value of T_{100}, we need the value of T_{99}, which in turn requires T_{98}, and so on, all the way back down to $T_0 = 0$. A better solution would give some kind of useful formula for T_n that only depends on algebraic operations involving n.

Do we recognize the sequence of numbers $0, 1, 3, 7, 15, \ldots$? Aha! They are each one less than a power of 2. Specifically, it seems that

$$T_n = 2^n - 1, \qquad \text{for } n \geq 0.$$

(At least it works for $n \leq 6$.) How can we verify this formula in general?

Well, we know it works when $n = 0$, since $2^0 - 1 = 1 - 1 = 0$, and if we suppose that $T_n = 2^n - 1$ for some particular $n \geq 0$, then, via the recurrence relation we find:

$$\begin{aligned} T_{n+1} &= 2T_n + 1, \\ &= 2(2^n - 1) + 1, \\ &= 2^{n+1} - 2 + 1 = 2^{n+1} - 1, \text{ as desired.} \end{aligned}$$

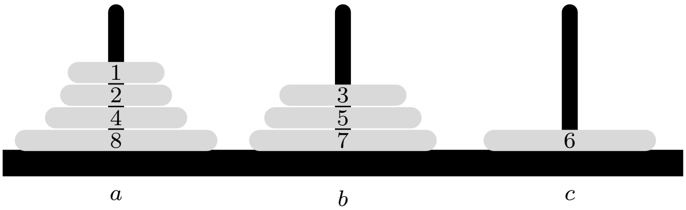
So if our formula works for some value of n, it works for the next value of n as well. Because we know it works for $n = 0$, it must work for $n = 1$, hence for $n = 2$, hence for $n = 3$, and so on.[1] It works for any value of n we like!

Now that we have complete faith in our formula, it is trivial to compute T_n. Going back to the original question, we see that $T_8 = 2^8 - 1 = 255$ moves are required to solve the Tower of Hanoi puzzle.

We've answered our original question and then some. But as mathematicians we may want a deeper understanding of the structure of the problem. We have only counted the minimal number of moves required to solve the puzzle; we haven't explicitly described how to solve it. If we really want to knock a problem out of the park, we want a *characterization of solutions*. Not only "how many are there?" but "what are they?"

To do this, it often helps to draw pictures or somehow encode the information in the problem.

MODELING: DISTILL THE ESSENTIALS.



$$\begin{array}{r|cccccccc} \text{disk:} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \text{peg:} & a & a & b & a & b & c & b & a \end{array} \longrightarrow \text{“}aababcba\text{”}$$

Fig. 0.1 Encoding a stage of the game.

For the Tower of Hanoi, we can label the pegs a, b, c and write a word like *aababcba* to mean the disks $1, 2, 4, 8$ are on peg a, disks $3, 5, 7$ are on peg b, and disk 6 is on peg c, as shown in Figure 0.1. We know that the word *aababcba* could only encode this configuration because it tells us precisely which disks go with which pegs; on each peg the disks must be stacked biggest to smallest from the bottom up. (In some sense we now think about the disks first, rather than the pegs first.) This is a big leap forward, as far as bookkeeping goes. It becomes much easier to record a sequence of moves than drawing pictures

[1] A mental image some people like is of a line of dominoes. The case $n = 0$ just shows that you can knock down the first domino. Then we prove that the dominoes are close enough together: if you knock down domino n, then domino $n+1$ will fall as well. This kind of argument is known as a PROOF BY INDUCTION; induction is often an easy way to verify facts about mathematical objects that have some sort of recursive structure.

showing what has happened. For example, to demonstrate how to solve the three disk puzzle, we could simply write down the following steps:

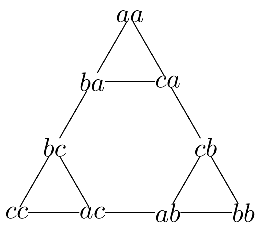
$$aaa \to baa \to bca \to cca \to ccb \to acb \to abb \to bbb.$$

This is certainly simpler than sketching the disks and pegs![2]

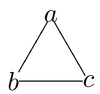
Now that we've got a way to encode the possible states of the game as strings of letters, we want to understand how one string of letters gets transformed into another. In particular, we want to know the best way to transform a string of all a's to a string of all b's (or all c's). It's probably good to start with small cases again. If there is only one disk, there are three possible strings, each with one letter: a, b, and c. At any point we can move the disk from one peg to another, so we can swap any of these strings for another. Let's move on to two disks.

There are now nine possible strings: $aa, ab, ac, ba, bb, bc, ca, cb,$ and cc. It is no longer possible to get from any string to another with just one move. For instance, aa indicates that both disks are on peg a, so it is impossible to move the bigger disk without first moving the smaller one. Hence we can't transform aa into something like ac with only one move.

After a little bit of thought however, we can sketch the following diagram[3] to indicate which one-step moves are allowed:



Let's compare this with the corresponding diagram for one disk:



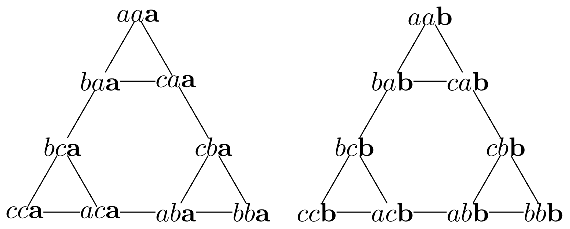
Hmm. It looks like the diagram for two disks is built out of the one-disk diagram by gluing three copies of the one-disk diagram together in a certain way. What is going on here? If we look at the smaller triangle at the top of

[2] This "encoding" is what's called a BIJECTION between the set of states of the n-disk game, and the set of (a, b, c)-strings of length n. This lets us easily see that there are 3^n possible game states with three pegs, since there are 3^n (a, b, c)-strings. The BIJECTION PRINCIPLE for enumeration is to count a set by creating a bijection (a reversible, one-to-one correspondence) with another, easier-to-count set.

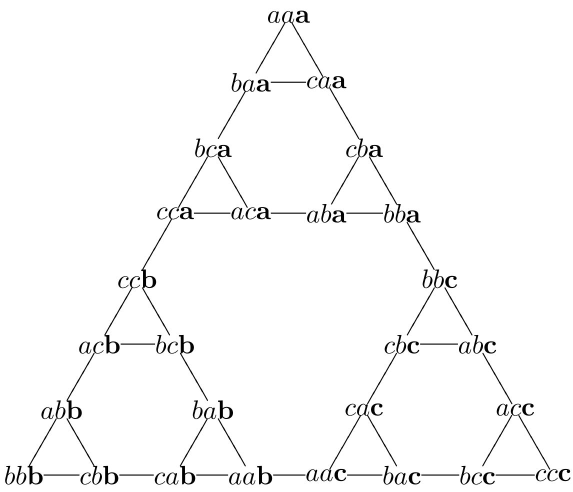
[3] This sort of diagram is called a GRAPH. In general, a graph is a collection of points, called "vertices" or "nodes", and edges between them. There is an entire branch of combinatorics devoted to the study of graphs.

the two-disk diagram, it looks *exactly* like the one-disk diagram if we were to add the letter a to the end of each string. But this makes sense, because these are just all the states of the game in which the biggest disk is left untouched on peg a. Similarly, the small triangle on the bottom left shows what happens when the largest disk is on peg c, and the bottom right triangle shows what happens when the largest disk is on peg b.

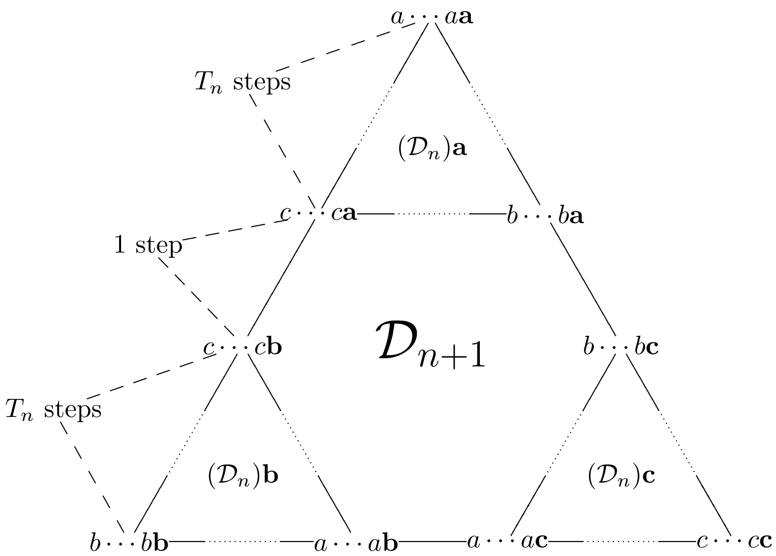
So if we want to figure out the diagram for three disks, we can get the major components of it by taking the diagram for two disks and attaching either an a, a b, or a c to the end of every string in the diagram (emphasized in bold):



All that's left is to tie these pieces together. But when can we move that biggest disk? Only when all the smaller disks are on the same peg, different from the location of the big guy. This occurs at one of the corners of the triangle. We connect these corners where possible ($cca \leftrightarrow ccb$, $bba \leftrightarrow bbc$, $aab \leftrightarrow aac$) and get the following picture:



All right, now we're cooking. We are seeing a kind of *structural recurrence* happening here. If we let $\mathcal{D}_n$ denote the nth diagram, we can continue our reasoning to see how to build $\mathcal{D}_{n+1}$ from $\mathcal{D}_n$. We append the letters a, b, and c to three separate copies of $\mathcal{D}_n$, then join up the appropriate corners:



This picture now lets us read off the optimal strategy for solving the Tower of Hanoi! To move the disks in the tower, we just do the sequence of moves indicated by the edges between the game states down one side of the triangle, from $a \cdots aa$ to $b \cdots bb$, say. Why is this optimal? Because the shortest distance between two points is a straight line! By the way, notice that our numeric recurrence, $T_{n+1} = 2T_n + 1$, is built right into the picture. Awesome!

Now are we done? Only if we want to be. The great thing about mathematics is that there's always room to improvise, always a new question that can be asked. We solved the "3-peg" Tower of Hanoi puzzle. What about the "4-peg" puzzle? The "k-peg" puzzle? What if there are k pegs, and two towers of different colors, and the goal is to relocate both towers? What if there are k pegs and ℓ towers? What if you put more (or fewer) restrictions on the kinds of moves that you allow?

The possibilities are only limited by your imagination.

An inquiry based approach

I LOVE MATHEMATICS, and combinatorics in particular, for the creative freedom it allows. When a class asks you to do nothing but sit and listen to a lecture, you are not experiencing the creative side of mathematics. Furthermore, listening to a lecture is not enough to learn. You must be active in the learning process. Who hasn't thought "I understood this when the professor was going over it, but now that I'm on my own..."? Don't get me wrong. Lecturing has its place, and I enjoy lecturing. When I teach this course it is not uncommon for me to give brief lectures at certain points in the book to highlight key topics and to make connections with other parts of mathematics. But no lecture can give you the joy of personal discovery.

To promote a more active participation in your learning, this book is designed for a course that incorporates ideas from an educational philosophy known as Inquiry-Based Learning, or IBL. The IBL approach is student-centered. It is a method of teaching and learning in which students are given tasks that require them to solve problems, ask questions, explore, create, and communicate effectively. In other words, the best way to *learn* mathematics is to *do* mathematics.

Rather than laying out a smooth and polished path to a solution, the IBL instructor guides and mentors student learning through well-crafted problems. I like to think of the difference between a lecture-driven classroom and an IBL classroom like the difference between taking a cable car to the summit of a mountain and climbing to the top of the same mountain. See Figure 0.2.

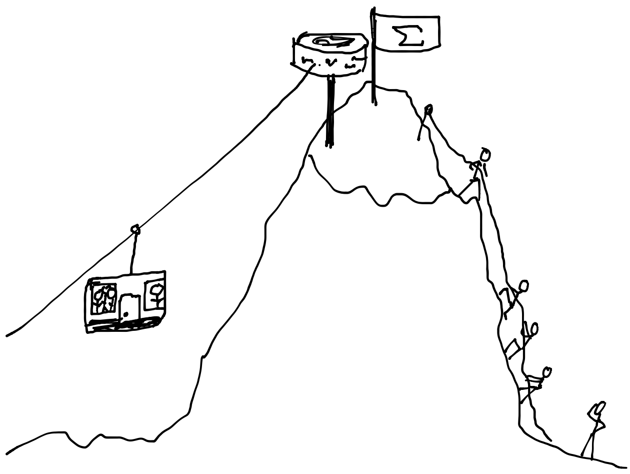


Fig. 0.2 Two ways to reach the summit.

One could argue that both approaches can get you to the top of the mountain, but clearly the experiences are different. And when you come to a new mountain, for which there is no cable car, who is better prepared to reach the summit?

Effective IBL courses can come in many different forms, but they all possess two essential ingredients. Students, as much as possible, are responsible for:

- guiding the acquisition of knowledge, and
- validating the ideas presented.

In particular, you should not look to the instructor as the sole authority on the mathematical content.

Of course, the instructor will usually have an advanced degree in mathematics and many years of experience, so their voice is important too. However, I like to think of the instructor in this setting as a "guide on the side" as opposed to a "sage on the stage," as is typical in the lecture format. The prominent role of student voices in the classroom means the instructor has more time to listen. This in turn allows the instructor to provide, in a dynamic fashion, those hints and suggestions that the students actually need. By constrast, a prepared lecture provides only those hints and suggestions that the instructor has anticipated the students need. While the intersection of the sets {anticipated student needs} and {actual student needs} is usually nonempty, these sets are rarely identical.

When I teach this course, our in-class time is mostly devoted to student presentations of the problems, and students are rewarded for how much mathematics they produce. My students are expected to:

- independently read and interact with the book,
- attempt all problems in good faith, solving as many of them as possible,
- present oral arguments during class for the solved problems,
- participate in discussions about each problem presented in class, and
- write clear and logical solutions for the presented problems.

These expectations foster self-reliance, diligence, and good communication skills.[4]

The toolbox

Good problems take time, and you should expect to be stuck much of the time. If you have solved a problem in under 10 minutes, that's great, but you aren't being challenged. Find a new problem. I hope you get stuck sometimes, so that you can learn what it takes to get un-stuck. Some of the skills I hope you develop are tenacity, flexibility, and persistence in your problem-solving. In other words, I hope you develop answers to the following question:

What do you do when you don't know what to do?

Here is an incomplete list of suggestions that might help you get started:

[4] Some instructors reading this may be starting to wonder about assessment: how do I assign grades in a class like this? There is no simple answer to this question. I think it depends heavily on context (the students, the institution, the goals of the course as it fits in the larger curriculum, and so on). I have used different strategies at different times. For those seeking advice on implementation, feel free to email me with questions. I am happy to share my thoughts, as they apply to the context in which you plan to run your course.

- reread any relevant definitions,
- reread the problem statement,
- work out some small examples by hand,
- ask a classmate,
- ask your professor during office hours, or
- go for a walk to get coffee.

With high probability, there will be problems that you are unable to solve on your own, even with good coffee. This is natural, but it is important that you understand the solution to *every* (yes, every!) problem that is presented during class. This means you have to learn to ask good questions, such as:

- Can you give me the broad outline of your reasoning?
- How did you know to set up that [equation, set, whatever]. . . ?
- How did you figure out what this question was asking?
- Was this the first thing you tried?
- Can you explain how you went from this line to that one?
- Could we have . . . instead?
- Would it be possible to . . . ?
- What if . . . ?

Another way to process a solution presented by another student is to make observations that connect their ideas to past work, or that compare approaches used by different people. Sample observations include:

- When I tried this problem, I thought I needed to . . . But I didn't need to, because. . .
- I think . . . is important to this argument because . . .
- When I read this problem, I thought it was the same as this earlier problem. Now I see that it is/isn't because. . .
- I really liked it when. . .

Let me also say a few words about resources external to this book and the people in your class. Much of the material in the book (particularly Chapters 1 to 4) can be easily found in other books and in online resources. Resist the temptation to look at these other sources!

The problems in this book are meant for *you*. You can do them, and you will be much more satisfied when you produce the solutions independently. Once a problem is already presented, then it might be interesting to find the result in a different location, for the purposes of comparing different approaches. But until then, resist the urge to look online!

Having said not to do online research, let me now immediately recommend breaking that rule in the case of one special website. From Chapter 5 onward, we will often be dealing with very interesting and important sequences of numbers, and here (after some initial engagement with the material yourself) I *do* recommend that you seek out another perspective via the

On-line Encyclopedia of Integer Sequences
oeis.org

This is a living, breathing database containing hundreds of thousands of entries that are constantly added to by researchers from around the world. By the time you get to Chapter 5, you should be ready to get a glimpse at the real world of research in enumerative combinatorics, and this website can be a window on that world.

Structure of this book

This book is different by design. Each chapter has a main topic, and there is a list of problems meant to guide you toward understanding of that topic. There are **Definitions**, **Theorems**, **Warmups**, and **Problems**. The **Definitions** and **Theorems** appear in special shaded bubbles like the one shown below, along with some brief exposition. The **Warmups** are meant to be easy problems to reinforce the ideas just introduced. Here's some sample text:

Unicorns

Definition. A *unicorn* is a mythical creature much like a horse, but with a single horn in its forehead.

Here is a picture of a unicorn:

Asian and European unicorns are generally smaller and more docile than South American unicorns. North American unicorns have shorter legs and come in many colors. The African unicorn has been extinct since the late 19th century.

Warmup. Can a unicorn have two horns? Which weighs more, a ton of delicious unicorn meat, or a ton of feathers?

In my classes, we briefly discuss the content of these text blocks together, which usually run less than a page. Sometimes we might split into small groups first, sometimes we just share as a whole class. I try not to use more than ten or fifteen minutes of class time on any one text block. I expect my students to be reading ahead, so these sessions are just to check and make sure we are all on the same page. This is a chance for them to ask basic questions, and for that student who confused a unicorn with a pegasus ("horn, no wings" versus "wings, no horn") to clarify their understanding. Whether you go over all the material in the expository text during classtime, understanding the **Definitions** and **Theorems** is essential for doing the problems successfully.

Each **Problem** is meant to be presented by you or one of your classmates, ideally in the order listed in the book. Sometimes these problems are guiding you to a bigger, more general result. Other times, the problems are about applying the ideas in a new way. To visually separate **Problems** from background material I use a blue side bar like this:

Problem. Show that in any collection of unicorns there are four times as many legs as horns.

These problems are the core of the book. Some problems are easier than others, but remember:

EVERY PROBLEM CAN BE DONE BY YOU, THE STUDENT.

I have tried to keep exposition to a minimum, since I don't want to spoil too many of the secrets that you are perfectly capable of revealing by yourselves. At the end of each chapter I have included a short essay on a relevant topic that can often lead to further exploration for interested students.

There are also additional exercises and problems at the back of the book, corresponding to each chapter. These can be deployed if it seems there is a need to review a topic from the chapter, or if there is a desire to do homework problems in addition to the presentation problems. With a few notable exceptions (especially in later chapters), these additional problems do not break new ground with respect to conceptual understanding.

A word about proofs

You may have noticed that there was no explicit "Theorem" stated in the Tower of Hanoi discussion, and nowhere was a "proof" clearly delineated. In the problems and theorems that constitute this book, you will be asked to provide solutions, to give explanations, and, occasionally, to "prove" that a certain result is true. What do I mean by "prove"?

Your instructor will help your class develop its own norms for mathematical discourse, but to me, a proof is a clear, logical explanation for why something is true. Nothing more or less.

If you've taken a class about proofs, you may have mental templates for how to write proofs, based on labels like "direct proof," "contrapositive," "contradiction," "induction," and so on. That's fine, but unless I clearly ask for a certain type of argument, I would prefer if you not waste energy worrying about the *form* your explanations take. Invest your energy in the *content* of your explanations, and the form will follow.[5]

When working on the problems put to you in this book, focus on the following two questions. If you can answer them both satisfactorily, you will be just fine.

WHAT IS THE TRUTH?

and

WHY IS IT TRUE?

[5] This idea is not too different from what Chicago architect Louis Sullivan famously wrote: "It is the pervading law of all things organic and inorganic . . . that form ever follows function. This is the law."

Rogues' Gallery of Integer Sequences

Eric Temple Bell

Eugène Catalan

Leonhard Euler

Fibonacci

Édouard Lucas

Percy MacMahon

T. V. Narayana

Blaise Pascal

Neil Sloane

Fig. 0.3 Portraits of some famous names encountered in this book (For Image Credits, refer p. xi).

Chapter 1
First principles

" 'Begin at the beginning,' the King said gravely, 'and go on till you come to the end: then stop.' "

–Charles Dodgson, a.k.a., Lewis Carroll

The dawn of counting.

One, two, three ... This is the *caveman's counting algorithm*. The only truly trivial counting problem is one for which the things you wish to count are obviously in a one-to-one correspondence with the first n counting numbers. Otherwise, most counting problems are best approached by breaking them down into smaller, more manageable pieces that can be counted "caveman style."

In this chapter, you'll explore basic definitions and counting principles: The Sum Principle, The Product Principle, and The Bijection Principle. The first allows you to count things separately on a case-by-case basis, provided your cases don't overlap (odds/evens, boys/girls, red/white/blue, ...). The second allows you to count the number of outcomes of a sequence of events. The third allows you to count new objects by relating them to objects you already know how to count.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_1

Sets

Definition 1. A *set* is a collection of distinct objects known as the *elements* or *members* of the set.

When possible, we may list the members of a set between a pair of curly braces, as in the following set, called "S."

$$S = \{3, \star, !, 7, \pi\}.$$

Membership in a set is denoted by the symbol "$\in$," so, for example, we can write $3 \in S$, which when read aloud says "3 is a member of set S" or simply "3 is in S."

Warmup 1. Using the membership symbol "$\in$," write statements identifying every member of the set S above.

Some sets are a bit too big to conveniently list all the members. Nonetheless, it might be easy to define membership in the set. For this, we use the so-called *set builder notation*, in which the set is given a name, followed by a definition of its elements. For example, to define the set of two-digit even numbers, we might write

$$E = \{x \in \mathbb{Z} : x = 2k \text{ for some } k \in \mathbb{Z} \text{ and } -100 < x < 100\}.$$

Read aloud, this says "E is the set of integers x such that $x = 2k$ for some integer k and x is strictly greater than -100 and strictly less than 100." There is no strict dogma about how to use set builder notation. For example, the following would also be a perfectly fine way to define E:

$$E = \{x = 2k : k \in \mathbb{Z} \text{ and } x \text{ has two digits } \}.$$

Warmup 2. Write the set of positive odd numbers using set builder notation.

Problem 1.
Write the set of perfect squares, $\{1, 4, 9, 16, \ldots\}$, using set builder notation. Next, translate the following set defined with set builder notation into English and list a few of its elements:

$$P = \{x \in \mathbb{N} : x = 2^n \text{ for some } n \in \mathbb{N}\}.$$

In this book nearly all sets we encounter are finite, in which case we want to know how many members the set has.

Cardinality

Definition 2. For a finite set S, the number of elements of S is called the *cardinality* of S, denoted by $|S|$.

The set $S = \{3, \star, !, 7, \pi\}$ has cardinality 5, or $|S| = 5$. The sets $\{-1\}$, $\{0\}$, and $\{S\}$ all have cardinality 1.

For convenience, we declare that there is a unique set with zero elements, called the "empty set" and denoted $\emptyset$, or $\{\}$. Other special sets include the set of integers, denoted $\mathbb{Z}$, and the set of positive integers, denoted $\mathbb{N}$.

Warmup 3. Describe three of your favorite finite sets of cardinality three. The author's favorites include the set $\{1, 2, 3\}$, the set {Larry, Curly, Moe}, and the set containing his three children.

Problem 2.
How many even numbers are positive and less than 100? Phrase your answer in terms of the cardinality of a set.

Problem 3.
Let $S = \{n \in \mathbb{Z} : 1 \leq n \leq 49\}$, i.e., the set of all positive integers from 1 to 49. What is $|S|$? In your own words, why does S have the same number of elements as the set in Problem 2?

Union

Definition 3. The *union* of two sets S and T, written $S \cup T$, is defined to be the set of all elements of S together with all elements of T, i.e., $x \in S \cup T$ if and only if $x \in S$ or $x \in T$ (or both).

The union of $\{a, b, 7\}$ and $\{3, \star, !, 7, \pi\}$ is

$$\{a, b, 7\} \cup \{3, \star, !, 7, \pi\} = \{a, b, 3, \star, !, 7, \pi\}.$$

Warmup 4. Come up with a pair of sets A and B such that $|A| = 3$, $|B| = 5$, and $|A \cup B| = 6$.

Problem 4.
With S as in Problem 3 and

$$T = \{n \in \mathbb{Z} : 0 < n < 100 \text{ and } n \text{ is even}\}$$

(just as in Problem 2), what is $|S \cup T|$?

Intersection

Definition 4. The *intersection* of two sets S and T, written $S \cap T$, is defined to be the set of all elements of both S and T, i.e., $x \in S \cap T$ if and only if $x \in S$ and $x \in T$.

The intersection of $\{a, b, 7\}$ and $\{3, \star, !, 7, \pi\}$ is

$$\{a, b, 7\} \cap \{3, \star, !, 7, \pi\} = \{7\}.$$

The intersection of $\{a, b, 7\}$ and $\{1, 2, 3\}$ is the empty set, i.e.,

$$\{a, b, 7\} \cap \{1, 2, 3\} = \emptyset.$$

Warmup 5. Come up with a pair of sets A and B such that $|A| = 8$, $|B| = 6$, and $|A \cap B| = 3$.

Problem 5.
With S and T as in Problems 3 and 4, what is $|S \cap T|$?

Problem 6.
For any two finite sets S and T, explain why

$$|S \cup T| + |S \cap T| = |S| + |T|.$$

If $S \cap T = \emptyset$, i.e., $|S \cap T| = 0$, then we say the sets S and T are *disjoint*. A special case of Problem 6 can be stated as follows.

Sum Principle

Theorem 1. Suppose S and T are disjoint sets. Then

$$|S \cup T| = |S| + |T|.$$

This result is known as the *sum principle* and it has many intuitive applications.

Warmup 6. Describe two different pairs of disjoint sets A and B whose union has cardinality $|A \cup B| = 9$. Is it possible that $|A| = |B|$?

Problem 7 (Starfolks).
I want to get a coffee from Starfolks. The nearest one is four blocks East and three blocks North from here. See Figure 1.1. Assuming I only walk East or North, how many different routes can I take to get there:

1. if the last block I walk is heading North?
2. if the last block I walk is heading East?
3. if I don't care whether the last block I walk is North or East?

How can the sum principle be applied here?

Problem 8.
A *complete graph* on n vertices, denoted K_n, is the graph in which every vertex is connected to every other vertex. In Figure 1.2 we see K_5; the complete graph on five vertices. It has ten edges. How many edges does K_{10} have?

Problem 9.
How many dominoes are there in a set that includes double blank through

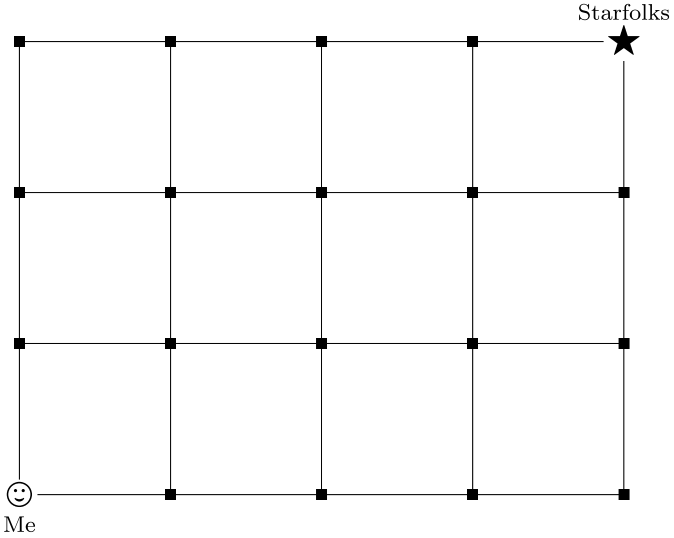


Fig. 1.1 Going for coffee.

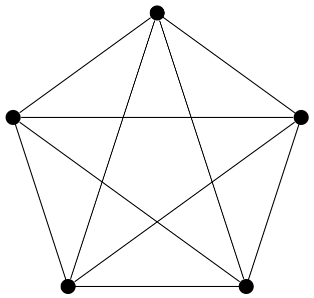


Fig. 1.2 The complete graph K_5.

double nine? Some of these dominoes are shown in Figure 1.3. What about double blank through double twelve? Through double n?

Problem 10.
Let's talk dominoes again. How many *dots* are there on a full set of dominoes? (You may want to work some small examples before finding a general formula.)

Problem 11.
Suppose we roll a six-sided die and then flip a coin. How many distinct outcomes are possible? (Assume we care only about the number on the die together with the side of the coin.) If we roll a six-sided die and then flip two

coins, how many outcomes are there? Does it matter if we toss two identical coins versus, say, flipping a penny and a nickel?

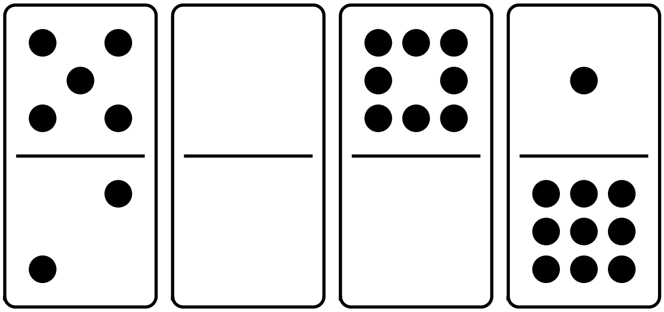


Fig. 1.3 Some dominoes.

Cartesian product

Definition 5. Suppose S and T are sets. Their *cartesian product*, denoted $S \times T$, is the set of all ordered pairs (s, t) such that $s \in S$ and $t \in T$:

$$S \times T = \{(s, t) : s \in S, t \in T\}.$$

In general, if $S_1, S_2, \ldots, S_m$ are sets, their cartesian product is the set of all ordered m-tuples:

$$S_1 \times S_2 \times \cdots \times S_m = \{(s_1, s_2, \ldots, s_m) : s_i \in S_i \text{ for all } i\}.$$

It is often helpful to think of the cartesian product of two sets as a rectangular array. For example, if $S = \{1, 2, 3\}$ and $T = \{a, b\}$, we might visualize their cartesian product as follows:

$$S \times T = \left\{ \begin{matrix} (1, a) & (1, b) \\ (2, a) & (2, b) \\ (3, a) & (3, b) \end{matrix} \right\}.$$

Warmup 7. Draw the cartesian product of the sets $\{1, x, \pi\}$ and $\{0, y, z\}$ in a rectangular array. If you haven't already, draw the outcomes from Problem 11 as a rectangular array as well.

Problem 12.
Let $S = \{n \in \mathbb{Z} : -5 \leq n \leq 5\}$ and let $T = \{n \in \mathbb{Z} : 1 \leq n \leq 6\}$. Describe the set $S \times T$. What is $|S \times T|$?

Problem 13.
Let $S_1 = \{1,2,3,4,5,6\}$, $S_2 = \{h,t\}$, and $S_3 = \{h,t\}$. Describe the set $S_1 \times S_2 \times S_3$. What is $|S_1 \times S_2 \times S_3|$?

The following theorem captures three progressively more general versions of the product principle.

The product principle

Theorem 2.
(Ver. I.) Suppose S and T are finite sets. Then

$$|S \times T| = |S| \cdot |T|.$$

(Ver. II.) Suppose $S_1, S_2, \ldots, S_m$ are finite sets. Then

$$|S_1 \times S_2 \times \cdots \times S_m| = |S_1| \cdot |S_2| \cdots |S_m|.$$

(Ver. III.) Suppose $\mathcal{S}$ is the set of outcomes of an m-step process, where for any $i \in \{1,2,\ldots,m\}$, there are a_i choices for step i, no matter what earlier choices were made. Then

$$|\mathcal{S}| = a_1 a_2 \cdots a_m.$$

The proof of the first version is straightforward when picturing the cartesian product with rectangular arrays. (One might even argue that we should take the statement of version I as our definition of multiplication for positive integers.) The second version follows from the first by induction on m, and the third follows if one can make a careful use of notation to identify the set $\mathcal{S}$ with a cartesian product.

The key distinction between version II and version III is that version III *does not assume that the set of choices for step i is independent of the previous choices*. If the choices at each step i are always the same, then we can simply let S_i denote the set of choices for step i and refer to version II.

Warmup 8. Amy, Bob, and Carol are three friends who are going to sit in seats $1,2,3$ of row ZZ at a concert. Here is a 3-step process for seating the friends. Step 1): choose who sits in seat 1. Step 2): choose who sits in seat 2. Step 3): choose who sits in seat 3. As there are three choices for step 1, two choices for step 2, and one choice for step 3, we find $3 \cdot 2 \cdot 1 = 6$ ways to seat the friends.

Explain why this counting argument uses Version III of the product principle, not Version II. Can you modify the counting argument so that Version II does apply?

Problem 14.
The following question refers to a standard deck of playing cards, in which there are 52 cards. The cards come in thirteen ranks: A (ace), $2, 3, \ldots, 10, J$ (jack), Q (queen), K (king), and four suits: $\heartsuit, \clubsuit, \diamondsuit, \spadesuit$. See Figure 1.4.

1. A friend takes a deck of cards and spreads them across a table, face down. You select one card at random and turn it over. There are many possible outcomes for the card you select. In how many ways can you select:
   a. a king?
   b. a black king?
   c. a red card?
   d. a face card (i.e., J, Q, K, or A)?
2. Now, after you select your first card (round 1), you give it back to your friend, who mixes all 52 cards and places them all back on the table for another card selection (round 2). Now, in how many ways can you select:
   a. a king in round 1 and a face card in round 2?
   b. a black king in round 1 and a red card in round 2?
   c. a black king in round 1 and a black card in round 2?
3. Now consider the situation in which you select your first card in round 1, but do not return it to your friend. Your friend mixes the remaining 51 cards and you select any one of them for round 2. In this situation, how many ways can you select:
   a. a black king in round 1 and a red card in round 2?
   b. a black king in round 1 and a black card in round 2?
   c. a black card in round 1 and a black king in round 2?
4. How many ways can you select a black king, then a red card, then a black 7, 8, or 9? Does your answer here depend on whether you return cards to your friend?

$A\heartsuit$	$2\heartsuit$	$3\heartsuit$	$4\heartsuit$	$5\heartsuit$	$6\heartsuit$	$7\heartsuit$	$8\heartsuit$	$9\heartsuit$	$10\heartsuit$	$J\heartsuit$	$Q\heartsuit$	$K\heartsuit$
$A\clubsuit$	$2\clubsuit$	$3\clubsuit$	$4\clubsuit$	$5\clubsuit$	$6\clubsuit$	$7\clubsuit$	$8\clubsuit$	$9\clubsuit$	$10\clubsuit$	$J\clubsuit$	$Q\clubsuit$	$K\clubsuit$
$A\diamondsuit$	$2\diamondsuit$	$3\diamondsuit$	$4\diamondsuit$	$5\diamondsuit$	$6\diamondsuit$	$7\diamondsuit$	$8\diamondsuit$	$9\diamondsuit$	$10\diamondsuit$	$J\diamondsuit$	$Q\diamondsuit$	$K\diamondsuit$
$A\spadesuit$	$2\spadesuit$	$3\spadesuit$	$4\spadesuit$	$5\spadesuit$	$6\spadesuit$	$7\spadesuit$	$8\spadesuit$	$9\spadesuit$	$10\spadesuit$	$J\spadesuit$	$Q\spadesuit$	$K\spadesuit$

Fig. 1.4 Playing cards in a standard deck come in thirteen ranks and four suits.

Problem 15.
Suppose you flip a coin five times in a row, recording the sequence of heads and tails you see, e.g., (h, h, t, h, t). How many different sequences of flips are possible?

Problem 16.
It's Halloween and five children arrive at your door, all hoping for candy. You have exactly five pieces of candy and you can give away none of the candy, all of the candy, or any amount in between. Supposing you don't give any child more than one piece, how many different ways can you distribute the candy? Does it matter if the pieces of candy are identical or not?

Subsets

Definition 6. Let S and T be sets. Then T is a *subset of* S, written

$$T \subseteq S,$$

if and only if every element of T is also an element of S.

Sometimes it is convenient to think of subsets in terms of intersections. That is, T is a subset of S if and only if

$$S \cap T = T.$$

Notice that it is always true that $\emptyset \subseteq S$ since $S \cap \emptyset = \emptyset$. Similarly, $S \subseteq S$ since $S \cap S = S$.

Warmup 9. Are there more two-element subsets of a four-element set, or more three-element subsets of a five-element set? Justify your response with examples.

Problem 17.
Let $S = \{a, b, c, d, e\}$. How many subsets does S have?

Problem 18.
Let n be a positive integer and suppose $S = \{i \in \mathbb{Z} : 1 \leq i \leq n\} = \{1, 2, 3, \ldots, n\}$. In terms of n, how many subsets does S have? Can you explain your formula with the product principle? Hint: Start small and work your way up. Look at the cases of $n = 1, 2, 3, 4$, and try to explain the pattern you see.

Functions

Definition 7. Suppose A and B are two finite sets. A *function* f from A to B is a subset of the Cartesian product $A \times B$ such that for each $a \in A$, there is exactly one $b \in B$ such that $(a, b) \in f$.

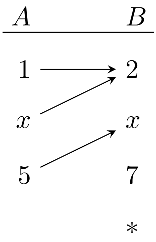
The set A is called the *domain* of f, while B is called the *codomain*. We denote a function from A to B by $f : A \to B$, and we use the common functional notation, $f(a) = b$, if and only if $(a, b) \in f$. This notation suggests f has "input" from the domain A and provides "output" in the codomain B.

The *range* of f, or *image* of f, denoted $f(A)$, is the subset of B consisting of second coordinates, i.e.,

$$f(A) = \{b \in B : f(a) = b \text{ for some } a \in A\}.$$

For example, if $A = \{1, x, 5\}$ and $B = \{2, x, 7, *\}$, then the set $f = \{(1,2), (x,2), (5,x)\}$ is a function, with range $f(A) = \{2, x\}$. We would write $f(1) = 2$, $f(x) = 2$, and $f(5) = x$.

For small functions, we can draw an "arrow diagram" to represent the function. For example, f as above could be drawn as follows.



Warmup 10. For the sets A and B above, give two subsets of $A \times B$ that are *not* functions.

Problem 19.
Count all the functions $f : \{a, b, c, d, e\} \to \{0, 1\}$. How could this result be related to Problem 17?

Injection/surjection/bijection

Definition 8. A function $f : A \to B$ is an *injection* if, for any two distinct elements of A, say $a \neq a'$, we have $f(a) \neq f(a')$. A function is a *surjection* if, for every $b \in B$, there exists an element $a \in A$ such that $f(a) = b$. A function that is both an injection and a surjection is called a *bijection*.

These concepts are easily described in terms of arrow diagrams: an injection has no two arrows pointing to the same element of B, while a surjection has at least one arrow pointing to each element of B. Thus, the arrow diagram for a bijection pairs off each element in A with a unique element in B.

Warmup 11. Draw arrow diagrams for three different functions. One of them should be an injection but not a surjection, one should be a surjection but not an injection, and the third should be a bijection. Choose different domains/codomains for each example.

Bijection principle

Theorem 3. *Suppose A and B are finite sets and f is a function $f : A \to B$. Then:*

- *if f is an injection, then $|A| \leq |B|$,*
- *if f is a surjection, then $|A| \geq |B|$, and*
- *if f is a bijection, then $|A| = |B|$.*

This last item describes a powerful counting technique that we have already used implicitly, and which we refer to as the *bijection principle*: If we know $|B| = n$, and we have a bijection between A and B, we can conclude $|A| = n$ as well.

Problem 20.
Here is an alternate approach to the result of Problem 16. Let A denote the set of ways to distribute candy in Problem 16, and let B denote the set of sequences of coin flips in Problem 15. Find a bijection $f : A \to B$. (Note that something similar could be done to relate these to Problem 19 and Problem 17.)

Organizing data

DATA COLLECTION SOUNDS BORING, and it can be. But this often underappreciated part of the problem-solving process can be crucial to developing true understanding. Throughout the problems you studied in Chapter 1, you probably noticed that working through small cases is often a key part of finding a complete solution, just like in the Tower of Hanoi problem from Chapter 0. Here we will discuss two straightforward methods for organizing data that can come in handy on all sorts of problems: the use of *lexicographic ordering* and the use of *decision trees*.

Lexicographic ordering

Let's return to Problem 7, which asked us to count paths to Starfolks as shown in Figure 1.1. How did you begin with this problem? Maybe you started by sketching some arbitrary paths as in Figure 1.5.

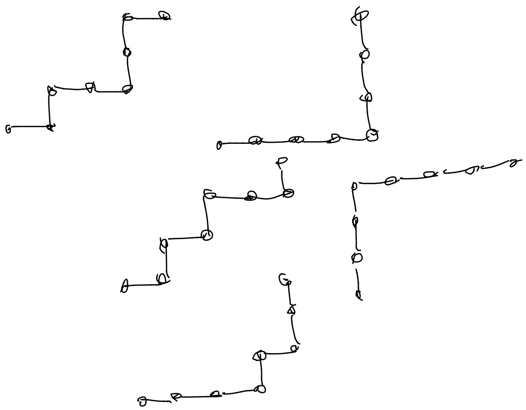


Fig. 1.5 Sketching some random paths for Problem 7.

But pretty soon you worry that you might be overlooking some of the paths, so you start to organize them a bit more carefully, like in Figure 1.6.

Intuitively, you are sorting the paths in a systematic manner, even if it is hard to articulate at first. You can now probably convince yourself that no path is left off the list, which is crucial, and you aren't accidentally duplicating any path.

All that drawing of paths has you worn out, though, so you give up on drawing all the pictures and you start translating the paths into "directions"

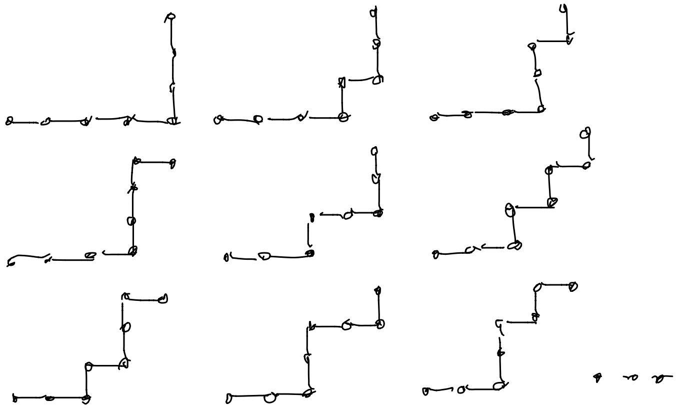


Fig. 1.6 Organizing the paths in Problem 7.

by way of a sequence of letters E and N that tell you how to traverse the path. (There is a bijection implicit here!) You start over:

$$EEEENNN, EEENENN, EEENNEN,$$
$$EEENNNE, EENEENN, EENENEN,$$
$$EENENNE, EENNEEN, EENNENE, \ldots$$

It turns out there is a name for this ordering. It is called *lexicographic ordering* and it appears in many guises. This is a generalization of alphabetical order that we all learn in grade school.

For the paths encoded as "words" on the "alphabet" $\{E, N\}$, we can compare paths in lexicographic order by reading from left to right. As soon as the two paths disagree, the one whose next letter is "E" is declared to be smaller in lexicographic order.

This ordering is natural even if we had used a different encoding for the paths. Instead of Ns and Es, maybe we just record which steps are East, with the understanding that all other steps are North. With this in mind, our paths correspond to the subsets of $\{1, 2, 3, 4, 5, 6, 7\}$ with exactly four elements. If we write the elements of our subsets in increasing order, then it is easy to order the subsets lexicographically:

$$\{1, 2, 3, 4\}, \{1, 2, 3, 5\}, \{1, 2, 3, 6\},$$
$$\{1, 2, 3, 7\}, \{1, 2, 4, 5\}, \{1, 2, 4, 6\},$$
$$\{1, 2, 4, 7\}, \{1, 2, 5, 6\}, \{1, 2, 5, 7\}, \ldots$$

While other methods might also work, always consider organizing your data with lexicographic order when generating small examples.

Decision trees

Another good way to organize small data is to look for recursive structure, and this can often be done with the help of a *decision tree*. Recall the subset counting that we needed to do in Problems 17 and 18. Here, lexicographic order works, but feels a bit awkward. In Table 1.1 we see all subsets of $\{1, \ldots, n\}$, for $n = 0, 1, 2, 3$. In each row the sets are listed in lexicographic order.

n	Subsets
0	$\{\}$
1	$\{\}, \{1\}$
2	$\{\}, \{1\}, \{1,2\}, \{2\}$
3	$\{\}, \{1\}, \{1,2\}, \{1,2,3\}, \{1,3\}, \{2\}, \{2,3\}, \{3\}$

Table 1.1 Subsets listed in lexicographic order.

Not exactly revealing, is it?

But if we think of creating new subsets from old, things look nicer, as in Figure 1.7. Here, each branch in the tree represents a step in a natural counting process using the product principle. At step i, we decide whether or not to include element i in the subset.

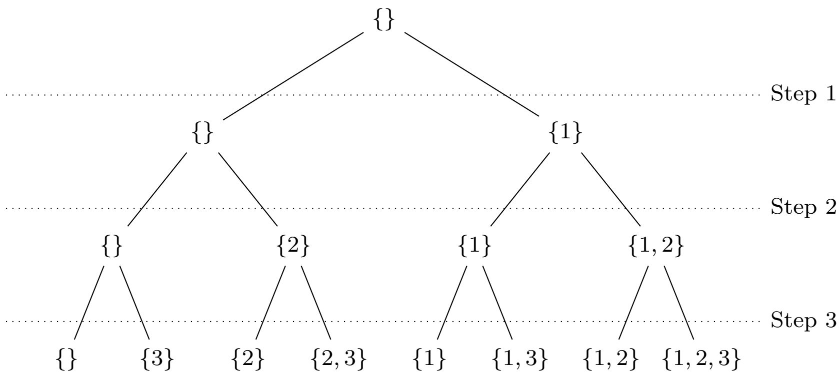


Fig. 1.7 Using a decision tree to generate subsets.

Listing out all possible orderings of a set can be done with a tree too. It would be pretty straightforward to list all the orderings of the set $\{a, b, c, d\}$ in lexicographic order, but we can also think of this process with a decision tree. Each step in the process has us choose where to place each letter, relative to any letters already placed. First we place letter a. Then we place b to the left or right of a. Then we place c in one of three positions relative to a and b, and so on. See Figure 1.8.

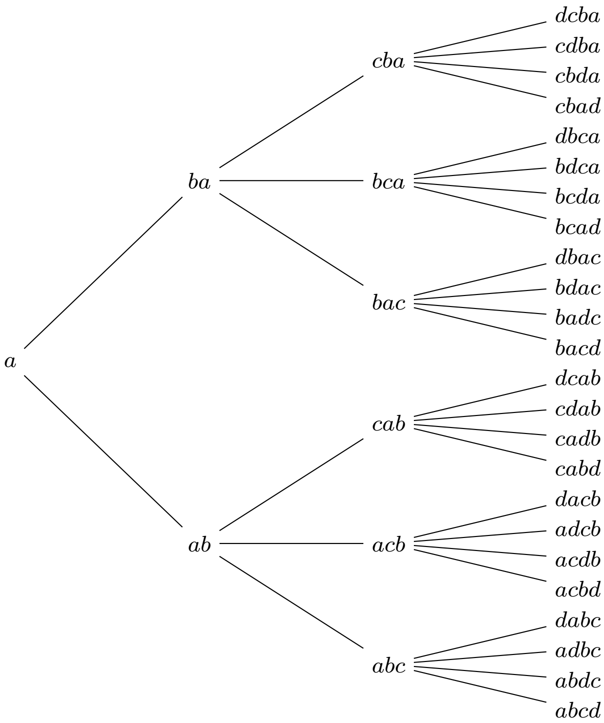


Fig. 1.8 A decision tree for orderings of the set $\{a, b, c, d\}$.

Notice how this process had one choice, then two choices, then three, then four. We could think of a kind of "inverse" decision process as follows. First pick which letter goes in position one, then which goes in position two, and so on. This yields the picture in Figure 1.9. Notice that now there are four choices first, then three choices, then two, then one.

More generally, decision trees can be a good way to organize your thinking about a problem and break it into conceptually smaller pieces. Figure 1.10 shows a sketch from a brainstorming session about Problem 14.

Whatever your method, when faced with a new problem you should take the time to write out small examples by hand. Hunt for structure. Look for

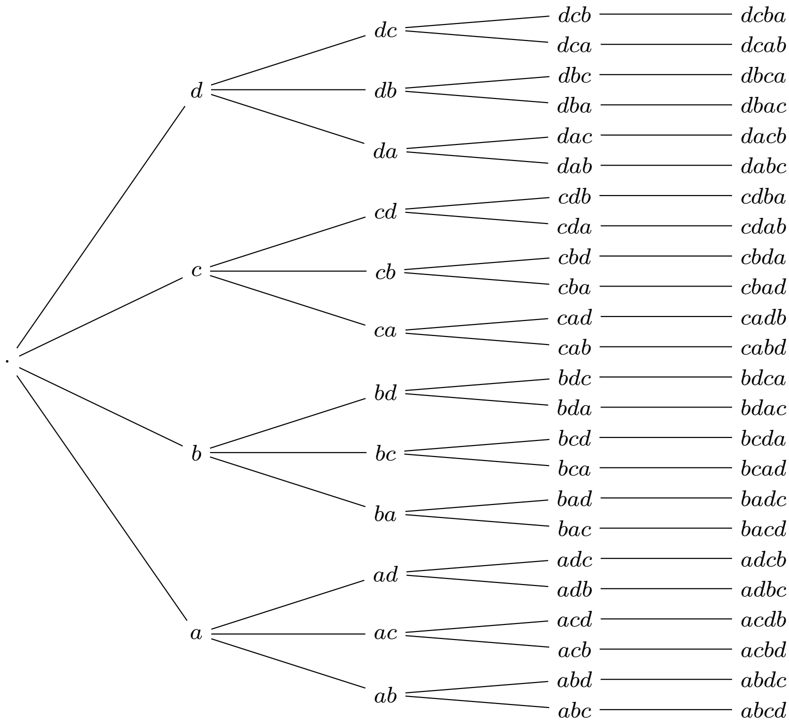


Fig. 1.9 Another decision tree for orderings of the set $\{a, b, c, d\}$.

patterns. You may have to do this more than once before you hit upon an idea that will generalize, but very often your patience will be rewarded.

Further reading

- George Pólya, "How to solve it," Princeton Science Library, 2004.
  This book, originally from 1945, is still a classic in the art of problem solving. There are four basic steps: 1) understand the problem, 2) make a plan, 3) carry out the plan, and 4) look back on your work.

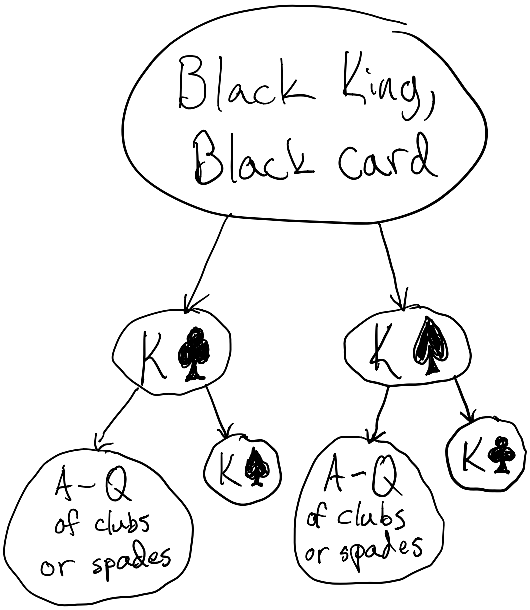


Fig. 1.10 Brainstorming about Problem 14.

Chapter 2
Permutations

"Words differently arranged have a different meaning and meanings differently arranged have a different effect."

–Blaise Pascal

53176842

$\updownarrow$

					X		
			X				
			•	X			
X							
•			•	•	•	X	
•	X						
•	•		•	•	•	•	X
•	•	X					

$\updownarrow$

$(4, 2, 0, 3, 2, 2, 1, 0)$

What is the best way to encode a permutation?

A PERMUTATION IS just about the most fundamental structure in all of enumerative combinatorics. Many, many, problems can be recast as a problem about counting certain ordered arrangements of a collection of objects.

In this chapter you will start to learn to count permutations.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_2

$n\backslash k$	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	2				
3	1	3	6	6			
4	1	4	12	24	24		
5	1	5	20	60	120	120	
6	1	6	30	120	360	720	720

Table 2.1 Triangle of the numbers $P(n, k)$, or the number of k-permutations of n.

Problem 21.
At a concert, you and three friends occupy seats 1, 2, 3, and 4 of row ZZ. How many different ways can you all be seated so that you have seat 1? How many different ways can you all be seated if you don't necessarily occupy seat 1?

Problem 22.
Generalizing the previous problem, suppose that there are n people sitting in seats $1, 2, 3, \ldots, n$ of a certain row. How many different ways can these people be arranged?

Problem 23.
We say an arrangement of rooks on a chessboard is *non-attacking* if no two of the rooks lie in the same row or column. For example, the left of Figure 2.1 shows an arrangement of four non-attacking rooks on a 4-by-4 chessboard. (We think of the rooks themselves as indistinguishable, so this is the only non-attacking arrangement of rooks in these cells.) How many different arrangements of four non-attacking rooks on a 4-by-4 chessboard are there with a rook in the bottom left corner (column a, row 1)? How many arrangements are there with a rook in column a, row 2? How many such arrangements are there in total?

Problem 24.
How many arrangements of n non-attacking rooks on an n-by-n chessboard are there?

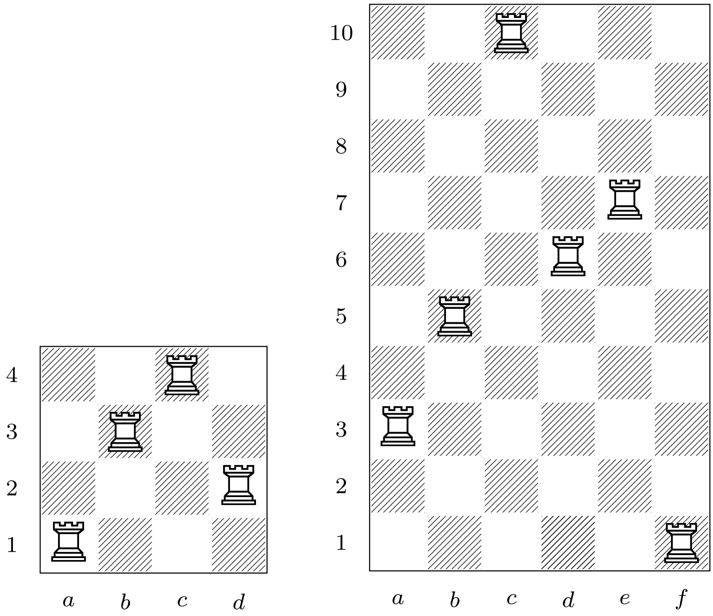


Fig. 2.1 An arrangement of four non-attacking rooks on a 4-by-4 chessboard and an arrangement of six non-attacking rooks on a 10-by-6 chessboard.

Factorial notation

Definition 9. The product of the first n consecutive natural numbers, $1 \cdot 2 \cdot 3 \cdots n$, is called n *factorial,* written $n!$ for short.

By convention, define $0! = 1$. Notice also that $n! = n \cdot (n-1)!$. The first few factorials are

$$1, 1, 2, 6, 24, 120, 720, \ldots.$$

Warmup 12. What is $\frac{7!}{3!}$? Explain why $\frac{a!}{b!}$ is an integer if $a \geq b$.

Problem 25.
How many bijections are there of the form $w : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$?

Problem 26.
Suppose ten friends go to a movie, but it is opening weekend and they can't find ten seats in a row. In one row, they find seats 1–6 unoccupied. How many different ways can six of the friends sit together in these seats? (Ignore what happens to the four friends who don't sit in these seats.)

Problem 27.
Generalizing the previous problem, suppose n friends find only k seats in a

row, where $k \le n$. Find a formula for the number of different ways k of the friends can sit together in these seats.

Problem 28.
How many ways are there to arrange six non-attacking rooks on a 10-by-6 chessboard?

Problem 29.
How many ways are there to arrange k non-attacking rooks on an n-by-k chessboard?

Problem 30.
Fix $n \ge k > 0$. How many injections are there of the form $w : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, n\}$?

123	132	213	231	312	321
124	142	214	241	412	421
125	152	215	251	512	521
134	143	314	341	413	431
135	153	315	351	513	531
145	154	415	451	514	541
234	243	324	342	423	432
235	253	325	352	523	532
245	254	425	452	524	542
345	354	435	453	534	543

Fig. 2.2 The members of $S_{5,3}$, the 3-permutations of $\{1, 2, 3, 4, 5\}$.

Permutations

Definition 10. A *k-permutation* of a set A is an injection

$$w : \{1, 2, \ldots, k\} \to A.$$

The set of all k-permutations of A is denoted $S_{A,k}$.

We usually write permutations as ordered lists of precisely k elements of A, i.e., $w = w(1)w(2)\ldots w(k)$. For example, if $A = \{\star, 4, a, b\}$, the set of 2-permutations of A is

$$S_{A,2} = \left\{ \begin{array}{l} \star 4, \star a, \star b, 4a, 4b, ab \\ 4\star, a\star, b\star, a4, b4, ba \end{array} \right\},$$

and, for example, $w = \star 4$ means $w(1) = \star$ and $w(2) = 4$.

The set of n-permutations of the set $\{1, 2, \ldots, n\}$ is denoted S_n. (The use of "S" is because this set is known in other contexts as *the symmetric group*.)

We use the notation $S_{n,k}$ to mean the set of k-permutations of $\{1, 2, \ldots, n\}$, with $S_n = S_{n,n}$.

The members of $S_{5,3}$, the 3-permutations of $\{1, 2, 3, 4, 5\}$, are displayed in Figure 2.2.

Warmup 13. Make a table listing the elements of S_4, i.e., all permutations of the set $\{1, 2, 3, 4\}$.

Warmup 14. If you haven't already, encode the non-attacking rook arrangements from Problem 23 as permutations. If you already have, find another logical encoding of the arrangements as permutations, e.g., from the perspective of your opponent.

Let $P(n, k)$ denote the number of k-permutations of a set with n elements, i.e., $P(n, k) = |S_{n,k}|$. By convention, $P(n, 0) = 1$, i.e., there is precisely 1 0-permutation of any set (even the empty set). Interpreting Problem 30 as a statement about permutations, we have the following result.

Permutation formula

Theorem 4. For any integers $n \geq k \geq 1$,

$$P(n, k) = n \cdot (n-1) \cdots (n+1-k) = \frac{n!}{(n-k)!}.$$

Warmup 15. Use this formula for $P(n, k)$ to fill out the next row of Table 2.1. Describe any patterns you see in the triangle of numbers.

Problem 31.
Prove that $P(n, n) = P(n, k)P(n-k, n-k)$, both by using the formula in Theorem 4, and by using the meaning of k-permutations and the bijection principle.

Here is a hint to get started. By definition $|S_{n,k}| = P(n, k)$. Moreover, the product principle gives us $|S_{n,k} \times S_{n-k}| = P(n, k)P(n-k, n-k)$. Thus to prove the identity with the bijection principle, it suffices to describe a bijection

$$f : S_n \to S_{n,k} \times S_{n-k}.$$

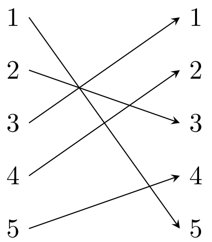
Problem 32.
Prove that $P(n, k) = P(n-1, k) + kP(n-1, k-1)$, both by using the formula in Theorem 4, and by using the meaning of k-permutations and the bijection principle. See Problem 31 for help setting up the desired correspondence.

The symmetric group

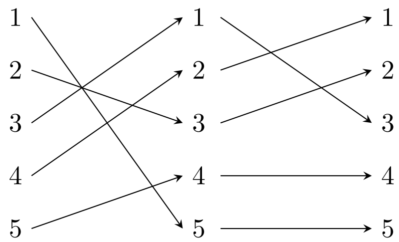
Permutations can be considered algebraic objects as well as combinatorial objects. It was mentioned that the set S_n is known as *the symmetric group*. We will discuss a little about what this means, though many details are best left to a course in abstract algebra.

Permutations as bijections

We have written permutations for the most part in "one-line" notation, e.g., $w = 53124$ is a member of S_5. In another context, it would be better to remember that each w is a bijective function, $w : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}$. The arrow diagram for $w = 53124$, which means $w(1) = 5$, $w(2) = 3$, and so on, would be drawn as follows.



It is easy to check that the composition of two such bijections is again a bijection, so S_n is closed under the operation of composition. For example, if we compose our permutation w above with the permutation $v = 31245$, we find the permutation $u = vw$, defined by $u(i) = v(w(i))$. (We write our composition of permutations from right to left as we would normally write the composition of functions.) With arrow diagrams, the composition u looks like:



Notice that $u(1) = v(w(1)) = v(5) = 5$, $u(2) = v(w(2)) = v(3) = 2$, and so on. We have $u = 52314$.

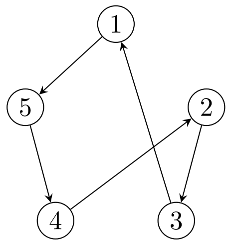
To say that S_n is a "group" means that:

- composition is associative, i.e., if $u, v, w \in S_n$, then $(uv)w = u(vw)$,
- there is an identity permutation, e, such that $we = ew = w$ for any $w \in S_n$, and
- every permutation w has an inverse, w^{-1}, such that $w^{-1}w = ww^{-1} = e$.

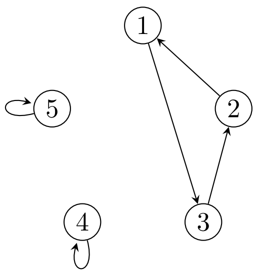
Each of these properties is easily verified. The identity permutation is the "do nothing" bijection $e = 12\cdots n$, i.e., $e(i) = i$ for all i.

Cycle notation

Another way to visualize a bijection on the set $\{1, 2, \ldots, n\}$ is to draw a directed graph whose nodes are the numbers 1 to n, with an arrow from i to j if $w(i) = j$. The example of $w = 53124$ would thus be drawn:



while $v = 31245$ is drawn:



This visualization helps us to see that permutations break up as collections of disjoint "cycles" on subsets of $\{1, 2, \ldots, n\}$. The standard way to write a permutation in terms of its cycles is to put the cycles in parentheses $(\cdots i\, w(i)\, w(w(i)) \ldots)$ and to concatenate the cycles. For example, we would write $w = (15423)$, $v = (132)(4)(5)$, and $u = (154)(2)(3)$. The cycles are usually ordered by their smallest elements as we've done here.

Cycle notation is very important for identifying algebraic properties of permutations. For example, the "order" of a permutation is the smallest nonnegative integer r such that $w^r = e$. For a cycle, r is just the size of the cycle. Thus in general, it is easy to argue that the order of any permutation is the least common multiple of the sizes of its cycles.

Symmetries of a simplex

Okay, so S_n is a group. But why the "symmetric" group? Where does symmetry come in to play? One nice answer to this question comes from a way of realizing the bijections in S_n as permutation matrices.

Given a permutation $w \in S_n$, define the $n \times n$ matrix M_w via

$$M_w[i,j] = \begin{cases} 1 & \text{if } w(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

For example, continuing with $w = 53124$, we have

$$M_w = \begin{bmatrix} 0&0&0&0&1 \\ 0&0&1&0&0 \\ 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&0&1&0 \end{bmatrix}$$

If we take a column vector $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ in $\mathbb{R}^5$, then

$$M_w\mathbf{x} = (x_5, x_3, x_1, x_2, x_4).$$

Look at the subscripts!

In general, if $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$ such that $M_w\mathbf{x} = \mathbf{y}$, then we find $y_i = x_{w(i)}$. Put simply, the permutation acts on $\mathbb{R}^n$ by permuting coordinates.

Now let Δ denote the set of points $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$ such that $x_i \geq 0$ and $\sum x_i = 1$. This is known as the *standard* $(n-1)$*-dimensional simplex.* (The dimension is one less than n because of the linear constraint on the coordinates.) Equivalently, the standard $(n-1)$-simplex can be described as the convex hull of the standard basis in $\mathbb{R}^n$.

For example, if $n = 2$, we find Δ is the one-dimensional simplex. It is defined as the set of points $\{(x,y) \in \mathbb{R}^2 : x, y, \geq 0, x + y = 1\}$, which is the line segment between $(0,1)$ and $(1,0)$. When $n = 3$, the two-dimensional simplex is the set $\{(x,y,z) \in \mathbb{R}^3 : x,y,z \geq 0, x+y+z = 1\}$, which is the triangle whose corners lie at $(0,0,1)$, $(0,1,0)$, and $(1,0,0)$. See Figure 2.3.

If $\mathbf{x}$ is any point in Δ, then linearity of matrix multiplication can be used to show $M_w\mathbf{x}$ is also a point in Δ. In other words, the image of Δ under the map M_w is Δ itself!

Since M_w is a linear transformation, many of the points inside Δ will likely get shuffled around. However, since $M_w(\Delta) = \Delta$, this transformation leaves us with an identical-looking object. So we say the action of w is a "symmetry"

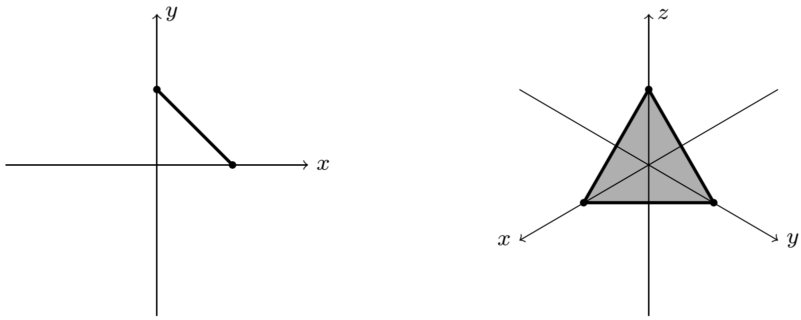


Fig. 2.3 The 1-dimensional simplex in $\mathbb{R}^2$ and 2-dimensional simplex in $\mathbb{R}^3$.

of Δ. For example, if we take $w = 312$ acting on the 2-simplex (triangle) in Figure 2.3, we have $M_w\mathbf{x} = (x_3, x_1, x_2)$. This action is 120 degree rotation around the line $x = y = z$, sending the corners to each other in a counterclockwise loop: $(1,0,0) \to (0,1,0) \to (0,0,1) \to (1,0,0)$. By contrast, if $v = 213$, then M_v keeps $(0,0,1)$ fixed, while swapping $(1,0,0)$ with $(0,1,0)$. The action of M_v reflects the triangle across the plane $x = y$.

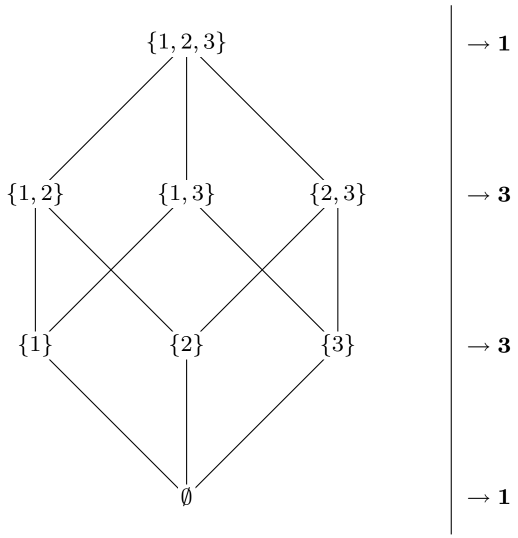
The simplex Δ is an example of a geometric object known as a "polytope." The study of symmetry groups of other polytopes is a vast subject involving interesting aspects of algebra, geometry, and combinatorics.

Further reading

- Marcus du Sautoy, "Symmetry: A Journey into the Patterns of Nature," Harper Perennial, 2009.
  This is an engaging and well-written book about symmetry and the life of a mathematician.
- Bruce Sagan, "The Symmetric Group," Springer Graduate Texts in Mathematics, 2001.
  For those who have already taken a first course in abstract algebra, this is a more advanced book on the algebraic combinatorics of permutations.
- Günter Ziegler, "Lectures on Polytopes," Springer Graduate Texts in Mathematics, 2006.
  Despite being a "graduate text," all that is needed to get started with this book is a background in linear algebra.

Chapter 3
Combinations

> "Mathematics is not a deductive science—that's a cliche. When you try to prove
> a theorem, you don't just list the hypotheses, and then start to reason. What you do is trial and error, experimentation, guesswork."
>
> –Paul Halmos



Organizing the subsets of the set $\{1, 2, 3\}$. What happens when we count them according to cardinality?

Binomial coefficients are among the most important whole numbers you will ever know. Most middle-school students have seen Pascal's triangle, and we will see that this triangle of numbers is one of the most flexible and durable objects in mathematics.

This chapter defines binomial coefficients and explores a few of their properties. In particular, we will find a nice formula for binomial coefficients, prove Pascal's recurrence, and discover some formulas for sums of binomial coefficients.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_3

$n\backslash k$	0	1	2	3	4	5	6
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

Table 3.1 Triangle of the binomial coefficients $\binom{n}{k}$, or the number of k-subsets of an n-element set.

Binomial coefficients

Definition 11. If A is a set and B is a subset of A with $|B| = k$, we refer to B as a *k-subset* of A. The set of k-subsets of a finite set A is denoted $\binom{A}{k}$, i.e.,

$$\binom{A}{k} = \{B \subseteq A : |B| = k\}.$$

The *binomial coefficient* $\binom{n}{k}$ is defined to be the number of k-subsets of an n-element set. By definition, if $|A| = n$, then $|\binom{A}{k}| = \binom{n}{k}$.

For example, if $A = \{\star, 4, a, b\}$, we have

$$\binom{A}{2} = \big\{\{\star, 4\}, \{\star, a\}, \{\star, b\}, \{4, a\}, \{4, b\}, \{a, b\}\big\},$$

which, since $|A| = 4$, establishes $\binom{4}{2} = 6$. Contrast this example with the example of $S_{A,2}$ following Definition 10.

Aloud we read $\binom{n}{k}$ as "n choose k", since each k-subset represents a different way to choose k elements from the set. In some other books the binomial coefficient $\binom{n}{k}$ is denoted $C(n, k)$, and k-subsets are sometimes called "k-combinations."

Notice that $\binom{n}{0} = 1$ since the empty set is a subset of every set, or in the language of choice, there is precisely one way to choose nothing.

Warmup 16. Let $A = \{1, 2, 3, 4, 5\}$, and list all the elements of $\binom{A}{3}$.

Caution! This definition of binomial coefficients contains no formula, though you may have learned about a formula for $\binom{n}{k}$ (involving a quotient of factorials) in another class. We will derive the formula in Theorem 5. Until then, rely only on what we have already proved and the definition given above.

Problem 33.
Five friends are outside a hip LA dance club, and they all want to get inside. However, the bouncer will only take two of them. In how many different ways can the bouncer choose a set of two of the friends to admit? Your answer here will establish the value of $\binom{5}{2}$.

Problem 34.
Generalizing the previous problem, suppose n friends are outside a hip LA dance club, but the bouncer will only admit 2 of them. In how many different ways can the bouncer choose a set of 2 of the n friends? (Hint: be careful not to double count!)

Problem 35.
Now the bouncer consents to admitting 3 of them. How many different ways can the bouncer choose 3 of the n friends? 4 of the friends?

Problem 36.
Generalizing the previous problem, suppose the bouncer will now take k of the n friends, where $k \leq n$. Find a formula for the number of ways the bouncer can choose k of the n friends. (Hint: it may help to relate this problem to the result of Problem 30. How is sitting in a movie theater different from entering a dance club?)

Problem 37.
Using the meanings of k-subset and k-permutation (i.e., using a bijection), explain why

$$P(n,k) = \binom{n}{k} \cdot k!.$$

From Problem 37 and the formula for $P(n,k)$ in Theorem 4 it follows that we get a nice formula for binomial coefficients using factorial notation.

Binomial formula

Theorem 5. For any $n \geq k \geq 0$,

$$\binom{n}{k} = \frac{P(n,k)}{k!} = \frac{n!}{k!(n-k)!}.$$

Now, for example, we can compute $\binom{8}{5} = 56$ by

$$\binom{8}{5} = \frac{8!}{5!3!} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = 8 \cdot 7 = 56.$$

Warmup 17. Use Theorem 5 to compute the next row of Table 3.1. Describe any patterns you see in the triangle of numbers.

Problem 38 (Pascal's identity)**.**
For any integers $n \geq k \geq 1$,

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

(This identity is why we sometimes call the array in Table 3.1 *Pascal's triangle*.) Prove this both by using the formula in Theorem 5 and by using the definition of $\binom{n}{k}$ in terms of k-subsets. Can you connect this problem to Problem 7?

Problem 39.
Prove that

$$\binom{n}{k} = \binom{n}{n-k},$$

both by using the formula in Theorem 5 and by using the definition of $\binom{n}{k}$ in terms of k-subsets.

Problem 40.
What are the row sums in Table 3.1? That is, for any $n \geq 0$, find a formula for

$$\sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.$$

Explain your formula using the meaning of $\binom{n}{k}$ and the result of Problem 18.

Problem 41.
What are the alternating row sums in Table 3.1? That is, for any $n \geq 0$, find a formula for

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n}.$$

Explain your formula using the meaning of $\binom{n}{k}$. Hint: this is easier to explain when n is odd.

Problem 42.
What are the diagonal sums in Table 3.1? For $n = 1, \ldots, 10$, find the numbers

$$\sum_{k \geq 0} \binom{n-k}{k} = \binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots.$$

(The sum ends after about $n/2$ terms because $\binom{a}{b} = 0$ if $b > a$.) Can you think of a general way to compute the nth such sum? You don't need to have an explicit formula for these numbers, but see if you can observe any patterns.

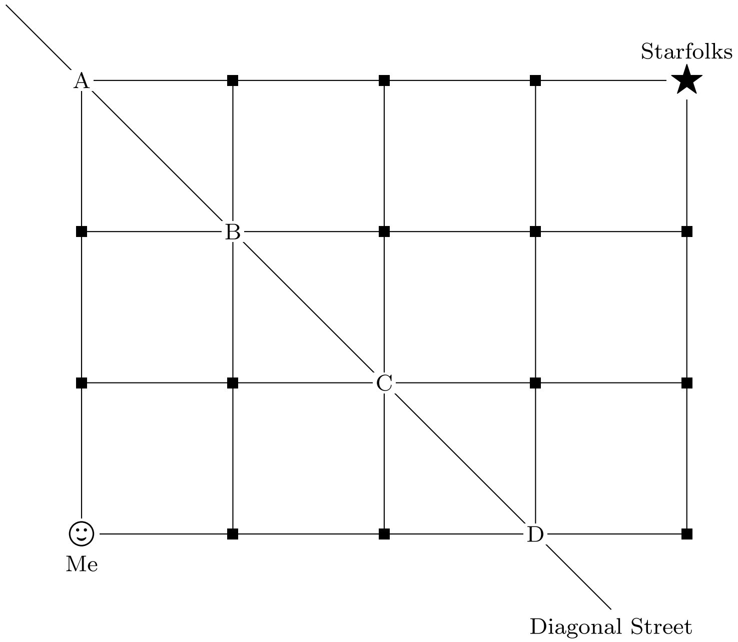


Fig. 3.1 Going for coffee again.

Problem 43 (Starfolks, II).

I'm headed to Starfolks again. It's still four blocks East and three blocks North, but there's a diagonal street as indicated in Figure 3.1. How many ways can I get to Starfolks by walking only seven blocks? Explain your answer in two different ways.

1. For your first explanation, ignore the diagonal street, since you won't walk along it anyway.
2. For your second explanation, break the set of paths into four subsetss, according to where you cross the diagonal street.

Problem 44

A university committee is composed of 3 men and 4 women. Three of the committee members must serve on an executive committee. How many ways can the executive committee be chosen? Explain your answer in two different ways.

1. For your first explanation, ignore gender.

2. For your second explanation, break the question into cases, according to the number of women on the executive committee.

Problem 45

Can you explain why, for any k and m less than or equal to n,

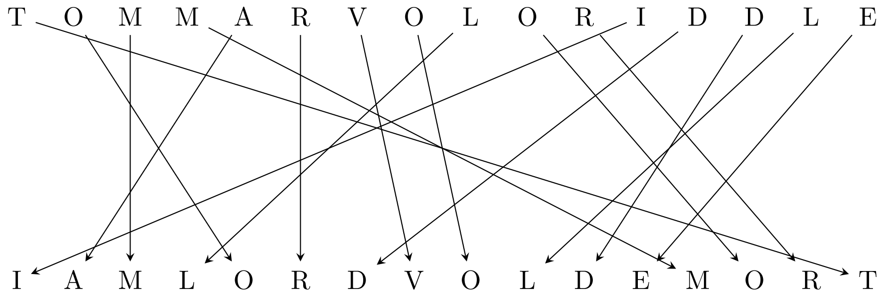
$$\binom{n}{k} = \sum_{j=0}^{k} \binom{n-m}{j}\binom{m}{k-j}?$$

For example, with $n = 7$, $k = 3$, and $m = 3$,

$$\binom{7}{3} = \binom{4}{0}\binom{3}{3}+\binom{4}{1}\binom{3}{2}+\binom{4}{2}\binom{3}{1}+\binom{4}{3}\binom{3}{0} = 1{\cdot}1+4{\cdot}3+6{\cdot}3+4{\cdot}1 = 35.$$

Anagrams, or multiset permutations

ANYONE WHO HAS READ ALL THE HARRY POTTER BOOKS knows that Voldemort's real name is "Tom Marvolo Riddle" and at a crucial point in the story we learn that these letters can be rearranged to spell "I am Lord Voldemort":



Of course, there were many other options available to young Mr. Riddle. How might things have been different if instead of "LORD VOLDEMORT" he became known as

LORD MOLD VOTER

or perhaps

OLD MR LOVE TROD?

The different arrangements of a given set of letters are called "anagrams" and usually we only care about those rearrangements that have meaning as words in English. However, as a mathematical notion, an anagram can be any ordering of the set of letters. For example, a quite uninteresting anagram of TOM MARVOLO RIDDLE is

ADDEILLMMOOORRTV.

Distinguishing "meaningful" anagrams from arbitrary rearrangements like this one is an interesting problem, but outside the scope of this book. However, finding the number of all rearrangements is well within our grasp. How many anagrams did Tom Riddle have to choose from? If Tom had been named Tim, would he have had more or fewer anagrams of his name?

Before we answer these questions, let's introduce some new mathematical terminology.

Multisets

Sets, by definition, have no repeated elements. By contrast, words often have repeated letters. The notion of a "multiset" takes multiplicity into account. Strictly speaking, a multiset is a set of pairs (x, m_x), where x is an element and m_x is a positive integer representing the "multiplicity" of the element x. For example, the set of letters in the word "OKLAHOMA" is $\{A, H, K, L, M, O\}$, whereas its multiset of letters is $\{(A, 2), (H, 1), (K, 1), (L, 1), (M, 1), (O, 2)\}$. When there is no worry of confusion we use notation similar to usual set notation, e.g., $\{A, A, H, K, L, M, O, O\}$.

For another example, we might have reason to consider the standard set of 52 playing cards as a multiset in two different ways. We can ignore suits to get a multiset with thirteen different elements, each with multipicity four:

$$\{(A, 4), (2, 4), (3, 4), \ldots, (10, 4), (J, 4), (Q, 4), (K, 4)\}.$$

Alternatively, we can ignore ranks to get a multiset with four different elements, each with multiplicity thirteen:

$$\{(\heartsuit, 13), (\clubsuit, 13), (\diamondsuit, 13), (\spadesuit, 13)\}.$$

Multiset permutations

The question of "how many anagrams?" now becomes the question of "how many permutations of a multiset?" Let's continue with the example of "OKLAHOMA." There are eight letters in the word, and so there can be at most 8! anagrams. But immediately we see this is at least twice as many as we need, since any permutation that swaps the two letters O has no effect on the anagram. Similarly, swapping the letters A has no effect. Perhaps there are 8!/4 anagrams?

Here is one way to answer the question clearly. Let $\mathcal{A}$ denote the set of anagrams of the multiset $S = \{A, A, H, K, L, M, O, O\}$. We want to enumerate set $\mathcal{A}$.

Suppose we make each letter distinct by adding subscripts to each of the repeated letters, i.e., $S' = \{A_1, A_2, H, K, L, M, O_1, O_2\}$. We might call S' the "linearization" of S, since it breaks the "ties" between repeated letters. Now, set S' has eight distinct letters and clearly 8! permutations. On the other hand, given any anagram in set $\mathcal{A}$, we can create a permutation of set S' by choosing where to place the subscripts on the letters O and letters A, e.g., see Figure 3.2. Each choice of an element of $\mathcal{A}$ yields four different choices of permutation of set S'. We find $|\mathcal{A}| \cdot 4 = |S'|!$, and so $|\mathcal{A}| = 8!/4$.

In general, suppose S is a multiset $\{(s_1, m_1), \ldots, (s_k, m_k)\}$ and let set S' be the linearization of S obtained by making distinct copies of all repeated

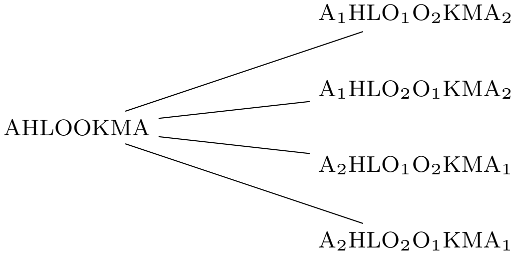


Fig. 3.2 The ways of distinguishing repeated letters.

elements of S. Let $\mathcal{A} = \mathcal{A}(S)$ denote the set of anagrams of S. For ease of notation, we will denote the cardinality of S' by $N = m_1 + \cdots + m_k = |S'|$.

For each anagram in $\mathcal{A}$, we can order the copies of the letter s_1 in $m_1!$ ways, we can order the copies of the letter s_2 in $m_2!$ ways, and so on. Thus

$$N! = |\mathcal{A}| \cdot m_1! m_2! \cdots m_k!.$$

Solving for $|\mathcal{A}|$ gives this nice result:

$$|\mathcal{A}| = \frac{N!}{m_1! \cdots m_k!}.$$

If we apply this formula to the "Tom Marvolo Riddle" multiset

$$\{(A,1),(D,2),(E,1),(I,1),(L,2),(M,2),(O,3),(R,2),(T,1),(V,1)\},$$

we find that there are

$$\frac{16!}{3!2!2!2!2!} = 217,945,728,000$$

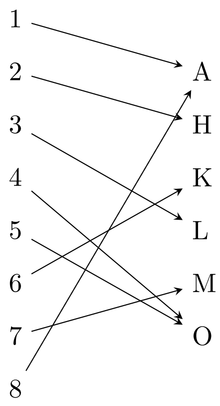
or about 217 billion anagrams in all.

If the character was named "*Tim* Marvolo Riddle" then there would be over 326 billion anagrams:

$$\frac{16!}{2!2!2!2!2!2!} = 326,918,592,000.$$

Multinomial coefficients

There is another way to think about counting anagrams. Let's return to the "OKLAHOMA" example. Any anagram corresponds to a surjection from the set of positions of letters, $\{1,2,3,4,5,6,7,8\}$, onto the set of letters, $\{A,H,K,L,M,O\}$, such that two positions map to letter O and two positions map to letter A. For example, “AHLOOKMA” corresponds to this function:



How many such functions are there?

We can count these through a six-step process of choosing which arrows point to the letters A, H, K, and so on:

- **Step 1:** choose which two arrows point to A.
- **Step 2:** choose which arrow points to H.
- **Step 3:** choose which arrow points to K.
- **Step 4:** choose which arrow points to L.
- **Step 5:** choose which arrow points to M.
- **Step 6:** choose which two arrows point to O.

In step 1, there are $\binom{8}{2}$ ways to choose which arrows point to A. Having completed step 1, there $\binom{6}{1}$ ways to choose an arrow that points to H. After making the choices in steps 1 and 2, there are $\binom{5}{1}$ ways to decide which arrow points to K. Continuing this line of reasoning for the remaining letters, we find the total number of such functions is:

$$\binom{8}{2}\binom{6}{1}\binom{5}{1}\binom{4}{1}\binom{3}{1}\binom{2}{2}.$$

Thus we can use our factorial formula for binomial coefficients to express the number of ways to write an anagram of OKLAHOMA as:

$$\frac{8!}{2!6!}\frac{6!}{1!5!}\frac{5!}{1!4!}\frac{4!}{1!3!}\frac{3!}{1!2!}\frac{2!}{2!0!},$$

which simplifies to

$$\frac{8!}{2!2!},$$

in agreement with our earlier formula.

In general, this line of reasoning applied to a multiset $S = \{(s_1, m_1), \ldots, (s_k, m_k)\}$, with $N = m_1 + \cdots + m_k$, yields the formula

$$\binom{N}{m_1}\binom{N-m_1}{m_2}\cdots\binom{m_k}{m_k} = \frac{N!}{m_1!\cdots m_k!}.$$

This kind of counting problem is common enough that there is a generalization of the notation for binomial coefficients that is sometimes used. We write

$$\binom{N}{m_1,\ldots,m_k} = \frac{N!}{m_1!\cdots m_k!}$$

to denote the number of permutations of a multiset with multiplicities $m_1,\ldots,m_k$ and $N = m_1 + \cdots + m_k$. These numbers are called "multinomial coefficients" since they generalize binomial coefficients, which are the case of $k = 2$.

For example, the number of anagrams in the OKLAHOMA example is

$$\binom{8}{2,2,1,1,1,1},$$

and the number of anagrams in the TOM MARVOLO RIDDLE example is

$$\binom{16}{3,2,2,2,2,1,1,1,1,1}.$$

When we read this aloud we say "sixteen choose three, two, two, two, two, one, one, one, one, one."

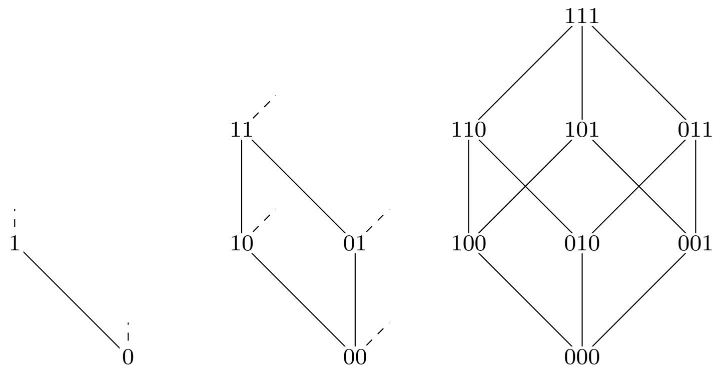
Further reading

- Howard Bergerson, "Palindromes and Anagrams," Dover (1973).
  This is just for fun! It is a collection of interesting palindromes and anagrams. There are hundreds of word puzzle books out there that are worth having fun with too.

Chapter 4
The Binomial Theorem

"Mathematics is the cheapest science. Unlike physics or chemistry, it does not require any expensive equipment. All one needs for mathematics is a pencil and paper."

–George Pólya



How is the structure of a cube like expanding the polynomial $(1+t)^3$?

YOU MAY HAVE WONDERED why the numbers $\binom{n}{k}$ are called binomial coefficients, and not the "choice numbers" or "combination numbers" or something related to subsets. Why "binomial"? We'll see why in this chapter, which begins a theme for us: encoding combinatorial results algebraically.[1]

The main result you will prove in this section is the BINOMIAL THEOREM. Then you will use this theorem to deduce several results for binomial coefficients in a way that is relatively sweat-free. Some of these results were already found in the previous chapter, while some would probably be much harder to guess and prove without the Binomial Theorem.

[1] Some might make the case that it's the other way around: that combinatorics exists to capture algebraic information. This is also true, but harder to appreciate until one has seen a good deal of abstract algebra.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_4

Binomial Theorem

Theorem 6. If $n \geq 0$, then

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

For example, we can expand $(a+b)^2 = a^2+2ab+b^2$, and row 2 of Pascal's triangle has $\binom{2}{0} = 1$, $\binom{2}{1} = 2$, and $\binom{2}{2} = 1$.

The proof of the Binomial Theorem can be done by induction using Pascal's recurrence. You will prove the theorem in the next few problems, but for the moment try these warmup exercises to familiarize yourself with the idea.

Warmup 18. Expand $(a+b)^3$ and $(a+b)^4$ and compare with rows 3 and 4 of Pascal's triangle.

Warmup 19. Carefully expand the product

$$(a_1+tb_1)(a_2+tb_2)(a_3+tb_3),$$

and group the terms according to powers of t, i.e., how many letters "b" appear in each monomial. For each monomial, write below it the set of subscripts you see on the "b" letters. Compare with the subsets of the set $\{1,2,3\}$.

There is a straightforward generalization of the Binomial Theorem to the "Trinomial Theorem" as follows:

$$(a+b+c)^n = \sum_{i+j+k=n} \binom{n}{i,j,k} a^i b^j c^k,$$

or indeed to the "Multinomial Theorem" for any $r \geq 2$

$$(a_1+\cdots+a_r)^n = \sum_{m_1+\cdots+m_r=n} \binom{n}{m_1,\ldots,m_r} a_1^{m_1}\cdots a_r^{m_r},$$

where

$$\binom{n}{m_1,\ldots,m_k} = \frac{n!}{m_1!\cdots m_k!}$$

is the multinomial coefficient discussed at the end of Chapter 3, but we will keep our focus on the Binomial Theorem in this chapter.

Problem 46.
Prove the Binomial Theorem by induction. Hint: suppose

$$(a+b)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} a^k b^{n-1-k}$$

and carefully expand $(a+b)^n = (a+b) \cdot (a+b)^{n-1}$.

Problem 47.
The steps in this problem will prove the Binomial Theorem without induction. Consider the following family of polynomials defined for any $n \geq 1$:

$$f_n(t) = (a_1 + tb_1)(a_2 + tb_2) \cdots (a_n + tb_n).$$

Hint: In Warmup 19, you looked at f_3. Maybe you should also look at $f_1(t)$, $f_2(t)$, and $f_4(t)$ before you try to answer the questions below.

1. What is the coefficient of t in $f_n(t)$?
2. What is the coefficient of t^2 in $f_n(t)$? Try to describe this coefficient by way of choosing terms in the product. (The mnemonic "a" for "absent", "b" for "be there" might help.)
3. Try to describe the coefficient of t^5 in $f_n(t)$ by way of choosing terms in the product.
4. Describe a bijection between the terms in the coefficient of t^k in $f_n(t)$ and the set of all k-subsets of $\{1, 2, \ldots, n\}$.
5. Specialize the variables $t, a_1, b_1, a_2, b_2, \ldots$ to conclude the Binomial Theorem.

Problem 48.
In Problem 40 you proved a formula for the row sum

$$\sum_{k=0}^{n} \binom{n}{k}.$$

Derive this formula as a corollary of the Binomial Theorem.

Problem 49.
Using the Binomial Theorem, find and prove a formula for

$$\sum_{k=0}^{n} 2^k \binom{n}{k}.$$

Problem 50.
Using the Binomial Theorem, find and prove a formula for

$$\sum_{k=0}^{n} (-2)^k \binom{n}{k}$$

Problem 51.

If we set $a = 1$ and $b = t$ in the Binomial Theorem, we get $(1 + t)^n = \sum_{k=0}^{n} \binom{n}{k} t^k$.

Differentiate both sides of this identity with respect to the variable t to get a new identity, and use it to help compute the average cardinality of a subset of $\{1, 2, \ldots, n\}$, i.e., use it to compute

$$\frac{1}{2^n}\left(\sum_{S\subseteq\{1,2,\ldots,n\}} |S|\right).$$

Problem 52.

Using the Binomial Theorem, prove that the number of subsets of n with an odd number of elements equals the number of subsets with an even number of elements, i.e.,

$$\sum_{k\geq 0}\binom{n}{2k} = \sum_{k\geq 0}\binom{n}{2k+1},$$

or

$$\binom{n}{0}+\binom{n}{2}+\binom{n}{4}+\cdots=\binom{n}{1}+\binom{n}{3}+\binom{n}{5}+\cdots.$$

Hint: this identity is equivalent to the identity in Problem 41.

Problem 53.

Using the Binomial Theorem, and the fact that $(1+t)^n = (1+t)^{n-m}(1+t)^m$ for any $0 \leq m \leq n$, show that

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

Your primary technique here should be careful algebraic bookkeeping. Compare this way of proving the identity with the method you used in Problem 45.

Counting permutations according to cycles

THE POLYNOMIAL FROM THE BINOMIAL THEOREM can be thought of as a handy algebraic device for counting subsets of $\{1, 2, \ldots, n\}$ according to cardinality, i.e.,

$$\sum_{k=0}^{n} \binom{n}{k} t^k = \sum_{S \subseteq \{1,2,\ldots,n\}} t^{|S|}.$$

The fact that this polynomial can also be expressed as $(1+t)^n$ is just a happy coincidence from this point of view—a compact encoding of data.

Something similar occurs in a natural counting problem related to permutations. We can replace the idea of

> "counting subsets according to cardinality"

with

> "counting permutations according to number of cycles."

We discussed a bit about cycle structure for permutations at the end of Chapter 2, and now we will see how many permutations have a given number of cycles.

Cycle notation revisited

How do we generate all permutations in cycle notation? We could first generate our permutations in one-line notation (in lexicographic order, say) and then convert each of these into cycle notation. This works, but we can also work directly with cycle notation. Recall that standard cycle notation has the cycles written so that the smallest element of the cycle appears first, and we order the cycles according to which cycle has the smallest element. For example, we write $w = (135)(24)(6)(78)$, not $(351)(24)(6)(78)$ or $(6)(24)(135)(78)$.

Now, given a permutation w in S_{n-1} written in cycle notation, we can form n different permutations in S_n by considering the insertion of "n" into w: either n is inserted somewhere in an existing cycle of w, or not. There are $n-1$ ways to insert n into an existing cycle—for each element $i < n$, we can insert "n" immediately following "i" in the cycle containing i. If n is not added to an existing cycle, then we insert n into a new singleton cycle that is placed at the end of w.

For example, $w = (135)(24)(6)(78)$ gives rise to the following nine permutations:

$$\begin{array}{lll}
(135)(24)(6)(78)(9) & (135)(24)(6)(789) & (135)(24)(6)(798) \\
(135)(24)(69)(78) & (135)(249)(6)(78) & (135)(294)(6)(78) \\
(1359)(24)(6)(78) & (1395)(24)(6)(78) & (1935)(24)(6)(78)
\end{array}$$

This recursive structure can be illustrated with a decision tree as in Figure 4.1. We can compare this method of generating permutations with the one appearing at the end of Chapter 1.

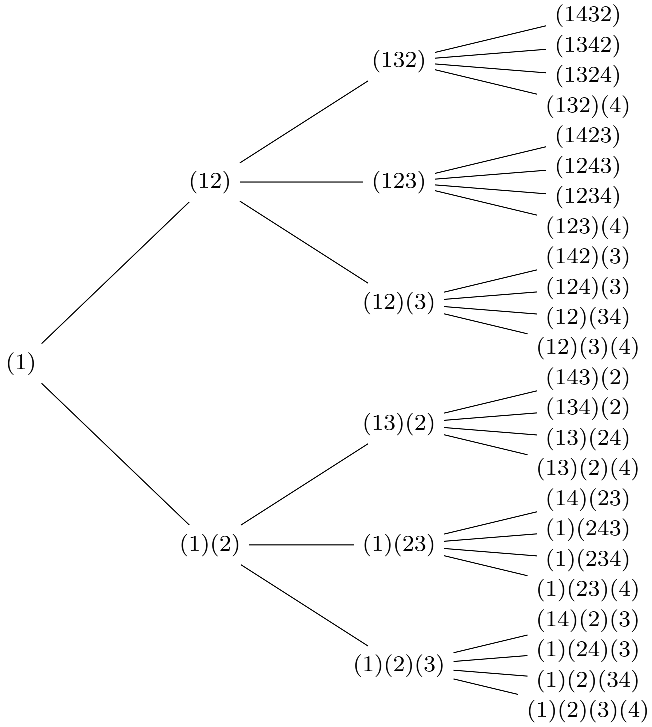


Fig. 4.1 The decision tree for permutations in cycle notation.

Stirling numbers of the first kind

We are now going to try to organize permutations according to the number of cycles. For a permutation $w \in S_n$, let us denote by $\mathrm{cyc}(w)$ the number of cycles in the permutation. For example the permutation $w = (135)(2)(6)(8)$ has four cycles, i.e., $\mathrm{cyc}(w) = 4$. In Table 4.1 we see all the permutations in S_1, S_2, S_3, S_4 grouped according to the number of cycles.

The *Stirling numbers of the first kind*, denoted $S(n,k)$, count the number of permutations in S_n with k cycles. That is,

$$S(n,k) = |\{w \in S_n : \mathrm{cyc}(w) = k\}|.$$

Group	$\mathrm{cyc}(w) = 1$	$\mathrm{cyc}(w) = 2$	$\mathrm{cyc}(w) = 3$	$\mathrm{cyc}(w) = 4$
S_1:	(1)			
S_2:	(12)	(1)(2)		
S_3:	(123) (132)	(1)(23) (12)(3) (13)(2)	(1)(2)(3)	
S_4:	(1234) (1243) (1324) (1342) (1423) (1432)	(1)(234) (1)(243) (12)(34) (13)(24) (14)(23) (123)(4) (132)(4) (124)(3) (142)(3) (134)(2) (143)(2)	(1)(2)(34) (1)(23)(4) (1)(24)(3) (12)(3)(4) (13)(2)(4) (14)(2)(3)	(1)(2)(3)(4)

Table 4.1 Permutations in cycle notation, grouped according to number of cycles.

$n\backslash k$	1	2	3	4	5	6	7
1	1						
2	1	1					
3	2	3	1				
4	6	11	6	1			
5	24	50	35	10	1		
6	120	274	225	85	15	1	
7	720	1764	1624	735	175	21	1

Table 4.2 Triangle of the numbers $S(n, k)$, or the number of permutations in S_n with k cycles.

The triangle of Stirling numbers for $n \leq 7$ is shown in Table 4.2.

After some thought, it is easy to see that

$$S(n, 1) = (n-1)!, \quad S(n, n) = 1,$$

and $S(n, k) = 0$ if $k < 1$ or $k > n$.

Moreover, the Stirling numbers satisfy a two-term recurrence relation

$$S(n, k) = (n - 1)S(n - 1, k) + S(n - 1, k - 1). \tag{4.1}$$

This follows immediately if we consider our recursive procedure for generating permutations illustrated in Figure 4.1. Each permutation in S_n with k cycles either comes from a permutation in S_{n-1} with $k - 1$ cycles and (n) is a singleton cycle, or it comes from some permutation in S_{n-1} that already has k cycles. For any of these permutations that already have k cycles, we have $n - 1$ options for where to place n within an existing cycle.

The Rising Factorial Theorem

Now, given Equation (4.1), we can prove a simple formula analogous to the Binomial Theorem.

Let

$$(t)^{(n)} = t(t + 1)(t + 2) \cdots (t + (n - 1)) = \prod_{i=0}^{n-1} (t + i).$$

This polynomial is sometimes called the "rising factorial."

The recurrence (4.1) implies (via induction) that

$$(t)^{(n)} = \sum_{w \in S_n} t^{\mathrm{cyc}(w)} = \sum_{k=1}^{n} S(n, k)t^k. \tag{4.2}$$

Note the second equality follows by the definition of the Stirling numbers of the first kind. By analogy with the Binomial Theorem we might call Equation (4.2) the Rising Factorial Theorem.

As a cool consequence of this result, we can give a slick answer to the question:

What is the expected number of cycles in a permutation in S_n?

Our approach here mirrors that of Problem 51.

First, note that the expected value of the number of cycles is by definition

$$\frac{\sum_{w \in S_n} \mathrm{cyc}(w)}{|S_n|}.$$

We know that $S(n, k)$ is number of permutations with $\mathrm{cyc}(w) = k$, so

$$\sum_{w \in S_n} \mathrm{cyc}(w) = \sum_{k=1}^{n} kS(n, k).$$

But simply using the power rule for derivatives we find

$$\frac{d}{dt}\left[\sum_{k=1}^{n} S(n,k)t^k\right] = \sum_{k=1}^{n} kS(n,k)t^{k-1},$$

and setting $t = 1$ gives us the sum that we want. For example, if $n = 4$, we have can see from Table 4.1 that $\sum_{w\in S_4} \operatorname{cyc}(w) = 50$, which agrees with

$$\frac{d}{dt}\left[6t + 11t^2 + 6t^3 + t^4\right]_{t=1} = \left[6\cdot 1 + 11\cdot 2t + 6\cdot 3t^2 + 4t^3\right]_{t=1} = 50.$$

There are 24 permutations in S_4, so we conclude the expected number of cycles is $50/24 = 25/12$.

As $|S_n| = n!$, we compute the expectation in general as

$$\frac{\sum_{w\in S_n} \operatorname{cyc}(w)}{|S_n|} = \frac{1}{n!}\frac{d}{dt}\left[\sum_{k=1}^{n} S(n,k)t^k\right]_{t=1}.$$

This is great, but in fact we can do better since the polynomial in question is $\sum_{k=1}^{n} S(n,k)t^k = (t)^{(n)}$. Now from the definition of $(t)^{(n)}$ and the product rule for derivatives, we have the following expression:

$$\begin{aligned}\frac{d}{dt}\left[\sum_{k=1}^{n} S(n,k)t^k\right] &= \frac{d}{dt}\left[(t)^{(n)}\right]\\ &= \sum_{i=0}^{n-1} \frac{\prod_{j=0}^{n-1}(t+j)}{(t+i)}.\end{aligned}$$

For example, when $n = 4$, the derivative of $(t)^{(4)} = t(t+1)(t+2)(t+3)$ is

$$(t+1)(t+2)(t+3) + t(t+2)(t+3) + t(t+1)(t+3) + t(t+1)(t+2).$$

Multiplying by $1/n!$ and setting $t = 1$ in this expression yields

$$\frac{1}{n!}\sum_{i=0}^{n-1} \frac{n!}{(1+i)} = \sum_{k=1}^{n} \frac{1}{k} = H_n,$$

where H_n denotes the nth harmonic number!

Looking back, we see that when $n = 4$, the expectation was $25/12 = 1+1/2+1/3+1/4$. Since the harmonic number grows logarithmically, we get a rough estimate that a random permutation in S_n should have about $\ln(n)$ cycles.

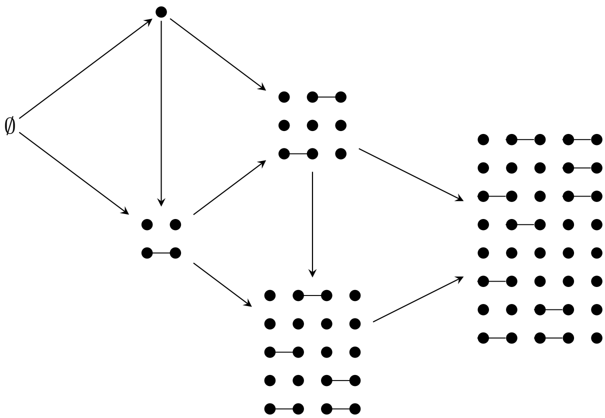
Further reading

- Miklós Bóna, "Combinatorics of Permutations, 2nd. Ed." CRC Press (2012).
  This book is a good introduction to many topics in permutation enumeration, including some discussed in this book. Chapter 3 is all about cycle structure for permutations.

Chapter 5
Recurrences

"I have had my results for a long time: but I do not yet know how I am to arrive at them."

–Carl Friedrich Gauss



How does this diagram grow from left to right?

THE SEQUENCE OF FIBONACCI NUMBERS,

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

is one of the most famous sequences in all of mathematics. It is defined by the initial values $f_0 = 1$, $f_1 = 1$, and the identity $f_n = f_{n-1} + f_{n-2}$ that holds for all $n \geq 2$.

Other sequences satisfy similar identities, called RECURRENCE RELATIONS. In this chapter we will see several different recurrence relations arising from enumeration problems. The Pascal identity for binomial coefficients is another sort of recurrence relation. It is a two-dimensional recurrence since it generates a two-dimensional array of numbers rather than a one-dimensional line of numbers like a sequence.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_5

Problem 54.
Generate the first few terms for the sequences defined by the following recurrences. Give an explicit formula for the terms of the sequence if possible.

1. $a_0 = 1$, $a_n = 2a_{n-1}$ for $n \geq 1$.
2. $a_0 = 1$, $a_n = na_{n-1}$ for $n \geq 1$.
3. $a_0 = 0$, $a_n = a_{n-1} + n$ for $n \geq 1$.
4. $a_0 = 1$, $a_n = \sum_{i=0}^{n-1} a_i a_{n-1-i}$ for $n \geq 1$.
5. $a_0 = 0$, $a_n = a_{n-1} + \sum_{i=0}^{n} (i+n)$ for $n \geq 1$.
6. $a_0 = 1$, $a_n = \sum_{k=0}^{n-1} \binom{n-1}{k} a_k$ for $n \geq 1$.

Compositions

Definition 12. A *composition* of n is an ordered list of positive integers whose sum is n, i.e., $\alpha = (\alpha_1, \ldots, \alpha_k)$ is a composition of n if and only if all $\alpha_i > 0$ are positive integers and $\alpha_1 + \cdots + \alpha_k = n$.

Notice that $(2,2,1)$ and $(1,2,2)$ are two *different* compositions of 5. (If we ignore the ordering on the list we get "partitions" rather than compositions. This is the subject of Chapter 12.)

A composition made up of k positive integers, e.g., $\alpha = (\alpha_1, \ldots, \alpha_k)$ is said to have k *parts*. For example, $(3,2)$ is a composition with two parts, while $(3,1,1)$ has three parts.

Warmup 20. List all the compositions of 3, then list all compositions of 4. Group the compositions in your lists by the number of parts.

Problem 55.
This problem explores some properties of compositions.

1. Let c_n denote the number of compositions of n. Express c_n in terms of c_{n-1}.
2. Write a formula for c_n in terms of n.
3. How many compositions of n have k parts? Can you describe a bijection between subsets and compositions?

Problem 56.
How many ways can you write n as a sum of ones and twos, i.e., how many compositions of n have $\alpha_i \in \{1, 2\}$?

Problem 57.
How many ways can you write n as a sum of odd numbers, i.e., how many compositions of n have $\alpha_i \in \{1, 3, 5, 7, 9, \ldots\}$?

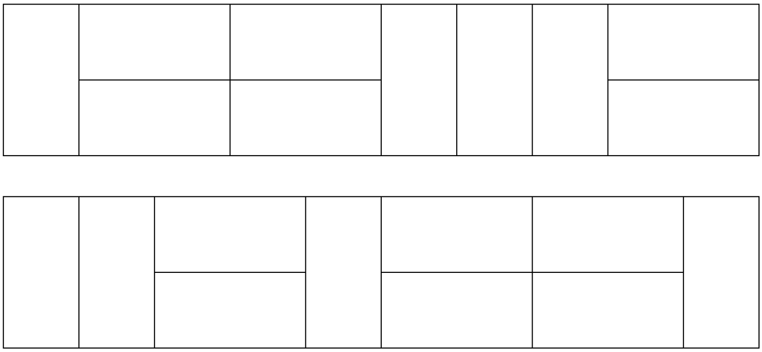


Fig. 5.1 Two domino tilings of a 2×10 rectangle.

Problem 58 (Domino tilings).
A *domino tiling* is a way to cover a rectangle with 1×2 or 2×1 rectangles so that the rectangles cover the larger rectangle with no overlapping and no hanging over the edges.

How many domino tilings of a 2×10 rectangle are there? We see two such tilings in Figure 5.1.

What about a $2 \times n$ rectangle?

Problem 59 (Handshake problem).
When the math club executive board meets, each of the members shakes hands with every other board member exactly once.

1. If there are three members, how many handshakes occur?
2. If a fourth person joins the board, how many more handshakes occur?
3. Let T_n denote the number of handshakes if the executive board has n members. Express T_n in terms of T_{n-1}.
4. Write a formula for T_n in terms of n. (The numbers T_n are called the *triangular numbers*. Can you think why?)
5. Can you relate the handshake problem to the complete graph K_n defined in Problem 8?

Problem 60.
Suppose n lines are drawn so that no two lines are parallel and no three lines intersect at any one point. (Such a set of lines is called *generic*, or is said to be in *general position*.) Into how many regions is the plane divided by n such lines?

How many of the regions are unbounded? bounded? For example, with $n = 3$ lines there are six unbounded regions and one bounded region. Try to use a recurrence to help guide your answer.

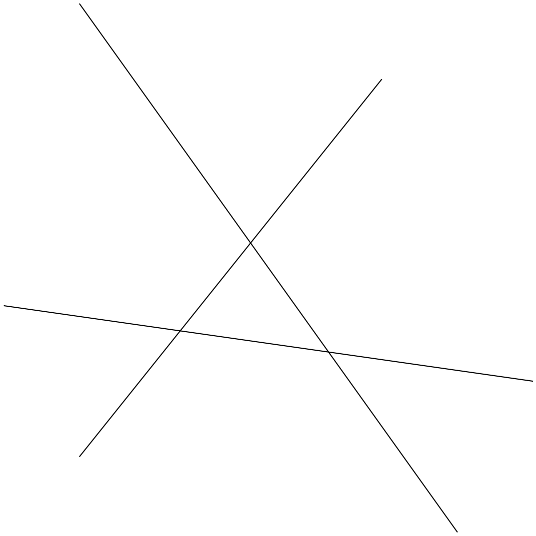


Fig. 5.2 Three lines in general position.

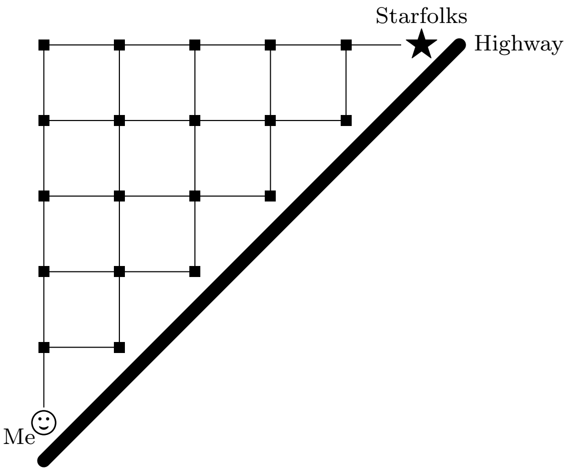


Fig. 5.3 Going for coffee one more time.

Problem 61 (Starfolks, III).
Okay, now I'm in a corner of town that abuts the highway, so my grid of streets is incomplete, as seen in Figure 5.3. Starfolks is five blocks East and five blocks North of me, but the highway runs on a straight diagonal line between me and my coffee.

Without walking along (or crossing) the highway, how many ways are there to get to Starfolks?

What if Starfolks was n blocks East and North of me? (Hint: try to show the recurrence from Problem 54, part 4, holds.)

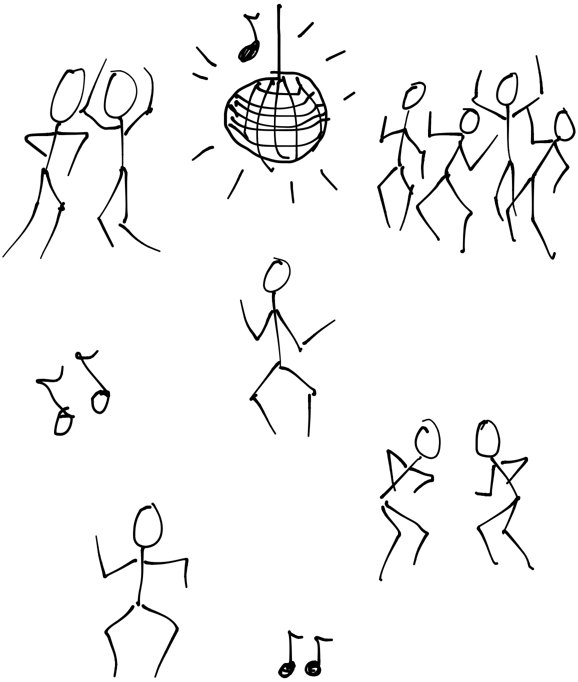


Fig. 5.4 A set partition of ten people.

Problem 62.

There are ten people at a party. As will happen, some smaller groups start to form, possibly with some people standing quietly by themselves. Perhaps you look around the room and see a group of four, two pairs of people, and two different people standing all by themselves, as in Figure 5.4.

The mathematical term for this splitting into groups is a *set partition*, i.e., a way to break up a set into a collection of nonoverlapping subsets. This can range from having each element in a set by itself, to having all the elements together in a single group.

How many ways can these ten people be split into smaller groups? (Hint: try to show that for a party of n people, the number of set partitions obeys the recurrence in part 6 of Problem 54.)

Lucas numbers and polynomials

THE FIBONACCI NUMBERS have attracted a lot of popular attention. People seem to find them everywhere! There could be many reasons why this is. One reason might be because the recurrence relation they satisfy,

$$f_n = f_{n-1} + f_{n-2},$$

is so darn simple. Add two terms together. A grade-school kid can do it. Hmmm. "Add two terms together" also describes Pascal's recurrence—is that why binomial coefficients are ubiquitous too?

Before starting the next chapter, we will explore a sequence of numbers called the "Lucas numbers" that are close cousins of the Fibonacci numbers. We will also generalize both sequences of numbers to sequences of polynomials.

The golden ratio

One of the things that novelists and movie producers like to drop into their stories to sound smart is "the golden ratio" often denoted by φ. The definition of the golden ratio is as the limit of ratios of consecutive Fibonacci numbers:

$$\varphi = \lim_{n\to\infty} \frac{f_n}{f_{n-1}}.$$

The sequence of ratios begins:

$$\frac{1}{1} = 1, \frac{2}{1} = 2, \frac{3}{2} = 1.5, \frac{5}{3} \approx 1.666, \frac{8}{5} = 1.6, \frac{13}{8} = 1.625, \frac{21}{13} \approx 1.615,$$

and, assuming the limit exists, it is not too hard to compute.

Using the recurrence, we have

$$\begin{aligned}
\varphi &= \lim_{n\to\infty} \frac{f_n}{f_{n-1}}, \\
&= \lim_{n\to\infty} \frac{f_{n-1} + f_{n-2}}{f_{n-1}}, \\
&= 1 + \lim_{n\to\infty} \frac{f_{n-2}}{f_{n-1}}, \\
&= 1 + \frac{1}{\varphi}.
\end{aligned}$$

Rewriting, we find

$$\varphi^2 - \varphi - 1 = 0,$$

and solving for φ, we have

$$\varphi = \frac{1+\sqrt{5}}{2} = 1.6180339887\ldots.$$

We know the answer is the larger root since the smaller root, $(1-\sqrt{5})/2$, is a negative number.

Lucas numbers

The sequence of *Lucas numbers* is similar to the Fibonacci sequence. It begins

$$1, 3, 4, 7, 11, 18, 29, 47, \ldots$$

and its recursive definition is nearly identical to the one for Fibonacci numbers. The only difference is in the starting values. That is, we define $L_0 = 1$, $L_1 = 3$, and if $n \geq 2$, $L_n = L_{n-1} + L_{n-2}$.

Since the limit of ratios of consecutive Fibonacci numbers gave us the golden ratio, we might wonder what the limit of ratios of consecutive Lucas numbers looks like. Experimentally, we find

$$\frac{3}{1} = 3, \frac{4}{3} \approx 1.333, \frac{7}{4} = 1.75, \frac{11}{7} \approx 1.571, \frac{18}{11} \approx 1.636, \frac{29}{18} \approx 1.611, \ldots.$$

This looks pretty close to the golden ratio. What is going on?

If we look back at the derivation of the golden ratio φ above, we see that all we used was the recurrence relation for the Fibonacci numbers. Since the Lucas numbers satisfy the same recurrence, we will have $\lim_{n\to\infty} L_{n+1}/L_n = \varphi$ as well.

This makes some sense, because in another context, we would call the limit of ratios of consecutive values the "growth rate" of the sequence. If two sequences are generated by the same recurrence relation, it seems clear that they should have the same growth rate.

Generalized Lucas numbers

The Fibonacci numbers and the Lucas numbers are but two sequences in an infinite family of sequences that all have φ as their growth rate. Define the *generalized Lucas numbers*, or the (a,b)-*Lucas numbers* to be the sequence that begins $L_0 = a$ and $L_1 = b$, and satisfying $L_n = L_{n-1} + L_{n-2}$. The classical Fibonacci numbers are the $(1,1)$-Lucas numbers, and the Lucas numbers of the previous section are the $(1,3)$-Lucas numbers. Interestingly, any (a,b)-

Lucas number is a linear combination of classical Fibonacci numbers. That is, the (a,b)-Lucas sequence begins

$$a, b, a+b, a+2b, 2a+3b, 3a+5b, 5a+8b, \ldots, f_{n-2}a + f_{n-1}b, \ldots.$$

It can be fun to make up your own favorite sequence by choosing the values of a and b to suit your fancy, e.g., you can choose a to be a month (1, 2, ..., 12) and b to be a date (from the first to the last of the month). I happen to be fond of the "April fools" sequence:

$$4, 1, 5, 6, 11, 17, 28, 45, 73, 118, 191, \ldots.$$

Matching polynomials

There is another interesting generalization of Fibonacci and Lucas numbers that comes from graph theory.

Define a *matching* of a graph to be a subgraph in which each vertex is connected to at most one other vertex. The graph with zero edges is always considered a matching, the "empty" matching. For example, the complete graph on five vertices has twenty-six matchings: the empty matching and twenty-five matchings with at least one edge, as shown in Figure 5.5.

The *matching polynomial* of a graph G is defined to be the sum

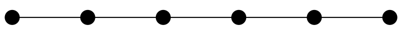
$$M_G(t) = \sum_{\text{matchings } m \text{ of } G} t^{|m|},$$

where $|m|$ denotes the number of edges in the matching. For example, with the complete graph K_5 we have

$$M_{K_5}(t) = 1 + 10t + 15t^2.$$

Fibonacci polynomials and Lucas polynomials

In Figure 5.6 we see the thirteen matchings of the graph



We call a graph that has its nodes connected in this way a *path graph*, denoted P_n, where n is the number of vertices. This example is P_6. A sharp-eyed reader might notice the connection between matchings on P_n and domino tilings of a $2 \times n$ rectangle as described in Problem 58.

From Figure 5.6 we compute the matching polynomial for P_6 to be

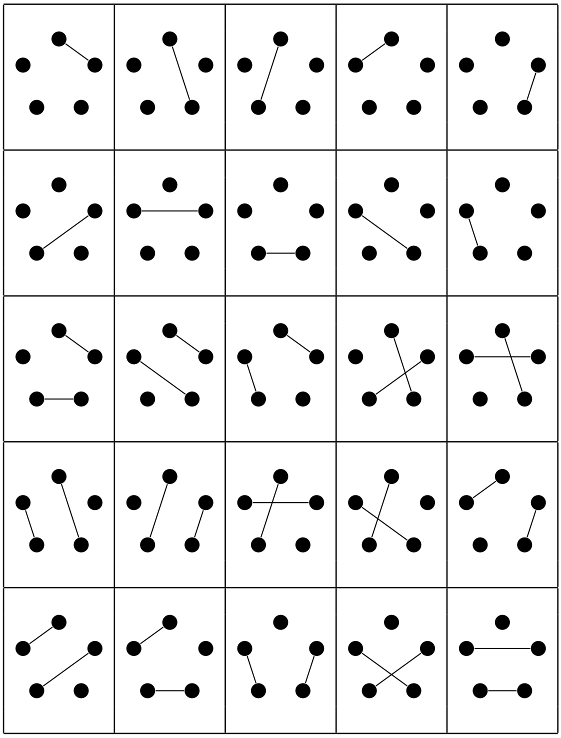


Fig. 5.5 The nonempty matchings of the complete graph K_5.

$$M_{P_6}(t) = 1 + 5t + 6t^2 + t^3.$$

Looking at the smaller cases, we find

$$\begin{aligned} M_{P_1}(t) &= 1, \\ M_{P_2}(t) &= 1 + t, \\ M_{P_3}(t) &= 1 + 2t, \\ M_{P_4}(t) &= 1 + 3t + t^2, \end{aligned}$$

and

$$M_{P_5}(t) = 1 + 4t + 3t^2.$$

Since a matching either has an edge connected to its rightmost vertex or it doesn't, we have

$$M_{P_n}(t) = M_{P_{n-1}}(t) + tM_{P_{n-2}}(t).$$

In pictures this says matchings either look like

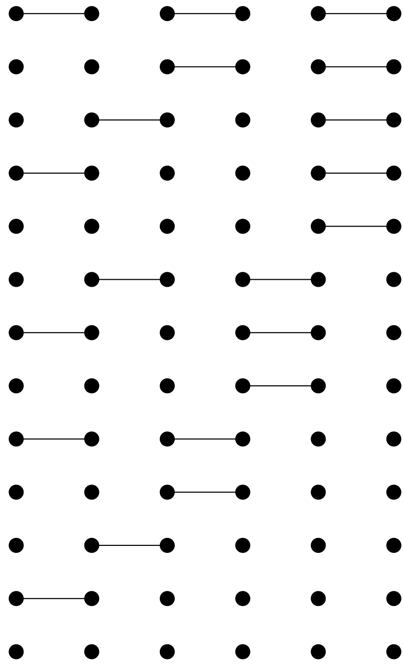
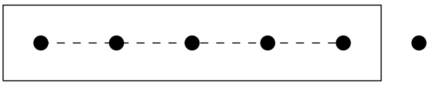
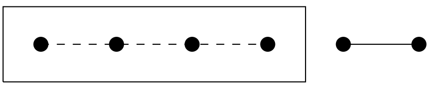


Fig. 5.6 The thirteen matchings of the path P_6.



(matching on P_{n-1})

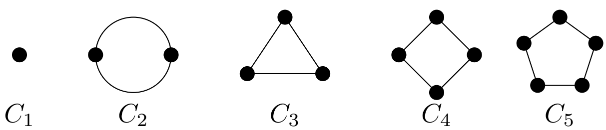
or



(matching on P_{n-2})

Since setting $t = 1$ recovers the Fibonacci sequence, the polynomials $M_{P_n}(t)$ are sometimes known as the *Fibonacci polynomials*.

Similarly, the *Lucas polynomials* are the matching polynomials for the cycle graphs:



We can see that $M_{C_1}(t) = 1$, $M_{C_2}(t) = 1 + 2t$, and (with a little work)

$$M_{C_n}(t) = M_{C_{n-1}}(t) + tM_{C_{n-2}}(t). \tag{5.1}$$

Thus, setting $t = 1$ in this recurrence generates the Lucas numbers.

Proving (5.1) is not quite as obvious as in the case of the path graph, but by considering whether a fixed node is matched to the left, right, or neither, we can prove the following recurrence:

$$M_{C_n}(t) = 2tM_{P_{n-2}}(t) + M_{P_{n-1}}(t),$$

or, considering whether a fixed edge is present or not,

$$M_{C_n}(t) = M_{P_n}(t) + tM_{P_{n-2}}(t).$$

(Can you draw pictures to illustrate these recurrences?)

Applying these recurrence relations, we find:

$$\begin{aligned}
M_{C_{n-1}}(t) + tM_{C_{n-2}}(t) &= M_{P_{n-1}}(t) + tM_{P_{n-3}}(t) + tM_{P_{n-3}}(t) + 2t^2M_{P_{n-4}}(t)\\
&= M_{P_{n-1}}(t) + 2t(M_{P_{n-3}}(t) + M_{P_{n-4}}(t))\\
&= M_{P_{n-1}}(t) + 2tM_{P_{n-2}}(t)\\
&= M_{C_n}(t).
\end{aligned}$$

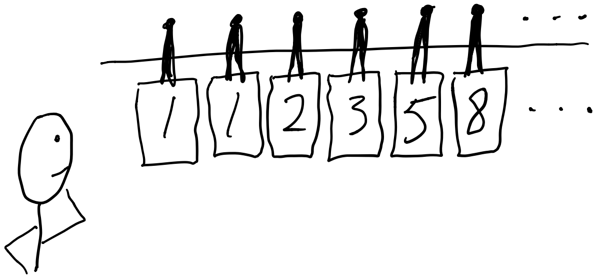
Thus we have established (5.1).

Further reading

- Art Benjamin and Jennifer Quinn, "Proofs that really count: The Art of Combinatorial Proof," Mathematical Association of America (2003).
  This book is a great read and gives a good way to think about bijective style proofs. Fibonacci numbers feature prominently.
- Martin Griffiths, "The golden string, Zeckendorf representations, and the sum of a series," American Mathematical Monthly, **118** (2011), 497–507.
  This article discusses Fibonacci numbers and a result due to Zeckendorf which says any number can be written uniquely as a sum of nonconsecutive Fibonacci numbers. Cool!

Chapter 6
Generating functions

"A generating function is a clothesline on which we hang up a sequence of numbers for display."
–Herb Wilf



A clothesline of numbers.

YOU MAY RECALL from a calculus class the *geometric series*:

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \cdots + z^k + \cdots = \sum_{k \geq 0} z^k.$$

The coefficients in the (MacLaurin) series expansion of a function $F(z)$ define a sequence of numbers. Generally, if $F(z)$ has MacLaurin series

$$\sum_{k \geq 0} a_k z^k = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k + \cdots ,$$

we can relate the function F with the sequence $a_0, a_1, a_2, \ldots$. So the function $1/(1-z)$ encodes the rather boring sequence $1, 1, 1, \ldots$.

What about working the other way around? Given a sequence, is there a function that encodes the sequence? Yes! And this function is called the GENERATING FUNCTION for the sequence. In this chapter we will explore some properties of generating functions. We will practice both (1) how to extract a sequence from a generating function, and (2) how to find the generating function for a sequence. The first task is fairly mechanical, whereas the second can be very tricky (and very interesting!) in general.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_6

Power series generating functions

Definition 13. Given a sequence of numbers $a_0, a_1, a_2, \ldots, a_k, \ldots$, we define its *formal power series* by:

$$F(z) = \sum_{k \geq 0} a_k z^k = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k + \cdots .$$

We also refer to F as the *generating function* for the sequence.

We typically use capital letters like F, G, or A to denote the names of generating functions, while we use lower case letters like q, t, x, y, or z for the arguments. Feel free to choose notation to suit your taste.

Formal power series obey much the same arithmetic operations as polynomials, listed here.

- $c \cdot \sum_{k \geq 0} a_k z^k = \sum_{k \geq 0} (c \cdot a_k) z^k$ for any constant c,
- $\left(\sum_{k \geq 0} a_k z^k\right) + \left(\sum_{l \geq 0} b_l z^l\right) = \sum_{m \geq 0} (a_m + b_m) z^m$,
- $$\left(\sum_{k \geq 0} a_k z^k\right)\left(\sum_{l \geq 0} b_l z^l\right) = \sum_{j \geq 0} \left(\sum_{k+l=j} a_k b_l\right) z^j, = \sum_{j \geq 0} \left(\sum_{k=0}^{j} a_k b_{j-k}\right) z^j.$$

We can also compute derivatives of power series using the power rule:

$$\frac{d}{dz}\left[\sum_{k \geq 0} a_k z^k\right] = \sum_{k \geq 0} k a_k z^{k-1} = \sum_{l \geq 0} (l+1) a_{l+1} z^l,$$

and this derivative obeys all the usual rules of calculus.

Logically, one can take two points of view for these rules for power series manipulation. With analysis (calculus) in mind, we could only wish to consider functions whose series expansions converge absolutely in some neighborhood of the origin, in which case all three properties hold easily. Alternatively, we can take an algebraic approach in which we define the ring of formal power series. Here the second and third properties *define* the ring operations of addition and multiplication. Either way, let's not sweat the details right now.

Note: You don't need an infinite sequence to have a generating function. If you like, finite sequences are just sequences with infinitely many zeroes attached at the end. This is just the same as considering polynomials of finite degree to be special kinds of formal power series.

For example, the Binomial Theorem shows that the generating function for the sequence $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n-1}, \binom{n}{n}$ is the function $(1+z)^n$. We will see other polynomial generating functions in Chapters 8 and 9.

Problem 63.
In this problem, we derive the geometric series identity as a fact about multiplicative inverses in the ring of formal power series. Let $F(z) = \sum_{k\geq 0} z^k$, and let $G(z) = 1 - z$. Show that $F(z)G(z) = 1$.

We conclude that F is the multiplicative inverse of G, which we write as a reciprocal $F(z) = 1/G(z)$. In other words,

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

Problem 64.
Taking the geometric series identity in Problem 63 as a starting point, what sequences are defined by the following generating functions?

1. $\dfrac{1}{1+z}$
2. $\dfrac{1}{1-2z}$
3. $\dfrac{1}{z-3}$
4. $\dfrac{1}{1-z^2}$
5. $\dfrac{1}{(1-z)^2}$

Problem 65.
Multiplying a power series by $1/(1-z)$ has a nice effect on a power series. If $F(z) = \sum_{j\geq 0} a_j z^j$, what sequence has generating function $F(z)/(1-z)$? That is, if

$$\frac{F(z)}{1-z} = \sum_{k\geq 0} b_k z^k,$$

what are the numbers b_k in terms of the a_j?

Problem 66.
Show that

$$\frac{1}{(1-z)^3} = \sum_{n\geq 0} T_{n+1} z^n,$$

where T_n is the nth triangular number discussed in Problem 59.

Problem 67.
With the usual differentiation rules from calculus applied to rational functions, find the sequences defined by the following series, where $n \geq 1$.

1. $\dfrac{d^n}{dz^n}\left[\dfrac{1}{1-z}\right]$
2. $\dfrac{1}{(1-z)^n}$

Can you find these sequences in Tables 2.1 and 3.1?

Problem 68.
What sequence is defined by the following generating function?

$$\frac{1}{1-5z+6z^2}$$

Problem 69.
Suppose α and β are nonzero real numbers. What sequences are defined by the following generating functions?

1. $\dfrac{1}{1-\alpha z}$
2. $\dfrac{1}{1-\beta z}$
3. $\dfrac{1}{(1-\alpha z)(1-\beta z)}$

Now find constants A and B (in terms of α and β) such that

$$\frac{1}{(1-\alpha z)(1-\beta z)} = \frac{A}{(1-\alpha z)} + \frac{B}{(1-\beta z)}$$

and derive another expression for the sequence given in part 3. You may also want to compare with the example in Problem 68.

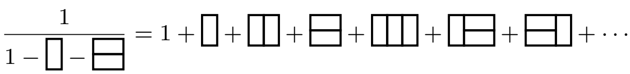


Fig. 6.1 A generating function for domino tilings of a $2\times n$ rectangle (see Problem 58).

Problem 70.
Recall that the Fibonacci numbers are defined by $f_0 = 1, f_1 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$. Find a formula for the generating function for the Fibonacci sequence:

$$F(z) = \sum_{k\geq 0} f_k z^k = 1 + z + 2z^2 + 3z^3 + 5z^4 + 8z^5 + 13z^6 + \cdots$$

A hint is suggested by Figure 6.1.

Problem 71.
Using the formula from Problem 70, along with the results of Problem 69, find a (non-recursive) formula for the nth Fibonacci number.

Problem 72.
We can encode two-dimensional arrays of numbers with bivariate generating functions in many circumstances. Here is one example. Define the generating function $F(t, z)$ by

$$F(t, z) = \sum_{n \geq k \geq 0} a_{n,k} t^k z^n = \frac{1}{1 - (1 + t)z}.$$

What is $a_{n,k}$?

Problem 73.
Use your expressions from Problems 70 and 72 to prove the following Fibonacci identity (which we first glimpsed in Problem 42):

$$f_n = \sum_{k \geq 0} \binom{n-k}{k}.$$

Problem 74.
Let $a_0, a_1, a_2, \ldots$ be the sequence defined by part 4 of Problem 54, i.e., $a_0 = 1$ and $a_n = \sum_{i=0}^{n-1} a_i a_{n-1-i}$ for $n \geq 1$. Find an expression for the generating function of this sequence.

Problem 75.
This problem explores a family of generating functions.

1. Find a formula for the generating function for the sequence of squares:

$$V_2(z) = \sum_{k \geq 0} k^2 z^k = z + 4z^2 + 9z^3 + 16z^4 + \cdots + k^2 z^k + \cdots .$$

2. Find a formula for the generating function for the sequence of cubes:

$$V_3(z) = \sum_{k \geq 0} k^3 z^k = z + 8z^2 + 27z^3 + 64z^4 + \cdots + k^3 z^k + \cdots .$$

3. Let $V_n(z)$ denote the generating function for the sequence of nth powers:

$$V_n(z) = \sum_{k \geq 0} k^n z^k = z + 2^n z^2 + 3^n z^3 + 4^n z^4 + \cdots + k^n z^k + \cdots .$$

   Find a formula for $V_n(z)$ in terms of $V_{n-1}(z)$.

Problem 76.
With $V_n(z)$ as in Problem 75, let $A_n(z) = (1-z)^{n+1}V_n(z)$. Find a formula for $A_n(z)$ in terms of $A_{n-1}(z)$, and use this recurrence (perhaps with the aid of a computer) to make a table of the coefficients of these polynomials (Wait! These are polynomials?!) for $n = 1, 2, \ldots, 8$.

What happens when you set $z = 1$ in $A_n(z)$?

Money problems

HERE IS A QUESTION that any third grader can understand:

> How many ways can you give someone a dollar using only pennies, nickels, dimes, and quarters?

In mathematical terms, how many nonnegative integer solutions are there to the equation

$$1 \cdot a + 5 \cdot b + 10 \cdot c + 25 \cdot d = 100?$$

If we are systematic about it, we can probably get to the answer through careful bookkeeping. How many ways can we do it with zero quarters? With one quarter? two? three?

But that seems inelegant.

A slick answer to the question comes from the generating function approach. Let's first think about using pennies and nickels only.

The generating function for amounts of money made with pennies only is

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \cdots = \sum_{a \geq 0} z^a,$$

and with nickels only the generating function is

$$\frac{1}{1-z^5} = 1 + z^5 + z^{10} + z^{15} + \cdots = \sum_{b \geq 0} z^{5b}.$$

Thus, the product of these two functions will be

$$\frac{1}{(1-z)(1-z^5)} = \sum_{a,b \geq 0} z^{a+5b}.$$

This means the coefficient of z^k will be the number of ways to express k as a nonnegative linear combination of 1 and 5, i.e., the number of ways to write $k = a + 5b$.

Adding in factors for dimes and quarters, we can answer the original question by saying that it is the coefficient of z^{100} in the function

$$F_{\text{coins}}(z) = \frac{1}{(1-z)(1-z^5)(1-z^{10})(1-z^{25})}.$$

Computer software tells us this coefficient happens to be 242. For example, using the computer algebra software Maple I type

```
Fcoins:=1/(1-z)/(1-z^5)/(1-z^(10))/(1-z^(25)):
```

and then

```
series(Fcoins,z,101);
```

to see the output

$$1 + z + z^2 + z^3 + z^4 + 2z^5 + 2z^6 + \cdots + 213z^{99} + 242z^{100} + O(z^{101}).$$

To just get the coefficient alone, I can type

```
coeff(series(Fcoins,z,101), z^(100));
```

and it spits out the number 242.

But what if we have other constraints?

Small purse solutions

Now suppose we have a small change purse, and it can hold no more than twenty coins. Then we need to restrict our solutions.

How many ways can you give someone a dollar using at most twenty coins?

We can get the answer to this question by modifying our generating function to keep track of the total number of coins used. Now, let

$$F_{\text{coins}}(t,z) = \frac{1}{(1-tz)(1-tz^5)(1-tz^{10})(1-tz^{25})}.$$

Then we can see now that each time we use a power of (say) z^{10} from the third term in the denominator we get a power t to indicate which power of z^{10} we used, i.e., how many dimes we are going to take. We get

$$\begin{aligned} F_{\text{coins}}(t,z) &= \sum_{a,b,c,d\geq 0} (tz)^a (tz^5)^b (tz^{10})^c (tz^{25})^d \\ &= \sum_{a,b,c,d\geq 0} t^{a+b+c+d} z^{a+5b+10c+25d}. \end{aligned}$$

This means the coefficient of $t^n z^k$ is the number of ways to choose n coins that add up to k cents. Let $\alpha_{n,k}$ denote this coefficient. This is almost what we want.

What we would really like is $\alpha_{20,100} + \alpha_{19,100} + \cdots + \alpha_{1,100}$. We can either extract each of these from F_{coins} separately and add them up, or we can use the trick of Problem 65 and multiply by $1/(1-t)$ to get

$$\begin{aligned} \frac{F_{\text{coins}}(t,z)}{1-t} &= \sum_{n,k\geq 0} \left(\sum_{0\leq m\leq n} \alpha_{m,k} \right) t^n z^k \\ &= \sum_{n,k\geq 0} \beta_{n,k} t^n z^k, \end{aligned}$$

where $\beta_{n,k}$ is the number of ways to make k cents using at most n coins.

We find the answer to our "small purse" question is the coefficient of $t^{20}z^{100}$ in this generating function. Computer software tells us the answer is 65.

Attainable amounts

If we look at the coefficient of z^k in $F_{\text{coins}}(t,z)$, it is a polynomial in t that records the number of ways to make k cents according to the number of coins used. For example, the coefficient of z^{10} is $t+t^2+t^6+t^{10}$, which reflects the fact that we can make 10 cents with one dime, two nickels, one nickel and five pennies, or ten pennies.

On the other hand, if we look at the coefficient of t^n in $F_{\text{coins}}(t,z)$, it is a polynomial in z that records the different amounts of money that can be made with n coins. For example, the coefficient of t^2 is

$$z^2+z^6+z^{10}+z^{11}+z^{15}+z^{20}+z^{26}+z^{30}+z^{35},$$

which tells us there are nine different amounts that can be made from two coins, and each of these amounts can be made in one way.

Likewise, the coefficient of t^n in $F_{\text{coins}}(t,z)/(1-t)$ is a polynomial in z that records the number of ways to make different amounts of money from at most n coins. The coefficient of t^2 here records the number ways to make different amounts of money using at most two coins. It is

$$1+z+z^2+z^5+z^6+2z^{10}+z^{11}+z^{15}+z^{20}+z^{25}+z^{26}+z^{30}+z^{35}.$$

We see there are now two ways to make ten cents (two nickels or one dime) and many more values are attainable. This leads us to a natural question.

Which values are *not* attainable?

The postage stamp problem

This version of the question is sometimes called the "postage stamp problem" since it was originally phrased in terms of sending a letter with restrictions on the number and type of stamps.

> You can fit at most n stamps on the letter. Given the values of the different stamps, what is the smallest amount of postage you cannot put on the letter?

Suppose for illustration purposes our stamps are like our coins and come in denominations of $1, 5, 10$, and 25. Then $F_{\text{coins}}(t,z)/(1-t)$ contains the answer to our question in some sense. Indeed, the smallest unattainable value is the first power of z *not* to appear in the coefficient of t^n. With $n = 2$ stamps, we see that three cents is the smallest unattainable value.

As a more interesting example, consider $n = 5$ stamps. The coefficient of t^5 from $F_{\text{coins}}(t,z)/(1-t)$ is shown in Figure 6.2. Can you find the smallest missing power of z?

$$
\begin{aligned}
&1 + z + z^2 + z^3 + z^4 + 2z^5 + z^6 + z^7 + z^8 + z^9 + 2z^{10} + 2z^{11} \\
&+2z^{12} + 2z^{13} + z^{14} + 2z^{15} + 2z^{16} + 2z^{17} + z^{18} + 3z^{20} + 3z^{21} \\
&+2z^{22} + z^{23} + 4z^{25} + 3z^{26} + 2z^{27} + z^{28} + z^{29} + 4z^{30} + 3z^{31} \\
&+2z^{32} + z^{33} + 4z^{35} + 3z^{36} + 2z^{37} + z^{38} + 4z^{40} + 3z^{41} + z^{42} \\
&+4z^{45} + 2z^{46} + z^{47} + 4z^{50} + 2z^{51} + z^{52} + z^{53} + 3z^{55} + 2z^{56} \\
&+z^{57} + 3z^{60} + 2z^{61} + z^{62} + 3z^{65} + z^{66} + 2z^{70} + z^{71} + 2z^{75} \\
&+z^{76} + z^{77} + 2z^{80} + z^{81} + 2z^{85} + z^{86} + z^{90} + z^{95} + z^{100} \\
&+z^{101} + z^{105} + z^{110} + z^{125}
\end{aligned}
$$

Fig. 6.2 The coefficient of t^5 in $F_{\text{coins}}(t,z)/(1-t)$.

Frobenius numbers

The postage stamp problem is about unattainable values where the number of coins is constrained. A different kind of problem arises when we allow any number of coins.

Let's state the general problem, sometimes known as the *coin problem* or the *Frobenius problem*.

> You have an unlimited supply of coins in n different (fixed) denominations: $d_1, \ldots, d_n$. What is the largest amount of money that cannot be obtained using these coins?

With enough pennies, any value is attainable, so let's assume all the denominations are greater than one: $d_i > 1$.

Further, if all the denominations have a common divisor, d, greater than 1, then every amount we create will be a multiple of d, so it is more interesting to assume $\gcd(d_1, \ldots, d_n) = 1$.

In this situation, it turns out (though it's not immediately obvious) that every *sufficiently large* amount of money will be attainable. The largest unattainable amount is known as the *Frobenius number* of the set $\{d_1, \ldots, d_n\}$, sometimes denoted $g(d_1, \ldots, d_n)$.

As a very small example, $g(3,5) = 7$. The number 7 is unattainable with threes and fives, since we cannot get 7 with only threes, and using a 5 would leave 2, which is clearly unattainable since it is smaller than 3. On the other hand, we can get $8 = 5 + 3$, $9 = 3 \cdot 3$, and $10 = 2 \cdot 5$, and every number larger than 10 can be obtained by adding sufficiently many threes.

For the case of two denominations, there is a formula for the Frobenius number. If $\gcd(d_1, d_2) = 1$, then

$$g(d_1, d_2) = d_1 d_2 - d_1 - d_2.$$

This formula can be proved with a little bit of elementary number theory. For more than two denominations, however, there is no such formula except in special cases.

Our generating function approach gives one way to find the Frobenius number. Let's consider the case of $d_1 = 6, d_2 = 9$, and $d_3 = 20$ to illustrate. These are the sizes of boxes of Chicken McNuggets from the McDonald's of my youth. (Nowadays the boxes come in sizes 4, 6, 10, and 20, and since there are no odd box sizes, the story is very different.) What is the largest unattainable number of nuggets, i.e., what is $g(6, 9, 20)$? First, we will define the relevant generating McFunction

$$F_{\text{nugget}}(z) = \frac{1}{(1 - z^6)(1 - z^9)(1 - z^{20})} = \sum_{a,b,c \geq 0} z^{6a+9b+20c}.$$

For this question we don't really care about the coefficients in the series, only which coefficients are nonzero. Since 6 is our smallest denomination, once we see six consecutive nonzero terms, we know that all larger terms will be nonzero as well, just by adding sixes. With a little computer assistance, we expand $F_{\text{nugget}}(z)$ as shown in Figure 6.3.

$$
\begin{array}{lrrrrrr}
F_{\text{nugget}}(z) = & & & & & & 1 \\
 & & & & & & +z^{6} \\
 & & & +z^{9} & & & +z^{12} \\
 & & & +z^{15} & & & +2z^{18} \\
 & & +z^{20} & +z^{21} & & & +2z^{24} \\
 & & +z^{26} & +2z^{27} & & +z^{29} & +2z^{30} \\
 & & +z^{32} & +2z^{33} & & +z^{35} & +3z^{36} \\
 & & +2z^{38} & +2z^{39} & +z^{40} & +z^{41} & +3z^{42} \\
 & & +2z^{44} & +3z^{45} & +z^{46} & +2z^{47} & +3z^{48} \\
 & +z^{49} & +2z^{50} & +3z^{51} & +z^{52} & +2z^{53} & +4z^{54} \\
 & & & \vdots & & &
\end{array}
$$

Fig. 6.3 Expanding the Chicken McNugget generating function.

The coefficient of z^{43} is zero, but the next six terms have nonzero coefficients. Hence, $g(6, 9, 20) = 43$, i.e., 43 is the largest "non-nuggetable" number.

Further reading

- Matthias Beck and Sinai Robins, "Computing the Continuous Discretely: Integer Point Enumeration," Springer, 2007.
  This book is an interesting blend of ideas from geometry and combinatorics. It begins by developing a geometric understanding of some of the problems in this chapter.
- Herb Wilf, "generatingfunctionology," AK Peters, 2006.
  As the title suggests, this book is all about generating functions. Wilf employs both a combinatorial and an analytic perspective.

Chapter 7
Exponential generating functions and Bell numbers

"The art of doing mathematics consists in finding that special case which contains all the germs of generality."
–David Hilbert

How many ways can nine people group themselves?

How to partition a finite set? This question is almost as fundamental as how to select a subset of elements, or how to permute the elements in the set. However, the answer to the question turns out to be rather more delicate, which is why we have deferred the question until now.

In this chapter we will introduce a new kind of generating function, known as an exponential generating function. This type of generating function will help us to study the number of ways to partition a set, both with and without order on the blocks.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_7

Exponential generating functions

Definition 14. Given a sequence of numbers $a_0, a_1, a_2, \ldots, a_k, \ldots$, we define its *exponential generating function* by

$$F(z) = \sum_{k\geq 0} a_k \frac{z^k}{k!} = a_0 + a_1 z + a_2 \frac{z^2}{2} + \cdots + a_k \frac{z^k}{k!} + \cdots .$$

For example, the sequence $1, 1, 1, \ldots$ has exponential generating function

$$e^z = \sum_{k\geq 0} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2} + \cdots + \frac{z^k}{k!} + \cdots ,$$

where $e^z = \exp(z)$ is the usual exponential function from calculus. To avoid confusion, we refer to the generating function of Definition 13 as an *ordinary generating function*.

Any particular sequence has both an ordinary and an exponential generating function. There is no rule that says we must use the ordinary generating function or that we must use the exponential generating function. We simply use the one that works best for us. So for example, the sequence $1, 1, 1, \ldots$ has exponential generating function e^z and ordinary generating function $1/(1-z)$.

Warmup 21. What is the exponential generating function for the sequence of factorials: $1, 1, 2, 6, 24, \ldots, k!, \ldots$?

If a sequence of numbers grows too quickly, e.g., if $\sqrt[k]{|a_k|}$ is unbounded, a nice closed-form expression for the ordinary generating function is unlikely since the series will diverge for $z \neq 0$. An added bonus of using the exponential generating function is that we are more likely to find a convergent power series.

Problem 77.
Find the power series expansions for $\sin(z)$ and $\cos(z)$. (It is not important whether you do this by hand or look it up; just find the correct series expansions.) What sequences have these as exponential generating functions? What are the corresponding ordinary generating functions?

Problem 78.
What is the exponential generating function for the sequence $0, 1, 2, 3, \ldots$? What is the exponential generating function for the sequence $1, 2, 3, 4, \ldots$ (i.e., starting at 1 instead of 0)?

Problem 79 (Multiplication for exponential generating functions).
Prove that if

$$F(z) = \sum_{k\geq 0} a_k \frac{z^k}{k!} \quad \text{and} \quad G(z) = \sum_{l\geq 0} b_l \frac{z^l}{l!},$$

are the exponential generating functions for the sequences $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$, then their product is

$$\sum_{m\geq 0}\left(\sum_{k=0}^{m}\binom{m}{k}a_k b_{m-k}\right)\frac{z^m}{m!}.$$

Problem 80.
Suppose $F(z)$ is the exponential generating function for a sequence $a_0, a_1, a_2, \ldots$. What sequence has $F'(z)$ as its exponential generating function? In other words, if

$$F'(z) = \sum_{l\geq 0} b_l \frac{z^l}{l!},$$

what is the sequence $b_0, b_1, b_2, \ldots$?

Problem 81.
Let $a_0, a_1, a_2, \ldots$ be the sequence from part 6 of Problem 54, defined by $a_0 = 1$ and the recurrence

$$a_{n+1} = \sum_{k=0}^{n}\binom{n}{k}a_k,$$

for $n \geq 0$. Find the exponential generating function for the sequence $a_0, a_1, a_2, \ldots$, i.e., find an expression for

$$F(z) = \sum_{n\geq 0} a_n \frac{z^n}{n!}.$$

Your answer should be expressible in terms of well-known functions.

Problem 82.
A *derangement* is a permutation such that for all i, $w(i) \neq i$. For example, 231 and 312 are derangements of $\{1, 2, 3\}$. Let d_n denote the number of derangements of $\{1, 2, \ldots, n\}$, with $d_0 = 1$.

1. Show that for $n \geq 0$,
$$n! = \sum_{k=0}^{n}\binom{n}{k}d_{n-k}.$$
(Hint: count all permutations in cases according to the number of fixed points.)
2. Use the identity above to find a formula for the exponential generating function
$$\sum_{n\geq 0} d_n \frac{z^n}{n!}.$$

$n\backslash k$	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	3	1				
4	1	7	6	1			
5	1	15	25	10	1		
6	1	31	90	65	15	1	
7	1	63	301	350	140	21	1

Table 7.1 Triangle of the Stirling numbers ${n \brace k}$, or the number of set partitions of $\{1, 2, \ldots, n\}$ with k blocks.

Set partitions

Definition 15. A *set partition* of a finite set S is a collection of pairwise disjoint, nonempty subsets whose union is the whole set.

For example, here are three different set partitions of the set $S = \{a, 2, !, \pi\}$:

$$\{\{a\}, \{2\}, \{!\}, \{\pi\}\},$$
$$\{\{a, 2, !, \pi\}\},$$
$$\{\{a, !\}, \{2, \pi\}\}.$$

Notice that each set partition is a set of subsets. The sets in the partition are often referred to as “blocks”. Thus we would say the partitions above have 4 blocks, 1 block, and 2 blocks, respectively.

Warmup 22. List all the set partitions of the set $\{1, 2, 3\}$ and all set partitions of the set $\{1, 2, 3, 4\}$. Group your set partitions according to the number of blocks in the partition.

Stirling numbers

Definition 16. The *Stirling numbers* of the second kind, denoted ${n \brace k}$, count the number of set partitions of an n-element set with k blocks.

In the essay after Chapter 4 we saw other kinds of Stirling numbers, Stirling numbers "of the first kind" that count permutations with a given number of "cycles," but we won't discuss these here.

Problem 83.
Show that the Stirling numbers satisfy the following recurrence for $n \geq k \geq 1$:

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}.$$

Problem 84.
For any $k \geq 1$, let $S_k(z)$ denote the exponential generating function for the number of set partitions with k blocks, i.e.,

$$S_k(z) = \sum_{n \geq k} \left\{ {n \atop k} \right\} \frac{z^n}{n!}.$$

Prove that

$$S_k(z) = \frac{(e^z - 1)^k}{k!}.$$

Bell numbers and Bell polynomials

Definition 17. The nth *Bell number* is the number of set partitions of an n-element set, denoted B_n. The nth *Bell polynomial* is the generating function for the Stirling numbers of the second kind:

$$B_n(t) = \sum_{k=1}^{n} \left\{ {n \atop k} \right\} t^k.$$

For $n = 0$, we define $B_0 = 1$ and $B_0(t) = t$.

For example, $B_3(t) = t + 3t^2 + t^3$ and $B_4(t) = t + 7t^2 + 6t^3 + t^4$. Note in particular that $B_n(1) = B_n$.

Problem 85.
Prove that the Bell numbers satisfy the following recurrence for $n \geq 1$:

$$B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k.$$

Note this gives a combinatorial interpretation to the recurrence in part 6 of Problem 54.

Problem 86.
Prove that the Bell polynomials satisfy the following identity for $n \geq 1$:

$$B_n(t) = t(B_{n-1}(t) + B'_{n-1}(t)),$$

where $B'_{n-1}(t) = \frac{d}{dt} B_{n-1}(t)$ is the derivative of the $(n-1)$st Bell polynomial.

Problem 87.
Let $S(t,z)$ denote the bivariate generating function for *all* the Stirling numbers, i.e.,

$$\begin{aligned} S(t,z) &= 1 + \sum_{n,k\geq 1} \left\{ {n \atop k} \right\} \frac{t^k z^n}{n!} \\ &= 1 + tz + (t+t^2)\frac{z^2}{2} + (t + 3t^2 + t^3)\frac{z^3}{6} + \cdots . \end{aligned}$$

1. Explain why
$$S(t,z) = 1 + \sum_{n\geq 1} B_n(t) \frac{z^n}{n!},$$
where $B_n(t)$ is the nth Bell polynomial.
2. Explain why
$$S(t,z) = 1 + \sum_{k\geq 1} S_k(z) t^k,$$
where $S_k(z)$ is the generating function from Problem 84.
3. Prove
$$S(t,z) = e^{t(e^z-1)}.$$
4. Conclude that the generating function for Bell numbers is
$$\sum_{n\geq 0} B_n \frac{z^n}{n!} = e^{(e^z-1)}.$$
Compare this result with Problem 81.

Set compositions

Definition 18. A *set composition* of a set S is a set partition with an ordering on its blocks.

For example, for the set $S = \{a, 2, !, \pi\}$, the following are two *different* set compositions:

$$(\{a,!\},\{2,\pi\}) \quad \text{and} \quad (\{2,\pi\},\{a,!\}).$$

Warmup 23. Explain why the number of set compositions of $\{1,2,\ldots,n\}$ with k blocks equals $k!\left\{ {n \atop k} \right\}$. This number could be called an "ordered Stirling number."

In practice, we abbreviate notation when considering set compositions of the set $\{1,2,\ldots,n\}$ by writing all members of each block in increasing

$n\backslash k$	1	2	3	4	5	6	7
1	1						
2	1	2					
3	1	6	6				
4	1	14	36	24			
5	1	30	150	240	120		
6	1	62	540	1560	1800	720	
7	1	126	1806	8400	16800	15120	5040

Table 7.2 Triangle of the ordered Stirling numbers $k!\left\{{n \atop k}\right\}$, or the number of set compositions of $\{1, 2, \ldots, n\}$ with k blocks.

order, and separating each block with a vertical bar. For example, the set composition $(\{3\}, \{4, 6\}, \{1, 5, 2\})$ would be written $3|46|125$.

The key thing about this notation is that it helps to identify set compositions with permutations. That is, each set composition has a permutation (346125 in the example), while each permutation can be associated with several different set compositions. We call the permutation obtained by ignoring the bars in set composition the "underlying" permutation of that set composition. For $n = 1$, there is only one permutation and one set composition. For $n = 2$, there are three set compositions: $12, 1|2$, and $2|1$. The permutation 12 corresponds to the set compositions 12 and $1|2$, while the permutation 21 corresponds to the set composition $2|1$.

Warmup 24. List all the set compositions of $\{1, 2, 3\}$; group them first according to the number of blocks, then according to the underlying permutation.

Problem 88.

Explain why the permutation 2351467 is the underlying permutation for 32 different set compositions of $\{1, 2, 3, 4, 5, 6, 7\}$.

Which permutation corresponds to the greatest number of set compositions?

Which permutation corresponds to the fewest set compositions?

Ordered Bell numbers and ordered Bell polynomials

Definition 19. The nth *ordered Bell polynomial*, denoted $\overline{B}_n(t)$, is the generating function for set compositions $\{1, 2, \ldots, n\}$ counted according to the number of blocks, i.e.,

$$\overline{B}_n(t) = \sum_{k=1}^{n} k! \left\{ n \atop k \right\} t^k.$$

The nth *ordered Bell number*, denoted $\overline{B}_n$, is the total number of ordered set compositions of $\{1, 2, \ldots, n\}$, i.e., $\overline{B}_n = \overline{B}_n(1)$. We declare $\overline{B}_0(t) = t$ and $\overline{B}_0 = 1$.

Problem 89.
Show the ordered Bell numbers satisfy the following identity for all $n \geq 1$:

$$\overline{B}_n = \sum_{k=0}^{n-1} \binom{n}{n-k} \overline{B}_k \, .$$

Problem 90.
Let $b_{n,k} = k! \left\{ n \atop k \right\}$ denote the number of ordered set compositions of $\{1, 2, \ldots, n\}$ with k blocks. Show that

$$b_{n,k} = k b_{n-1,k-1} + k b_{n-1,k}.$$

Problem 91.
Prove the ordered Bell polynomials satisfy the following identity for $n \geq 1$:

$$\overline{B}_n(t) = t\,\overline{B}_{n-1}(t) + t(1+t)\,\overline{B}'_{n-1}(t).$$

Problem 92.
Emulating our approach to Problem 87, let

$$\begin{aligned} \overline{S}(t,z) &= 1 + \sum_{n,k \geq 1} k! \left\{ n \atop k \right\} \frac{t^k z^n}{n!}, \\ &= 1 + \sum_{n \geq 1} \overline{B}_n(t) \frac{z^n}{n!}, \\ &= 1 + tz + (t + 2t^2)\frac{z^2}{2} + (t + 6t^2 + 6t^3)\frac{z^3}{6} + \cdots . \end{aligned}$$

Prove that

$$\overline{S}(t,z) = \frac{1}{1 - (e^z - 1)t}$$

and conclude that

$$\sum_{n\geq 0} \overline{B}_n \frac{z^n}{n!} = 1 + z + 3\frac{z^2}{2} + 13\frac{z^3}{6} + \cdots = \frac{1}{2 - e^z}$$

OEIS

As a researcher in combinatorics, one of my favorite tools is the On-Line Encyclopedia of Integer Sequences, or OEIS. This database was started by the mathematician Neil Sloane, who first started keeping an index of popular sequences of integers that came up in his work. At the time, Sloane was a graduate student at Cornell University. A photo of the first page of Sloane's notebook is shown in Figure 7.1. Recognize any of these sequences?

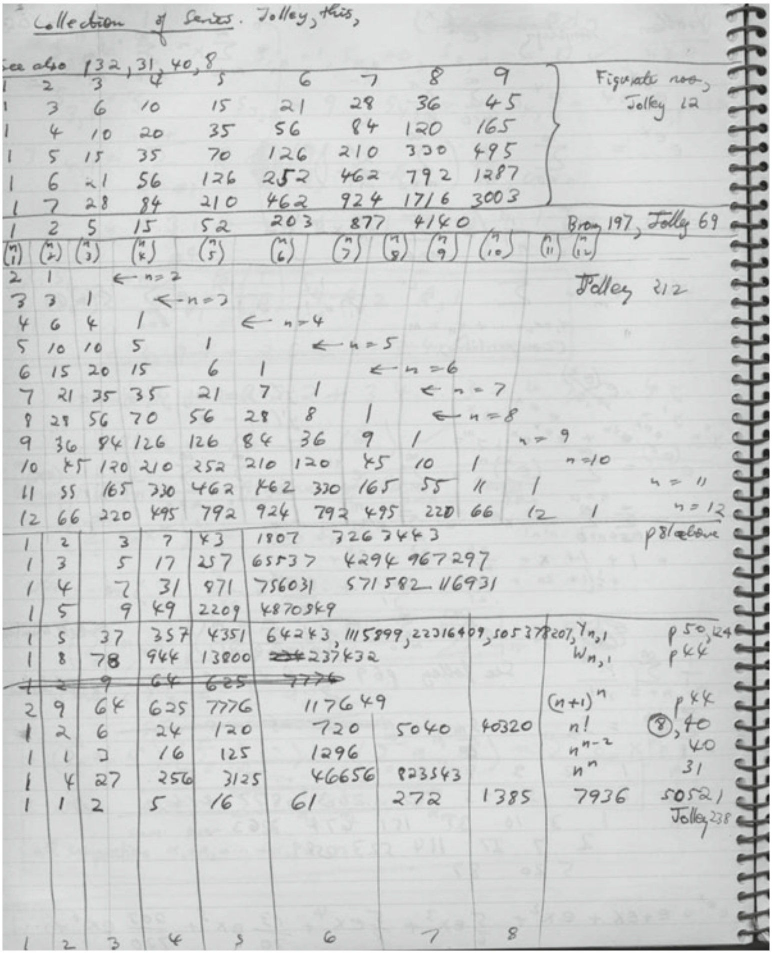
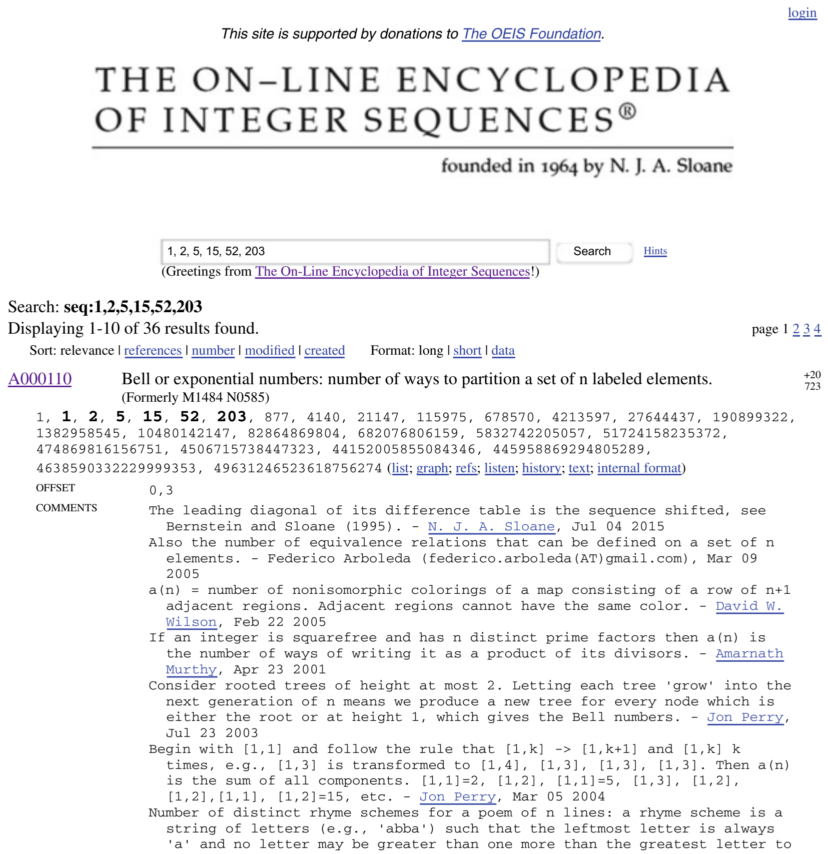


Fig. 7.1 A photograph of Neil Sloane's first notebook of integer sequences.

By 1973, Sloane had collected 2372 sequences and published them in a book: "The Handbook of Integer Sequences." The book was very popular with researchers, and in 1995, Sloane was joined by Simon Plouffe in producing a follow-up, the "Encyclopedia of Integer Sequences," with a then-astounding 5487 sequences.

The On-Line Encyclopedia of Integer Sequences was launched in 1996, and initially hosted on Sloane's personal web page at AT&T Labs, where he worked. In the early 2000s, Sloane started enlisting the help of other researchers in maintaining the Encyclopedia, and it is now run in a moderated wiki format with a team of editors who review edits and new submissions. There are currently hundreds of thousands of entries. A search on the terms $1, 2, 5, 15, 52, 203$ gives the result shown in Figure 7.2


login

This site is supported by donations to The OEIS Foundation.

THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES®

founded in 1964 by N. J. A. Sloane

1, 2, 5, 15, 52, 203 Search Hints

(Greetings from The On-Line Encyclopedia of Integer Sequences!)

Search: **seq:1,2,5,15,52,203**

Displaying 1-10 of 36 results found. page 1 2 3 4

Sort: relevance | references | number | modified | created Format: long | short | data

A000110 Bell or exponential numbers: number of ways to partition a set of n labeled elements. +20 723

(Formerly M1484 N0585)

1, **1**, **2**, **5**, **15**, **52**, **203**, 877, 4140, 21147, 115975, 678570, 4213597, 27644437, 190899322, 1382958545, 10480142147, 82864869804, 682076806159, 5832742205057, 51724158235372, 474869816156751, 4506715738447323, 44152005855084346, 445958869294805289, 4638590332229999353, 49631246523618756274 (list; graph; refs; listen; history; text; internal format)

OFFSET 0,3

COMMENTS The leading diagonal of its difference table is the sequence shifted, see Bernstein and Sloane (1995). - N. J. A. Sloane, Jul 04 2015

Also the number of equivalence relations that can be defined on a set of n elements. - Federico Arboleda (federico.arboleda(AT)gmail.com), Mar 09 2005

a(n) = number of nonisomorphic colorings of a map consisting of a row of n+1 adjacent regions. Adjacent regions cannot have the same color. - David W. Wilson, Feb 22 2005

If an integer is squarefree and has n distinct prime factors then a(n) is the number of ways of writing it as a product of its divisors. - Amarnath Murthy, Apr 23 2001

Consider rooted trees of height at most 2. Letting each tree 'grow' into the next generation of n means we produce a new tree for every node which is either the root or at height 1, which gives the Bell numbers. - Jon Perry, Jul 23 2003

Begin with [1,1] and follow the rule that [1,k] -> [1,k+1] and [1,k] k times, e.g., [1,3] is transformed to [1,4], [1,3], [1,3], [1,3]. Then a(n) is the sum of all components. [1,1]=2, [1,2], [1,1]=5, [1,3], [1,2], [1,2],[1,1], [1,2]=15, etc. - Jon Perry, Mar 05 2004

Number of distinct rhyme schemes for a poem of n lines: a rhyme scheme is a string of letters (e.g., 'abba') such that the leftmost letter is always 'a' and no letter may be greater than one more than the greatest letter to

Fig. 7.2 Search results for $1, 2, 5, 15, 52, 203$ on https://oeis.org.

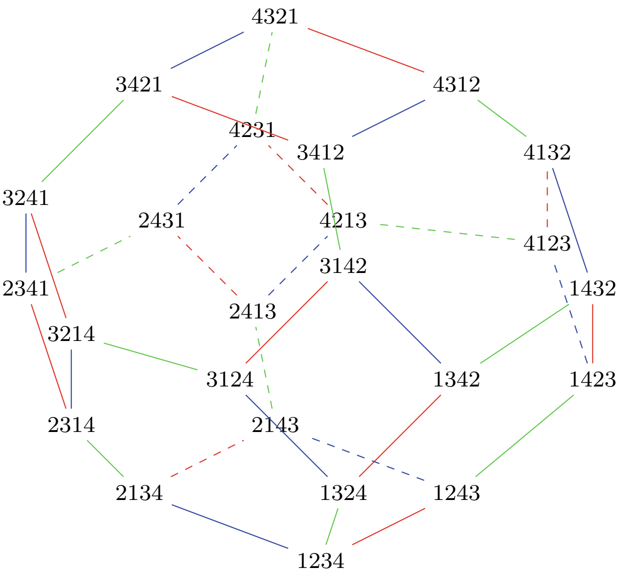
The best way to really learn about OEIS is to check it out for yourself. Go to https://oeis.org and poke around!

Further reading

- Alexander Yong, *The Joseph Greenberg problem: Combinatorics and Comparative Linguistics*, Mathematics Magazine, Vol. 91, 192–197 (2018)
  This short article talks about an application of Stirling numbers and Bell numbers in linguistics.

Chapter 8
Eulerian numbers

"If you don't make mistakes, you're not working on hard enough problems. And that's a big mistake."
–Frank Wilczek



Can you think of a rule for organizing permutations this way?

Leonhard Euler certainly didn't have permutations on his mind when in 1755 he wrote about the numbers bearing his name in this chapter. (He was trying to find an easy way to compute sums of the form $1^k + 2^k + 3^k + \cdots$ for any value of k.) Nowadays, however, these numbers are usually defined combinatorially: in terms of permutations.

In this chapter we will study Eulerian numbers and relate their generating functions to the generating functions for the ordered Stirling numbers. Among other things, we will see that these numbers satisfy a Pascal-like recurrence.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_8

$n\backslash k$	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	4	1				
4	1	11	11	1			
5	1	26	66	26	1		
6	1	57	302	302	57	1	
7	1	120	1191	2416	1191	120	1

Table 8.1 Triangle of the numbers $\left\langle {n \atop k} \right\rangle$, or the number of permutations in S_n with $k-1$ descents.

Recall S_n denotes the set of permutations of $\{1, 2, \dots, n\}$. A typical permutation is denoted $w = w(1) \cdots w(n)$, e.g., $w = 321546$ is an element of S_6 with $w(1) = 3$, $w(2) = 2$, $w(3) = 1$, $w(4) = 5$, $w(5) = 4$, and $w(6) = 6$.

Descents of a permutation

Definition 20. A *descent* of a permutation $w = w(1) \cdots w(n)$ is an index $i \in \{1, 2, \dots, n-1\}$ such that $w(i) > w(i+1)$. We let $\mathrm{Des}(w)$ denote the set of descents of w, and $\mathrm{des}(w)$ denote the number of descents in w, i.e.,

$$\mathrm{des}(w) = |\,\mathrm{Des}(w)| = |\{i : w(i) > w(i+1)\}|.$$

For example, if $w = 51243$, $\mathrm{Des}(w) = \{1, 4\}$ and $\mathrm{des}(w) = 2$ since $w(1) > w(2)$ and $w(4) > w(5)$. Counting descents is one way to measure how "mixed up" a permutation is.

Warmup 25. How many permutations in S_4 have no descents? exactly one descent? two descents? three descents? four?

Runs of a permutation

Definition 21. A *run* of a permutation $w = w(1) \cdots w(n)$ is a maximal sequence of consecutive entries that increase, i.e., for some $i \le j$, we have $w(i) \le \cdots \le w(j)$, but $w(i-1) \not\le w(i)$ and $w(j) \not\le w(j+1)$. We let $\mathrm{runs}(w)$ denote the number of runs in w.

For example, if $w = 51243$, $\operatorname{runs}(w) = 3$, since there are three runs of w, separated here with vertical bars: $5|124|3$. Runs and descents are essentially the same, in that descent positions mark the gaps between runs.

Warmup 26. How many permutations in S_4 have one run? two runs? three runs? four? Show that $\operatorname{runs}(w) = \operatorname{des}(w) + 1$.

Problem 93.
An *ascent* of a permutation w is an index i such that $w(i) < w(i+1)$. Show that the number of permutations with k ascents equals the number of permutations with k descents.

Problem 94.
A *return* in a permutation is a number i such that $i+1$ appears to the left of i in the permutation. For example, if $w = 2714365$, there are returns at 1 (since 2 is to the left of 1), at 3 (since 4 appears to the left of 3), at 5 (since 6 is to the left of the 5), and at 6 (since 7 appears to the left of 6).

Count all the elements of S_4 according to the number of returns each permutation has.

Can you prove in general that the number of permutations in S_n with k returns equals the number of permutations in S_n with k descents?

Problem 95.
Find and prove a formula for the number of permutations in S_n that have exactly one descent.

Eulerian numbers

Definition 22. For any $n \geq k \geq 1$, the *Eulerian number* $\left\langle {n \atop k} \right\rangle$ is the number of permutations in S_n with k runs, or equivalently, $k-1$ descents.

For example, the Eulerian numbers for $n = 3$ are $\left\langle {3 \atop 1} \right\rangle = 1$, $\left\langle {3 \atop 2} \right\rangle = 4$, and $\left\langle {3 \atop 3} \right\rangle = 1$, since we can gather the following data about descents in S_3:

w	$\operatorname{runs}(w)$	$\operatorname{des}(w)$
123	1	0
132	2	1
213	2	1
231	2	1
312	2	1
321	3	2

Eulerian polynomials

Definition 23. The nth *Eulerian polynomial*, denoted $A_n(t)$, is the generating function for permutations in S_n counted according to the number of runs, i.e.,

$$A_n(t) = \sum_{w \in S_n} t^{\mathrm{runs}(w)} = \sum_{w \in S_n} t^{1+\mathrm{des}(w)} = \sum_{k=1}^{n} \left\langle {n \atop k} \right\rangle t^k.$$

For example, our table for S_3 shows $A_3(t) = t + 4t^2 + t^3$.

Warmup 27. What is $A_4(t)$?

Problem 96.
Use the meaning of Eulerian numbers to prove

$$\left\langle {n \atop k} \right\rangle = \left\langle {n \atop n+1-k} \right\rangle.$$

Problem 97 (Eulerian recurrence).
Prove the Eulerian numbers satisfy the following Pascal-like recurrence:

$$\left\langle {n \atop k} \right\rangle = k\left\langle {n-1 \atop k} \right\rangle + (n+1-k)\left\langle {n-1 \atop k-1} \right\rangle,$$

for $n \geq 2$ and $1 \leq k \leq n$. Use the recurrence to compute the next row of Table 8.1.

Problem 98.
Prove the Eulerian polynomials satisfy the following recurrence for all $n \geq 2$:

$$A_n(t) = ntA_{n-1}(t) + t(1-t)A'_{n-1}(t).$$

This is the same recurrence as the polynomials in Problem 76, so conclude

$$\frac{A_n(t)}{(1-t)^{n+1}} = \sum_{k \geq 1} k^n t^k. \tag{8.1}$$

Problem 99. (Worpitzky's Identity).
For any $n, k \geq 1$,

$$k^n = \sum_{i=1}^{n} \left\langle {n \atop i} \right\rangle \binom{k+n-i}{n}.$$

For example, $k^3 = \binom{k+2}{3} + 4\binom{k+1}{3} + \binom{k}{3}$. You can use Equation 8.1 along with the series for $t^i/(1-t)^{n+1}$ to prove this in general.

Problem 100.
Use Worpitzky's identity or Equation (8.1) to get the following formulas for Eulerian numbers:

$$\left\langle {n \atop 1} \right\rangle = 1, \quad \left\langle {n \atop 2} \right\rangle = 2^n - (n+1), \quad \left\langle {n \atop 3} \right\rangle = 3^n - \binom{n+1}{1} 2^n + \binom{n+1}{2}.$$

Conjecture and prove a general formula for $\left\langle {n \atop k} \right\rangle$ as an alternating sum of powers and binomial coefficients.

Problem 101.
Let us define the exponential generating function for Eulerian polynomials as follows:

$$\begin{aligned} E(t,z) &= 1 + \sum_{n \geq 1} A_n(t) \frac{z^n}{n!} \\ &= 1 + \sum_{n \geq k \geq 1} \left\langle {n \atop k} \right\rangle \frac{t^k z^n}{n!} \\ &= 1 + tz + (t + t^2)\frac{z^2}{2} + (t + 4t^2 + t^3)\frac{z^3}{6} + \cdots . \end{aligned}$$

Prove

$$E(t,z) = \frac{1-t}{1 - te^{z(1-t)}}.$$

Hint: let $u = z/(1-t)$ and manipulate $E(t,u)/(1-t)$, keeping in mind Equation (8.1).

In the next few problems, we will derive $E(t,z)$ another way.

Problem 102.
Let's do a plausibility check on our last result. If we set $t = 1$ in $E(t,z)$, we should get

$$E(1,z) = 1 + \sum_{n \geq 1} n! \frac{z^n}{n!} = \sum_{n \geq 0} z^n = 1/(1-z).$$

But plugging in $t = 1$ gives

$$\left[\frac{1-t}{1-te^{z(1-t)}} \right]_{t=1} = \frac{1-1}{1-e^{z \cdot 0}} = \frac{0}{0}.$$

This seems bad. Resolve the issue. (Hint: instead of "plugging in" $t = 1$, take a limit.)

Problem 103.
Recall from Problem 88 that the permutation $w = 2351467$ corresponds to 32 different set compositions of $\{1,2,3,4,5,6,7\}$, such as $235|1467$, $2|35|1467$, $23|5|1467$, and so on. In general, let $\mathcal{C}(w)$ denote the set of set compositions

whose underlying permutation (when we write the elements in each block in order and remove bars) is w.

Show that

$$\sum_{C\in\mathcal{C}(w)} t^{|C|} = t^{\mathrm{runs}(w)}(1+t)^{n-\mathrm{runs}(w)},$$

where $|C|$ denotes the number of blocks in the set composition C.

Problem 104.
Use Problem 103 to show that for $n \geq 1$,

$$\overline{B}_n(t) = (1+t)^n A_n\left(\frac{t}{1+t}\right),$$

where $\overline{B}_n(t)$ is the nth ordered Bell polynomial from Definition 19. For example,

$$\begin{aligned}
(1+t)^3 A_3\left(\frac{t}{1+t}\right) &= (1+t)^3\left(\frac{t}{1+t} + 4\frac{t^2}{(1+t)^2} + \frac{t^3}{(1+t)^3}\right)\\
&= t(1+t)^2 + 4t^2(1+t) + t^3\\
&= t + 6t^2 + 6t^3 = \overline{B}_3(t).
\end{aligned}$$

Problem 105.
Use Problem 104 to relate the generating function for the Eulerian polynomials, $E(t,z)$, to the generating function for the ordered Bell polynomials, $\overline{S}(t,z)$. (This is defined in Problem 92.) This should give another proof of the formula for $E(t,z)$ found in Problem 101.

Euler (not Eulerian) numbers

LEONHARD EULER DID A LOT OF THINGS, and many quantities are named after him. There are of course the "Eulerian numbers" of this chapter, but also we have:

- "Euler's number"
$$e = \sum_{k\geq 0} \frac{1}{k!} = 2.71828\ldots,$$
- "Euler characteristic"
$$\chi = V - E + F,$$
where V, E, and F count vertices, edges, and faces of a polyhedron, respectively,
- and "Euler's constant" (or "Euler-Mascheroni constant")

$$\gamma = \lim_{n\to\infty}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln(n)\right) = 0.57721566\ldots$$

Another, more combinatorial set of numbers is the sequence of "Euler numbers" which we will denote E_n. The sequence begins:

$$1, 1, 1, 2, 5, 16, 61, 272, 1385, \ldots.$$

It is sequence $A000111$ in OEIS.

The Euler numbers can be given the following recursive definition:

$$E_0 = E_1 = 1, \qquad E_{n+1} = \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}E_k E_{n-k} \quad (n \geq 1). \tag{8.2}$$

From this recursive description, we can show that the exponential generating function for the sequence of Euler numbers is $\sec(z) + \tan(z)$, i.e.,

$$E(z) = \sum_{k\geq 0} E_k \frac{z^k}{k!} = \sec(z) + \tan(z).$$

This identity amounts to showing that $E(z)$ satisfies

$$E(z)^2 + 1 = 2E'(z).$$

We can prove this using the recurrence (8.2) on the one hand, and by computing the derivative of $\sec(z) + \tan(z)$ on the other hand.

Let's now discuss a combinatorial interpretation of the sequence.

Zig-zag permutations

A combinatorial interpretation for Euler numbers seems to first appear in work of Désiré André from 1888. We will define an *up-down* permutation as one for which the entries alternately increase and decrease:

$$w(1) < w(2) > w(3) < w(4) > \cdots .$$

Similarly, the *down-up* permutations are those for which the entries alternately decrease and increase:

$$w(1) > w(2) < w(3) > w(4) < \cdots .$$

The two families together are collectively called *alternating permutations* or *zig-zag permutations*. In Table 8.2 we see the zig-zag permutations for $n \leq 5$.

n	up-down	down-up
1	1	1
2	12	21
3	132 231	312 213
4	1324 1423 2314 2413 3412	4231 4132 3241 3142 2143
5	13254 14253 14352 15243 15342 23154 24153 24351 25143 25341 34152 34251 35142 35241 45132 45231	53412 52413 52314 51423 51324 43512 42513 42315 41523 41325 32514 32415 31524 31425 21534 21435

Table 8.2 The zig-zag permutations for $n \leq 5$.

There is a simple correspondence between up-down permutations and down-up permutations implicit in Table 8.2: for any up-down permutation of

$\{1, 2, \ldots, n\}$ replacing the letter j with $n+1-j$ will yield a corresponding down-up permutation (and vice-versa). Notice the permutations 3517264 and 5371624 can be obtained from each other in this way. Thus the set of zig-zag permutations splits evenly into these two types and we shall soon see that for $n \geq 1$,

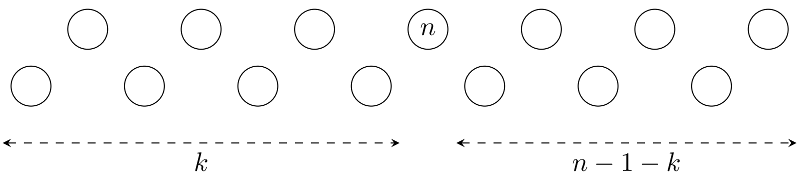
$$\begin{aligned} E_n &= |\{\text{up-down permutations in } S_n\}|, \\ &= |\{\text{down-up permutations in } S_n\}|. \end{aligned}$$

That is, $E(z) = \sec(z) + \tan(z)$ is the exponential generating function for the number of up-down (or down-up) permutations.

To prove this result, we need only show that the zig-zag permutations satisfy the following identity for $n \geq 2$:

$$2E_n = |\{\text{zig-zag permutations in } S_n\}| = \sum_{k=0}^{n-1} \binom{n-1}{k} E_k E_{n-1-k}.$$

Here is a bijective argument to prove the identity above. Consider how to create the $2E_n$ zig-zag permutations (up-down or down-up) of length n. First suppose that we put n in postion $k+1$, for some $k = 0, 1, \ldots, n-1$. Then to the left of n we will have an alternating permutation of length k that ends in a down step, and to the right of n we will have an up-down permutaton of length $n-1-k$. We can sketch the situation like this:



Note that while we have drawn an up-down alternating permutation, it could also be down-up. It depends on whether k is even or, as in this case, k is odd.

To fill in the rest of the picture, we need to do two things:

1. choose the k elements that go to the left of n and arrange them as an appropriately alternating permutation, and
2. arrange the remaining $n-1-k$ elements to the right of n as an appropriately alternating permutation.

Step 1 can be done in $\binom{n-1}{k} E_k$ ways. Indeed, we are choosing k of $n-1$ elements, and these can be ordered in E_k ways, according to the reversal of any up-down permutation of length k. For step 2, we also want to form an up-down permutation, this time written left to right. Since there are $n-1-k$ elements to the right of n, this can be done in E_{n-1-k} ways.

Now we can conclude that there are $\binom{n-1}{k} E_k E_{n-1-k}$ alternating permutations with n in position $k+1$. Summing over all k, we have proved the recurrence.

Other generating functions

The "classical" Euler numbers are actually the odd-indexed Euler numbers as we have defined them, with a sign:

$$0, 1, 0, -2, 0, 16, 0, -272, 0, \ldots, (-1)^n E_{2n+1}, \ldots.$$

Since secant is an even function and tangent is an odd function, calculus tells us that this sequence has exponential generating function

$$\frac{\tan(iz)}{i} = z - 2\frac{z^3}{3!} + 16\frac{z^5}{5!} - \cdots .$$

Some basic identities allow us to rewrite $\tan(iz)/i$ as

$$\tanh(z) = \frac{-1 + e^{2z}}{1 + e^{2z}}.$$

Interestingly, we can connect this result to the generating function for Eulerian polynomials found in Problem 101:

$$E(t, z) = \sum_{n \geq 0} A_n(t) \frac{z^n}{n!} = \frac{1-t}{1 - te^{z(1-t)}}.$$

By setting $t = -1$, we find

$$E(-1, z) = \frac{2}{1 + e^{2z}} = 1 - \tanh(z) = 1 - z + 2\frac{z^3}{3!} - 16\frac{z^5}{5!} + \cdots .$$

But then by coefficient comparison, we have for $n > 0$:

$$A_n(-1) = \begin{cases} (-1)^k E_{2k-1} & \text{if } n = 2k-1, \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

On the one hand, this gives us an expression for the Euler number E_{2k-1} as an alternating sum of Eulerian numbers:

$$E_{2k-1} = (-1)^k A_{2k-1}(-1) = \sum_{i=1}^{2k-1} (-1)^{k+i} \left\langle {2k-1 \atop i} \right\rangle.$$

On the other hand, it tells us combinatorial information. First, since $A_{2k}(-1) = 0$, it tells us that the number of permutations in S_{2k} with an even number of descents equals the number of permutations with an odd number of descents. Second, it tells us that in the case n is odd, the difference between the number of permutations with an even number of descents and the number of permutations with an odd number of descents is (up to sign) the number of up-down permutations:

$$|\{w \in S_{2k-1} : \mathrm{des}(w) \text{ is even}\}| - |\{w \in S_{2k-1} : \mathrm{des}(w) \text{ is odd}\}| = (-1)^k E_{2k-1}.$$

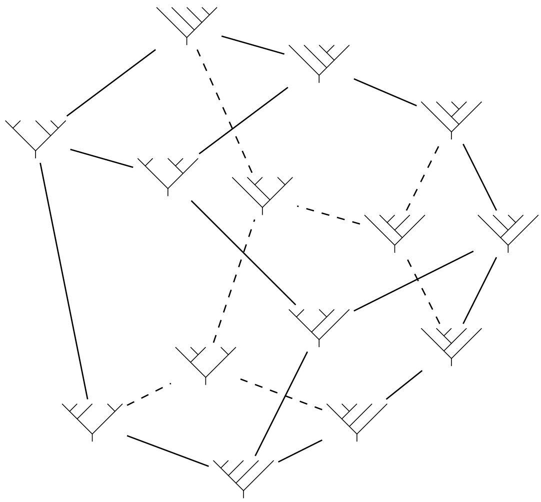
It is fun to find a bijective explanation for this identity.

Further reading

- Joe Buhler, David Eisenbud, Ron Graham, Colin Wright, *Juggling drops and descents*, American Mathematical Monthly, **101**, 507–519, (1994).
  This fun and engaging article shows how Eulerian numbers arise the mathematical study of juggling!
- Ira Gessel, *The Smith College Diploma Problem*, American Mathematical Monthly, **108**, 55–57, (2001).
  Gessel's short note uses a clever bijection to explain how Eulerian numbers crop up in a very different-sounding problem.
- Kyle Petersen, "Eulerian Numbers," Birkhäuser Advanced Texts, (2015).
  This book is all about Eulerian numbers from a combinatorial, geometric, and algebraic point of view.

Chapter 9
Catalan and Narayana numbers

"I am interested in mathematics only as a creative art."
–Godfrey Harold Hardy



How are these trees related to each other?

THE FIBONACCI NUMBERS ARE PRETTY COOL, but in modern algebraic combinatorics, the most interesting sequence is the sequence of CATALAN NUMBERS:

$$1, 1, 2, 5, 14, 42, 132, 429, \ldots$$

The remarkable ability of these numbers to pop up in surprising locations has led some to joke that a combinatorics paper is not complete until the Catalan numbers have made an appearance. There are now more than two hundred distinct combinatorial interpretations for Catalan numbers!

In this chapter we will study some families of objects counted by Catalan numbers, as well as a triangle of numbers known as the NARAYANA NUMBERS that possess many of the properties of the Eulerian numbers.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_9

$n\backslash k$	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	3	1				
4	1	6	6	1			
5	1	10	20	10	1		
6	1	15	50	50	15	1	
7	1	21	105	175	105	21	1

Table 9.1 Triangle of the numbers $N(n, k)$, or the number of 231-avoiding permutations in S_n with $k-1$ descents.

231-avoiding permutations

Definition 24. A permutation $w = w(1)\cdots w(n)$ in S_n is said to *contain the pattern 231* if there is a triple of indices $i < j < k$ such that $w(k) < w(i) < w(j)$. Otherwise we say w is *231-avoiding*. Let $S_n(231)$ denote the set of 231-avoiding permutations.

Pattern containment and avoidance is best thought of in terms of pictures. We can graph a permutation w by plotting the pairs $(i, w(i))$ in cartesian coordinates. The pattern 231 is then just a triple of dots, which appear at heights "middle-high-low" when viewed from left to right. See Figure 9.1(a). For example, the permutation 53412 contains the pattern 231 in two different ways, with $(i, j, k) = (2, 3, 4)$ and with $(i, j, k) = (2, 3, 5)$, as outlined in Figure 9.1(b) with dashed lines.

Warmup 28. The permutations 1234567 and 7654321 both avoid the pattern 231. Come up with three different members of $S_7(231)$.

Catalan numbers

Definition 25. The nth *Catalan number*, denoted C_n, counts the number of 231-avoiding permutations in S_n, i.e., $C_n = |S_n(231)|$. By convention $C_0 = 1$, since there is one empty permutation that trivially avoids the pattern.

The Catalan numbers are named for Eugène Catalan, a 19th-century mathematician. Although these are now known as Catalan numbers, Catalan called them Segner numbers, after Johann Segner. Segner, in turn, learned of the

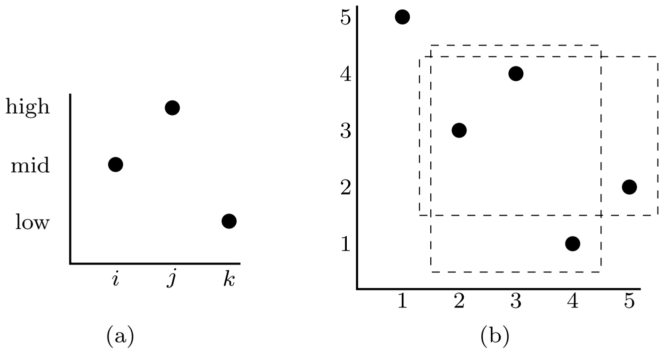


Fig. 9.1 Containing the pattern 231 in a permutation.

sequence from Euler! But even Euler's work was predated by that of Mongolian mathematician Minggatu, though it seems unlikely that Euler was aware of Minggatu.[1]

Warmup 29. List all the elements of $S_n(231)$, for $n \leq 4$ and compute the corresponding Catalan numbers. For each n, count the number of permutations with 0, 1, 2, and 3 descents.

Problem 106.
By considering different values of i such that $w(i+1) = n$, show that $S_n(231)$ can be partitioned into subsets that are each in bijection with $S_i(231) \times S_{n-i-1}(231)$. Use your correspondence to show that for any $n \geq 1$,

$$C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1},$$

where $C_0 = 1$.

This shows the Catalan numbers form the sequence from Problem 54, part 4. We found the ordinary generating function for this sequence in Problem 74, which we now denote by $C(z)$:

$$C(z) = \sum_{k \geq 0} C_k z^k = 1 + z + 2z^2 + 5z^3 + 14z^4 + \cdots = \frac{1 - \sqrt{1 - 4z}}{2z}.$$

(You may want to re-derive this formula from the recurrence as practice.)

[1] This phenomenon of naming scientific discoveries for someone other than the original discoverer is known as "Stigler's law of eponymy," a name coined somewhat facetiously by Stephen Stigler in 1980. Of course, Stigler knew that others had made this observation before, which makes "Stigler's law" an example of Stigler's law.

Keeping track of descents or runs in 231-avoiding permutations gives the definition of Narayana numbers, by analogy with our definition of the Eulerian numbers.

Narayana numbers and Narayana polynomials

Definition 26. For any $n \geq k \geq 1$, the *Narayana number* $N(n,k)$ is defined to be the number of 231-avoiding permutations with $k-1$ descents, or equivalently, k runs. The nth *Narayana polynomial*, denoted $C_n(t)$, is the generating function for permutations in $S_n(231)$ according to the number of runs, i.e.,

$$C_n(t) = \sum_{w \in S_n(231)} t^{\mathrm{runs}(w)} = \sum_{w \in S_n(231)} t^{1+\mathrm{des}(w)} = \sum_{k=1}^{n} N(n,k)t^k.$$

For convenience we set $C_0(t) = t$.

Warmup 30. Compute $C_n(t)$ for $n \leq 4$.

Problem 107.
Returning to the bijection constructed in Problem 106, so that the Narayana polynomials satisfy the following recurrence:

$$C_n(t) = C_{n-1}(t) + \sum_{i=0}^{n-2} C_i(t)C_{n-i-1}(t).$$

Use this recurrence to find an expression for the bivariate generating function for the Narayana numbers,

$$\begin{aligned} C(t,z) &= \sum_{n \geq 0} C_n(t)z^n, \\ &= t + tz + (t+t^2)z^2 + (t+3t^2+t^3)z^3 + \cdots . \end{aligned}$$

As a reality check, note that setting $t = 1$ should specialize to the generating function from Problem 106.

Problem 108.
Use bijections to show that the number of permutations in S_n avoiding the pattern 231 equals the number of permutations in these sets:

1. The set of 132-avoiding permutations, denoted $S_n(132)$, which are characterized by having no indices $i < j < k$ with $w(i) < w(k) < w(j)$.
2. The set of 213-avoiding permutations, denoted $S_n(213)$, which are characterized by having no indices $i < j < k$ with $w(j) < w(i) < w(k)$.
3. The set of 312-avoiding permutations, denoted $S_n(312)$, which are characterized by having no indices $i < j < k$ with $w(j) < w(k) < w(i)$.

Problem 109.
Show $N(n,k) = N(n,n+1-k)$. Try to do this both with a bijection and algebraically.

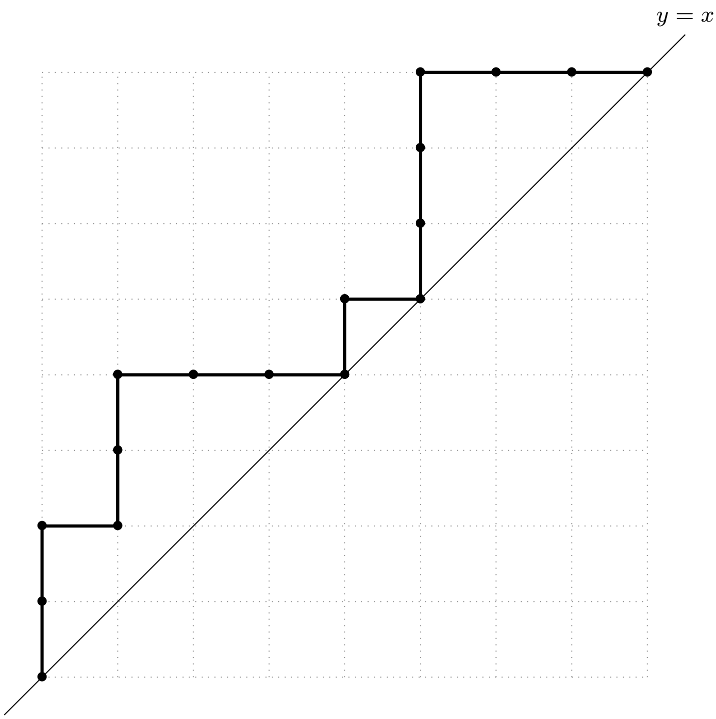


Fig. 9.2 One of the 4862 paths in Dyck(8).

Dyck paths

Definition 27. A *Dyck path* of length $2n$ is a lattice path from $(0,0)$ to (n,n) that takes n steps "East" from (i,j) to $(i+1,j)$ and n steps "North" from (i,j) to $(i,j+1)$, such that all points on the path satisfy $i \le j$. In other words, when drawn in the cartesian plane, a Dyck path lies on or above the line $y = x$. The set of all Dyck paths of length $2n$ is denoted Dyck(n).

We first saw Dyck paths in one of our Starfolks problems (Problem 61). When discussing Dyck paths, we can draw a picture or record the list of steps taken in the path. For example, the path given by

$$p = NNENNEEENENENNNEEE$$

is shown in Figure 9.2.

Peaks of a Dyck path

Definition 28. A *peak* in a Dyck path is a coordinate (i,j) in the path such that both $(i,j-1)$ and $(i+1,j)$ are also on the path. In other words, a peak corresponds to a North step followed by an East step. We let $\mathrm{pk}(p)$ denote the number of peaks of the path p.

For example, the path in Figure 9.2 has four peaks, $\mathrm{pk}(p)=4$.

Warmup 31. Draw all paths in $\mathrm{Dyck}(n)$ for $n \leq 4$, grouping them according to the number of peaks.

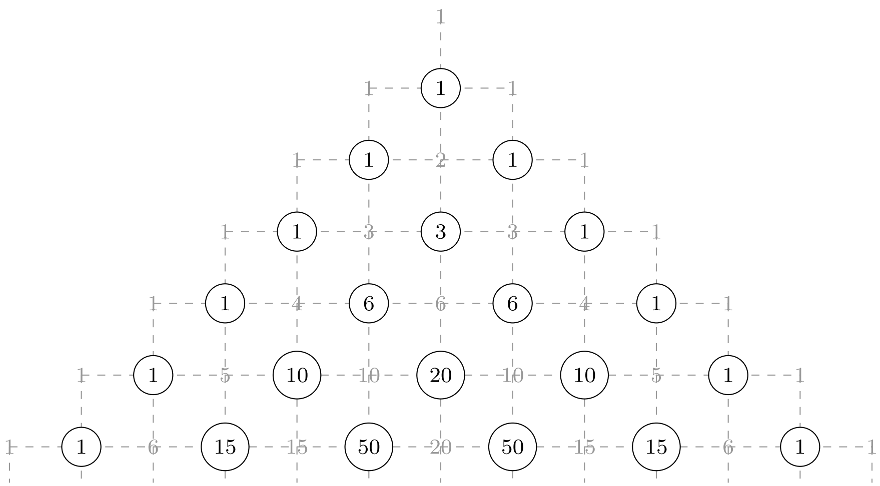


Fig. 9.3 The triangle of Narayana numbers obtained as 2×2 minors of Pascal's triangle.

Problem 110.
For $n \geq 1$, show that $|\mathrm{Dyck}(n)| = C_n$ and moreover,

$$C_n(t) = \sum_{p \in \mathrm{Dyck}(n)} t^{\mathrm{pk}(p)}.$$

Problem 111.
Show that

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

Hint: one way to do this is to consider the set of all paths from $(0,0)$ to (n,n), and show that there are $\binom{2n}{n-1}$ paths that fall below the line $y = x$.

Problem 112.
Show that

$$N(n,k) = \frac{1}{k}\binom{n}{k-1}\binom{n-1}{k-1} = \binom{n-1}{k-1}\binom{n+1}{k} - \binom{n}{k-1}\binom{n}{k}.$$

This shows we can think of Narayana numbers as 2×2 minors of Pascal's triangle, if we think of the triangle as large matrix. See Figure 9.3.

Problem 113 (Catalania = Catalan Mania)**.**
There are over two hundred (!) different sets of combinatorial objects that are enumerated by Catalan numbers. It is great fun to find bijections between these sets, and to try to count them in a manner that gives Narayana numbers. Try to make these connections with the following sets. In each case, the five objects corresponding to $n = 3$ are listed.

1. (123-avoiding permutations) A permutation is 123-*avoiding* if there is no triple of indices $i < j < k$ such that $w(i) < w(j) < w(k)$. The set of 123-avoiders is denoted $S_n(123)$. We have $S_3(123) = \{132, 213, 231, 312, 321\}$.
2. (321-avoiding permutations) A permutation is 321-*avoiding* if there is no triple of indices $i < j < k$ such that $w(k) < w(j) < w(i)$. The set of 321-avoiders is denoted $S_n(321)$. We have $S_3(321) = \{123, 132, 213, 231, 312\}$.
3. (Noncrossing partitions) A *noncrossing partition* is a set partition of $\{1, 2, \ldots, n\}$ such that no two of its blocks, say A and B, contain members $a, c \in A$ and $b, d \in B$ such that $a < b < c < d$. Here are the noncrossing partitions on three elements: $\{\{1\},\{2\},\{3\}\}$, $\{\{1,2\},\{3\}\}$, $\{\{1,3\},\{2\}\}$, $\{\{1\},\{2,3\}\}$, and $\{\{1,2,3\}\}$. These are all set partitions of $\{1,2,3\}$. For $n = 4$, there are 15 set partitions, and they are all noncrossing except for $\{\{1,3\},\{2,4\}\}$.
4. (Balanced parenthesizations) A sequence of parentheses is *balanced* if it can be parsed syntactically. In other words, there should be the same number of open parentheses "(" and closed parentheses ")", and when reading from left to right there should never be more closed parentheses than open. Here are the five balanced parenthesizations containing three pairs: ()()(), ()(()), (()()), (())(), ((())).
5. (Two-row standard Young tableaux) A *standard Young tableau* is a two-dimensional array of numbers (from 1 to the number of entries in the array) that increases across rows and down columns. Let $SYT(2, n)$ denote the number of standard Young tableaux in a $2 \times n$ rectangular array. For $n = 3$, these are:

1	2	3
4	5	6

1	2	4
3	5	6

1	2	5
3	4	6

1	3	4
2	5	6

1	3	5
2	4	6

6. (Triangulations of a polygon) See Figure 9.4. Here we dissect an $(n+2)$-gon with n triangles. Notice the polygon is fixed in space, so one might as well label the vertices. (Incidentally, this is the problem that Euler was interested in when he studied the Catalan numbers!)
7. (Decreasing binary trees) See Figure 9.5. Here n is the number of vertices of degree 2. Notice that how these are drawn in the plane matters, i.e., left and right subtrees matter.

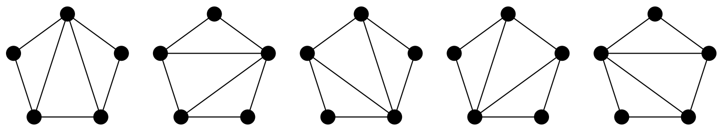


Fig. 9.4 Triangulations of a pentagon.

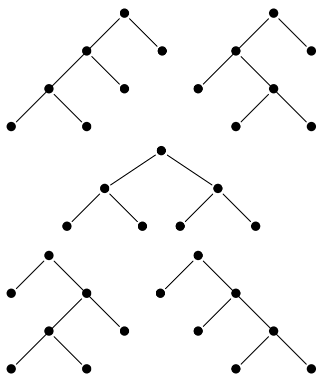


Fig. 9.5 Binary trees.

Weak order and the Tamari lattice

A CURIOUS READER may be wondering what is going on with the illustrations at the beginning of the last two chapters. These are shown in miniature in Figure 9.6.

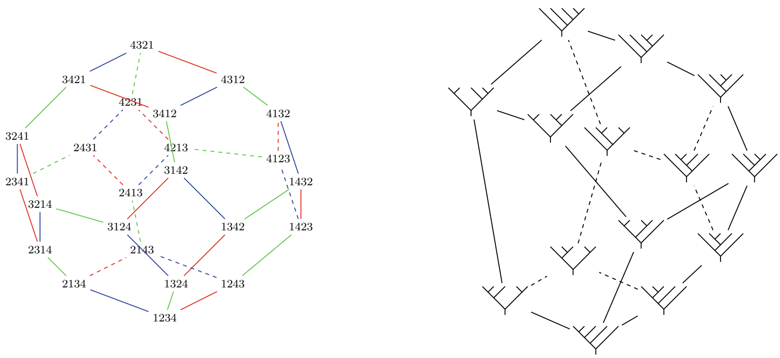


Fig. 9.6 The weak order and the Tamari lattice.

These are illustrations of certain partially ordered sets of combinatorial objects that also happen to have a geometric structure, in that the combinatorial objects can be placed at the vertices of polyhedra. The picture on the left shows the partial ordering known as the *weak order* on the set of all permutations in S_4. The picture on the right shows the partial ordering known as the *Tamari lattice* on the set of binary trees with four internal nodes. The corresponding geometric objects are known as the *permutahedron* and the *associahedron*. Let's learn a little more about these things.

Inversions

Pretend you have some books on a shelf, like in Figure 9.7. Just about the simplest way to sort the books into alphabetical order (by author) is to look along the shelf, from left to right, and when you come to two books that are out of order, you swap them. The new ordering of books might not yet be in alphabetical order, but it is closer than it was. Now look at the books again and see if anything is still out of order. Swap again if necessary. Eventually, after some number of swaps, all the books will be sorted in alphabetical order.

We will model the "sorted" order by the permutation $123\cdots n$, since the numbers appear in their natural order. Any other permutation has at least one pair of numbers that are out of order. The *inversion number* of a permutation

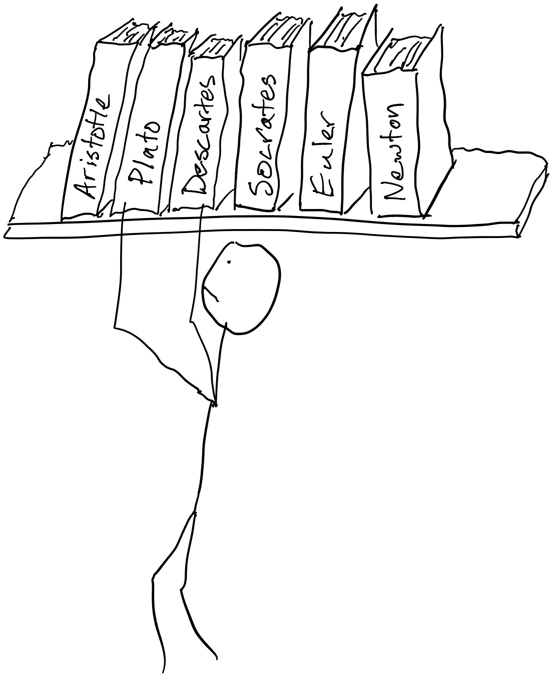


Fig. 9.7 Sorting books on a bookshelf.

w, denoted $\mathrm{inv}(w)$, is the total number of pairs of numbers that are out of order relative to one another.

For example, the permutation $w = 23514$ has $\mathrm{inv}(w) = 4$. The four inversion pairs are

$$\begin{aligned} w(1) &= 2 > 1 = w(4), \\ w(2) &= 3 > 1 = w(4), \\ w(3) &= 5 > 1 = w(4), \\ w(3) &= 5 > 4 = w(5). \end{aligned}$$

Notice that the inversion pairs don't need to be adjacent.

The number of inversions provides a natural way to quantify how close a permutation is to being sorted. That is, if a permutation has k inversions, then:

1. there is some sequence of k adjacent swaps that can be used to sort the permutation, and
2. no sequence of fewer than k adjacent swaps will sort the permutation.

This assertion can be proved by induction, by considering the effect of swapping any adjacent inversion pair. For example, if we swap the 5 and the 1 in $w = 23514$, we get $w' = 23154$, which has only three inversions since the "$5 > 1$" inversion has been resolved.

The full "scan and swap" sorting procedure for $w = 23514$ runs as follows (swaps are highlighted in bold):

$$23\mathbf{51}4 \to 2\mathbf{31}54 \to \mathbf{21}354 \to 123\mathbf{54} \to 12345.$$

Note we are always choosing the leftmost adjacent inversion in this procedure.

It turns out that, much like counting permutations according to number of cycles as we did at the end of Chapter 4, the generating function for counting permutations by inversions has a nice closed form. We have

$$\sum_{w \in S_n} q^{\mathrm{inv}(w)} = (1)(1+q)\cdots(1+q+\cdots+q^{n-1}) = \prod_{i=1}^{n} \frac{(1-q^i)}{(1-q)},$$

where the second equation follows from the algebraic identity $(1 - q^k) = (1-q)(1+q+\cdots+q^{k-1})$. This result is known as Rodrigues' Theorem.

One proof of the theorem is via induction and can be understood with an edge-labeled decision tree as shown in Figure 9.8. This is analogous to the tree in Figure 1.9. Level n of this tree corresponds to the permutations in S_n, and multiplying the edge weights on the path from 1 to w in the tree gives $q^{\mathrm{inv}(w)}$. We can see that the sum of the weights in each level corresponds to $(1+q+\cdots+q^{n-1})$ times the sum of the weights in the previous level, and so Rodrigues' Theorem follows.

Weak order

Setting aside the enumerative aspects of inversion numbers, let's return to the original problem of sorting with adjacent swaps. Draw all your permutations in S_n on a piece of paper, and draw an arrow from one permutation to another if you can swap two adjacent entries and reduce the number of inversions. For example, we would have an arrow

$$23\mathbf{51}4 \longrightarrow 23\mathbf{15}4$$

as we can swap the 5 and the 1 in 23514 to get 23154, and the latter permutation has fewer inversions.

There are $n-1$ adjacent positions in any permutation of S_n, so each permutation w will have $n-1$ arrows connected to it: some arrows will be coming in and some arrows will be going out.

Notice that $123\cdots n$ has no inversions, so all its arrows are incoming, while $n\cdots 321$ has nothing but inversion pairs, so all arrows are outgoing. In graph theory terminology, $n\cdots 321$ is a *source* and $123\cdots n$ is a *sink*. In general, the

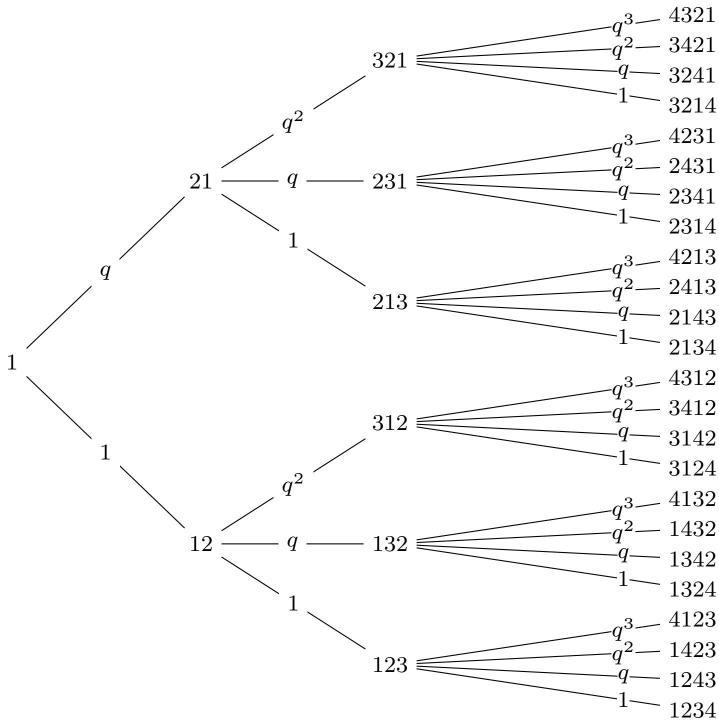


Fig. 9.8 The permutation growth tree with edge labels corresponding to the inversion number.

number of outgoing arrows for the permutation w is the number of descents, since descents are precisely adjacent inversions.

The illustration of the weak order on the left of Figure 9.6 is oriented so that all edges are pointing downward. In the smaller case of S_3, we would draw the directed graph in Figure 9.9.

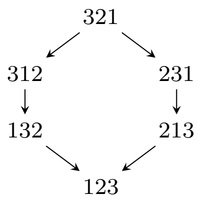


Fig. 9.9 The weak order on S_3.

Since the arrows are defined to strictly decrease the number of inversions, we know that no loops can occur, and that the arrows define a transitive relation. This gives a partial (not necessarily total) ordering on the set S_n which is known as the *weak order* (or sometimes the *weak Bruhat order*).

Permutahedron

We can see from our small-dimensional pictures that it appears as though there is a geometry to the weak order. This geometric structure is captured in the *permutahedron*. This polytope is defined by taking the convex hull of the orbit of a generic point in $\mathbb{R}^n$. Its vertices obviously correspond to permutations, but what takes a little more effort to see is that the edges correspond to swapping adjacent entries.

For example, take the point $\mathbf{x} = (0, 1, 2)$ in $\mathbb{R}^3$. Then by permuting coordinates we get a total of six points

$$(0, 1, 2), (0, 2, 1), (1, 0, 2), (1, 2, 0), (2, 0, 1), (2, 1, 0),$$

all of which sit in the plane defined by $x + y + z = 3$. The convex hull of these is a hexagon. This is illustrated in Figure 9.10. Compare with Figure 2.3 from Chapter 2.

One detail to notice about the permutahedron versus the weak order is the labeling of the points. Each point can be identified with a permutation by considering the relative ordering of the coordinates. That is, the point $(x_1, x_2, \ldots, x_n)$ is identified with the permutation w such that $x_{w(1)} < x_{w(2)} < \cdots < x_{w(n)}$. For example, the point $(3, 2, 5, 1, 4, 0)$ in $\mathbb{R}^6$ would correspond to the permutation $w = 642153$. In Figure 9.10, we have put the corresponding permutations below each point.

Tamari lattice

The Tamari lattice is named after Dov Tamari who first described it in 1962, though it was independently discovered by James Stasheff around 1963. It is a partial ordering on parenthesizations of a string of $n + 1$ symbols.

For example, consider all valid parenthesizations of $wxyz$, which we can think of as an associative product of four elements. They are

$$(((wx)y)z), ((wx)(yz)), ((w(xy))z), (w((xy)z)), (w(x(yz))).$$

These parenthesizations can be given a partial order by declaring that $((fg)h) < (f(gh))$ for any sub-expressions f, g and h. We then extend this relation by transitivity.

We can visualize these parenthesizations of $n + 1$ elements with planar binary trees with n internal nodes and $n + 1$ leaves. For example, we see the

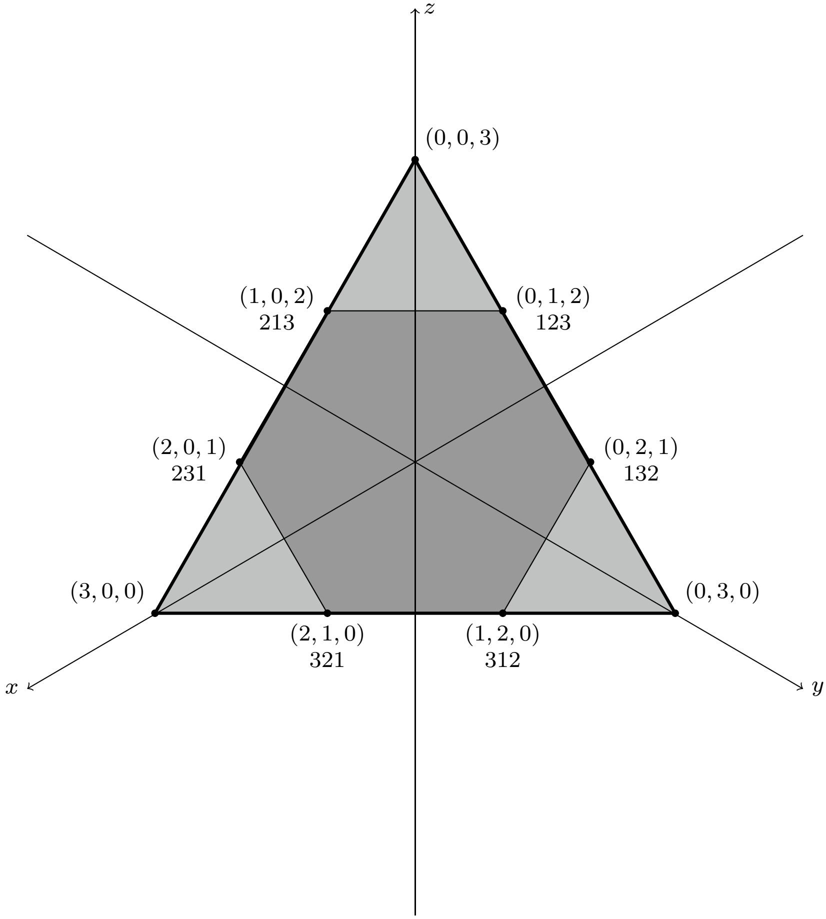


Fig. 9.10 The convex hull of permutations of the point $\mathbf{x} = (0, 1, 2)$ is a hexagon.

Tamari lattice for $n + 1 = 4$ symbols in Figure 9.11. The illustration on the right of Figure 9.6 shows the Tamari lattice with $n+1 = 5$ symbols in terms of trees.

Associahedron

Just as the weak order can be realized with the polytope known as the permutahedron, the Tamari lattice can be realized with a polytope known as the *associahedron*. Realizing the associahedron in Euclidean space is not as simple

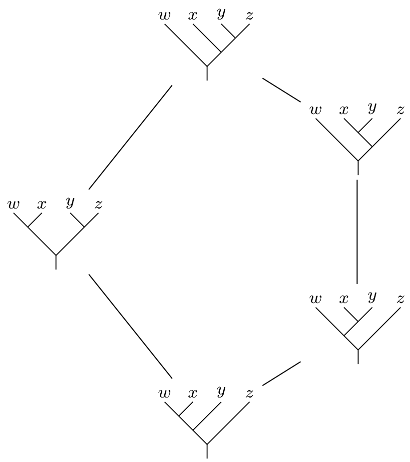


Fig. 9.11 The Tamari lattice for $n+1=4$ symbols.

as realizing the permutahedron, although it was immediately conjectured in the 1960s that the "Stasheff polytope" (as the associahedron is sometimes known) is, in fact, a polytope. Despite this early guess, it was not until the 1980s that a proof of this fact was first published. Today there are many different constructions and we will not attempt to describe them here.

However, it is worth noting that there is a perspective due to Alex Postnikov that puts both the permutahedron and the associahedron into a continuous family of polytopes known as "generalized permutahedra." In this way the associahedron is obtained by a certain deformation of the permutahedron. For the partially ordered sets, this means the Tamari lattice can be obtained from the weak order. One way to state this in purely combinatorial terms is that the Tamari lattice is equivalent to the weak order restricted to the 231-avoiding permutations! See Figure 9.12.

Further reading

- Richard Stanley, "Catalan numbers," Cambridge University Press, (2015). Richard Stanley has the definitive catalogue of combinatorial objects counted by Catalan numbers. There are over two hundred different examples!

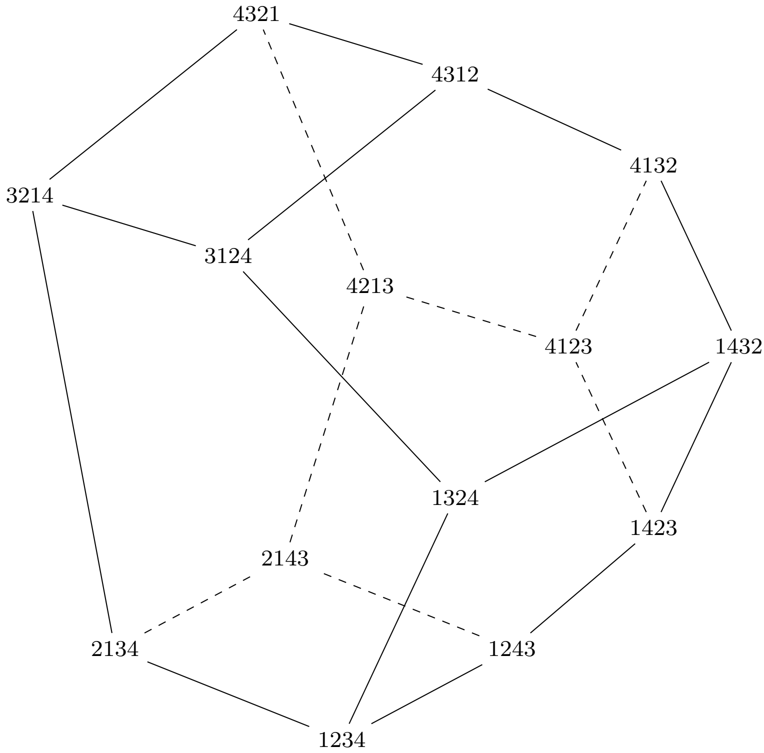
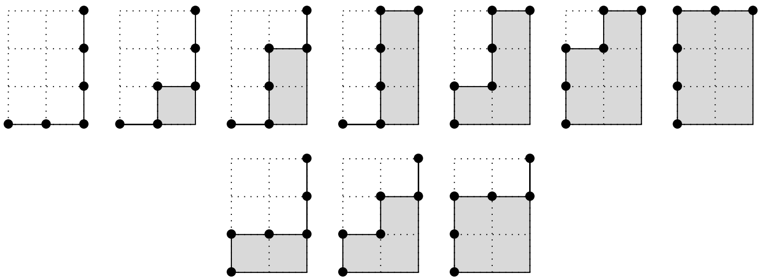


Fig. 9.12 The Tamari lattice for $n+1=5$ symbols, given by the weak order on 231-avoiding permutations in S_4.

Chapter 10
Refined Enumeration

"The whole of science is nothing more than a refinement of everyday thinking."
–Albert Einstein



Not all paths are created equal.

MAJOR PERCY ALEXANDER MACMAHON was an early 20th-century mathematician who wrote one of the first major books on enumerative combinatorics. Among many other things, he studied permutations and certain permutation statistics that we will investigate in this chapter. We can view this work as a variation on a theme: when counting a set of objects (like permutations) we may be able to keep track of more refined properties of these objects along the way. In this chapter we will revisit some earlier counting problems with an eye toward refinement.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_10

$n\backslash k$	0	1	2	3	4	5	6	7	8	9	10
1	1										
2	1	1									
3	1	2	2	1							
4	1	3	5	6	5	3	1				
5	1	4	9	15	20	22	20	15	9	4	1

Table 10.1 Triangle of the Mahonian numbers $M(n,k)$. Or, the inversion numbers $I(n,k)$, the number of permutations in S_n with k inversions.

Inversions of a permutation

Definition 29. An *inversion* of a permutation $w = w(1)\cdots w(n)$ is a pair (i,j) such that $i < j$ and $w(i) > w(j)$. We let $\mathrm{Inv}(w)$ denote the set of inversions, and we let $\mathrm{inv}(w)$ denote the number of inversions of a permutation w, i.e.,

$$\mathrm{inv}(w) = |\,\mathrm{Inv}(w)| = |\{(i,j) : i < j \text{ and } w(i) > w(j)\}|.$$

For example, $w = 31542$ has

$$\mathrm{Inv}(w) = \{(1,2),(1,5),(3,4),(3,5),(4,5)\}$$

and $\mathrm{inv}(w) = 5$. Notice that descents $w(i) > w(i+1)$ correspond to inversions of the form $(i, i+1)$. Here is a table with the number of inversions and descents for permutations in S_3.

w	$\mathrm{inv}(w)$	$\mathrm{des}(w)$
123	0	0
132	1	1
213	1	1
231	2	1
312	2	1
321	3	2

Inversion generating functions and inversion numbers

Definition 30. Let

$$S_n^{\mathrm{inv}}(q) = \sum_{w \in S_n} q^{\mathrm{inv}(w)} = \sum_{k \geq 0} I(n,k) q^k,$$

where $I(n,k)$ denotes the number of permutations in S_n with k inversions.

We have $S_1^{\mathrm{inv}}(q) = 1$, $S_2^{\mathrm{inv}}(q) = 1 + q$, and the table for S_3 above shows

$$S_3^{\mathrm{inv}}(q) = 1 + 2q + 2q^2 + q^3.$$

Warmup 32. Make a table showing the number of inversions and descents for permutations in S_4 and use it to write down $S_4^{\mathrm{inv}}(q)$.

Problem 114.
What is the inversion number $I(5,4)$? That is, how many permutations w in S_5 have $\mathrm{inv}(w) = 4$?

Problem 115.
Let

$$\mathrm{Inv}'(w) = \{(w(i), w(j)) : i < j \text{ and } w(i) > w(j)\}.$$

Compute $\mathrm{Inv}'(w)$ for $w = 31542$ and explain its connection to $\mathrm{Inv}(w)$.

Prove that $|\mathrm{Inv}(w)| = |\mathrm{Inv}'(w)|$ for any permutation w in S_n.

Problem 116.
Prove that for each permutation w in S_n,

$$0 \leq \mathrm{inv}(w) \leq \binom{n}{2},$$

and these bounds are sharp. Describe the (unique) permutation w in S_n such that $\mathrm{inv}(w) = 0$ and the (unique) permutation w in S_n such that $\mathrm{inv}(w) = \binom{n}{2}$.

Problem 117.
Prove that the number of permutations in S_n with k inversions equals the number of permutations with $\binom{n}{2} - k$ inversions. That is, for any $n \geq 1$ and any $0 \leq k \leq \binom{n}{2}$,

$$I(n,k) = I\left(n, \binom{n}{2} - k\right).$$

Problem 118.
Prove that for any $n \geq 1$,

$$S_{n+1}^{\text{inv}}(q) = \sum_{k=0}^{n} q^k S_n^{\text{inv}}(q) = (1 + q + q^2 + \cdots + q^n) S_n^{\text{inv}}(q),$$

and conclude by induction that

$$S_n^{\text{inv}}(q) = \prod_{i=1}^{n} \left(\sum_{j=0}^{i-1} q^j \right).$$

(See Figure 9.8 from the end of the last chapter for a hint.) This is sometimes known as *Rodrigues' Theorem*, named for nineteenth-century mathematician Benjamin Olinde Rodrigues.

Major index of a permutation

Definition 31. The *major index* of a permutation w in S_n, denoted $\text{maj}(w)$, is the sum of the elements in the descent set (with the empty set having sum zero). That is,

$$\text{maj}(w) = \sum_{i \in \text{Des}(w)} i,$$

where we recall the descent set is $\text{Des}(w) = \{i : w(i) > w(i+1)\}$.

For example, the permutation $w = 31542$ has descent set $\text{Des}(w) = \{1, 3, 4\}$, so $\text{maj}(w) = 1 + 3 + 4 = 8$.

The major index gets its name from Major Percy Alexander MacMahon, a late 19th- and early 20th-century combinatorialist. MacMahon was an officer in the British army until 1898, retiring at the rank of major. When MacMahon studied the statistic in Definition 31 he called it the "greater index" of a permutation.

Here is a table with the major index and descent number for permutations in S_3.

w	$\text{maj}(w)$	$\text{des}(w)$
123	0	0
132	2	1
213	1	1
231	2	1
312	1	1
321	3	2

Major index generating function and Mahonian numbers

Definition 32. Let

$$S_n^{\mathrm{maj}}(q) = \sum_{w \in S_n} q^{\mathrm{maj}(w)} = \sum_{k \geq 0} M(n,k) q^k,$$

where $M(n,k)$ denotes the number of permutations in S_n with major index k. We call the coefficients $M(n,k)$ the *Mahonian numbers.*

We have $S_1^{\mathrm{maj}}(q) = 1$, $S_2^{\mathrm{maj}}(q) = 1 + q$, and the table for S_3 above shows

$$S_3^{\mathrm{maj}}(q) = 1 + 2q + 2q^2 + q^3.$$

Warmup 33. Make a table showing the major index and descent number for permutations in S_4, and use it to write down $S_4^{\mathrm{maj}}(q)$. Compare with Warmup 32.

Problem 119.
What is the Mahonian number $M(n,k)$? How many permutations w in S_5 have $\mathrm{maj}(w) = 4$?

Problem 120.
Prove that for each permutation w in S_n,

$$0 \leq \mathrm{maj}(w) \leq \binom{n}{2},$$

and these bounds are sharp. Describe the (unique) permutation w in S_n such that $\mathrm{maj}(w) = 0$ and the (unique) permutation w in S_n such that $\mathrm{maj}(w) = \binom{n}{2}$.

Problem 121.
Prove that the number of permutations in S_n with major index k equals the number of permutations with major index $\binom{n}{2} - k$. That is, for any $n \geq 1$ and any $0 \leq k \leq \binom{n}{2}$,

$$M(n,k) = M\left(n, \binom{n}{2} - k\right).$$

Problem 122.
Prove that for any $n \geq 1$,

$$S_{n+1}^{\mathrm{maj}}(q) = \sum_{k=0}^{n} q^k S_n^{\mathrm{maj}}(q) = (1 + q + q^2 + \cdots + q^n) S_n^{\mathrm{maj}}(q),$$

and conclude by induction that

$$S_n^{\mathrm{maj}}(q) = \prod_{i=1}^{n} \left(\sum_{j=0}^{i-1} q^j \right).$$

Further, by Problem 118, we now know that $S_n^{\mathrm{maj}}(q) = S_n^{\mathrm{inv}}(q)$ and hence $M(n,k) = I(n,k)$.

Problem 123.
Prove that $M(n,k) = I(n,k)$ by establishing a bijection between permutations with major index k and permutations with k inversions.

To simplify notation in formulas like those found in Problem 122, we use a special notation.

q-integer notation and q-factorials

Definition 33. Let

$$[n] = 1 + q + q^2 + \cdots + q^{n-1},$$

which is known as a *q-integer*. Similarly, we write $[n]!$ to mean the product $[n][n-1]\cdots[2][1]$.

Now we can write the result of Problems 118 and 122 as follows.

q-analogues of $n!$

Theorem 7. For any $n \geq 1$,

$$S_n^{\mathrm{inv}}(q) = S_n^{\mathrm{maj}}(q) = [n]!$$

This theorem is what is sometimes known as a *q-analogue*. That is, we have taken the formula for counting permutations, $n!$, and "q-ified" it. This is a refinement that has the parameter q keeping track of interesting information, while if we set $q = 1$ we obtain our original formula.

Another example of a q-analogue comes from the Binomial Theorem. We know there are 2^n subsets of an n-element set, and the q-analogue $[2]^n$ counts these subsets according to cardinality:

$$[2]^n = (1+q)^n = \sum_{S \subseteq \{1,2,\ldots,n\}} q^{|S|}.$$

In the remaining problems of this chapter we investigate some similar results.

Problem 124 (q-Binomial Coefficients).
This problem investigates the q-binomial coefficients, which are obtained by taking the q-analogue of the binomial formula. That is, for any $0 \leq k \leq n$,

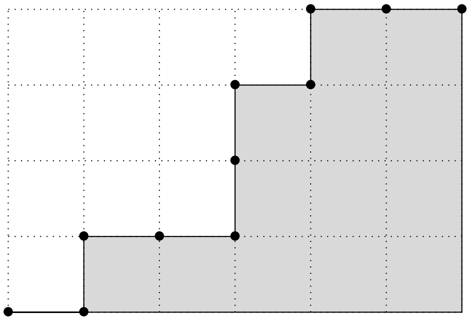


Fig. 10.1 The area beneath a lattice path.

define

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{[n]!}{[k]![n-k]!},$$

where $[0]! = 1$. For example

$$\begin{aligned} \begin{bmatrix} 4 \\ 2 \end{bmatrix} &= \frac{[4]!}{[2]![2]!} \\ &= \frac{[4][3]}{[2][1]} \\ &= \frac{(1+q+q^2+q^3)(1+q+q^2)}{(1+q)(1)} \\ &= (1+q^2)(1+q+q^2) = 1+q+2q^2+q^3+q^4. \end{aligned}$$

1. One of the first surprising things about q-binomial coefficients is that they are polynomials. (They are obviously rational functions of q, but why should the denominators always cancel?) For all $0 \le k \le n \le 4$, expand $\begin{bmatrix} n \\ k \end{bmatrix}$ as a polynomial in q, and arrange these polynomials in an array like Pascal's triangle.
2. Show $\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n \\ n-k \end{bmatrix}$.
3. Show that q-binomial coefficients satisfy a refinement of Pascal's recurrence. That is, $\begin{bmatrix} n \\ 0 \end{bmatrix} = \begin{bmatrix} n \\ n \end{bmatrix} = 1$ and for all $0 < k < n$,

$$\begin{bmatrix} n \\ k \end{bmatrix} = q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + \begin{bmatrix} n-1 \\ k \end{bmatrix}.$$

   Use induction to conclude that $\begin{bmatrix} n \\ k \end{bmatrix}$ is indeed a polynomial.
4. Show the degree of the polynomial $\begin{bmatrix} n \\ k \end{bmatrix}$ is $k(n-k)$.
5. Show that for any $0 \le k \le n$, the polynomial $\begin{bmatrix} n \\ k \end{bmatrix}$ has symmetric coefficients. That is, if

$$\begin{bmatrix} n \\ k \end{bmatrix} = \sum_{j=0}^{k(n-k)} c_j q^j,$$

then $c_j = c_{k(n-k)-j}$ for all j. (Hint: show replacing q by $1/q$ in $\begin{bmatrix} n \\ k \end{bmatrix}$ gives the same formula, up to a factor of $q^{k(n-k)}$.)

Problem 125 (q-Binomial Coefficients revisited).
We now give a combinatorial interpretation to the q-binomial coefficients. We know that $\binom{n}{k}$ counts lattice paths (to Starfolks, say) on a grid that is k blocks East by $(n-k)$ blocks North. Let $L(k, n-k)$ denote the set of such paths. For a given path p in $L(k, n-k)$, let area(p) denote the area under the path (or the number of blocks southeast of the path). For example, Figure 10.1 shows a path on a six by four grid ($n = 10, k = 6$) with area$(p) = 13$.

Define the generating function for paths in $L(k, n-k)$ by

$$L_{n,k}(q) = \sum_{p \in L(k,n-k)} q^{\text{area}(p)}.$$

1. Compute $L_{n,k}(q)$ for all $0 \le k \le n \le 4$ and arrange these polynomials in an array like Pascal's triangle.
2. Show that $L_{n,k}(q) = \begin{bmatrix} n \\ k \end{bmatrix}$. (Hint: show both sets of polynomials satisfy the same recurrence with the same boundary conditions.)
3. The definition of $L_{n,k}(q)$ makes it obvious that $L_{n,k}(q)$ is a polynomial, whereas it is not obvious that $\begin{bmatrix} n \\ k \end{bmatrix}$ is a polynomial from its definition. In a similar way, use combinatorial properties of $L_{n,k}(q)$ to deduce parts 2), 4), and 5) of Problem 124.

Problem 126 (q-Catalan numbers).
Using the q-binomial coefficients as a starting point, let $C_n(q) = \frac{1}{[n+1]}\begin{bmatrix} 2n \\ n \end{bmatrix}$ denote the q-analogue of the Catalan numbers.

1. Show that
$$C_n(q) = \begin{bmatrix} 2n \\ n \end{bmatrix} - q\begin{bmatrix} 2n \\ n+1 \end{bmatrix},$$
and conclude that $C_n(q)$ is a polynomial.
2. Show that the degree of $C_n(q)$ is $n(n-1)$.
3. For a lattice path p, we can define a *valley* as positions in which an East step is followed immediately by a North step. For example, the path in Figure 10.1 has $p = ENEENNENEE$, so its set of valleys is $\{1, 4, 7\}$. The *major index* of a path is the sum of its valleys. With this example, we would have maj$(p) = 1 + 4 + 7 = 12$.
Show that
$$C_n(q) = \sum_{p \in \text{Dyck}(n)} q^{\text{maj}(p)}.$$

Problem 127 (Euler-Mahonian distribution).
In this problem we will study a q-analogue of the Eulerian numbers via the joint distribution of permutation runs and major index. This joint distribution is known as the *Euler-Mahonian* distribution.

That is, define the q-Eulerian numbers $\left\langle {n \atop k} \right\rangle^{\mathrm{maj}}$ as

$$\left\langle {n \atop k} \right\rangle^{\mathrm{maj}} = \sum_{w \in S_n, \mathrm{runs}(w)=k} q^{\mathrm{maj}(w)}.$$

By analogy with the Eulerian polynomials, denote the generating function for the joint distribution of major index and runs (or descents) as follows:

$$S_n^{\mathrm{maj}}(q,t) = \sum_{w \in S_n} q^{\mathrm{maj}(w)} t^{\mathrm{runs}(w)} = \sum_{k=1}^{n} \left\langle {n \atop k} \right\rangle^{\mathrm{maj}} t^k.$$

For example,

$$S_3^{\mathrm{maj}}(q,t) = t + (2q + 2q^2)t^2 + q^3 t^3.$$

Note that $S_n^{\mathrm{maj}}(q,1) = [n]!$ and $S_n^{\mathrm{maj}}(1,t) = A_n(t)$ (the Eulerian polynomial from Chapter 8).

1. For each $n \leq 4$, compute $S_n^{\mathrm{maj}}(q,t)$ and group terms according to powers of t to see some small examples of q-Eulerian numbers.
2. Explain why

$$\left\langle {n \atop k} \right\rangle^{\mathrm{maj}} = [k] \left\langle {n-1 \atop k} \right\rangle^{\mathrm{maj}} + q^{k-1}[n+1-k] \left\langle {n-1 \atop k-1} \right\rangle^{\mathrm{maj}}.$$

   Compare with Problem 97.
3. Show that for $n \geq 0$,

$$\frac{1}{(1-t)(1-qt)\cdots(1-q^n t)} = \sum_{k \geq 0} \begin{bmatrix} n+k \\ n \end{bmatrix} t^k.$$

   As hint, use induction and the fact that

$$\begin{bmatrix} n+k \\ n \end{bmatrix} = \begin{bmatrix} n+k-1 \\ n-1 \end{bmatrix} + q^n \begin{bmatrix} n+k-1 \\ n \end{bmatrix}.$$

4. Show that for any $n \geq 0$,

$$\frac{S_n^{\mathrm{maj}}(q,t)}{(1-t)(1-qt)\cdots(1-q^n t)} = \sum_{k \geq 0} [k+1]^n t^{k+1}.$$

   Compare with Problem 98.

Continued fractions

WHAT IS

$$x = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}}? \tag{10.1}$$

If you like mathematical puzzles, you may have encountered this sort of question before.

How are we to make sense of an expression like this? It seems that x is defined as the limit of some sort of process of repeated division, but how exactly is it defined? And assuming the process is well-defined, does the limit exist? These are valid concerns, but for now, let us dismiss such worries and push forward with the assumption that there is a real number satisfying (10.1).

An intuitive idea for finding x is to recognize that x seems to be expressed in terms of itself, in that everything below the first fraction bar is the same as x itself. That is,

$$x = 1 + \frac{1}{x},$$

which is equivalent to $x^2 = x + 1$, or $x^2 - x - 1 = 0$. This equation has roots

$$\frac{1+\sqrt{5}}{2} \approx 1.618 \quad \text{and} \quad \frac{1-\sqrt{5}}{2} \approx -.618.$$

Since x is clearly bigger than 1 (it is 1 plus positive stuff), we know $x = \frac{1+\sqrt{5}}{2} = \varphi$. This is our friend from the end of Chapter 5: the golden ratio!

An expression like Equation (10.1), with a collection of nested fractions, is called a *continued fraction*. Continued fractions have been studied for centuries. They provide a representation of real numbers that is quite distinct from the usual decimal representation. It can be fun to try to show that

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}}}$$

or

$$e = 2 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}},$$

where $a_1, a_2, a_3, \ldots$ is the sequence

$$1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, 1, 12, 1, \ldots .$$

Continued fractions provide a sequence of rational approximations to a real number. We can see this by truncating the expression after one fraction bar, after two fraction bars, three fraction bars, and so on. (Of course decimal representations do this too, by truncating after one decimal place, two decimal places, three decimal places, and so on.) For example, Equation (10.1) tells us that there is a sequence of rational numbers that converge to $(1+\sqrt{5})/2$ in the limit. The sequence begins

$$\begin{aligned}
x_0 &= 1, \\
x_1 &= 1 + \frac{1}{1} = 1 + \frac{1}{x_0} = 2, \\
x_2 &= 1 + \frac{1}{1+\frac{1}{1}} = 1 + \frac{1}{x_1} = \frac{3}{2}, \\
x_3 &= 1 + \frac{1}{1+\frac{1}{1+\frac{1}{1}}} = 1 + \frac{1}{x_2} = \frac{5}{3}, \\
&\vdots
\end{aligned}$$

Based on our discussion at the end of Chapter 5 about the golden ratio, we are not too surprised to find this expression:

$$x_{n+1} = 1 + \frac{1}{x_n} = \frac{f_{n+1}}{f_n},$$

where the f_n are the Fibonacci numbers beginning with $f_0 = f_1 = 1$.

For $\sqrt{2}$, we get the sequence

$$1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \frac{99}{70}, \frac{239}{169} \approx 1.4142, \ldots,$$

and for e, we get

$$2, 3, \frac{8}{3}, \frac{11}{4}, \frac{19}{7}, \frac{87}{32}, \frac{106}{39} \approx 2.7179, \ldots .$$

Continued fractions for generating functions

In Chapter 9, we found the generating function for the Catalan numbers,

$$C(z) = \sum_{n\geq 0} C_n z^n = 1 + z + 2z^2 + 5z^3 + 14z^4 + 42z^5 + \cdots ,$$

is given by

$$C(z) = \frac{1 - \sqrt{1-4z}}{2z}. \tag{10.2}$$

While the closed-form expression for $C(z)$ is not rational, we can use the idea of partial fractions to find a sequence of rational generating functions that approximate $C(z)$.

Recall that we obtained the closed-form expression for $C(z)$ by first showing $C(z)$ satisfies the identity

$$zC(z)^2 - C(z) + 1 = 0,$$

or equivalently,

$$C(z)(1 - zC(z)) = 1.$$

Dividing by $(1 - zC(z))$ on both sides yields

$$C(z) = \frac{1}{1 - zC(z)}, \tag{10.3}$$

which we can then "unravel" recursively to get

$$\begin{aligned} C(z) &= \frac{1}{1 - zC(z)}, \\ &= \cfrac{1}{1 - \cfrac{z}{1 - zC(z)}}, \\ &= \cfrac{1}{1 - \cfrac{z}{1 - \cfrac{z}{1 - zC(z)}}}, \\ &= \cfrac{1}{1 - \cfrac{z}{1 - \cfrac{z}{1 - \cfrac{z}{1 - \cdots}}}}, \end{aligned} \tag{10.4}$$

This suggests that we ought to consider the following sequence of functions:

$$
\begin{aligned}
C^{(1)}(z) &= \frac{1}{1-z}, \\
C^{(2)}(z) &= \cfrac{1}{1-\cfrac{z}{1-z}} &&= \frac{1-z}{1-2z}, \\
C^{(3)}(z) &= \cfrac{1}{1-\cfrac{z}{1-\cfrac{z}{1-z}}} &&= \frac{1-2z}{1-3z+z^2}, \\
C^{(4)}(z) &= \frac{1}{1-zC^{(3)}(z)} &&= \frac{1-3z+z^2}{1-4z+3z^2}, \\
C^{(5)}(z) &= \frac{1}{1-zC^{(4)}(z)} &&= \frac{1-4z+3z^2}{1-5z+6z^2-z^3}, \\
&\vdots
\end{aligned}
$$

Expanding, we find

$$
\begin{aligned}
C^{(1)}(z) &= \mathbf{1} + \mathbf{z} + z^2 + z^3 + z^4 + z^5 + z^6 + \cdots, \\
C^{(2)}(z) &= \mathbf{1} + \mathbf{z} + \mathbf{2z^2} + 4z^3 + 8z^4 + 16z^5 + 32z^6 + \cdots, \\
C^{(3)}(z) &= \mathbf{1} + \mathbf{z} + \mathbf{2z^2} + \mathbf{5z^3} + 13z^4 + 34z^5 + 89z^6 + \cdots, \\
C^{(4)}(z) &= \mathbf{1} + \mathbf{z} + \mathbf{2z^2} + \mathbf{5z^3} + \mathbf{14z^4} + 41z^5 + 122z^6 + \cdots, \\
C^{(5)}(z) &= \mathbf{1} + \mathbf{z} + \mathbf{2z^2} + \mathbf{5z^3} + \mathbf{14z^4} + \mathbf{42z^5} + 131z^6 + \cdots, \\
&\vdots
\end{aligned}
$$

The bold terms are those that agree with the Catalan number generating function, and we can see that each iteration gives us more terms that agree with $C(z)$. In particular, $C^{(n)}(z)$ is accurate up to the coefficient of z^n.

Before we move on, we briefly mention a curiosity that some alert readers may have noticed. If we take the limit as $z \to -1^+$ in Equation (10.4), we get:

$$\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}},$$

which is precisely one less than the expression in Equation (10.1). Moreover, taking the same limit in Equation (10.2) yields

$$\lim_{z\to -1^+} \frac{1-\sqrt{1-4z}}{2z} = \frac{1-\sqrt{5}}{-2} \approx .618,$$

which is one less than the golden ratio, as expected. It turns out that this is also the reciprocal of the golden ratio! That is, $1/\varphi = \varphi - 1$, so

$$\lim_{n\to\infty} \frac{f_n}{f_{n+1}} = \varphi - 1 \approx .618.$$

Hmm... what happens when we take $z \to -1^+$ in the functions $C^{(n)}(z)$?

Refinements via statistics for Dyck paths

We can give a combinatorial argument for the continued fraction expression for $C(z)$ by using Dyck paths, and moreover, this argument allows us to track several statistics for Dyck paths along the way.

Recall from Chapter 9 that Dyck paths of length $2n$ are lattice paths from $(0,0)$ to (n,n) that do not pass below the line $y = x$. We let $\mathrm{Dyck}(n)$ denote the set of all such paths, and if $p \in \mathrm{Dyck}(n)$, we write $|p| = n$. Further, we let $\mathrm{Dyck} = \bigcup \mathrm{Dyck}(n)$ denote the set of all Dyck paths. We have $C_n = |\,\mathrm{Dyck}(n)|$, so

$$C(z) = \sum_{n\geq 0} C_n z^n = \sum_{p\in \mathrm{Dyck}} z^{|p|}.$$

Call a Dyck path *prime* if the only points at which it touches the line $y = x$ occur at $(0,0)$ and (n,n). For convenience, we do *not* consider the empty path that starts and ends at $(0,0)$ to be prime. Notice that each nonempty path $p \in \mathrm{Dyck}$ has a unique "prime decomposition" into concatenated prime paths, $p_1, p_2, \ldots$. (If the empty path was prime, then such factorizations wouldn't be unique.) See, for example, Figure 10.2.

Write $\mathrm{Dyck}'(n)$ for the set of all prime Dyck paths of length n and let $\mathrm{Dyck}' = \bigcup \mathrm{Dyck}'(n)$ denote the set of all prime Dyck paths. We let $P(z)$ denote the generating function for prime Dyck paths according to size, i.e.,

$$P(z) = \sum_{p\in \mathrm{Dyck}'} z^{|p|}.$$

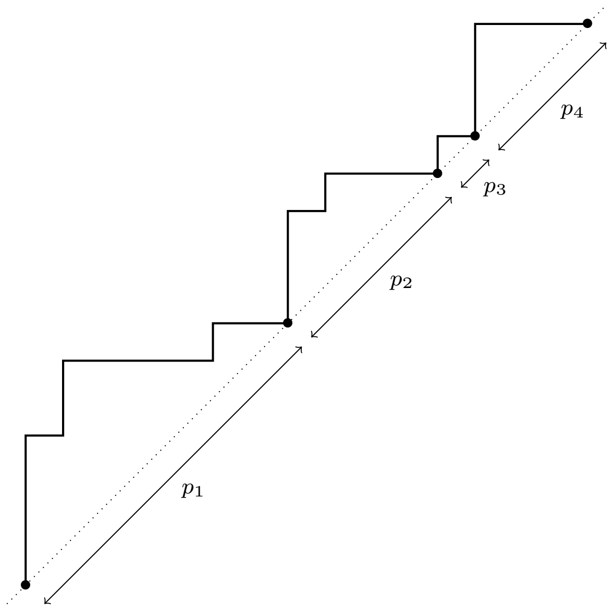


Fig. 10.2 A Dyck path decomposed into prime Dyck paths.

Notice that if we concatenate two paths, size is additive. That is, if $p_1p_2 = p$, we have $|p| = |p_1| + |p_2|$. This means $P(z)^2$ is the generating function for all Dyck paths with exactly two prime factors, $P(z)^3$ is the generating function for all Dyck paths with exactly three prime factors, and so on.

Grouping paths according to the number of prime factors, we have the following identity:

$$C(z) = 1 + P(z) + P(z)^2 + P(z)^3 + \cdots = \frac{1}{1 - P(z)}. \tag{10.5}$$

Here is the interesting part: apart from the first North step and the final East step, a prime path does not go below the line $y = x + 1$. This means that a prime path can be uniquely written as $p' = NpE$ for some (not necessarily prime) path p of length n. The correspondence $p' \leftrightarrow p$ gives a bijection between $\mathrm{Dyck}'(n+1)$ and $\mathrm{Dyck}(n)$ for each n. In terms of generating functions, $P(z) = zC(z)$. Thus, Equation (10.5) amounts to $C(z) = 1/(1 - zC(z))$, as in (10.3).

But there is more insight to glean from Equation (10.5) if we keep track of some statistics other than size.

For one, there is a natural q-analogue of the Catalan numbers obtained by keeping track of the area below the lattice path. Since Dyck paths do not go below the line $y = x$, we will normalize the area statistic, so for a Dyck path p, $\mathrm{area}(p)$ counts the number of unit squares below the path and above the line $y = x$. This implies $0 \le \text{area}(p) \le \binom{n}{2}$ for any path p in Dyck(n). For example, the path in Figure 10.3 is a member of Dyck(8) with area 7.

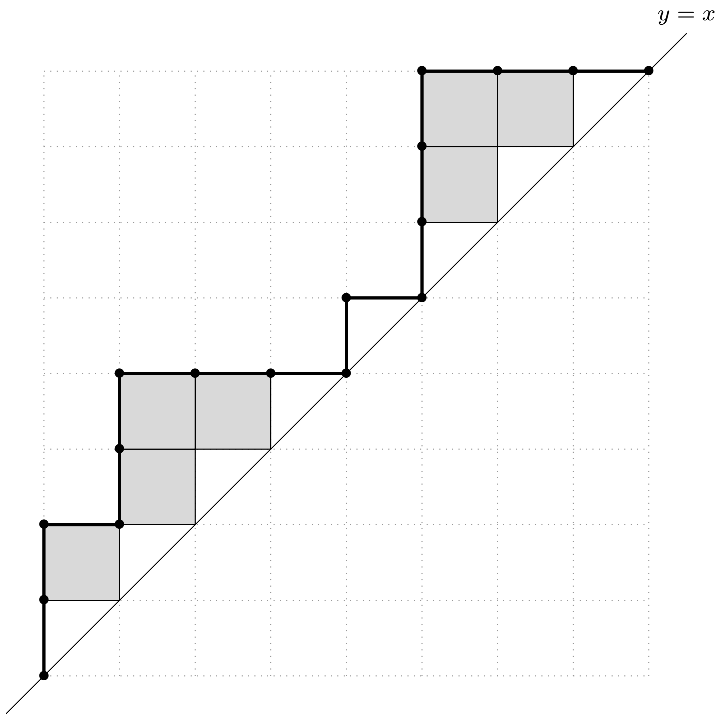


Fig. 10.3 A path in Dyck(8) that is the concatenation of three prime paths. It has four peaks and area seven. Note that length, area, and number of peaks are all additive.

Since area is easily seen to be additive when concatenating paths, area $(p_1p_2) = \text{area}(p_1) + \text{area}(p_2)$, we immediately get

$$C(q,z) = 1 + P(q,z) + P(q,z)^2 + P(q,z)^3 + \cdots = \frac{1}{1 - P(q,z)},$$

where

$$C(q,z) = \sum_{p \in \text{Dyck}} q^{\text{area}(p)} z^{|p|} \quad \text{and} \quad P(q,z) = \sum_{p' \in \text{Dyck}'} q^{\text{area}(p')} z^{|p'|}.$$

Moreover, if $p \in \text{Dyck}(n)$ is any path, then NpE is prime with area $\text{area}(NpE) = \text{area}(p) + n = \text{area}(p) + |p|$. This shows

$$\begin{aligned} P(q,z) &= \sum_{p\in\mathrm{Dyck}} q^{\mathrm{area}(NpE)} z^{|p|+1}, \\ &= z \sum_{p\in\mathrm{Dyck}} q^{\mathrm{area}(p)+|p|} z^{|p|}, \\ &= z \sum_{p\in\mathrm{Dyck}} q^{\mathrm{area}(p)} (qz)^{|p|}, \\ &= zC(q,qz). \end{aligned}$$

Now we have a new, refined identity:

$$C(q,z) = \frac{1}{1-zC(q,qz)},$$

which leads to a refined continued fraction generating function:

$$C(q,z) = \cfrac{1}{1-\cfrac{z}{1-\cfrac{qz}{1-\cfrac{q^2z}{1-\cfrac{q^3z}{1-\cdots}}}}}$$

But wait... there's more! We can also produce a continued fraction for counting Dyck paths by area and *peaks*, since the number of peaks is additive with respect to concatenation as well: $\mathrm{pk}(p_1p_2) = \mathrm{pk}(p_1) + \mathrm{pk}(p_2)$. Let

$$C(q,t,z) = \sum_{p\in\mathrm{Dyck}} q^{\mathrm{area}(p)} t^{\mathrm{pk}(p)} z^{|p|},$$

and

$$P(q,t,z) = \sum_{p'\in\mathrm{Dyck}'} q^{\mathrm{area}(p')} t^{\mathrm{pk}(p')} z^{|p'|}.$$

Our key observation is (still!) that C is the geometric series in P:

$$C(q,t,z) = 1 + P(q,t,z) + P(q,t,z)^2 + P(q,t,z)^3 + \cdots = \frac{1}{1-P(q,t,z)}.$$

It remains to write P in terms of C, and this is a little bit delicate.

Notice the number of peaks in $p' = NpE$ is $\mathrm{pk}(p) = \mathrm{pk}(p')$, which is good! ... unless p is the empty path, in which case $\mathrm{pk}(NE) = 1$ but $\mathrm{pk}(\emptyset) = 0$. This means

$$\begin{aligned}P(q,t,z) &= tz + \sum_{\emptyset \neq p \in \mathrm{Dyck}} q^{\mathrm{area}(NpE)} t^{\mathrm{pk}(p)} z^{|p|+1},\\ &= tz + z \sum_{p \neq \emptyset \in \mathrm{Dyck}} q^{\mathrm{area}(p)} t^{\mathrm{pk}(p)} (qz)^{|p|},\\ &= tz + z(C(q,t,qz) - 1),\\ &= z(t-1) + zC(q,t,qz).\end{aligned}$$

and

$$C(q,t,z) = \frac{1}{1 - z(t-1) - zC(q,t,qz)}.$$

Although this is not quite as nice as our first refinement, we can still plow ahead to find the following three-variable generating function for area, number of peaks, and size:

$$C(q,t,z) = \cfrac{1}{1 - z(t-1) - \cfrac{z}{1 - qz(t-1) - \cfrac{qz}{1 - q^2 z(t-1) - \cfrac{q^2 z}{1 - q^3 z(t-1) - \cdots}}}}$$

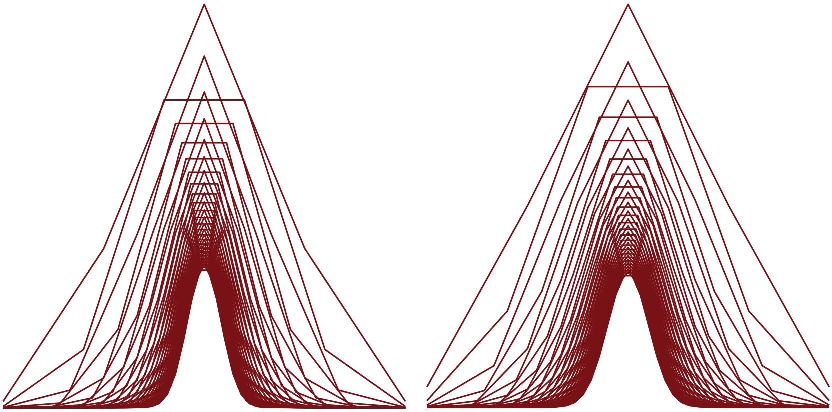
There are probably other statistics that we can throw in to the mix, but at this point it seems like we might be reaching the point of diminishing returns.

Further reading

- Art Benjamin, Francis Su, and Jennifer Quinn, *Counting on Continued Fractions* Mathematics Magazine, **73**, 98–104, (2000). This playful article gives a combinatorial interpretation to continued fractions in terms of tilings. This greatly generalizes the case of the continued fraction for the golden ratio.
- Andrew Odlyzko and Herb Wilf, *The Editor's Corner: n Coins in a Fountain* American Mathematical Monthly, **95**, 840–843, (1988). This short note discusses a counting problem whose generating function is a continued fraction. The analysis by Odlyzko and Wilf inspired the idea for much of the discussion of Dyck paths here.

Chapter 11
Applications to probability

"Creativity is the ability to introduce order into the randomness of nature."
–Eric Hoffer



(a) (b)

Rescaling the Eulerian and Narayana numbers to make probability distributions. Which is which?

Probability and statistics are the central components of many ideas in science and industry. The goal of this chapter is not to give a general introduction to probability, but to give a taste for how generating functions can unlock probabilistic results for combinatorial problems. In this chapter we will explore how the mean and variance of probability distributions can be extracted from generating functions, with sometimes surprising results about the behavior of "random" combinatorial models.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_11

Discrete probability

Definition 34. A *sample space* S is any finite set. An *event* relative to S is any subset A of S. The *probability* of event A relative to the sample space S is defined to be

$$\Pr(A) = \frac{|A|}{|S|}.$$

There are more sophisticated definitions of probability, but this one will serve our purposes. In practice we describe sample spaces intuitively as "experiments" whose outcome cannot be known with any certainty ahead of time. We may want to know the likelihood that our experiment has a certain type of outcome and this is the "event" mentioned above.

For example, if our experiment is rolling a die and recording which side is face up, it is natural to define the sample space as the set

$$S = \{1, 2, 3, 4, 5, 6\}.$$

If we want to know the probability of rolling a number greater than 4, the event set A is $A = \{5, 6\}$, and $\Pr(A) = 2/6$.

Random variables

Definition 35. Given a sample space S, a *random variable* is a function $X : S \to \mathbb{R}$ that is used to help define events. For any set Y of real numbers, we write

$$\Pr(X \in Y) = \frac{|\{s \in S : X(s) \in Y\}|}{|S|}.$$

Since S is finite, X only achieves finitely many values. Moreover, the values of X can usually be assumed to be nonnegative integers, so that we can encode this distribution in a polynomial generating function

$$p(t) = \sum_{k \geq 0} \Pr(X = k) t^k.$$

The function X gives rise to a *probability distribution* on the set of nonnegative integers.

Let us return to the dice rolling example. In the language of random variables, we would say X denotes the random variable that gives the number showing on the top of the die. Then we would write $\Pr(X > 4) = 2/6$. The full distribution has generating function

$$\frac{1}{6}t + \frac{1}{6}t^2 + \frac{1}{6}t^3 + \frac{1}{6}t^4 + \frac{1}{6}t^5 + \frac{1}{6}t^6.$$

Warmup 34. Suppose an experiment is to flip a coin five times in a row and record the sequence of outcomes (heads or tails). What is the sample space of this experiment? Let X denote the random variable that records the number of heads that occur. What is the probability of getting more heads than tails?

Problem 128.
Let X_n denote the random variable that counts how many times a coin comes up heads in a sequence of n tosses. What is $\Pr(X_n = k)$? Let

$$p_n(t) = \sum_{k \geq 0} \Pr(X_n = k) t^k.$$

What is $p_n(t)$? This is called the *binomial distribution*.

Find expressions for the ordinary and exponential generating functions for the binomial distribution, i.e.,

$$F(t,z) = \sum_{n \geq 0} p_n(t) z^n = 1 + \left(\frac{1}{2} + \frac{t}{2}\right) z + \left(\frac{1}{4} + \frac{2t}{4} + \frac{t^2}{4}\right) z^2 + \cdots ,$$

and

$$G(t,z) = \sum_{n \geq 0} p_n(t) \frac{z^n}{n!}$$

Problem 129.
This problem refers to cycle notation for permutations, as discussed at the end of Chapter 4.

Let X_n denote the random variable that counts how many cycles occur in a random permutation in S_n.

What is the formula for

$$p_n(t) = \sum_{k \geq 0} \Pr(X_n = k) t^k?$$

What is the probability that a permutation in S_9 has 4 cycles?

Problem 130.
What is the probability that 1 and 2 are in the same cycle in a permutation? For example, the permutation (152)(34) has 1 and 2 in the same cycle, whereas (1345)(2) has them in different cycles. Hint: consider the decision tree from the essay at the end of Chapter 4.

Problem 131.
What is the probability that a permutation has no fixed points, i.e., what is the probability that $w(i) \neq i$ for all i? Hint: look back at Problem 82.

Problem 132.
This problem relies on Rodrigues' Theorem for the generating function for inversions in permutations, which is given at the end of Chapter 9.

Let X_n denote the random variable that counts how many inversions occur in a random permutation in S_n.

What is the formula for

$$p_n(t) = \sum_{k\geq 0} \Pr(X_n = k)t^k?$$

What is the probability that a permutation in S_9 has 21 inversions?

Expectation

Definition 36. The *mean* or *expectation* of a random variable X on a sample space S is denoted $E(X)$ and is defined to be

$$E(X) = \sum_r r \Pr(X = r).$$

For example, if we are rolling a six-sided die and considering the side that lands face up, our expectation would be

$$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5.$$

For a different example, suppose our sample space is a set of standard playing cards,

$$S = \{A\heartsuit, A\clubsuit, A\diamondsuit, A\spadesuit, \ldots, K\heartsuit, K\clubsuit, K\diamondsuit, K\spadesuit\},$$

and X is the random variable that gives the point value of each card in the game of "Hearts." In this game most cards have no point value, but hearts each have a value of 1, and the queen of spades has a value of 13.

Thus the expected point value of a randomly selected card is

$$E(X) = 0 \cdot \frac{38}{52} + 1 \cdot \frac{13}{52} + 13 \cdot \frac{1}{52} = 1/2.$$

Warmup 35. If X has probability generating function

$$p(t) = \sum_{k\geq 0} \Pr(X = k)t^k,$$

show that

$$E(X) = p'(1).$$

Problem 133.
What is the expected number of fixed points of a permutation in S_n? For example (134)(2)(5)(67) has two fixed points, while (1)(2)(3574)(6) has three fixed points.

Problem 134.
What is the probability that a permutation in S_n begins with an increasing run of length k? In other words, what is the probability that

$$w(1) < \cdots < w(k),$$

but $w(k)$ is not less than $w(k+1)$?

Problem 135.
Continuing with Problem 134, what is the expected length of the first increasing run of a permutation in S_n? What is the limit as $n \to \infty$ of this quantity?

Problem 136.
Consider the set of paths in the first Starfolks problem (Problem 7) where we walk four blocks East and three blocks North. We can think of each path as part of the boundary of a region in the plane whose bottom-left corner sits at the origin, and with Starfolks at the point $(4, 3)$. The lower boundary is the x-axis and the rightmost boundary is the vertical line at $x = 4$. For example, the path shown in Figure 11.1 has area seven.

1. Let X denote the random variable that counts the area under a path. Compute the probability generating function $p(t) = \sum_k \Pr(X = k)t^k$. (Hint: Think back to Problem 125.)
2. Now suppose $q(t)$ is the generating function for paths that end at $(3, 3)$ and $r(t)$ is the generating function for paths that end at $(4, 2)$. Express $p(t)$ in terms of $q(t)$ and $r(t)$ in a way that is reminiscent of Pascal's recurrence for binomial coefficients.
3. What is the expected area under a path from $(0, 0)$ to $(4, 3)$?
4. What are some ways that you can generalize this problem?

While the expected value of a random variable is good to know, the spread of values is also important to know. Variance is one way to measure this.

Variance

Definition 37. The *variance* of a random variable is defined to be

$$\mathrm{Var}(X) = E((X - E(X))^2).$$

In other words, it is the expected value of the square of the distance between X and the mean.

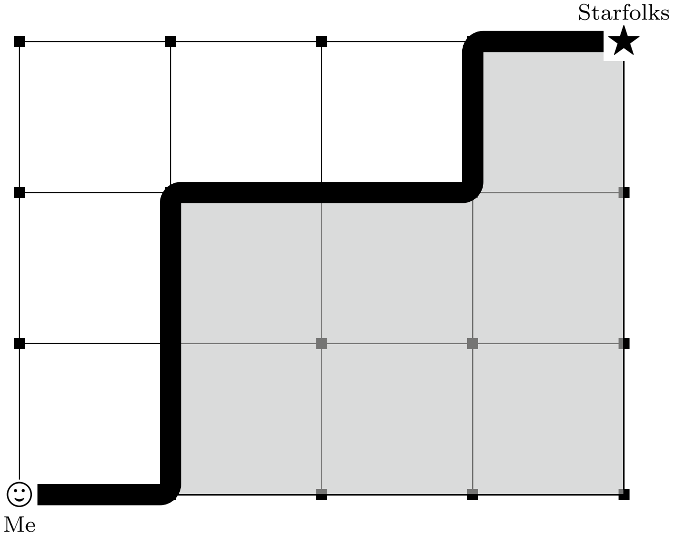


Fig. 11.1 Area under a path.

It turns out that variance can be computed as a difference of expectations:

$$\mathrm{Var}(X) = E(X^2) - E(X)^2. \tag{11.1}$$

For example, to compute the variance of point values in a deck of cards related to the game of Hearts, we first compute

$$E(X^2) = 0^2 \cdot \frac{38}{52} + 1^2 \cdot \frac{13}{52} + 13^2 \cdot \frac{1}{52} = 182/52 = 7/2.$$

Combined with $E(X) = 1/2$ from earlier, we have

$$\mathrm{Var}(X) = 7/2 - (1/2)^2 = 13/4.$$

Warmup 36. Compute the mean and variance of the binomial distribution $p_n(t)$ for $n \le 5$.

Problem 137.
If X has probability generating function

$$p(t) = \sum_{k\ge 0} \Pr(X = k)t^k,$$

show that

$$E(X^2) = \frac{d}{dt}[tp'(t)]_{t=1} = p''(1) + p'(1),$$

and hence

$$\mathrm{Var}(X) = p''(1) + p'(1) - p'(1)^2.$$

Problem 138.
Let X_n denote the random variable that counts the number of cycles in a random permutation in S_n. At the end of Chapter 4 the generating function for X_n was derived, and the expectation of X_n was shown to be $E(X_n) = H_n$, where H_n is the nth harmonic number

What is $\mathrm{Var}(X_n)$?

Problem 139.
As in Problem 128, let X_n denote the random variable that counts how many times a coin comes up heads in a sequence of n tosses, and let $p_n(t)$ denote its distribution, i.e., the binomial distribution.

1. Use the generating function formula you found for $\sum_{n\geq 1} p_n(t) z^n$ in Problem 128 to derive the generating function for the expectation $E(X_n)$. Show
$$\sum_{n\geq 1} E(X_n) z^n = \frac{z}{2(1-z)^2}.$$
2. What is $E(X_n)$ as a function of n? Does this formula have a more combinatorial explanation?
3. Derive the generating function for the variance $\mathrm{Var}(X_n)$. Show
$$\sum_{n\geq 1} \mathrm{Var}(X_n) z^n = \frac{z}{4(1-z)^2}.$$
4. What is $\mathrm{Var}(X_n)$ as a function of n?

Problem 140.
Now let's do for the Eulerian distribution what Problems 128 and 139 did for the binomial distribution.

Recall that the Eulerian polynomial $A_n(t)$ is the generating function for permutations in S_n according to the number of runs. Let X_n be the random variable that counts the number of runs in a random permutation of S_n. Then the polynomial $p_n(t) = A_n(t)/n!$ is the probability generating function for X_n.

1. What is the ordinary generating function for the sequence of polynomials $p_n(t)$?
2. Derive the generating function for the expectation $E(X_n)$ from part 1. Show
$$\sum_{n\geq 1} E(X_n) z^n = \frac{z(2-z)}{2(1-z)^2}.$$
3. What is $E(X_n)$ as a function of n? Does this formula have a more combinatorial explanation?

4. Derive the generating function for the variance $\mathrm{Var}(X_n) = E(X_n^2) - E(X_n)^2$. Show

$$\sum_{n\geq 1} \mathrm{Var}(X_n) z^n = \frac{z^2(3-2z)}{12(1-z)^2}.$$

5. What is $\mathrm{Var}(X_n)$ as a function of n? For large n, the variance is approximately what proportion of n?

You may want the aid of a computer to prevent errors on parts 2 and 4. Also keep in mind the lesson of Problem 102, i.e., to specialize some generating functions to $t = 1$ it is not enough to plug in. We should really take a limit as $t \to 1$.

Problem 141.
We will repeat the steps in Problem 140 for the set of 231-avoiding permutations. Here we let X_n denote the random variable that counts the number of runs in a random 231-avoiding permutation in S_n. Then $p_n(t) = C_n(t)/C_n$, where $C_n(t)$ is the Narayana polynomial from Definition 26 and $C_n = C_n(1) = \frac{1}{n+1}\binom{2n}{n}$ is the nth Catalan number.

For this problem you will probably want the generating function for the Narayana numbers in Problem 107, denoted $C(t,z) = \sum_{n\geq 0} C_n(t) z^n$.

1. Show that

$$\frac{d}{dt}[C(t,z)]_{t=1} = \sum_{n\geq 1} C_n'(1) z^n = \sum_{n\geq 1} \binom{2k-1}{k-1} z^n.$$

2. What is $E(X_n)$ as a function of n? Does this formula have a more combinatorial explanation?
3. Compute $E(X_n^2)$, and hence $\mathrm{Var}(X_n)$ as a function of n. For large n, the variance is approximately what proportion of n?
4. Comparing with Problem 140, which distribution is more disperse, the Eulerian distribution or the Narayana distribution? (Look at Figure on Page 147. One is a picture of the Narayana distribution; the other is a picture of the Eulerian distribution. Which is which?)

Longest increasing subsequences and random partitions

SOME PERMUTATION STATISTICS ARE EASIER to analyze than others. Descents, inversions, and cycles are straightforward when compared to the permutation statistic known as the *longest increasing subsequence*. The study of this statistic leads to a large body of work at the intersection of probability theory and mathematical physics—in particular the study of random matrices! This is a lot to cover briefly, but we will introduce some of the combinatorial models and give hints about their large-scale behavior.

Patience sorting

Consider the following solitaire card game, known as *patience sorting*. We turn the cards over one at a time and place them in piles. At each stage in the game, the card may be placed on top of a pile or it can start a new pile. However the card may only be placed on top of a pile if the card is smaller in value than the current top card, e.g., a 3 can go on top of an 8, but not on top of a 2. The object of the game is to end up with the fewest piles possible.

For example, if we have only nine cards ordered in the deck as 318925647 then first 3 is placed in a new pile, then the 1 can either go on top of the 3: $\begin{array}{c}1\\3\end{array}$ or it can start a new pile: 3 1.

The *greedy strategy* is to always place a card on top of the leftmost pile possible, only starting a new pile if the card is greater than the top card on each existing pile. It is not difficult to argue that the greedy strategy is optimal in that it produces the fewest possible piles. Carrying out this strategy with 318925647 we have the following sequence of piles:

$$
\begin{array}[b]{c}\mathbf{3}\end{array}\quad
\begin{array}[b]{c}\mathbf{1}\\3\end{array}\quad
\begin{array}[b]{cc}1&\\3&\mathbf{8}\end{array}\quad
\begin{array}[b]{ccc}1&&\\3&8&\mathbf{9}\end{array}\quad
\begin{array}[b]{ccc}1&\mathbf{2}&\\3&8&9\end{array}\quad
\begin{array}[b]{ccc}1&2&\mathbf{5}\\3&8&9\end{array}\quad
\begin{array}[b]{cccc}1&2&5&\\3&8&9&\mathbf{6}\end{array}\quad
\begin{array}[b]{cccc}&&\mathbf{4}&\\1&2&5&\\3&8&9&6\end{array}\quad
\begin{array}[b]{ccccc}&&4&&\\1&2&5&&\\3&8&9&6&\mathbf{7}\end{array}
$$

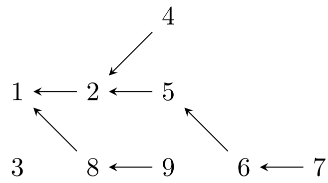
Thus we see the minimal number of piles for this ordering of nine cards is five.

Longest increasing subsequences

It turns out that the optimal number of piles in patience sorting for the ordering w is precisely equal to the permutation statistic we want to study. To be clear, an *increasing subsequence* is just what it sounds like: a sequence of values $w(i_1) < w(i_2) < \cdots < w(i_l)$ for some $i_1 < i_2 < \cdots < i_l$. The length of the longest such subsequence is denoted $l(w)$.

For example, with $w = 318925647$ we have $l(w) = 5$ since $\cdot\,1\cdot\cdot\,2\,5\,6\cdot 7$ is an increasing subsequence and no other increasing subsequence is longer. It turns out that greedy patience sorting gives a clever way to describe an algorithm of Hammersley for finding this longest subsequence.

To make this connection, we augment the steps in the patience sorting by drawing an arrow $a \leftarrow b$ whenever b is placed on top of the pile immediately to the right of a and $a < b$. Notice that in the greedy algorithm, the top card in the pile to the left of b is necessarily smaller, so each card in a column other than the first has at least one arrow pointing left. Doing this with $w = 318925647$ we get the directed graph below:



By construction, there will always be a path from the leftmost pile to the rightmost pile: $1 \leftarrow 2 \leftarrow 5 \leftarrow 6 \leftarrow 7$. This is a longest increasing subsequence of our permutation.

For a larger example, consider the permutation in S_{20} given by

$$u = 13\ \ 6\ \ 2\ \ 12\ \ 4\ \ 5\ \ 17\ \ 8\ \ 18\ \ 20\ \ 11\ \ 19\ \ 10\ \ 14\ \ 15\ \ 9\ \ 16\ \ 1\ \ 7\ \ 3$$

Here the directed graph from the greedy algorithm is shown in Figure 11.2.

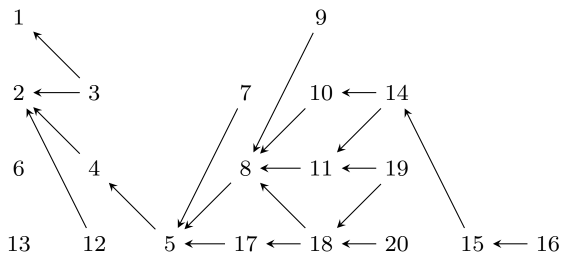


Fig. 11.2 Patience sorting graph for an element of S_{20}.

We can see that the length of the longest increasing subsequence is $l(u) = 8$, e.g., we have

$$2 \leftarrow 4 \leftarrow 5 \leftarrow 8 \leftarrow 10 \leftarrow 14 \leftarrow 15 \leftarrow 16.$$

In this example there is more than one longest subsequence: we could go $\cdots 8 \leftarrow 11 \leftarrow 14 \cdots$ instead of $\cdots 8 \leftarrow 10 \leftarrow 14 \cdots$. The directed graph contains all increasing subsequences.

Table 11.1 shows the distribution of longest increasing subsequences across all permutations. There doesn't appear to be standard terminology for this array of numbers in the literature, but we will call it the triangle of *Ulam numbers*, or the *Ulam distribution*, since Stanislaw Ulam initiated the study of longest increasing subsequences.

$n\backslash k$	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	4	1				
4	1	13	9	1			
5	1	41	61	16	1		
6	1	131	381	181	25	1	
7	1	428	2332	1821	421	36	1

Table 11.1 Triangle of the number of permutations in S_n with longest increasing subsequence of length $l(w) = k$.

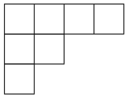
Young tableaux and Schensted insertion

In 1961, Craige Schensted developed a correspondence between permutations and pairs of what are known as *Young tableaux* as another way to tackle the increasing subsequence problem. In this approach we recursively record a pair of arrays of numbers that increase across rows and down columns.

To be a bit more precise, define a *standard Young tableau* of size n to be an array of numbers $1, 2, \ldots, n$ that increase across rows and down columns. For example,

$$T = \begin{array}{l} 1\ 3\ 5\ 6 \\ 2\ 4 \\ 7 \end{array}$$

is a standard Young tableau of $n = 7$. The *shape* of a tableau is the list of lengths of the rows, $\lambda = (\lambda_1, \ldots, \lambda_k)$. Such lists of weakly decreasing numbers are known as *integer partitions*, and are studied further in Chapter 12. In this example, the shape is $\lambda = (4, 2, 1)$. Often the shape of a Young tableau is drawn with an array of empty boxes known as a *Young diagram*, e.g.,

$\lambda =$ 

Young diagrams give a nice way to visualize integer partitions in many contexts.

Let's now see how to do Schensted insertion. The process is similar to the placing of cards on piles in patience sorting. We will have two tableaux of the same shape, P and Q, that are updated recursively to keep track of the values we insert and the positions that get filled. Suppose $w \in S_n$. Then for each $i = 1, 2, \dots n$, we do the following with letter $w(i) = j$.

1. Insert j in the first row of P. Either j is placed at the end of the row, or, if the first row contains elements larger than j, we displace the smallest such element.
2. If j displaces an element from row one, insert the displaced element in row two. This element goes either at the end of the row or it displaces the smallest element that is larger than it.
3. Repeat this process row by row until some element is placed at the end of a row.
4. Add a new box to Q in the same position. This box is filled with i, the current number of values inserted.

We illustrate with the example permutation $w = 318925647$ from before. To begin, we place the first letter of w, "3" in P and we record a "1" in Q to indicate where the new box was added:

$$P = \boxed{3} \qquad\qquad Q = \boxed{1}$$

Next we insert the second letter of w in P. Since $w(2) = 1$ is smaller than 3, this action displaces the 3, which moves to the second row. Now we have

$$P = \begin{array}{|c|}\hline 1 \\ \hline 3 \\ \hline\end{array} \qquad\qquad Q = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array}$$

The "2" added to Q reflects the position in which a new cell was added to the shape of P. The remaining steps in the algorithm are shown in Figure 11.3.

Notice that the top row in our final P tableau corresponds exactly to the cards on top of the piles in our greedy patience sorting algorithm! In fact, we can see by induction that this holds at every stage of the insertion process compared to every stage of the greedy patience sorting algorithm. (The tableau P doesn't make it so easy to read off the longest increasing subsequence itself, however.)

P	Q
$\begin{array}{l}3\end{array}$	$\begin{array}{l}1\end{array}$
$\begin{array}{l}1\\3\end{array}$	$\begin{array}{l}1\\2\end{array}$
$\begin{array}{ll}1&8\\3\end{array}$	$\begin{array}{ll}1&3\\2\end{array}$
$\begin{array}{lll}1&8&9\\3\end{array}$	$\begin{array}{lll}1&3&4\\2\end{array}$
$\begin{array}{lll}1&2&9\\3&8\end{array}$	$\begin{array}{lll}1&3&4\\2&5\end{array}$
$\begin{array}{lll}1&2&5\\3&8&9\end{array}$	$\begin{array}{lll}1&3&4\\2&5&6\end{array}$
$\begin{array}{llll}1&2&5&6\\3&8&9\end{array}$	$\begin{array}{llll}1&3&4&7\\2&5&6\end{array}$
$\begin{array}{llll}1&2&4&6\\3&5&9\\8\end{array}$	$\begin{array}{llll}1&3&4&7\\2&5&6\\8\end{array}$
$\begin{array}{lllll}1&2&4&6&7\\3&5&9\\8\end{array}$	$\begin{array}{lllll}1&3&4&7&9\\2&5&6\\8\end{array}$

Fig. 11.3 Schensted insertion for the permutation $w = 318925647$.

What Schensted proved was even more: not only does the length of the first row corresponds to the longest increasing subsequence, the length of the first column corresponds to the length of the longest decreasing subsequence! (Check this is true for the example above.)

Before moving on, observe that the insertion algorithm can be run in reverse. Thus Schensted insertion is really a bijection between elements $w \in S_n$ and pairs of standard Young tableaux (P, Q). This is more widely known as the Robinson–Schensted–Knuth correspondence, or RSK correspondence.

The hook length formula

If we let d_λ denote the number of standard Young tableau of shape λ, then Schensted's correspondence implies the curious formula

$$n! = \sum_{\lambda \vdash n} d_\lambda^2,$$

where $\lambda \vdash n$ means λ is a partition of n. Moreover, since the length of the first row of P corresponds to the longest increasing subsequence of w, then for each k we have

$$|\{w \in S_n : l(w) = k\}| = \sum_{\substack{\lambda \vdash n \\ \lambda_1 = k}} d_\lambda^2,$$

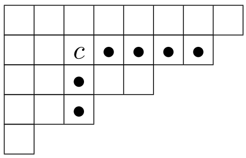
which suggests a way to calculate the entries in Ulam's distribution without referring to permutations at all.

But are the numbers d_λ easier to compute? Yes! There is a nice formula for d_λ, known as the *hook length formula*, due to Frame–Robinson–Thrall. Schensted knew this formula, and recognized that it would provide one way to compute the numbers in Ulam's distribution.

The hook length formula is

$$d_\lambda = \frac{n!}{\prod_{c \in \lambda} h_c},$$

where the product is over all *cells* in λ and h_c is the *hook length* of the cell c. This measures the number of cells to the right and below the cell c, including c itself. For example, cell c marked below has $h_c = 7$:



For the shape $\lambda = (5, 3, 1)$, the hook length formula gives

$$d_{(5,3,1)} = \frac{9!}{7 \cdot 5 \cdot 4 \cdot 2 \cdot 1 \cdot 4 \cdot 2 \cdot 1 \cdot 1} = 162,$$

so the permutation $w = 318925647$ is just one of $162^2 = 26244$ permutations whose Young tableaux have shape $(5, 3, 1)$.

If we want to find how many permutations in S_9 have longest increasing subsequence 9, we simply sum over all partitions of 9 with $\lambda_1 = 5$. Using the hook length formula several times we find:

$$\begin{aligned} |\{w \in S_9 : l(w) = 5\}| &= d_{(5,4)}^2 + d_{(5,3,1)}^2 + d_{(5,2,2)}^2 + d_{(5,2,1,1)}^2 + d_{(5,1,1,1,1)}^2 \\ &= 42^2 + 162^2 + 120^2 + 189^2 + 70^2 \\ &= 83029. \end{aligned}$$

Expectation and limit shapes

The connection between the longest increasing subsequence problem and Young tableaux means that instead of asking Ulam's question,

> What is the expected length of a longest increasing subsequence of a random permutation?

we can instead ask

> What is the expected length of the first part in a random partition?

We have to be a little careful about what we mean by a "random" partition, but the hook length discussion gives us the proper weighting to make the questions equivalent. We say the partition $\lambda \vdash n$ occurs with probability

$$\frac{d_\lambda^2}{n!} = \frac{n!}{\prod_{c\in\lambda} h_c^2}.$$

This is called the *Plancharel measure* on partitions. An incredible result due to Logan-Shepp-Vershik-Kerov answers the following much stronger question,

> What is the expected **shape** of a random partition?

The answer to this question is illustrated in Figure 11.4. Here we have rotated the Young diagram 135 degrees. The boundary of the partition (rescaled appropriately) is very close to the curve Ω given by

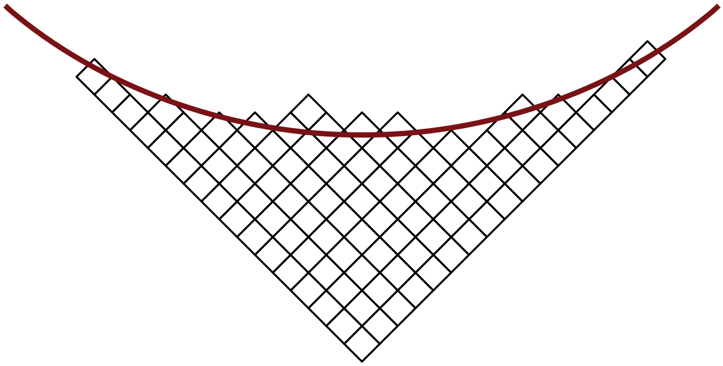
$$\Omega(x) = \frac{2}{\pi}\left(x \arcsin\left(\frac{x}{\sqrt{2}}\right) + \sqrt{2-x^2}\right).$$

Once we pass through the rescaling of our Young diagram (draw each box as a diamond with integer corners, then shrink each box by a factor of $1/\sqrt{2n}$) and compare with the limit curve, we find the expected length of the first row to be

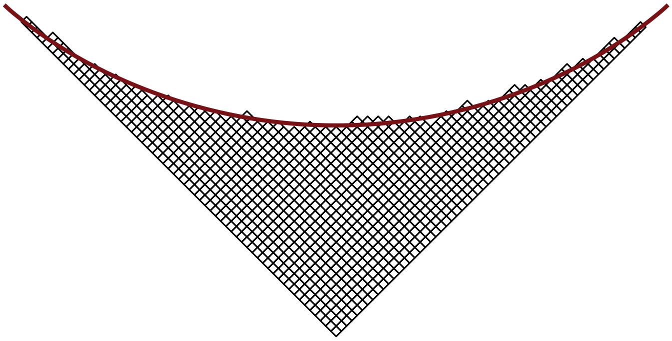
$$\lambda_1 \approx 2\sqrt{n},$$

answering Ulam's original question. So, for instance, a random permutation of $n = 100$ should have a longest increasing subsequence of about 20.

A completely surprising result known as the Baik–Deft–Johansson Theorem says that Ulam's distribution converges asymptotically to what is known as the Tracy–Widom distribution. This distribution comes from Mathematical Physics; it is the distribution of the largest eigenvalue in a random Hermitian matrix! Subsequent work of Fields medalist Andrei Okounkov and others shows that in fact the kth largest row in a random Young diagram is distributed like the kth largest eigenvalue. Few would have guessed at the depth of these results given Ulam's original question.



(a)



(b)

Fig. 11.4 Random Young diagrams of: (a) size $n = 100$, and (b) size $n = 1000$, as compared to the limit shape Ω.

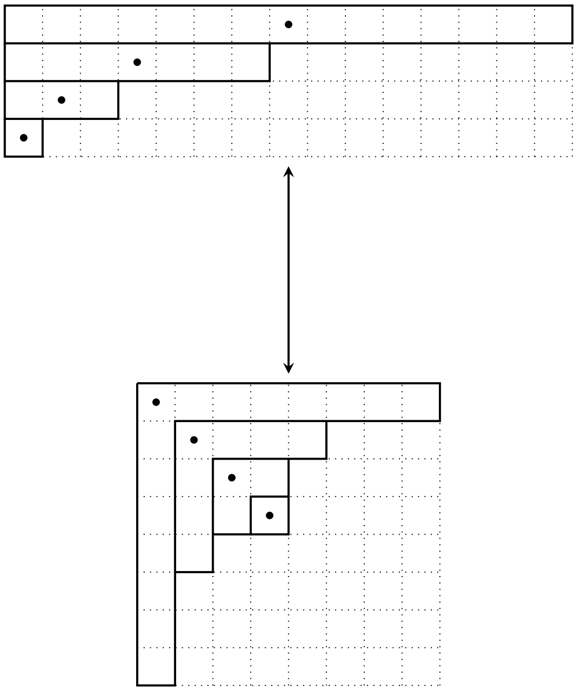
Further reading

- Noga Alon and Joel Spencer, "The Probabilistic Method," Wiley, 2008.
  This book shows how to use probabilistic arguments to obtain combinatorial results, particularly in graph theory.
- Dan Romik, "The surprising mathematics of longest increasing subsequences," Cambridge University Press, 2015.
  This book tells the story of the longest increasing subsequence problem in its full detail.

Chapter 12
Some partition theory

"Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk"
("God made the integers, all else is the work of man.")

–Leopold Kronecker



A correspondence between partitions (15,7,3,1) and (8,5,4,4,2,1,1,1).

How many ways can you write a number as a sum of smaller numbers? This simple question is not as easy to answer as you might think. The study of integer partitions has a long history in Number Theory and Combinatorics, with contributions from Leonhard Euler, Srinivasa Ramanujan, Godfrey Harold Hardy, and more recently George Andrews and Ken Ono. In this chapter we will ask some of the first questions about counting partitions.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_12

Integer partitions

Definition 38. A *partition* of n is an unordered collection of positive integers whose sum is n. By convention, we list the elements in a partition in decreasing order, e.g., $\lambda = (\lambda_1, \ldots, \lambda_k)$ is a partition of n if and only if the λ_i are positive integers, $\lambda_1 \geq \cdots \geq \lambda_k \geq 1$, and $\lambda_1 + \cdots + \lambda_k = n$. Denote the the number of parts, or *length* of a partition by $\ell(\lambda) = k$. The *size* of the partition is the sum of the parts, denoted $|\lambda| = n$.

Thus while $(2, 2, 1)$, $(2, 1, 2)$, and $(1, 2, 2)$ all represent different compositions (in the sense of Chapter 5), they represent the same partition of 5. The default way to write this partition is $(2, 2, 1)$.

Warmup 37. List all the partitions of 3, 4, 5, and 6. Organize your data by the number of parts.

Throughout this chapter we denote the number of partitions of n by p_n. The sequence of partition numbers $p_n, n \geq 1$ begins:

$$1, 2, 3, 5, 7, 11, 15, 22, 30, 42, 56, 77, 101, \ldots$$

and we usually take $p_0 = 1$ for convenience.

Problem 142.
How many ways can you write n as a sum of ones and twos? This question can be interpreted with or without ordering the summands.

1. How many partitions of n use only parts of size 1 and size 2, i.e., have $\lambda_i \in \{1, 2\}$?
2. How many compositions of n use only parts of size 1 and size 2? (See Problem 56.)

Problem 143.
In this problem you will get an expression for the ordinary generating function for p_n with an infinite product.

For any n and k, let

$$p_{n,k}(t) = \sum_{\substack{|\lambda|=n \\ \lambda_1 \leq k}} t^{\ell(\lambda)}.$$

This is the generating function for partitions of n according to the number of parts, where the parts are at most k. For example if $n = 5$ and $k = 3$, we are counting the partitions $(3, 2)$, $(3, 1, 1)$, $(2, 2, 1)$, $(2, 1, 1, 1)$, and $(1, 1, 1, 1, 1)$ according to length, and we find $p_{5,3}(t) = t^2 + 2t^3 + t^4 + t^5$.

1. Let $P_k(t, z) = \sum_{n \geq 0} p_{n,k}(t) z^n$ denote the generating function for all partitions whose parts are at most k. Show that for any $k \geq 1$,

$$P_k(t,z) = \prod_{i=1}^{k} \frac{1}{1-tz^i}.$$

2. Let $p_n(t) = \sum_{|\lambda|=n} t^{\ell(\lambda)}$ and let $P(t,z) = \sum_{n\geq 0} p_n(t)z^n$ denote the generating for all partitions. Show that

$$P(t,z) = \prod_{i\geq 1} \frac{1}{1-tz^i}.$$

3. Conclude that the generating function for the partition numbers, $P(z) = \sum_{n\geq 0} p_n z^n$, is

$$P(z) = \prod_{i\geq 1} \frac{1}{1-z^i}.$$

   This is known as *Euler's product formula.*

Problem 144.
Prove that for $n \geq 3$,

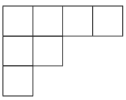
$$p_n \leq p_{n-1} + p_{n-2}.$$

Use this inequality to conclude that for all $n \geq 1$, $p_n \leq f_n$, where f_n denotes the nth Fibonacci number, with initial values $f_1 = 1$ and $f_2 = 2$. In particular, $p_n \leq \varphi^n$, where $\varphi = (1+\sqrt{5})/2$ is the golden ratio.

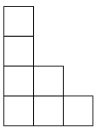
Young diagrams

Definition 39. A *Young diagram* is a visual illustration for partitions. Given a partition $\lambda = (\lambda_1, \ldots, \lambda_k)$, we draw k rows of boxes, such that row i has λ_i boxes.

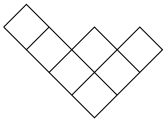
For example, the partition $(4, 2, 1)$ is drawn



Notice that our convention is to have the rows drawn from largest on top to smallest on bottom, and that the rows are left justified. This is the "English" convention for visualizing partitions. The "French" convention is to have k columns arranged from tallest on the left to shortest on the right. These are known as "Ferrers diagrams" or "Ferrers boards," e.g., the partition $(4, 2, 1)$ has Ferrers diagram



Yet another way of drawing Young diagrams is “Russian style” as in



which is simply obtained by rotating the Ferrers diagram another 45 degrees. (Russian style is also shown in Figure 11.4.)

Problem 145.
How many Young diagrams fit inside an $a \times b$ rectangle? Let

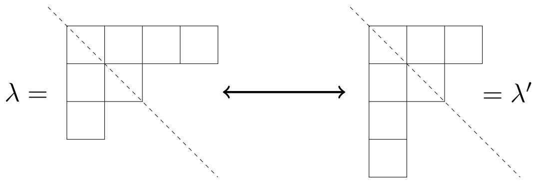
$$L_{a,b}(t) = 1 + \sum_{\lambda \subseteq a \times b} t^{|\lambda|},$$

where $\lambda \subseteq a \times b$ means the Young diagram for λ has at most a rows and at most b columns. Find a recurrence relation for the polynomial $L_{a,b}(t)$.

Conjugate partitions

Definition 40. The *conjugate* of a partition λ is the partition whose parts are the lengths of the columns of the Young diagram for λ. We denote the conjugate of λ by λ'. To put it another way, $\lambda'_i = |\{j : \lambda_j \geq i\}|$, so that λ'_1 is the number of parts of λ, λ'_2 is the number of parts of λ that are at least two, and so on.

For example, if λ is the partition $(4, 2, 1)$, its conjugate is $\lambda' = (3, 2, 1, 1)$. In pictures, the Young diagram of the conjugate is obtained by reflecting the Young diagram across the line $y = -x$:



Warmup 38. Use the notion of conjugate partitions to explain why $L_{a,b}(t) = L_{b,a}(t)$, where $L_{a,b}(t)$ is as defined in Problem 145.

Warmup 39. Use the notion of conjugate partitions to explain why

$$p_{n,k}(t) = \sum_{\substack{|\lambda|=n \\ \ell(\lambda) \leq k}} t^{\lambda_1},$$

where $p_{n,k}(t)$ is the function defined in Problem 143. In other words, counting partitions with restricted part size according to the number of parts is the same as counting partitions with restricted number of parts according to part size.

Problem 146.
The largest square that fits inside the Young diagram for λ is called the *Durfee square*. The side length of the Durfee square is the largest s such that $\lambda_s \geq s$. Use the notion of a Durfee square to prove

$$P(t, z) = 1 + \sum_{s \geq 1} t^s z^{s^2} \prod_{i=1}^{s} \frac{1}{(1 - z^i)(1 - tz^i)},$$

and hence

$$P(z) = 1 + \sum_{s \geq 1} z^{s^2} \prod_{i=1}^{s} \frac{1}{(1 - z^i)^2},$$

where $P(t, z)$ and $P(z)$ are the partition generating functions from Problem 143. Here's a hint. Decompose the Young diagram for λ into three Young diagrams: the Durfee square itself, the diagram to the right of the square, and the diagram below the square.

Problem 147.
A partition is *self-conjugate* if $\lambda = \lambda'$. A partition μ is said to have *distinct odd parts* if $\mu_1 > \mu_2 > \cdots > \mu_k$ and each of the μ_i are odd. Use a bijection to show the number of self-conjugate partitions of n equals the number of partitions of n with distinct odd parts.

Problem 148.
Let D_n denote the set of partitions of n whose parts are *distinct*, i.e., $\lambda_1 > \lambda_2 > \cdots > \lambda_k$. Let O_n denote the set of partitions whose parts are only odd numbers.

1. Compute $|D_n|$, for $n = 1, 2, \ldots, 8$.
2. Show
$$1 + \sum_{n \geq 1} |D_n| z^n = \prod_{i \geq 1} (1 + z^i).$$
3. Compute $|O_n|$, for $n = 1, 2, \ldots, 8$.
4. Show
$$1 + \sum_{n \geq 1} |O_n| z^n = \prod_{i \geq 1} \frac{1}{(1 - z^{2i-1})}.$$
5. Compare generating functions to show $|O_n| = |D_n|$.

Problem 149.
Obtain the result of Problem 148, that $|D_n| = |O_n|$, with a bijection.

Problem 150.
Let $\xi(z) = \prod_{i\geq 1}(1 - z^i)$ denote the denominator in the generating function for all partitions found in Problem 143. This is sometimes called the *Euler function*. Show that

$$\xi(z) = 1 - z - z^2 + z^5 + z^7 - z^{12} - z^{15} + z^{22} + z^{26} - \cdots .$$

Find a formula for the exponents of the nonzero terms in the series expansion of $\xi(z)$, and show the only coefficients are $1, -1$, and 0. (Hint: interpret the left-hand side as running over all partitions into distinct parts, where if the partition has an odd number of parts it gets counted with a minus sign.)

Problem 151.
From Problem 150, we can imagine rewriting the partition generating function as

$$P(z) = \sum_{n\geq 0} p_n z^n = \frac{1}{1 - z - z^2 + z^5 + z^7 - z^{12} - z^{15} + z^{22} + z^{26} - \cdots}$$

Explain how to use this expression to give the recursive formula

$$p_n = p_{n-1} + p_{n-2} - p_{n-5} - p_{n-7} + p_{n-12} + p_{n-15} - \cdots ,$$

with the convention that $p_k = 0$ if $k < 0$. Use the recurrence to make a table of p_n for $n \leq 30$.

Plane Partitions

In this chapter we studied integer partitions, which we can think of as one-dimensional arrays of finitely many weakly decreasing positive integers:

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots .$$

A natural way to generalize this idea is to study two-dimensional arrays. Define a *plane partition* ρ to be a two-dimensional array of finitely many positive integers that weakly decrease across rows and down columns,

$$\rho = \begin{matrix} \rho_{1,1} & \rho_{1,2} & \cdots \\ \rho_{2,1} & \rho_{2,2} & \cdots \\ \rho_{3,1} & \rho_{3,2} & \cdots \\ \vdots & \vdots & \ddots \end{matrix}$$

i.e., $\rho_{i,j} \geq \rho_{i,j+1}$ and $\rho_{i,j} \geq \rho_{i+1,j}$. If the sum of the parts of ρ is n, $|\rho| = \sum \rho_{i,j} = n$, we say ρ is a plane partition of n. We can draw plane partitions much like we draw Young tableaux, except weakly decreasing across rows and down columns, e.g.,

$$\rho = \begin{array}{|c|c|c|c|} \hline 9 & 3 & 2 & 1 \\ \hline 7 & 3 & 1 \\ \cline{1-3} 2 & 2 \\ \cline{1-2} 2 \\ \cline{1-1} 1 \\ \cline{1-1} \end{array}$$

is a plane partition of $n = 33$. Apart from the sum of the entries, we can also keep track of the *shape* of ρ, which corresponds to some ordinary partition λ. We will write $\mathrm{sh}(\rho) = \lambda$. For example, the plane partition above has shape $\mathrm{sh}(\rho) = (4, 3, 2, 1, 1)$.

Another visual way to think of plane partitions is in terms of stacks of n cubes, in a kind of three-dimensional Young diagram, which we will call a *solid Young diagram*. For example, in Figure 12.1 we see the plane partition above as a stack of cubes.

MacMahon embarked on a thorough study of partitions and plane partitions in the early twentieth century. We will survey some of his results and some more recent developments. Of course if you can study one- and two-dimensional partitions, why not d-dimensional partitions? It transpires that beyond two dimensions things get much harder. While we shall soon see a generating function for plane partitions, as of this writing there is no comparable formula for the generating function for three-dimensional partitions, known as *solid partitions*.

$$\rho = \begin{array}{|c|c|c|c|}\hline 9 & 3 & 2 & 1 \\ \hline 7 & 3 & 1 \\ \cline{1-3} 2 & 2 \\ \cline{1-2} 2 \\ \cline{1-1} 1 \\ \cline{1-1} \end{array} =$$

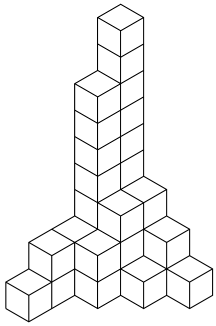


Fig. 12.1 A plane partition as a stack of cubes.

Generating function

Let pp_n denote the number of plane partitions of n, with $pp_0 = 1$. The sequence pp_n begins

$$1, 1, 3, 6, 13, 24, 48, 86, 160, 282, 500, 859, \ldots.$$

For example, here are the six plane partitions of 3:

$$\begin{array}{|c|}\hline 3 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 1 \\ \hline\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline\end{array} \qquad \begin{array}{|c|}\hline 2 \\ \hline 1 \\ \hline\end{array} \qquad \begin{array}{|c|}\hline 1 \\ \hline 1 \\ \hline 1 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 1 \\ \cline{1-1}\end{array}$$

MacMahon's formula for the generating function is a deceptively nice generalization of Euler's product formula:

$$\sum_{n \geq 0} pp_n z^n = \prod_{i \geq 1} \frac{1}{(1 - z^i)^i} \tag{12.1}$$

Compare with Problem 143.

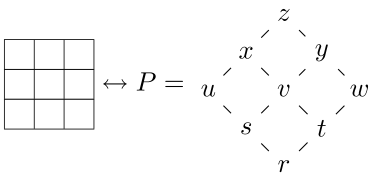
While there are combinatorial explanations of this identity, none are so quick and elementary as Euler's identity. One approach is similar to the one-dimensional case, in that we first count the plane partitions whose shape fits inside a $m \times m$, square. That is, if $\mathrm{sh}(\rho) = \lambda$, we want $\lambda_1 \leq m$ and $\ell(\lambda) \leq m$. This generating function turns out to be

$$\sum_{\mathrm{sh}(\rho) \subseteq m \times m} z^{|\rho|} = \prod_{i=1}^{m} \frac{1}{(1 - z^i)^i} \prod_{j=1}^{m-1} \frac{1}{(1 - z^{m+j})^{m-j}}, \tag{12.2}$$

and as $m \to \infty$, we get Equation (12.1).

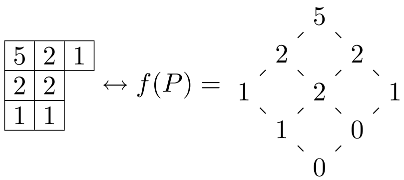
P-partitions

But how to prove the identity in (12.2)? One method comes from the theory of P*-partitions*, developed by Richard Stanley in the 1970s. Here, "P" stands for a partially ordered set. In this case, we think of the cells in an $m \times m$ square as being partially ordered from the bottom right to top left corner, e.g.,



We have drawn P so that $p < q$ if there is an upward path of edges from p to q. This is a partial ordering since since some elements are incomparable, e.g., s and t are incomparable, as are s and w.

A P-partition is an order-preserving map $f : P \to \{0, 1, 2, 3, \ldots\}$, i.e., if p is below q in P, then $f(p) \leq f(q)$. When the partially ordered set is a square grid as above, P-partitions correspond to plane partitions, e.g.,



corresponds to the function given by $f(z) = 5$, $f(x) = f(y) = f(v) = 2$, $f(w) = f(u) = f(s) = 1$, and $f(r) = f(t) = 0$.

We can now interpret (12.2) in terms of these maps:

$$\sum_{\mathrm{sh}(\rho) \subseteq m \times m} z^{|\rho|} = \sum_{f \in \mathcal{A}(P)} z^{|f|},$$

where $\mathcal{A}(P)$ denotes the set of P-partitions for the $m \times m$ grid and $|f| = \sum_{p \in P} f(p)$.

For arbitrary partially ordered sets P, one does not expect the generating function for P-partitions to be particularly simple. However, in the case where P resembles a rectangular grid, Stanley proved a formula that can be phrased in terms of the *hook lengths*, h_c, as discussed at the end of Chapter 11. This result is

$$\sum_{f \in \mathcal{A}(P)} z^{|f|} = \prod_{c \in P} \frac{1}{1 - z^{h_c}},$$

which we will not prove here.

For the 3×3 example above, the hook lengths are

5	4	3
4	3	2
3	2	1

so the generating function would be

$$\sum_{f \in \mathcal{A}(P)} z^{|f|} = \frac{1}{(1-z)(1-z^2)^2(1-z^3)^3(1-z^4)^2(1-z^5)}.$$

Thinking about the hook lengths in a general $m \times m$ square yields Equation (12.2).

Plane partitions in a box

Another way to approach Equation (12.1) is to count the number of plane partitions whose solid Young diagrams fit in an $a \times b \times c$ box. That is, if ρ has shape λ, we require $\lambda_1 \le a$, $\ell(\lambda) \le b$ and $\rho_{1,1} \le c$.

MacMahon also had a formula for this counting problem. The number of plane partitions in an $a \times b \times c$ box is

$$\prod_{i=1}^{a} \prod_{j=1}^{b} \prod_{k=1}^{c} \frac{i+j+k-1}{i+j+k-2},$$

or more precisely, the generating function for these according to area is

$$\sum_{\rho \subseteq a \times b \times c} z^{|\rho|} = \prod_{i=1}^{a} \prod_{j=1}^{b} \prod_{k=1}^{c} \frac{1-z^{i+j+k-1}}{1-z^{i+j+k-2}}.$$

Note that letting $a, b, c \to \infty$ implies Equation (12.1). This identity is not trivial to prove either, and there have been several techniques for proving it, including Stanley's P-partition approach and variations on the Schensted correspondence from Chapter 11.

For example, there are 116424 plane partitions that fit inside of a $3 \times 4 \times 5$ box, and the generating function is

$$\prod_{i=1}^{3} \prod_{j=1}^{4} \prod_{k=1}^{5} \frac{1-z^{i+j+k-1}}{1-z^{i+j+k-2}} = \frac{(1-z^7)^2(1-z^8)^3(1-z^9)^3(1-z^{10})^2(1-z^{11})}{(1-z)(1-z^2)^2(1-z^3)^3(1-z^4)^3(1-z^5)^2}$$

which expands as shown in Figure 12.2. In Figure 12.3 we find one of the 5199 such plane partitions comprised of $n = 26$ small cubes.

$$
\begin{aligned}
&1 + z + 3z^2 + 6z^3 + 12z^4 + 20z^5 + 36z^6 + 56z^7 + 91z^8 + 136z^9 + 203z^{10} \\
&+ 287z^{11} + 407z^{12} + 548z^{13} + 736z^{14} + 955z^{15} + 1225z^{16} + 1525z^{17} \\
&+ 1883z^{18} + 2259z^{19} + 2686z^{20} + 3116z^{21} + 3574z^{22} + 4010z^{23} \\
&+ 4454z^{24} + 4837z^{25} + 5199z^{26} + 5477z^{27} + 5704z^{28} + 5823z^{29} \\
&+ 5884z^{30} + 5823z^{31} + 5704z^{32} + 5477z^{33} + 5199z^{34} + 4837z^{35} \\
&+ 4454z^{36} + 4010z^{37} + 3574z^{38} + 3116z^{39} + 2686z^{40} + 2259z^{41} \\
&+ 1883z^{42} + 1525z^{43} + 1225z^{44} + 955z^{45} + 736z^{46} + 548z^{47} \\
&+ 407z^{48} + 287z^{49} + 203z^{50} + 136z^{51} + 91z^{52} + 56z^{53} \\
&+ 36z^{54} + 20z^{55} + 12z^{56} + 6z^{57} + 3z^{58} + z^{59} + z^{60}.
\end{aligned}
$$

Fig. 12.2 The generating function for plane partitions inside a $3 \times 4 \times 5$ box.

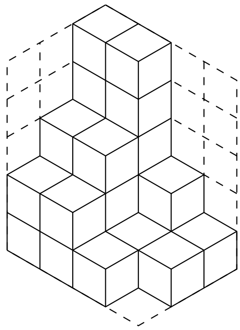
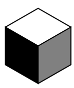


Fig. 12.3 One of the 5199 plane partitions of $n = 26$ that fits in a $3 \times 4 \times 5$ box.

Lozenge tilings of a hexagon

There are many other interesting questions one can ask about plane partitions, especially as relate to symmetries. We will finish the discussion by connecting back to the ideas of Chapter 11 and consider the "typical" features of a large random plane partition. First, notice that when the dimensions of the box are fixed, our solid Young diagrams are in bijection with so-called "lozenge tilings" of a hexagon. The lozenges are the three types of rhombi that correspond to the different exposed faces of the small cubes. We can shade these for emphasis: white on top, black on the front left face, gray on the front right face, as seen below.



Rather than think about the solid Young diagram, we can speak of the lozenge tiling instead.

There are robust methods for studying random tiling models like this, and an incredible result due to Henry Cohn, Michael Larsen, and Jim Propp says that as the size of the hexagon gets large, a typical lozenge tiling has its corners filled by mostly the same type of tile. Moreover, in the case of a regular hexagon (with $a = b = c$) the boundary between the corners and the middle is very close to an inscribed circle! See Figure 12.4. In other words, from a probabilistic standpoint, the corners are "frozen" whereas the interior is warmer. This is what is known as an *arctic circle theorem*.

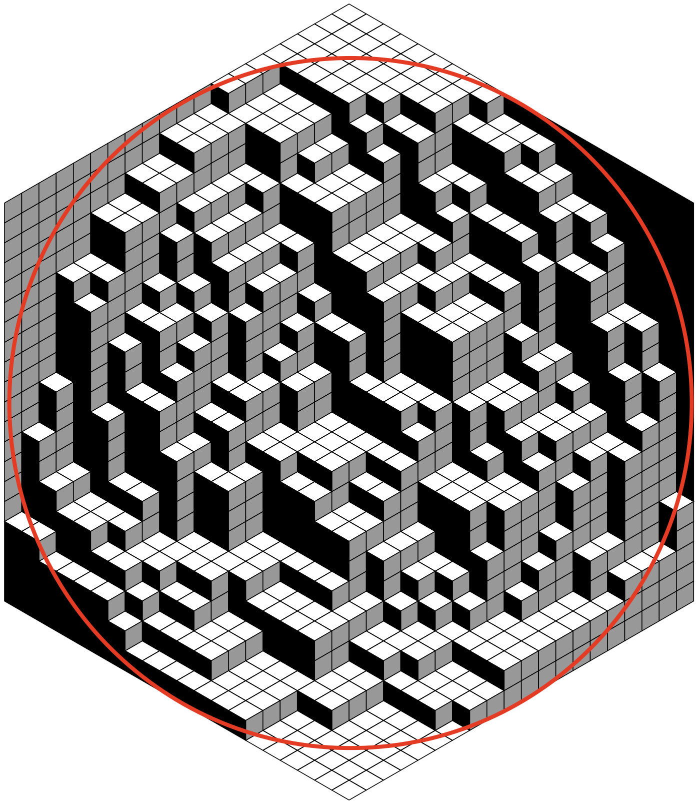


Fig. 12.4 A random plane partition in a $20 \times 20 \times 20$ box.

Further reading

- Amir Aczel and Ken Ono, "My Search for Ramanujan: How I Learned to Count," Springer, 2016.

Srinavasa Ramanujan is one of the fabled names in early 20th century mathematics. His best known contributions have to do with questions about partition numbers. Ken Ono is a prominent number theorist who has contributed a lot to the understanding of partition numbers in more recent years. This book is a memoir of Ono's life, woven together with Ramanujan's life story.

- David Bressoud, "Proofs and Confirmations," Cambridge University Press (1999).

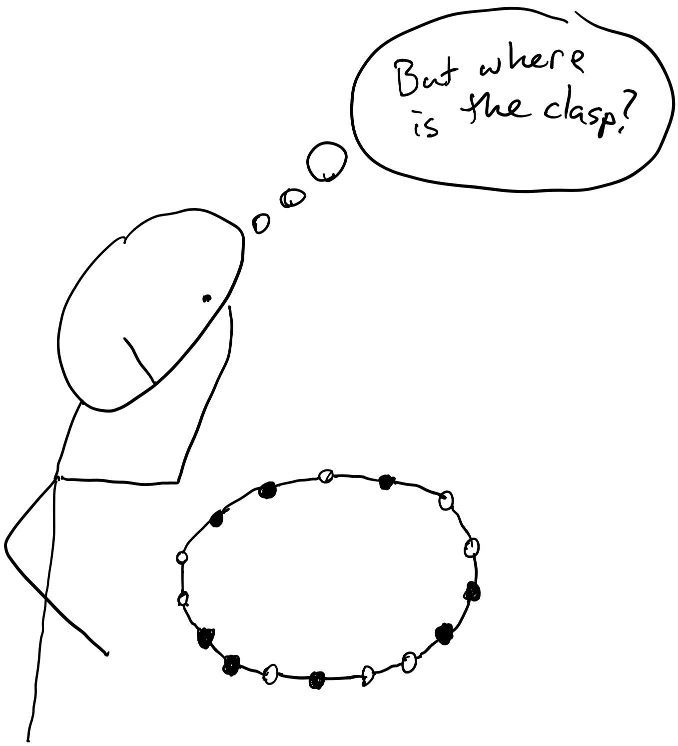
  This book brings mathematical research to life. It studies the problem of "alternating sign matrices" but there are connections to plane partitions.
- Robert Kanigel, "The Man Who Knew Infinity: A Life of the Genius Ramanujan," Washington Square Press, 1992.

  This biography of Ramanujan was the basis for the movie of the same name.

Chapter 13
A bit of number theory

"God created infinity, and man, unable to understand infinity, had to invent finite sets."

–Gian-Carlo Rota



How many necklaces are there with eight black beads and eight white beads?

We have seen ordinary generating functions, and we have seen exponential generating functions. In this chapter we will investigate Dirichlet generating functions. These are most useful for studying sequences that are related to divisibility.

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0_13

For most of the book, we have not needed any particular knowledge other than algebra skills and our wits. (Okay, maybe a tiny bit of calculus.) In this chapter we need to assume only a little more, all of which can be found in most undergraduate texts on number theory. The main focus here is the multiplicative structure of positive integers.

Divisibility

Definition 41. Suppose d and n are integers. If there exists an integer k such that $n = dk$, we say d *divides* n. In symbols we write $d|n$.

For example, $3|54$, since $54 = 3 \cdot 18$. Some synonyms for "d divides n" include "d is a divisor or n," "d is a factor of n," or "n is a multiple of d." While the definition of divisibility applies equally to negative numbers, e.g., $7|(-35)$, we will primarily restrict our attention to positive integers.

Greatest common divisor

Definition 42. The *greatest common divisor* of integers m and n, denoted $\gcd(m, n)$, is the largest integer d such that both m and n are multiples of d. If $\gcd(m, n) = 1$, we say the integers m and n are *relatively prime*.

For example, $\gcd(36, 24) = 12$, and $\gcd(14, 9) = 1$.

Warmup 40. Find $\gcd(72, 28)$ and $\gcd(310, 186)$. (There is a quick method for computing $\gcd(m, n)$ known as the *Euclidean algorithm*. You don't need to use this method to compute the greatest common divisor, but if you don't know the method, you may want to look it up.)

We assume you are familiar with the notion of a *prime* number, i.e., an integer greater than 1 that has no proper divisors. The first few of these are:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots.$$

We can write any integer greater than 1 as a product of primes, e.g., $84 = 2^2 \cdot 3 \cdot 7$. This is known as a *prime factorization*. We will rely heavily on the fact that prime factorizations are unique.

Fundamental Theorem of Arithmetic

Theorem 8. For any integer $n > 1$, there exists a unique set of primes $p_1 < p_2 < \cdots < p_k$ and positive integer exponents $e_1, e_2, \ldots, e_k$ such that

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}.$$

The fact that such factorizations exist is not too difficult to justify by mathematical induction. The uniqueness of the factorization is less obvious, but follows by appealing to "Euclid's Lemma": if a and b are integers and p is a prime that divides ab, then $p|a$ or $p|b$. Details can be found in most any book on elementary number theory.

Warmup 41. Find the prime factorizations of 342 and 510, then compute $\gcd(342, 510)$.

We have seen ordinary and exponential generating functions. Now we consider a new type of generating function, one that is built to deal with number-theoretic sequences. Loosely, these are sequences whose recursive structure is tied up with prime factorization.

Dirichlet generating function

Definition 43. Given a sequence of numbers $a_1, a_2, \ldots, a_k, \ldots,$ we define its *Dirichlet generating function* by:

$$F(s) = \sum_{n\geq 1} a_n \frac{1}{n^s} = a_1 + a_2 \frac{1}{2^s} + a_3 \frac{1}{3^s} + \cdots + a_n \frac{1}{n^s} + \cdots.$$

Addition and scalar multiplication of Dirichlet series work as you might expect:

$$c \cdot \sum_{n\geq 1} a_n \frac{1}{n^s} = \sum_{n\geq 1} c a_n \frac{1}{n^s},$$

and

$$\sum_{n\geq 1} a_n \frac{1}{n^s} + \sum_{n\geq 1} b_n \frac{1}{n^s} = \sum_{n\geq 1} (a_n + b_n) \frac{1}{n^s},$$

but multiplication for Dirichlet series is really nice for number-theoretic purposes. Suppose $F(s) = \sum a_n/n^s$ and $G(s) = \sum b_n/n^s$ are two Dirichlet series. Then

$$\begin{aligned}
F(s)G(s) &= \left(a_1 + a_2 \frac{1}{2^s} + a_3 \frac{1}{3^s} + \cdots\right)\left(b_1 + b_2 \frac{1}{2^s} + b_3 \frac{1}{3^s} + \cdots\right) \\
&= a_1 b_1 + (a_1 b_2 + a_2 b_1)\frac{1}{2^s} + (a_1 b_3 + a_3 b_1)\frac{1}{3^s} \\
&\quad + (a_1 b_4 + a_2 b_2 + a_4 b_1)\frac{1}{4^s} + (a_1 b_5 + a_5 b_1)\frac{1}{5^s} + \cdots \\
&= \sum_{n\geq 1} \left(\sum_{dm=n} a_d b_m\right) \frac{1}{n^s} \\
&= \sum_{n\geq 1} \left(\sum_{d|n} a_d b_{\frac{n}{d}}\right) \frac{1}{n^s}.
\end{aligned}$$

This is the key property of multiplication: the coefficients in a product of Dirichlet series can be expressed as a sum over all divisors.

The power series $1/(1-z)$ and e^z were fundamental building blocks in earlier chapters since they encoded the sequence $1, 1, 1, \ldots$ in an ordinary or exponential generating function. The Dirichlet generating function for the sequence $1, 1, 1, \ldots$ has a special name. It is known as the *Riemann zeta function*.

Riemann zeta function

Definition 44. We define the *Riemann zeta function* as the formal series

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} + \cdots .$$

As an analytic creature, $\zeta(s)$ is rather different from the other types of functions we have worked within this book. You may know some facts about special values of ζ from a calculus class. For example, $\zeta(1) = \sum 1/n$ is the harmonic series, which diverges, while $\zeta(2) = \frac{\pi^2}{6}$. The problem of computing the sum of this series was known as the *Basel problem* and one of Euler's first big achievements was its solution.

For us, it will be good enough to think of $\zeta(s)$ as a formal series, along with the following fundamental formula. This formula expresses $\zeta(s)$ as a product over all prime numbers:

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - (\frac{1}{p})^s}. \tag{13.1}$$

To see why this identity holds it helps to start small. For $p = 2$ we have

$$\frac{1}{1 - (\frac{1}{2})^s} = 1 + \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{8^s} + \cdots = \sum_{k \geq 0} \frac{1}{(2^k)^s}.$$

The only numbers appearing in the denominator are powers of 2. Likewise,

$$\frac{1}{1 - (\frac{1}{3})^s} = 1 + \frac{1}{3^s} + \frac{1}{9^s} + \frac{1}{27^s} + \cdots = \sum_{l \geq 0} \frac{1}{(3^l)^s}.$$

If we multiply these two series we get

$$\frac{1}{1 - (\frac{1}{2})^s} \cdot \frac{1}{1 - (\frac{1}{3})^s} = \sum_{k,l \geq 0} \frac{1}{(2^k 3^l)^s}.$$

Now the numbers showing up in the denominator are precisely those that have only 2 or 3 as prime factors. But of course to produce every positive integer n, we need to include all the primes. Moreover, the uniqueness of the Fundamental Theorem of Arithmetic says we won't get any integers n repeated in the sum.

Note: A cautious reader might squirm at this point, since we are passing from a finite product to an infinite product. As in Chapter 6 when we first introduced generating functions, we can either take a formal algebraic stance or a careful analytic accounting. Either way, the statement can be placed on solid ground and we should be unafraid to push forward.

Warmup 42. Show that $\zeta(s-1) = \sum_{n\geq 1} n\frac{1}{n^s}$. What sequence does $\zeta(s-k)$ encode?

Problem 152.
Let $\tau(n)$ denote the number of divisors of n. Show that

$$\zeta(s)^2 = \sum_{n\geq 0} \tau(n)\frac{1}{n^s}.$$

Problem 153.
Let $\sigma(n)$ denote the sum of divisors of n. Show that

$$\zeta(s)\zeta(s-1) = \sum_{n\geq 0} \sigma(n)\frac{1}{n^s}.$$

Problem 154.
Define

$$\sigma_k(n) = \sum_{d|n} d^k.$$

This function generalizes the number of divisors and sum of divisors: $\sigma_0(n) = \tau(n)$ and $\sigma_1(n) = \sigma(n)$. Express the Dirichlet generating function for $\sigma_k(n)$ in terms of the zeta function.

Multiplicative function

Definition 45. A function f is *multiplicative* if, for any pair of relatively prime positive integers m and n, $f(mn) = f(m)f(n)$.

By the Fundamental Theorem of Arithemetic, this means that if $n = p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}$, then

$$f(n) = f(p_1^{e_1})f(p_2^{e_2})\cdots f(p_k^{e_k}).$$

We will give several examples of functions that have this property.

First, the divisor counting function $\tau(n)$ from Problem 152 is multiplicative. To see this, suppose m and n are relatively prime. Then by the Fundamental Theorem of Arithmetic, any divisor of their product can be uniquely expressed as the product of a divisor d of m and a divisor d' of n. Hence $\tau(mn) = \tau(m)\tau(n)$. In fact, a similar argument shows the function $\sigma_k(n)$ from Problem 154 is multiplicative for any k.

Warmup 43. List all the divisors of $m = 3$ and $n = 4$, then list all the divisors of $mn = 12$. Verify $\sigma_k(12) = \sigma_k(3)\sigma_k(4)$ for $k = 0, 1, 2$.

Another important example counts how many numbers are relatively prime to a given number.

Totient function

Definition 46. The *totient function*, ϕ, counts the number of positive integers that are less than or equal, and relatively prime to, a given input, i.e.,

$$\phi(n) = |\{1 \leq i \leq n : \gcd(i, n) = 1\}|.$$

Notice that $\phi(1) = 1$ and if p is a prime number, then $\phi(p) = p - 1$. It takes a bit of thought to see why ϕ is multiplicative, but we will just take this fact for granted.

Warmup 44. Compute $\phi(16)$, $\phi(32)$, $\phi(20)$.

Our final example may seem rather strange at first, but it will turn out to play an important role in certain counting problems. It is multiplicative by fiat.

Möbius function

Definition 47. The *Möbius function*, μ, is the multiplicative function that records the presence or absence of distinct prime factors. For each power of a prime μ is defined by

$$\mu(p^e) = \begin{cases} 1 & \text{if } e = 0, \\ -1 & \text{if } e = 1, \\ 0 & \text{if } e \geq 2. \end{cases}$$

We define $\mu(n)$ in terms of the prime factorization of n. If $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$, then $\mu(n) = \mu(p_1^{e_1})\mu(p_2^{e_2}) \cdots \mu(p_k^{e_k})$.

Warmup 45. Compute $\mu(15)$, $\mu(30)$, $\mu(45)$. Try to come up with a simple way to say when $\mu(n) = 0$.

Problem 155.
Show that $\sigma_k(n)$ is multiplicative for any k. Hint: it suffices to consider the case $n = p^e \cdot m$ and $\gcd(m, p) = 1$.

Problem 156.
Show that

$$\sum_{d|n} \phi(d) = n.$$

Hint: write all the fractions $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$ in reduced form and count them according to denominators.

Problem 157.
Show that $\mu(1) = 1$ and if $n > 1$,

$$\sum_{d|n} \mu(d) = 0.$$

Hint: first consider which divisors have $\mu(d) = 0$, then, by considering subsets of the prime factors of n, show that the other divisors pair off in a natural way.

Problem 158.
Show that if $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$, then

$$\phi(n) = \prod_{i=1}^{k} (p_i^{e_i} - p_i^{e_i - 1}) = \prod_{i=1}^{k} p_i^{e_i} \left(1 - \frac{1}{p_i}\right) = n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right).$$

This is called *Euler's formula* for the totient function. Hint: since ϕ is multiplicative, it suffices to consider the case of a prime power.

Fundamental Theorem of multiplicative functions

Theorem 9. Suppose $f(n)$ is a multiplicative function and

$$F(s) = \sum_{n \geq 1} f(n) \frac{1}{n^s}$$

is the Dirichlet series for the function. Then

$$F(s) = \prod_{p \text{ prime}} \left(f(1) + f(p) \frac{1}{p^s} + f(p^2) \frac{1}{(p^2)^s} + \cdots \right).$$

This identity greatly generalizes our formula for the zeta function given in Equation 13.1, but the reasoning is similar.

For each prime p, we can write the series for the powers of p, call it F_p:

$$F_p(s) = \sum_{e \geq 0} f(p^e) \frac{1}{(p^e)^s} = f(1) + f(p) \frac{1}{p^s} + f(p^2) \frac{1}{(p^2)^s} + \cdots$$

If we multiply two of these, say $F_p(s)$ and $F_q(s)$, we get

$$\begin{aligned}
F_p(s)F_q(s) &= \left(\sum_{e\geq 0} f(p^e)\frac{1}{(p^e)^s}\right)\left(\sum_{k\geq 0} f(q^k)\frac{1}{(q^k)^s}\right), \\
&= \sum_{e,k\geq 0} f(p^e)f(q^k)\frac{1}{(p^e)^s(q^k)^s}, \\
&= \sum_{e,k\geq 0} f(p^eq^k)\frac{1}{(p^eq^k)^s},
\end{aligned}$$

where the final equation holds because f is multiplicative. In other words, the product is a sum over precisely all integers whose prime factors are p or q.

The same holds for any number of factors, and, given uniqueness of factorizations in the Fundamental Theorem of Arithmetic, the result follows.

Problem 159.
Show that

$$\sum_{n\geq 1} \mu(n)\frac{1}{n^s} = \frac{1}{\zeta(s)}.$$

Hint: first consider the case of powers of a fixed prime p, then use Theorem 9.

Problem 160.
Suppose $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ are sequences related by the identity

$$a_n = \sum_{d|n} b_d,$$

for all $n \geq 1$.

1. Let $F(s) = \sum_{n\geq 1} a_n \frac{1}{n^s}$ and $G(s) = \sum_{n\geq 1} b_n \frac{1}{n^s}$. Show that
$$F(s) = G(s)\zeta(s).$$
2. Conclude that
$$b_n = \sum_{d|n} \mu(d) a_{\frac{n}{d}}.$$
This technique for expressing the b_n sequence in terms of the a_n sequence is known as *Möbius inversion*.

Problem 161.
We will study the totient function again.

1. Use the result of Problem 156 to conclude that
$$\sum_{n\geq 1} \phi(n)\frac{1}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}$$

and by Möbius inversion,

$$\phi(n) = \sum_{d|n} \mu(d) \cdot \frac{n}{d} = n \cdot \sum_{d|n} \frac{\mu(d)}{d}.$$

2. Use the definition of the Möbius function to conclude that for any $n \geq 1$,

$$\sum_{d|n} \frac{\mu(d)}{d} = \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

where the product is taken over all distinct primes dividing n.

3. Deduce Euler's formula from Problem 158,

$$\phi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Problem 162.

Use the Dirichlet generating functions for $\tau(n)$, $\sigma(n)$, and $\phi(n)$ to deduce the identity

$$\sigma(n) = \sum_{d|n} \tau(d) \phi\left(\frac{n}{d}\right).$$

Problem 163.

A *primitive* binary string is one which cannot be written as a concatenation of smaller, identical binary strings. For example, 1101010110 is primitive, but 101101101 is not primitive.

1. List all primitive binary strings of length $n = 1, 2, 3, 4$.
2. Explain why a binary string of length n that is not primitive can be written uniquely as a concatenation of d identical copies of a primitive string of length $\frac{n}{d}$. To be more precise, prove the following: For any binary string of length n that is not primitive, call it s, there exists a unique integer $d|n$ and a primitive string t of length n/d, such that s is the concatenation of d copies of t.
3. Let $b(n)$ denote the number of primitive binary strings of length n, and show

$$2^n = \sum_{d|n} b(n).$$

4. Use Möbius inversion to give a formula for $b(n)$. What is $b(20)$?

Problem 164.

A *binary necklace* is a binary string drawn on a circle. In this visual drawing, we think of a white bead as a "0" and a black bead as a "1." See Figure 13.2 for the necklaces of length 5. Note that for a necklace, all that matters is the circular ordering of the beads.

If we "cut" a necklace, it becomes a binary string. For example suppose we cut the binary necklace below at the location indicated in Figure 13.1. If we read clockwise from the cut point, this corresponds to the binary string 000100011101. For another example, notice that each of the strings $00101, 10010, 01001, 10100, 01010$ correspond to the same 5-bead necklace (the leftmost picture in the second row of Figure 13.2).

We say a necklace is *primitive* if, when cut, the corresponding binary string is primitive in the sense of Problem 163.

1. Let $\eta(n)$ denote the number of binary necklaces with n beads and let $c(n)$ denote the number of primitive binary necklaces with n beads. Explain why
$$\eta(n) = \sum_{d|n} c(d).$$
2. Show
$$n \cdot c(n) = b(n),$$
where $b(n)$ is the number of primitive binary strings. Use the formula for $b(n)$ from Problem 163 to give a formula for $c(n)$, and hence $\eta(n)$. (Your formula for $\eta(n)$ should be a double sum involving the Möbius function and powers of 2.)
3. Let
$$B(s) = \sum_{n\geq 1} b(n)\frac{1}{n^s}, \quad C(s) = \sum_{n\geq 1} c(n)\frac{1}{n^s},$$
$$N(s) = \sum_{n\geq 1} \eta(n)\frac{1}{n^s}, \quad T(s) = \sum_{n\geq 1} 2^n\frac{1}{n^s},$$
and recall
$$\frac{1}{\zeta(s)} = \sum_{n\geq 1} \mu(n)\frac{1}{n^s}, \quad \frac{\zeta(s-1)}{\zeta(s)} = \sum_{n\geq 1} \phi(n)\frac{1}{n^s}.$$
Use identities between these generating functions to prove
$$\eta(n) = \frac{1}{n}\sum_{d|n} \phi(d) 2^{\frac{n}{d}}.$$

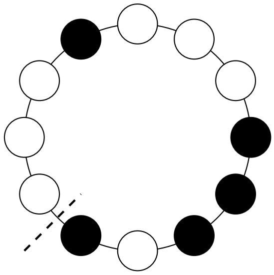


Fig. 13.1 A cut in a necklace.

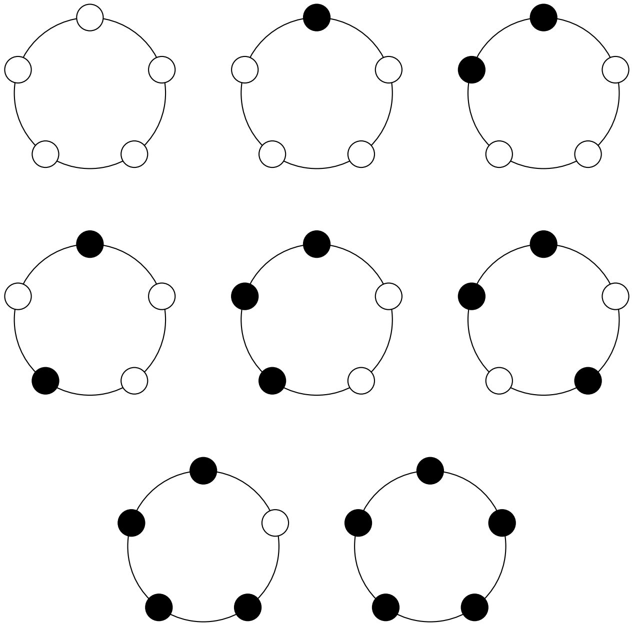


Fig. 13.2 The binary necklaces with five beads, demonstrating $\eta(5) = 8$.

General Möbius Inversion

Classical Möbius inversion can be viewed as a special case of a general technique for inverting formulas indexed by partially ordered sets. This idea was developed by Gian-Carlo Rota in the mid-twentieth century.

To get a feel for the general concept, consider the following example. Suppose we have some elements A, B, C, D, E, F, G related as indicated in Figure 13.3.

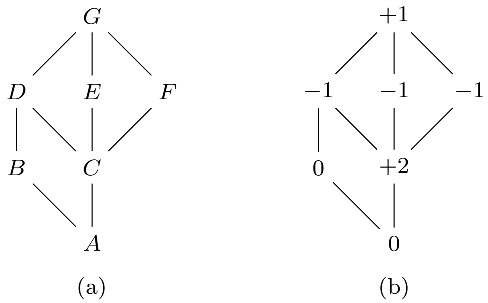


Fig. 13.3 (a) A partial ordering of the set $\{A, B, C, D, E, F, G\}$, and (b) values of the Möbius function $\mu(x, G)$.

Further, suppose α and β are functions defined on these elements. Further, suppose we know all values of α and we know that $\alpha(y) = \sum_{x \le y} \beta(x)$, where $x \le y$ means x is below y in the diagram. Concretely, we have

$$\begin{aligned}
\alpha(G) &= \beta(A) + \beta(B) + \beta(C) + \beta(D) + \beta(E) + \beta(F) + \beta(G),\\
\alpha(F) &= \beta(A) + \beta(C) + \beta(F),\\
\alpha(E) &= \beta(A) + \beta(C) + \beta(E),\\
\alpha(D) &= \beta(A) + \beta(B) + \beta(C) + \beta(D),\\
\alpha(C) &= \beta(A) + \beta(C),\\
\alpha(B) &= \beta(A) + \beta(B),\\
\alpha(A) &= \beta(A).
\end{aligned}$$

Can we solve for the values of β? In particular, what is $\beta(G)$?

Since $\alpha(A) = \beta(A)$, we can substitute and work recursively, solving $\beta(B) = \alpha(B) - \alpha(A)$, $\beta(C) = \alpha(C) - \alpha(A)$, and so on, until we find

$$\beta(G) = \alpha(G) - \alpha(F) - \alpha(E) - \alpha(D) + 2\alpha(C). \tag{13.2}$$

We can see that all that is required to invert a system of equations like this is a little linear algebra.

The idea behind Möbius inversion is to give a general method for solving for β in terms of α like this. But rather than thinking of it as a linear algebra problem, we come up with a description for the coefficients in terms of the combinatorics on the ordering of the elements $\{A, B, C, \ldots\}$. To make this notion more precise we need a little background.

Posets

We have mentioned partially ordered sets several times before in this book. Let's see the precise definition now. A partially ordered set P, or "poset," is a set with a relation "$\leq$" that is

- (reflexive) $x \leq x$ for all $x \in P$,
- (antisymmetric) if $x \leq y$ and $y \leq x$, then $x = y$, and
- (transitive) if $x \leq y$ and $y \leq z$, then $x \leq z$.

What distinguishes a partially ordered set from a totally ordered set is that elements can be incomparable.

One familiar partial order is the set of subsets of a finite set ordered by inclusion. This is illustrated in Figure 13.4(a) for the set $\{a, b, c\}$. Another is the lattice of divisors of a positive integer, as shown in Figure 13.5(b) for $n = 300$.

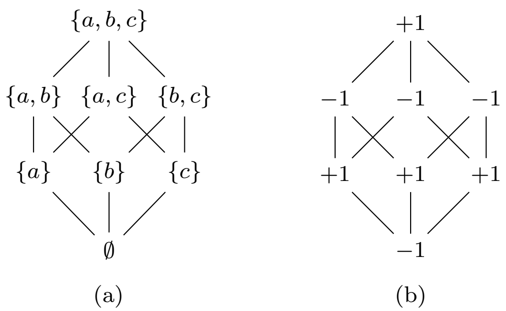


Fig. 13.4 (a) The poset of subsets of $\{a, b, c\}$ and (b) the Möbius function $\mu_P(I, \{a, b, c\})$.

The posets for which Möbius inversion works need not be finite, but they should be "locally finite" in the sense that every interval $[u, v] = \{p \in P : u \leq p \leq v\}$ is finite. It is also helpful to assume that the poset has a unique minimal element.

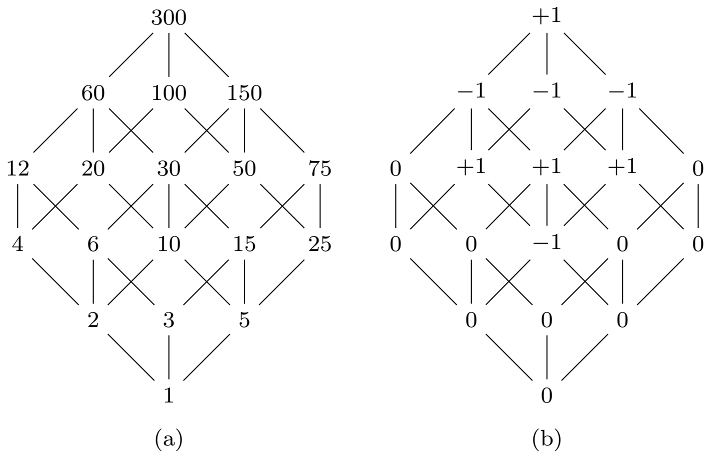


Fig. 13.5 (a) The poset of divisors of $n = 300$ and (b) the Möbius function $\mu_P(d, 300)$.

Möbius function of a poset

Let P be a locally finite poset. The *Möbius function* of P, denoted μ_P, is defined recursively as follows for any $u \leq v$:

$$\mu_P(u,v) = \begin{cases} 1 & \text{if } u = v, \\ -\sum_{u \leq x < v} \mu_P(u,x) & \text{if } u < v. \end{cases}$$

Another way to think about this is that we have $\mu_P(v,v) = 1$ for all v and

$$\sum_{u \leq x \leq v} \mu_P(x,v) = 0.$$

For example, if we fix $v = G$ in the poset of Figure 13.3(a) we have the values for $\mu_P(x, G)$ as shown in Figure 13.3(b). Compare these values with the coefficients we found in Equation (13.2).

The general version of Möbius inversion reads like this. Suppose P is a locally finite poset with a unique minimal element, and suppose α and β are functions defined on P that are related via

$$\alpha(y) = \sum_{x \leq y} \beta(x),$$

where the sum is over all elements below y in P. (Note that since we assume P has a unique minimum element, this is a finite sum.) Then a formula for β is given by

$$\beta(y) = \sum_{x \le y} \mu_P(x,y)\alpha(x). \tag{13.3}$$

We will now look at several manifestations of this formula.

Inclusion–exclusion and descent sets

If P is the set of all subsets of a finite set S, then induction shows that for any sets $I \subseteq J$, we have $\mu_P(I,J) = (-1)^{|J|-|I|}$. (See Figure 13.4(b).) This means if

$$\alpha(J) = \sum_{I \subseteq J} \beta(I),$$

then

$$\beta(J) = \sum_{I \subseteq J} (-1)^{|J|-|I|} \alpha(I). \tag{13.4}$$

This relation is sometimes known as "inclusion–exclusion."

Let's use inclusion–exclusion to count the number of permutations in S_n with descent set J. That is, let $\mathrm{Des}(w) = \{j : w(j) > w(j+1)\}$ denote the set of descents of the permutation w, and define

$$\alpha(J) = |\{w \in S_n : \mathrm{Des}(w) \subseteq J\}| \quad \text{and} \quad \beta(J) = |\{w \in S_n : \mathrm{Des}(w) = J\}|.$$

To form a permutation whose descent set is contained in $J = \{j_1 < j_2 < \cdots < j_k\}$, we can first choose j_1 elements and arrange them in increasing order, then choose another $j_2 - j_1$ elements and arrange them in increasing order, and so on. See Figure 13.6.

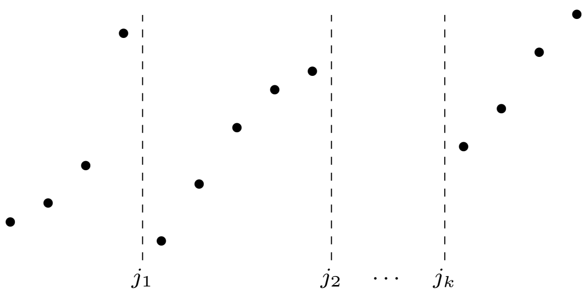


Fig. 13.6 Arranging increasing runs between elements of $J = \{j_1 < j_2 < \cdots < j_k\}$.

There might or might not be descents in the positions indexed by J, but there certainly cannot be descents elsewhere. Thus it is not too difficult to see that

$$\alpha(J) = \binom{n}{j_1, j_2 - j_1, j_3 - j_2, \ldots, n - j_k}.$$

By inclusion–exclusion, we find

$$\begin{aligned}\beta(J) &= \sum_{I \subseteq J} (-1)^{|J|-|I|} \alpha(I) \\ &= \sum_{i=0}^{k} (-1)^{k-i} \sum_{1 \leq a_1 < \cdots < a_i \leq k} \binom{n}{j_{a_1}, j_{a_2} - j_{a_1}, \ldots, n - j_{a_i}}.\end{aligned}$$

Derangements problem revisited

Notice that if $\alpha(J)$ and $\beta(J)$ depend only on the cardinality of J, then for any $|J| = n$, we can write $\alpha(J) = \alpha(n)$ and $\beta(J) = \beta(n)$. Collecting like terms in Equation (13.4) we get

$$\beta(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \alpha(k).$$

To see this idea applied, we can revisit Problem 82, the problem of counting derangements. Let $d_n = |\{w \in S_n : w(i) \neq i \text{ for any } i\}|$ denote the number of derangements, i.e., those permutations that have no fixed points.

Let $I \subseteq \{1, 2, \ldots, n\}$. To form a permutation whose fixed point set is exactly I, we must ensure that the remaining elements in $\{1, 2, \ldots, n\} - I$ are permuted according to a derangement. By summing over all I, we obtain all permutations in S_n. Thus,

$$n! = \sum_{I \subseteq \{1,2,\ldots,n\}} d_{n-|I|}.$$

By inclusion–exclusion,

$$d_n = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{n-|I|} |I|!$$

Each summand depends only on the cardinality of I, not I itself. Thus we can group these terms according to cardinality to find

$$
\begin{aligned}
d_n &= \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} k! \\
&= n! \sum_{k=0}^{n} (-1)^{n-k} \frac{1}{(n-k)!} \\
&= n! \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots (-1)^n \frac{1}{n!} \right).
\end{aligned}
$$

The poset of divisors

Next we show how number-theoretic Möbius inversion fits in this framework. The underlying poset P is all positive divisors of n, ordered by divisibility. Here, "$x \leq y$" in P means $x|y$, and $\mu_P(x,y) = \mu\left(\frac{y}{x}\right)$. In the poset of divisors, we see that if $n = p_1^{e_1} \cdots p_k^{e_k}$ then the poset has the stucture of a cartesian product totally ordered sets $1|p|p^2| \cdots |p^e$, e.g., in the case of $n = 300 = 2^2 \cdot 3 \cdot 5^2$, we have

$$
P = \{1|2|2^2\} \times \{1|3\} \times \{1|5|5^2\},
$$

which is also illustrated in Figure 13.7.

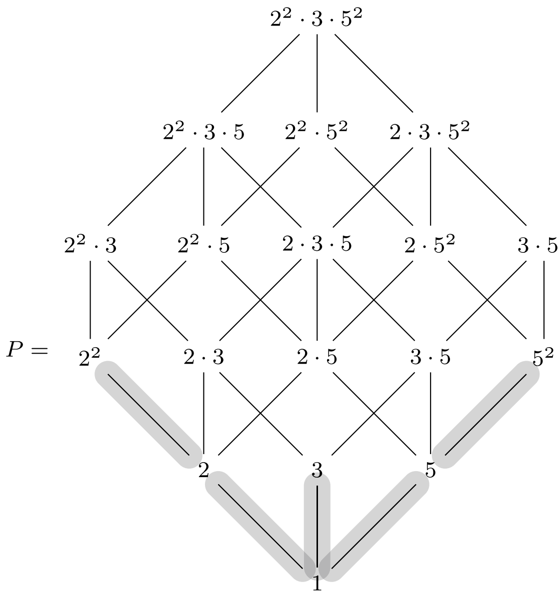


Fig. 13.7 The poset of divisors of $n = 300$.

Compare with Figure 13.5(b).

Further reading

- George Andrews, “Number Theory,” Dover, 1994.
  There are many good books on Number Theory out there. This is a good one for blending Number Theory with Combinatorics. It is inexpensive and well written.
- Paul Hoffman, “The Man Who Loved Only Numbers,” Hachette Books, 1998.
  This is an amazing biography of Paul Erdős, one of the greatest mathematicians of the 20th century.
- Simon Singh, “Fermat’s Enigma,” Anchor, 1998.
  This is a great journalistic account of one the biggest mathematical accomplishments of the last century.

Appendix: Supplementary exercises

T. K. Petersen, *Inquiry-Based Enumerative Combinatorics*, Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-18308-0

Exercises for Chapter 1

1. Use set builder notation to describe the set of all odd two-digit numbers.
2. Let $S = \{n \in \mathbb{Z} : 1 \leq n \leq 100\}$ and let T denote the set of positive multiples of 3 with at most three digits, i.e., $T = \{3, 6, 9, \ldots, 999\}$. Compute the following cardinalities.

   a. $|S|$
   b. $|T|$
   c. $|S \cap T|$
   d. $|S \cup T|$

3. A PIN (personal identification number) is a sequence of four digits used for security purposes by banks and other organizations to protect consumer information. Each digit is typically from 0 to 9 e.g., 0394 is a PIN.

   a. How many PIN numbers are there?
   b. How many PIN numbers have no repeated digits?
   c. How many PIN numbers repeat at least one digit?

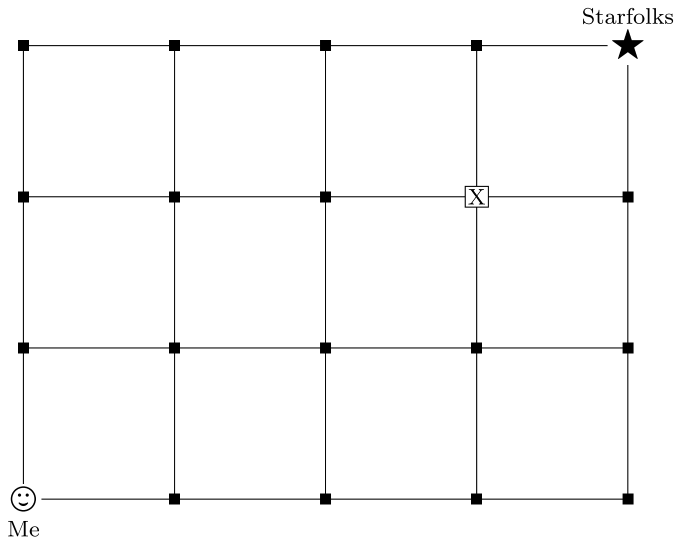


Fig. A.1 Going for coffee.

4. In Figure A.1 we see the same grid of streets from Problem 7, but there is construction at the intersection marked with an X. How many different routes can I take to get to Starfolks that avoid that intersection? (Like problem 7, assume I only walk East or North from one intersection to another.)

5. Consider the grid of streets shown in Figure A.2.

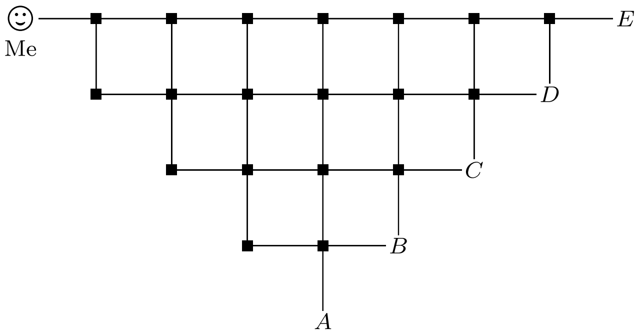


Fig. A.2 A new neighborhood of streets.

How many ways can I walk eight blocks on this grid of streets, assuming I only walk East and South? (Notice this means I have to reach either A, B, C, D or E.)
In what ways can you generalize this problem?

6. Let $S = \{1,2,3,4,5,6\}$ and let $T = \{(i,j) \in \mathbb{Z} \times \mathbb{Z} : 1 \leq i \leq j \leq 6\}$.

   a. How is the set $S \times S$ related to rolling a pair of dice?
   b. What is $|S \times S|$?
   c. How is the set T related to rolling a pair of dice?
   d. What is $|T|$?
   e. What circumstances would lead you to prefer to use set $S \times S$ versus set T to model the roll of two dice?

7. Consider drawing three cards from a standard deck as in Problem 14, part 4. How many ways can you:

   a. Select a ten in round 1, then a nine in round 2, then an eight in round 3?
   b. Select a nine, then an eight, then a ten?
   c. Select three cards whose ranks are $\{8, 9, 10\}$ (ignore suits, and ignore the order in which the cards are selected).
   d. Select three cards that form a run? (A *run* is any three consecutive cards of the *same suit*, where the ace can be played low, as in the run $A♣, 2♣, 3♣$, or high, as in $Q♣, K♣, A♣$, but a run cannot "wrap around" as in $K♣, A♣, 2♣$.)

   Explain your reasoning for each part.

8. Consider all possible subsets of $\{1,2,3\}$. How many unordered pairs of distinct subsets (i.e., $\{A, B\}$ with $A \neq B$) are there?
   Now among the pairs of distinct subsets, how many are there with the following properties?

a. $A \cup B = \{1, 2, 3\}$
b. $A \cap B = \emptyset$
c. $A \cap B = \{3\}$
d. $|A \cap B| = 1$

9. How many functions are there $f : \{a, b, c, d\} \to \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$? How many of these are injective?
10. How many injections $f : \{0, 1, 2\} \to \{a, b, c, d, e\}$ are there? How many surjections $g : \{a, b, c, d, e\} \to \{0, 1, 2\}$ are there?
11. Suppose $|A| = k$ and $|B| = n$, with $3 \le k < n$. Are there more functions $f : A \to B$ or are there more functions $g : B \to A$? Explain.

Exercises for Chapter 2

1. How many permutations of a deck of 52 cards are there? Is this number bigger or smaller than the number of atoms in the solar system?
2. The top students in the state take an exam, and the top three students receive a scholarship. First prize is $20,000, second prize is $10,000, and third prize is $5,000. There are 1729 students taking the test this year. How many ways are there to award the prizes, assuming there are no ties?
3. Which is bigger, $n!$ or 2^n? Explain.
4. Recall a PIN number is a sequence of four digits.

   a. How many PIN numbers use precisely the digits $\{1, 2, 7, 9\}$?
   b. How many PIN numbers use precisely the digits $\{2, 7, 9\}$? (Note one digit must be repeated.)
   c. How many PIN numbers use precisely the digits $\{7, 9\}$?

5. Suppose a bank allows its PIN numbers to include letters as well as numbers.

   a. How many PIN numbers are there, assuming the letters are only lowercase?
   b. How many such PIN numbers don't repeat any characters?
   c. If we allow letters to appear as upper- and lowercase (e.g., `a` and `A` are considered different characters), how many PIN numbers are there?
   d. Allowing both upper- and lowercase letters, how many PIN numbers don't repeat any characters?
   e. Some of the PIN numbers from part d. have the same letter appearing twice—once as a lowercase letter and once as an upper case letter. For example: `a32A`. How many of the PIN numbers from part d. don't have this property? In other words, how many case-sensitive alphanumeric PIN numbers have no repeated numbers and no repeated letters?

6. How many ways are there to rearrange the letters in the word $SAMMY$? Note this is like a permutation of the letters S, A, M, and Y except that the letter M needs to be used twice.
7. How many ways are there to arrange a deck of cards so that all the black cards appear before the red cards?
8. How many ways are there to arrange a deck of cards so that all the clubs appear before any of the spades, which appear before any of the hearts, which appear before any of the diamonds?
9. How many ways are there to arrange a deck of cards so that the face values on the cards are in increasing order, i.e., all the aces appear before the twos, the twos appear before the threes, and so on, with kings at the end?
10. How many non-attacking rook arrangements are there on a 7-by-5 chessboard?

11. Eight colleagues are sitting at a circular table during a banquet at a mathematics conference. If we only care about who sits next to whom in clockwise order, and not the exact seats in which they sit, how many ways can these people be arranged?
12. How many:
    a. permutations w of $\{1, 2, 3, 4, 5\}$ have $w(3) > w(4)$?
    b. permutations w in S_n have $w(3) > w(4)$? (Assuming $n \geq 4$)
13. How many:
    a. permutations w of $\{1, 2, 3, 4, 5\}$ have $w(2) < w(3) > w(4)$?
    b. permutations w in S_n have $w(2) < w(3) > w(4)$? (Assuming $n \geq 4$)
14. How many:
    a. permutations of $\{1, 2, 3, 4, 5\}$ have the number 3 appearing to the right of the number 4?
    b. permutations in S_n have the number 3 appearing to the right of the number 4? (Assuming $n \geq 4$)
15. How many:
    a. permutations of $\{1, 2, 3, 4, 5\}$ have the number 3 appearing to the right of both the number 4 and the number 2?
    b. permutations in S_n have the number 3 appearing to the right of the number 4 and the number 2? (Assuming $n \geq 4$)

Exercises for Chapter 3

1. The local elementary school chess league keeps rankings throughout the school year. The top four finishers will compete in a playoff during the last week of school. If there are 20 players in the league, how many possible combinations of players can compete in the playoff?
2. The math club executive board meets once a month. There are ten members of the board and they begin each meeting by sharing a special handshake. If each board member shakes hands with every other board member, how many handshakes are there in all?
3. In the 2014 FIFA World Cup there were 32 teams arranged into 8 groups of 4.

   a. If the groups are labeled A–H, how many ways could the groups have been formed?
   b. If we don't care about the labels of the groups, how many ways could the groups have been formed?
   c. The United States and Portugal were placed in the same group in 2014. How many ways could the groups have been formed so that these two teams were placed in the same group?

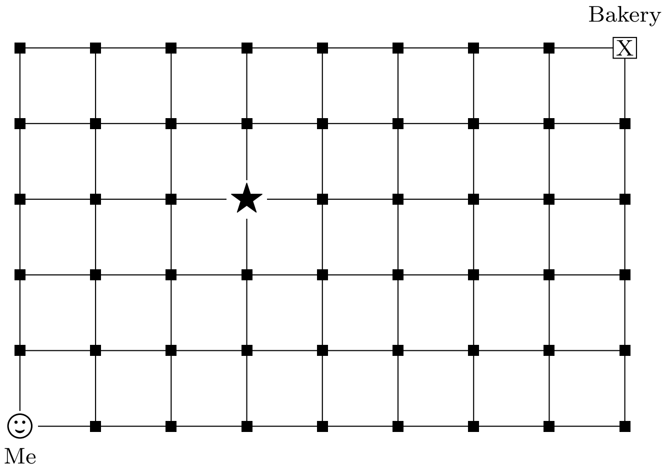
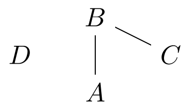


Fig. A.3 Going for coffee and a bagel.

4. This morning I want a coffee from Starfolks (indicated with a ★ on the map in Figure A.3) and a bagel from the bakery (indicated on the map with the X). If I only follow the grid of streets and walk the minimum total distance (e.g., six blocks to Starfolks and seven blocks from Starfolks to the bakery), how many ways can I:

a. Walk from home to Starfolks?
b. Walk from home to the bakery (whether or not I stop at Starfolks)?
c. Walk from Starfolks to the bakery?
d. Walk from home to the bakery, after first stopping for Coffee at Starfolks?
e. Make a round trip: home to Starfolks, Starfolks to bakery, then bakery back home?

5. A *graph* is a pair (V, E), where V is a set called the *vertex set* and E is a set of 2-element subsets of V known as the *edge set*. For example, the pair $(\{A, B, C, D\}, \{\{A, B\}, \{B, C\}\})$ is a graph, which might be drawn like this:



   a. Draw all the graphs with vertex set $\{A, B, C\}$.
   b. How many edges can you form from the vertex set $\{A, B, C, D, E\}$?
   c. How many graphs have vertex set $\{A, B, C, D, E\}$?
   d. What is the greatest number of edges that a graph with n vertices may have?
   e. Fix a finite set V with $|V| = n$. How many graphs have vertex set V?

6. Recall that an *anagram* is a rearrangement of the letters in a word.

   a. Come up with a five-letter name that has 120 anagrams.
   b. Come up with a five-letter name that has 60 anagrams.
   c. Which has more anagrams, ILLINOIS or OKLAHOMA?
   d. Which U.S. state has the most anagrams?

7. Suppose a bucket has a hundred red balls, a hundred blue balls, and a hundred white balls. If you draw out ten balls from the bucket, how many combinations of ball colors can you see? Note that all we care about is the *number* of each type of ball, i.e., how many different triples of nonnegative integers (r, b, w) have $r + b + w = 10$?
8. How many ways can a deck of cards be arranged so that all the clubs appear before the spades? (Hint: Choose where to place these cards first, then consider how many ways they might occupy those positions.)
9. Notice that $\binom{4}{2} = 1+2+3$ and $\binom{5}{2} = 1+2+3+4$. Use Pascal's identity repeatedly to prove that

$$\binom{n}{2} = 1 + 2 + \cdots + (n-2) + (n-1).$$

10. Explain why

$$\binom{2n}{n} = \sum_{j=0}^{n} \binom{n}{j}^2.$$

11. Consider the binomial coefficients $\binom{2}{1}$, $\binom{4}{2}$, $\binom{8}{4}$, $\binom{16}{8}$, and so on. What is the largest power of 2 that divides $\binom{2^{n+1}}{2^n}$?

Exercises for Chapter 4

1. Expand $(a+b)^7$ without any technology apart from Pascal's triangle.
2. Expand $(2+t)^5$ as a polynomial in t without the aid of technology apart from Pascal's triangle.
3. Expand $(x-y)^5$ as a polynomial in x and y without the aid of technology apart from Pascal's triangle.
4. Expand $(N+E)^5$ and explain how the terms relate to the paths in the grid of streets shown in Figure A.4.

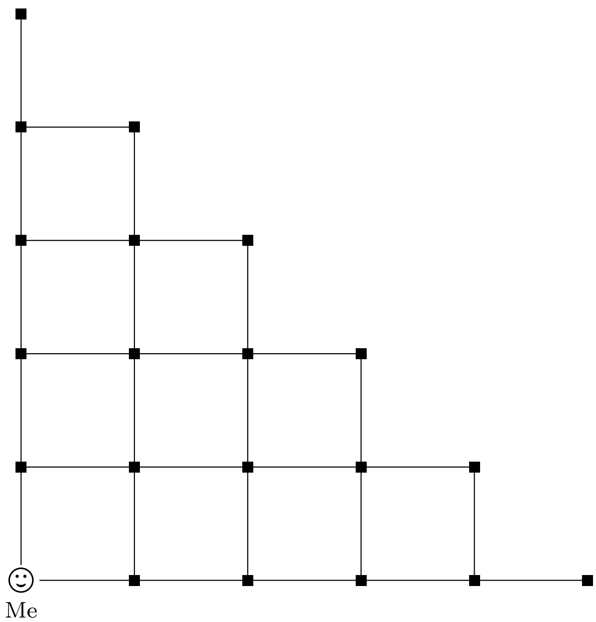


Fig. A.4 Lattice paths in a neighborhood for explaining $(N+E)^5$.

5. Compute $\displaystyle\sum_{k=0}^{100}(-1)^k\binom{100}{k}$.
6. Compute $\displaystyle\sum_{k=0}^{100}2^k(-3)^{n-k}\binom{100}{k}$.
7. Compute $\displaystyle\sum_{k=0}^{100}3^k(-2)^{n-k}\binom{100}{k}$.
8. In calculus, the *power rule* for derivatives states that $\frac{d}{dx}[x^n]=nx^{n-1}$. The definition of the derivative says

$$\frac{d}{dx}[x^n] = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h}.$$

Use the Binomial Theorem to help prove the power rule in the case where n is a positive integer.

9. Write down the first, say, ten rows of Pascal's triangle. Did you ever notice that (except for the far left and far right entries) some rows consist entirely of multiples of the row number? For example, when $n = 5$, the row is:

$$1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1$$

and we see both 5 and 10 are multiples of 5. Which rows have this property? In other words, for which values of n is it true that all values of $\binom{n}{k}$ (with $1 \le k \le n-1$) are multiples of n? What is your conjecture? Can you prove it?

Exercises for Chapter 5

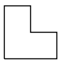
1. Generate the first five terms in each of the following sequences.

   a. $a_0 = 1$, $a_n = 3a_{n-1}$ for $n \geq 1$.
   b. $a_0 = 1$, $a_n = 2 + a_{n-1}$ for $n \geq 1$.
   c. $a_0 = 0$, $a_n = 1 + 2a_{n-1}$ for $n \geq 1$.
   d. $a_0 = 1$, $a_1 = 1$, $a_n = 2a_{n-1} + a_{n-2}$ for $n \geq 2$.
   e. $a_n = \frac{1}{n+1}\binom{2n}{n}$ for $n \geq 0$.

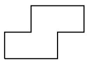
2. The *growth rate* of a sequence is the limit

$$\rho = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

   (The golden ratio discussed at the end of Chapter 5 is the growth rate for the Fibonacci sequence.) Compute the growth rate for each of the sequences in Exercise 1.
3. How many subsets of $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ contain no consecutive numbers? For example $\{1, 3, 7, 9\}$ contains no consecutive numbers, but $\{1, 3, 7, 8\}$ does contain consecutive numbers.
4. In Problem 56 we saw that the number of compositions of n using only parts of size 1 and 2 was given by a Fibonacci number. How many compositions of n use only parts of size 1 and 3? For example, there are three such compositions of four: $(1, 1, 1, 1)$, $(3, 1)$, and $(1, 3)$.
5. How many compositions of n use only parts of size 2 and 3? For example, there are three such compositions of seven: $(2, 2, 3)$, $(2, 3, 2)$, and $(3, 2, 2)$.
6. How many compositions of n have all their parts greater than one, except possibly the final part? For example, there are three such compositions of four: (4), $(2, 2)$, and $(3, 1)$.
7. How many compositions of n have an odd final part? For example, there are three such compositions of three: (3), $(2, 1)$, and $(1, 1, 1)$.
8. In Problem 58, we counted the number of domino tilings of a $2 \times n$ rectangle. How many ways can you tile a $3 \times n$ rectangle with dominoes?
9. An *L-tiling* of a $2 \times n$ rectangle is like a domino tiling except that it uses the pieces

 and 

   See Figure A.5. How many L-tilings of a $2 \times n$ rectangle are there?
10. An *S-tiling* of a $2 \times n$ rectangle is like a domino tiling except that it uses the pieces

 and 

   See Figure A.5. How many S-tilings of a $2 \times n$ rectangle are there?
11. A *T-tiling* of a $2 \times n$ rectangle is a tiling that uses the pieces

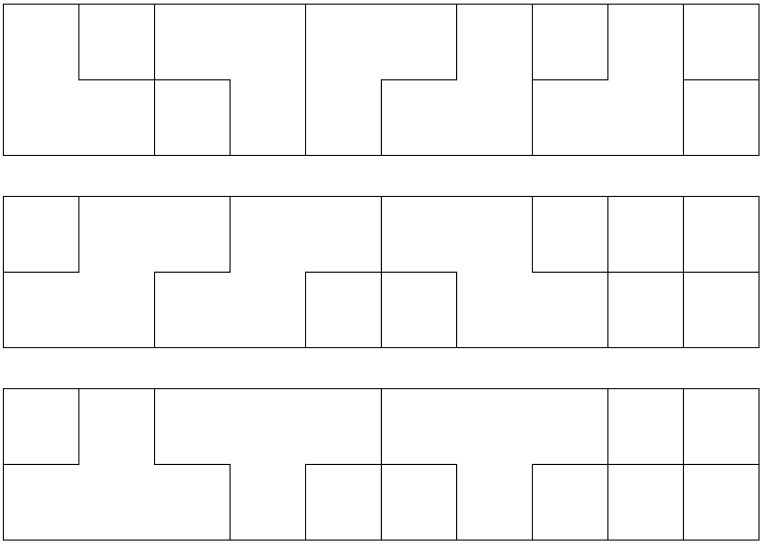
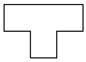


Fig. A.5 An L-tiling, an S-tiling, and a T-tiling of a 2×10 rectangle.

 and 

See Figure A.5. How many T-tilings of a $2 \times n$ rectangle are there?

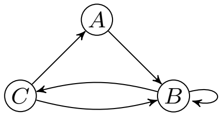


Fig. A.6 A directed graph on vertex set $\{A, B, C\}$.

12. This problem is about counting walks on directed graphs. A *directed graph* is a pair (V, E) such that V is a finite set known as the *vertex set* and E is a collection of ordered pairs of vertices in $V \times V$ known as the *edge set*. For example, the graph

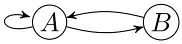
$$(\{A, B, C\}, \{(A, B), (B, B), (B, C), (C, B), (C, A)\})$$

is shown in Figure A.6. A *walk* on a directed graph is a sequence of vertices and edges of the graph such that there is a directed edge from each vertex to the subsequent vertex. The *length* of a walk is the number of edges in the sequence. Since there is at most one edge from any vertex to another, we can merely list the vertices in the walk. For example, $ABCBBCA$ records a walk of length 6 on the graph in Figure A.6.

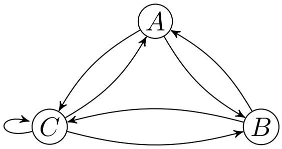
In each example below, you are asked to consider a sequence a_n, where a_n is the number of walks of length n that possess certain features. For each sequence compute the first five terms and then try to find a recurrence and/or formula for the sequence.

a. Let a_n denote the number of walks of length n, on the graph below, that begin and end at A.

b. Let a_n denote the number of walks of length n, on the graph below, that begin and end at A.



c. Let a_n denote the number of walks of length n, on the graph below, that begins at A but can end at any vertex.



Exercises for Chapter 6

1. Expand the following generating functions as series. What sequences do they encode?
   a. $\dfrac{1}{1-3z}$
   b. $\dfrac{1}{2-z}$
   c. $\dfrac{z}{2-3z}$
   d. $\dfrac{1}{1-z^3}$
   e. $\dfrac{1}{1+z^2}$
2. What sequence is encoded by the generating function
$$\frac{1}{1-7z+10z^2}?$$
   You should be able to give both a recursive formula and a closed-form expression for the nth term.
3. Let a_n be the sequence defined by $a_0 = 1$, $a_1 = 3$ and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$. What is the generating function for the sequence a_k?
4. The (a,b)-*Lucas numbers* are defined by $a_0 = a$, $a_1 = b$, and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$. What is the generating function for the (a,b)-Lucas numbers?
5. Let a_n be the sequence defined by $a_0 = a_1 = 1$ and $a_n = 2a_{n-1} + a_{n-2}$ for $n \geq 2$. What is the generating function for the sequence a_k?
6. Let a_n be the sequence defined by $a_0 = a_1 = 1$ and $a_n = a_{n-1} + 2a_{n-2}$ for $n \geq 2$. What is the generating function for the sequence a_k?
7. Let a_n be the sequence defined by $a_0 = a_1 = a_2 = 1$ and $a_n = a_{n-1} + a_{n-2} + a_{n-3}$ for $n \geq 2$. What is the generating function for the sequence a_k?
8. Let $b_k = \sum_{j=0}^{k} f_j$, where f_j is the jth Fibonacci number, defined by $f_0 = f_1 = 1$ and $f_j = f_{j-1} + f_{j-2}$ for $j \geq 2$. What is the generating function for the sequence b_k?
9. The Euro (€) is a currency used in much of Europe. Like the dollar, it has coins that are worth hundredths of a unit, also known as cents. The coins come in denominations of 1, 2, 5, 10, 20, and 50 cents. Let c_n denote the number of ways to make n cents out of some combination of these coins. Write an expression for the generating function
$$F_{€}(z) = 1 + \sum_{n\geq 1} c_n z^n,$$
   and use software to find c_{100}.

10. Chicken McNuggets traditionally came in boxes of $6, 9$, and 20. Nowadays, most American locations of McDonald's sell the nuggets in sizes of $4, 6, 10$, and 20, whereas in some locations outside the United States, you can find McNuggets sold in packs of $4, 6, 9$, and 20.
    Explain how the different collections of pack sizes impact the conclusions about "non-nuggetable numbers" discussed in the essay at the end of Chapter 6. In particular, compare the generating functions for these two versions of the question.

Exercises for Chapter 7

1. Expand the following generating functions as exponential series in z. What sequences do they encode?

   a. e^{2z}
   b. e^{1+z}
   c. $\cos(3z)$
   d. $\dfrac{\sin(\sqrt{z})}{\sqrt{z}}$

2. Expand $e^{(1+t)z}$ first as an exponential series in z, then as an ordinary series in t i.e.,
$$e^{(1+t)z} = \sum_{n\geq 0}\sum_{k\geq 0} a_{n,k} t^k \frac{z^n}{n!}.$$
   What are the numbers $a_{n,k}$?
3. Use Problem 79 to find the series expansions for $\sin(z)^2 = \sin(z)\cdot\sin(z)$ and $\cos^2(z) = \cos(z)\cdot\cos(z)$. Use the result to prove the Pythagorean Theorem, i.e., $\sin^2(z) + \cos^2(z) = 1$, the hard way.
4. Expand e^{iz}, where $i^2 = -1$, and use this to explain Euler's identity:
$$e^{i\pi} + 1 = 0.$$
   (Hint: When expanding e^{iz}, regroup the terms according to whether they have a factor of i in them.)
5. How many ways can we divide a group of 10 people into 4 subgroups:

   a. if we specify that the groups have sizes $6, 2, 1, 1$?
   b. if we specify that the groups have sizes $3, 3, 2, 2$?
   c. if we don't care how many people are in each group, so long as there is at least one person per group? (You can build off of Table 7.1 to save some work here.)

6. List all the set partitions of $\{a, b, c, d, e\}$ that have three blocks.
7. List all the set compositions of $\{1, 2, 3, 4\}$, grouped according to the underlying permutation in S_4.
8. Make a chart comparing the following five generating functions. Explain what each counts, and how they are related to each other:
$$B_n(t), \overline{B}_n(t), S_k(z), S(t, z), \overline{S}(t, z)$$
9. Let a_n denote the number of set partitions of an n-element set with an even number of blocks, minus the number of set partitions with an odd number of blocks. For example, with $n = 3$, there are five set partitions. Three of them have two blocks, one has one block, and one has three blocks. Thus $a_3 = 3 - 2 = 1$. What is the exponential generating function for the a_n?

10. Let a_n denote the number of set compositions of an n-element set with an even number of blocks, minus the number of set partitions with an odd number of blocks. For example, with $n = 3$, there are thirteen set compositions. Six of them have two blocks, one has one block, and six have three blocks. Thus $a_3 = 6 - 7 = -1$. What is the exponential generating function for the a_n?
11. This problem asks you to explore OEIS. First, generate several terms for each of the sequences in Problem 54, and look them up.
    a. What are their entry numbers?
    b. Did you uniquely identify each sequence on the first search, or did you have to generate some more terms before you were sure you had the correct entry?
    c. Which entry has the longest list of references?
    d. Which entry do you find the most surprising, either in terms of its combinatorial interpretations or in terms of its connections to other areas of mathematics?

Exercises for Chapter 8

1. How many permutations in S_9 have three ascents?
2. Martha's lab is getting together for a group photo. There are ten team members, no two of which are the same height. Let $h_1, h_2, \ldots, h_{10}$ denote their heights. How many ways can the team be arranged so that:

   a. the members are arranged in increasing order? In other words, how many permutations w have $h_{w(1)} < h_{w(2)} < \cdots < h_{w(10)}$?
   b. the heights increase to the middle, then decrease? In other words, how many permutations w have $h_{w(1)} < h_{w(2)} < h_{w(3)} < h_{w(4)} < h_{w(5)}$ and $h_{w(6)} > h_{w(7)} > h_{w(8)} > h_{w(9)} > h_{w(10)}$?
   c. the heights increase (not necessarily to the middle), then decrease?
   d. the two tallest team members are in the middle of the photo, with shorter people arranged in decreasing order moving away from the middle?

3. How many permutations in S_7 have at most one return? How many permutations w in S_n have at most one return? (See Problem 94 for the definition of a return in a permutation.)
4. An *inversion sequence* of length n is a list of numbers $s = (s_1, s_2, \ldots, s_n)$ such that $0 \leq s_i \leq i-1$ for all i. Let I_n denote the set of inversion sequences of length n. There are $n!$ inversion sequences of length n, so these elements are in bijection with permutations in various ways. In this problem you will describe the Eulerian numbers in terms of inversion sequences.
   Let $\mathrm{asc}(s)$ denote the number *ascents* of an inversion sequence, i.e.,
   $$\mathrm{asc}(s) = |\{i : s_i < s_{i+1}\}|.$$
   For example, if $s = (0, 0, 1, 2, 0, 1)$ then $\mathrm{asc}(s) = |\{2, 3, 5\}|$=3. Prove that
   $$\left\langle {n \atop k} \right\rangle = |\{s \in I_n : \mathrm{asc}(s) = k - 1\}|.$$
5. A *weak descent* of an inversion sequence is an entry i such that $s_i \geq s_{i+1}$, denoted
   $$\overline{\mathrm{des}}(s) = |\{i : s_i \geq s_{i+1}\}|.$$
   Prove that
   $$\left\langle {n \atop k} \right\rangle = |\{s \in I_n : \overline{\mathrm{des}}(s) = k - 1\}|.$$
6. Suppose we have n boxes and n balls. We will play a game of placing balls into the boxes, but there are some subtleties.
   First of all, the boxes look similar to one another but the openings at the top vary in size. We arrange the boxes from biggest opening to smallest opening and see that they are all distinct sizes.

The balls also appear to come in all different sizes. In fact, we find that for each i, the ith largest ball is just barely small enough to fit into the ith largest box, but it can't quite fit into any of the smaller boxes. See Figure A.7 for an illustration when $n = 5$.

a. How many ways are there to place the balls in the boxes? (Assume that each box has enough room to accommodate all the balls that fit through its opening.)
b. How many ways are there to place the balls in the boxes so that the smallest box is empty?
c. How many ways are there to place the balls in the boxes so that exactly one box is empty?
d. How many ways are there to place the balls in the boxes so that exactly k boxes are empty?

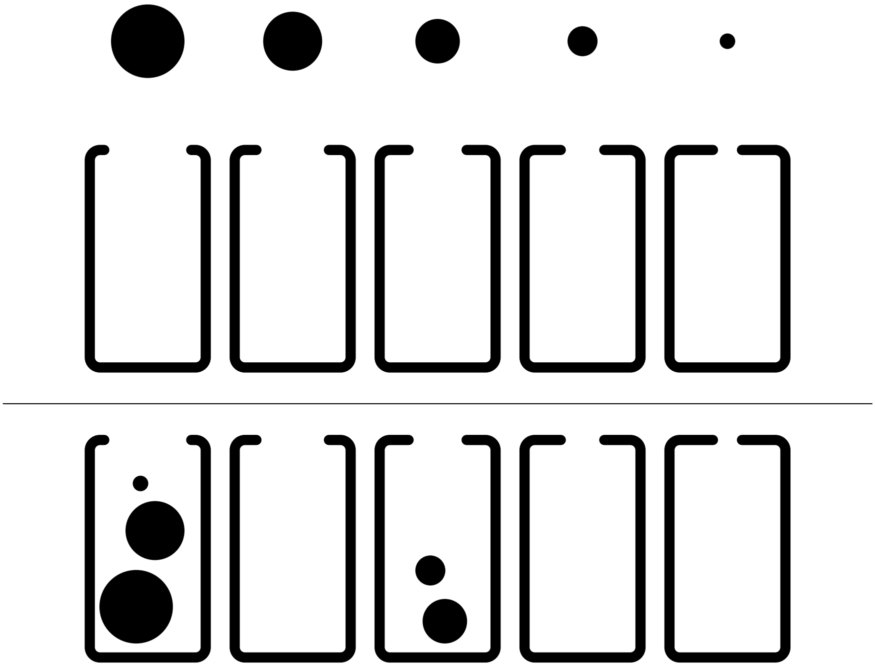


Fig. A.7 Putting some balls of different sizes into boxes with different size openings.

7. Recall the *descent set* of a permutation w, denoted $\mathrm{Des}(w)$, is the set of all descents of w, i.e., $\mathrm{Des}(w) = \{i : w(i) > w(i+1)\}$.

a. How many permutations $w \in S_n$ have $\mathrm{Des}(w) \subseteq \{1\}$?

b. How many permutations $w \in S_n$ have $\mathrm{Des}(w) \subseteq \{2\}$?
c. How many permutations $w \in S_n$ have $\mathrm{Des}(w) \subseteq \{j\}$?
d. How many permutations $w \in S_n$ have $\mathrm{Des}(w) \subseteq \{5, 8\}$?
e. How many permutations $w \in S_n$ have $\mathrm{Des}(w) \subseteq \{j_1, j_2\}$?

8. For this and the next two exercises, let $\alpha_J(n)$ denote the number of permutations in S_n with $\mathrm{Des}(w) \subseteq J$. Fix $n = 5$. For each subset $J \subseteq \{1, 2, 3, 4\}$, find $\alpha_J(5)$.
9. Find a formula for $\alpha_J(n)$ in general.
10. In terms of n, which subsets J seem to maximize $\alpha_J(n)$?
11. Use Problem 96 to prove that the Eulerian polynomials are self-reciprocal, i.e., that $t^{n+1} A_n(1/t) = A_n(t)$.
12. Use the identity of Problem 104 to prove the inverse relationship as follows:

$$A_n(s) = (1-s)^n\, \overline{B}_n\left(\frac{s}{1-s}\right).$$

13. Now use the self-reciprocity of $A_n(s)$ (from Exercise 11) to prove

$$A_n(s) = s(s-1)^n\, \overline{B}_n\left(\frac{1}{s-1}\right),$$

and conclude that the ordered Bell numbers can be extracted from the Eulerian polynomials, i.e., $\overline{B}_n = \frac{1}{2} A_n(2)$. Can you obtain this result directly from the exponential generating functions? (Hint: take care with the constant terms!)

Exercises for Chapter 9

1. Use the recursive idea of Problem 106 to list all the elements of $S_5(231)$.
2. Draw all the Dyck paths in Dyck(5).
3. Show that C_n counts the number of *nonnesting partitions* of $\{1, 2, \ldots, n\}$. A nonnesting partition is a set partition $\pi = \{R_1, \ldots, R_k\}$ such that if $\{a, d\} \subseteq R_i$ and $\{b, c\} \subseteq R_j$ with $a < b < c < d$, then $R_i = R_j$. The fourteen nonnesting partitions of $\{1, 2, 3, 4\}$ are shown in Figure A.8.

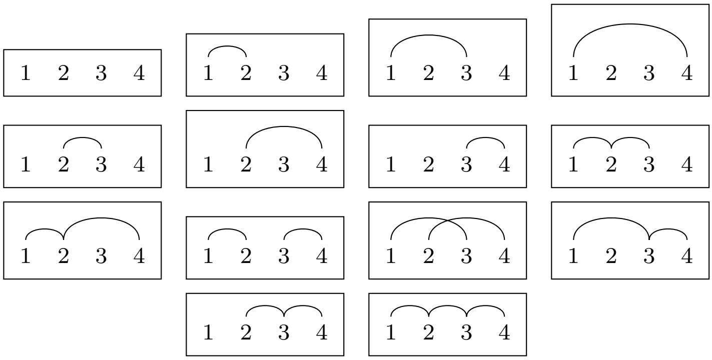


Fig. A.8 The fourteen nonnesting partitions of $\{1, 2, 3, 4\}$.

4. Show that counting nonnesting partitions by number of blocks gives the Narayana numbers.
5. Show that C_n counts the number of *noncrossing matchings* on $\{1, 2, \ldots, 2n-1, 2n\}$. A noncrossing matching is a noncrossing partition with all the blocks having size two. For example, when $n = 3$, the five noncrossing matchings on $\{1, 2, 3, 4, 5, 6\}$ are shown in Figure A.9.

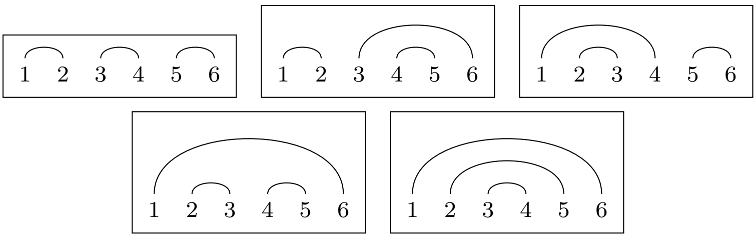


Fig. A.9 The five noncrossing matchings on $\{1, 2, 3, 4, 5, 6\}$.

6. Describe a statistic for noncrossing matchings so that the distribution of this statistic gives the Narayana numbers. (Hint: look for a bijection with noncrossing partitions.)

7. A *Motzkin path* of length n is a lattice path from $(0,0)$ to (n,n) that never passes below the line $y=0$ and uses only "up" steps from (i,j) to $(i+1,j+1)$, "down" steps from (i,j) to $(i+1,j-1)$, and "horizontal" steps from (i,j) to $(i+1,j)$. Note that Motzkin paths that contain no horizontal steps are in bijection with Dyck paths. For example, the nine Motzkin paths of length four are shown in Figure A.10.

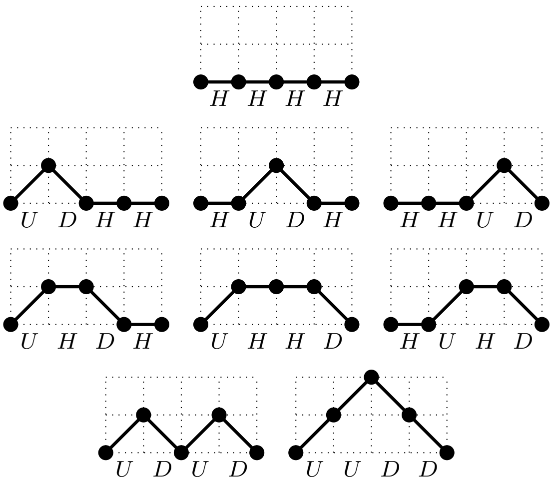


Fig. A.10 The nine Motzkin paths of length four.

Let M_n denote the number of Motzkin paths of length n, with $M_0=1$. Here are the first few values of M_n, sometimes called *Motzkin numbers*:

$$1,1,2,4,9,21,51,127,323,835,2188,5798,\ldots.$$

Let $M(z)=\sum_{n\geq 0} M_n z^n$. Show that

$$M(z)=\frac{1-z-\sqrt{1-2z-3z^2}}{2z^2}.$$

Hint: each Motzkin path is built from a Dyck path by inserting horizontal steps between the steps of the Dyck path. Use this fact to show

$$M(z)=\frac{1}{1-z}C\left(\frac{z^2}{(1-z)^2}\right),$$

where $C(z)$ is the Catalan generating function.

8. A *Schröder path* of size n is a lattice path from $(0,0)$ to (n,n) that never passes below the line $y=x$ and uses only steps "North" from (i,j) to $(i,j+1)$, "East" from (i,j) to $(i+1,j)$ and "Northeast" from (i,j) to

$(i+1, j+1)$. Note that Schröder paths with no northeast steps are Dyck paths. For example, the six Schröder paths of size 2 are shown in Figure A.11.

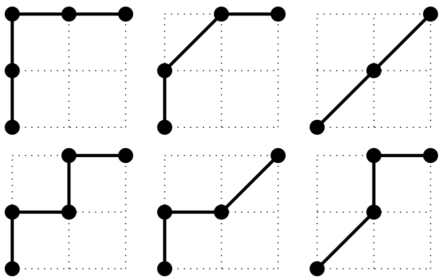


Fig. A.11 The six Schröder paths of size $n = 2$.

Let R_n denote the number of Schröder paths of size n, with $R_0 = 1$. We call the number R_n a *Schröder number*. Here are the first few values for R_n:

$$1, 2, 6, 22, 90, 394, 1806, 8558, 41586, 206098, \ldots.$$

Let $R(z) = \sum_{n\geq 0} R_n z^n$. Show that

$$R(z) = \frac{1 - z - \sqrt{1 - 6z + z^2}}{2z}.$$

Hint: Just as with Motzkin paths, each Schröder path can be built from a Dyck path by inserting northeast steps between the steps of the Dyck path. Use this fact to show

$$R(z) = \frac{1}{1-z} C\left(\frac{z}{(1-z)^2}\right),$$

where $C(z)$ is the Catalan generating function.

9. Show the Schröder numbers (apart from $R_0 = 1$) are always even. You can do this by manipulating the generating function from the previous problem, but try to explain it combinatorially.
Hint: find a bijection between the Schröder paths with a peak on the line $y = x+1$ and those without. The number of Schröder paths with no peak on the line $y = x+1$ are called *small Schröder numbers*, denoted r_n. Here are the first few values of the small Schröder numbers:

$$1, 1, 3, 11, 45, 197, 903, 4279, 20793, 103049, \ldots.$$

Given that $r_0 = 1$ and $r_n = R_n/2$ for $n \geq 1$, use the generating function found in the previous problem to conclude that

$$\sum_{n\geq 0} r_n z^n = \frac{1+z-\sqrt{1-6z+z^2}}{4z}.$$

10. Show that the small Schröder numbers r_n count the number of valid parenthesizations of $n+1$ symbols with at most $n-1$ pairs of parentheses. Parentheses around the entire expression are not allowed, and each pair of parentheses must enclose at least two sub-expressions. For example, here are the eleven parenthesizations of four symbols:

$$((wx)y)z \quad (w(xy))z \quad (wx)(yz) \quad w((xy)z) \quad w(x(yz))$$

$$(wx)yz \quad (wxy)z \quad w(xy)z \quad w(xyz) \quad wx(yz)$$

$$wxyz$$

Can you interpret these parenthesizations in terms of planar rooted trees of some kind?

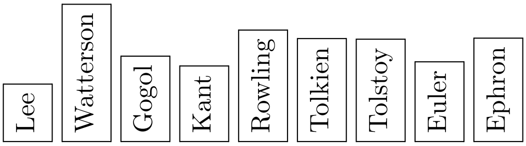
11. Come up with a combinatorial interpretation for the numbers $\binom{n}{k} - \binom{n}{k-1}$ with $0 \leq k \leq \lfloor n/2 \rfloor$:

$$\begin{array}{llll}
1 & & & \\
1 & & & \\
1 & 1 & & \\
1 & 2 & & \\
1 & 3 & 2 & \\
1 & 4 & 5 & \\
1 & 5 & 9 & 5 \\
1 & 6 & 14 & 14
\end{array}$$

These are sometimes called *ballot numbers*. (Note the Catalan numbers correspond to the special case where n is even and $k = n/2$.)

Exercises for Chapter 10

1. Regina has several books on her shelf:

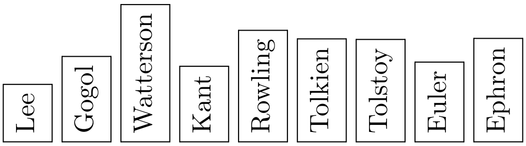


She would like them to be in alphabetical order by the author's last name.

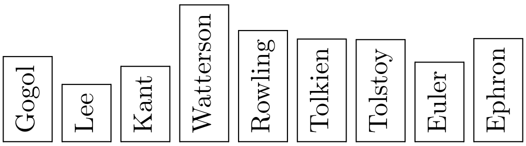
a. Thinking of the order of the books as a permutation, how many inversions does the shelf have right now? That is, how many pairs of books are out of order alphabetically?

b. Regina is going to use a simple procedure to sort the books on her shelf. She will scan the books from left to right, and when she sees two adjacent books that are out of order, she will swap them. She continues comparing and swapping until she reaches the right edge of the shelf, then returns to the left edge and scans the books again. If she scans all the way from left to right without swapping any books, she knows her task is complete!

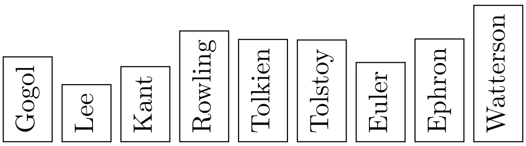
For example, "Lee" comes before "Watterson" but "Watterson" should come after "Gogol" so Regina swaps "Watterson" and "Gogol":



Now Regina compares "Watterson" and "Kant" and makes another swap:



So far, Regina has performed two book swaps. After several more comparisons and swaps, Regina has reached the right edge of the shelf and the books are in this order:



Now Regina returns to the left edge and starts the procedure over again. By the time Regina has sorted the books in alphabetical order, how many swaps will she have performed?

2. Regina's algorithm from Exercise 1 can be applied to any list of objects that are ordered, which we can model with permutations. To describe the algorithm in general, let w be an element of S_n. We scan the permutation w from left to right, comparing adjacent entries: $w(i)$ with $w(i+1)$. If $w(i) > w(i+1)$, then we swap these two numbers before moving to the next two. Upon comparing (and possibly swapping) $w(n-1)$ and $w(n)$, we return to the beginning of the list and start over. If we scan all the way through the permutation and find $w(i) < w(i+1)$ for each $i < n$, then we know our permutation is completely sorted and the algorithm terminates. For example, here is the algorithm applied to the permutation $w = 325146$ in S_6. Below we have underlined comparisons. If two adjacent entries swap places, we highlight them in bold:

$$\underline{\mathbf{32}}5146 \to 2\underline{35}146 \to 23\underline{\mathbf{51}}46 \to 231\underline{\mathbf{54}}6 \to 2314\underline{56} \to 231456$$

After the first pass, we return to the beginning and start comparing again:

$$\underline{23}1456 \to 2\underline{\mathbf{31}}456 \to 21\underline{34}56 \to 213\underline{45}6 \to 2134\underline{56} \to 213456$$

The comparisons and swaps continue until we can scan through the permutation without finding any adjacent numbers out of order. In this case, it takes two more scans through the permutation:

$$\underline{\mathbf{21}}3456 \to 1\underline{23}456 \to 12\underline{34}56 \to 123\underline{45}6 \to 1234\underline{56} \to 123456$$

and finally

$$\underline{12}3456 \to 1\underline{23}456 \to 12\underline{34}56 \to 123\underline{45}6 \to 1234\underline{56} \to 123456$$

For this example, we made a total of 5 swaps. Let's try to figure out how many swaps we do in general.
Let $w^{(1)}$ denote the permutation we obtain after the first swap, $w^{(2)}$ the permutation after the second swap, and so on. Then we see that the algorithm in general produces a sequence of permutations that terminates with the identity permutation $123\cdots n$ after some number of swaps, i.e.,

$$w \to w^{(1)} \to w^{(2)} \to \cdots \to w^{(k)} = 123\cdots n,$$

for some number k.

a. Show that each time we make a swap in this algorithm, the number of inversions drops by exactly one. That is, show

$$\mathrm{inv}(w^{(i+1)}) = \mathrm{inv}(w^{(i)}) - 1.$$

Conclude that if Regina's algorithm performs k swaps, then $\mathrm{inv}(w) = k$.

b. Show that no matter what, if we swap two numbers $w(i), w(i+1)$ in a permutation w to obtain the permutation w', then either $\mathrm{inv}(w') = \mathrm{inv}(w) + 1$ or $\mathrm{inv}(w') = \mathrm{inv}(w) - 1$.

c. Conclude that Regina's algorithm is optimal with respect to the number of swaps. That is, suppose w in S_n is a permutation for which $\mathrm{inv}(w)=k$ is the number of swaps performed in her algorithm. Then there is no way to sort w with fewer than k adjacent swaps.

3. Do an investigation of sorting algorithms. Regina's algorithm is known as "Bubble Sort" in the literature. What are some other common sorting algorithms? What are the pros and cons of each?
4. An *inversion sequence* of length n is a list of numbers $s = (s_1, \ldots, s_n)$ such that $0 \leq s_i \leq i - 1$ for each i. Let I_n denote the set of such sequences, which we recognize is simply a cartesian product defined as follows:

$$I_n = \{0\} \times \{0,1\} \times \{0,1,2\} \times \cdots \times \{0,1,2,\ldots,n-1\}.$$

For an inversion sequence s, denote the sum of its entries by $|s|$. That is, for s in I_n, let

$$|s| = s_1 + s_2 + \cdots + s_n.$$

For example, with $s = (0,0,1,2,0,4)$ in I_6, we have $|s| = 0+0+1+2+0+4 = 7$.

a. For $n \leq 4$, make a table that counts inversion sequences $s \in I_n$ according to $|s|$.

b. It is easy to see that $|I_n| = n!$. Now show that counting inversion sequences according to $|s|$ gives a q-analogue, i.e., show that

$$\sum_{s \in I_n} q^{|s|} = [n]!$$

c. Part b. and Theorem 7 imply that

$$I(n,k) = |\{w \in S_n : \mathrm{inv}(w) = k\}| = |\{s \in I_n : |s| = k\}|.$$

Give a bijection between permutations and inversion sequences that proves this identity directly.

5. The *weak descent set* of an inversion sequence is the set of all i such that $s_i \geq s_{i+1}$, i.e.,

$$\overline{\mathrm{Des}}(s) = \{i : s_i \geq s_{i+1}\}.$$

Define the *major index* for an inversion sequence to be the sum of the weak descents, i.e., if $s \in I_n$,

$$\text{maj}(s) = \sum_{i \in \overline{\text{Des}}(s)} i$$

For example, with $s = (0, 0, 1, 2, 0, 4)$ in I_6, we have $\overline{\text{Des}}(s) = \{1, 4\}$ and $\text{maj}(s) = 1 + 4 = 5$.

a. For $n \leq 4$, make a table that counts inversion sequences $s \in I_n$ according to $\text{maj}(s)$.
b. Find a bijection between permutations and inversion sequences that proves

$$|\{w \in S_n : \text{maj}(w) = k\}| = |\{s \in I_n : \text{maj}(s) = k\}|.$$

c. Part b. and Theorem 7 now imply that

$$\sum_{s \in I_n} q^{\text{maj}(s)} = [n]!$$

Can you prove this result without the bijection in part b.?

6. The following is a greedy algorithm for sorting a permutation with swaps that aren't necessarily adjacent: find the largest element that is out of place, move it to its proper place, and repeat. This algorithm is known as *straight selection sort.*
More precisely, if $w(n) = n$, do nothing and move on to sort $w(1) \cdots w(n-1)$.
Otherwise, if $w(i) = n$, with $i < n$, swap $w(i)$ and $w(n)$ to get the permutation w'. Then $w'(n) = n$, and we can now sort $w'(1) \cdots w'(n-1)$.
For example, here is the algorithm applied to the permutation $w = 3172546$. Below we denote a swap between positions i and j by (ij), and we highlight $w(i)$ and $w(j)$ in bold.

$$31\underbrace{\mathbf{7}254\mathbf{6}}_{(37)} \to 31\underbrace{\mathbf{6}254}_{(36)}7 \to 31\underbrace{\mathbf{42}}_{(34)}567 \to \underbrace{\mathbf{312}}_{(13)}4567 \to \underbrace{\mathbf{21}}_{(12)}34567 \to 1234567$$

Suppose $(i_1 j_1), \ldots, (i_k j_k)$ are the swaps used in straight selection sort for some w. Define the *sorting index* for w to be

$$\text{sor}(w) = \sum_{r=0}^{k} (j_r - i_r),$$

e.g., with $w = 3172546$ above, we get

$$\text{sor}(3172546) = (7-3) + (6-3) + (4-3) + (3-1) + (2-1) = 11.$$

Informally, the sorting index measures the "cost" of straight selection sort, with transpositions of elements that are far away costing more than elements closer by.

a. For $n \leq 4$, make a table that counts permutations $w \in S_n$ according to $\mathrm{sor}(w)$.
b. Show that the sorting index is *Mahonian*, i.e., that it has the same distribution as the Mahonian numbers:

$$\sum_{w \in S_n} q^{\mathrm{sor}(w)} = [n]!.$$

7. We will define yet another Mahonian statistic on permutations. This one comes from a searching procedure. We will scan our list of numbers from left to right, and we want to pull out the numbers $1, 2, \ldots, n$ in order. However, we are only allowed to pull out the numbers in strict succession: 2 after 1, 3 after 2, and so on. Once we scan through the list, we scan through the (now shorter) list again, repeating as often as necessary until we have selected all the numbers. For example, with the permutation w=3172546, we perform the following sequence of scans through the list. In bold, we have highlighted the numbers selected in each scan:

$$\begin{array}{ccccccc|l} 3 & \mathbf{1} & 7 & \mathbf{2} & 5 & 4 & 6 & \\ \mathbf{3} & & 7 & & 5 & \mathbf{4} & 6 & \mathbf{12} \\ & & 7 & & \mathbf{5} & & \mathbf{6} & \mathbf{1234} \\ & & \mathbf{7} & & & & & \mathbf{123456} \\ & & & & & & & \mathbf{1234567} \end{array}$$

What we will count here is the number of times a number is *not* selected when scanning through the list. For example, the number 3 was not selected in the first scan, and the number 7 was not selected in any of the first three scans. All told, there were nine instances in which we scanned, but did not select a number in the example above: 3 was overlooked once, 7 was overlooked three times, 5 was overlooked twice, 4 was overlooked once, and 6 was overlooked twice.
We call this total number the *disorder* of a permutation, and denote it by $\mathrm{dis}(w)$. In the example above, then, $\mathrm{dis}(w) = 9$.

a. For $n \leq 4$, make a table that counts permutations $w \in S_n$ according to $\mathrm{dis}(w)$.
b. Show that disorder is Mahonian, i.e.,

$$\sum_{w \in S_n} q^{\mathrm{dis}(w)} = [n]!.$$

8. We now have four distinct permutation statistics that give rise to the Mahonian numbers: inv, maj, sor, and dis.

a. Can you find a permutation w such that no two of these statistics agree?
b. Parts i)–vi). For each pair of statistics $\{\mathrm{stat}_1, \mathrm{stat}_2\} \subset \{\mathrm{inv}, \mathrm{maj}, \mathrm{sor}, \mathrm{dis}\}$, find a bijection $f : S_n \to S_n$ such that if $f(w) = w'$ then $\mathrm{stat}_1(w) = \mathrm{stat}(w')$. (Warning: many of these correspondences are not obvious!)

9. Using any of the combinatorial interpretations for the Mahonian numbers (pick your favorite), show that:

a. For $k < n$,
$$I(n,k) = I(n,k-1) + I(n-1,k).$$

b. For any n and k,
$$\begin{aligned} I(n,k) &= I(n-1,k-n+1) + I(n-1,k-n+2) + \cdots + I(n-1,k) \\ &= \sum_{j=1}^{n} I(n-1,k-n+j) \end{aligned}$$

(Note that the formula is valid for all n and k since $I(n-1,k) = 0$ if $k < 0$ or $k > \binom{n-1}{2}$.)

10. Let's revisit Problem 125, but give q-binomial coefficients a new combinatorial interpretation. As in that problem, let $L(k, n-k)$ denote the set of paths from $(0,0)$ to $(k, n-k)$ that take steps East and North, for a total of n steps.
Here we define the major index of a path, $\mathrm{maj}(p)$, by thinking of p as a word in $\{N, E\}$ with $E > N$. For example, if $p = NEENENE$, $\mathrm{maj}(p) = 3+5 = 8$. In terms of pictures, $\mathrm{maj}(p)$ is adding the positions of the valleys of p, since a valley is an East step followed by a North step. In Table A.1 we see the ten paths in $L(2,3)$ grouped according to major index.

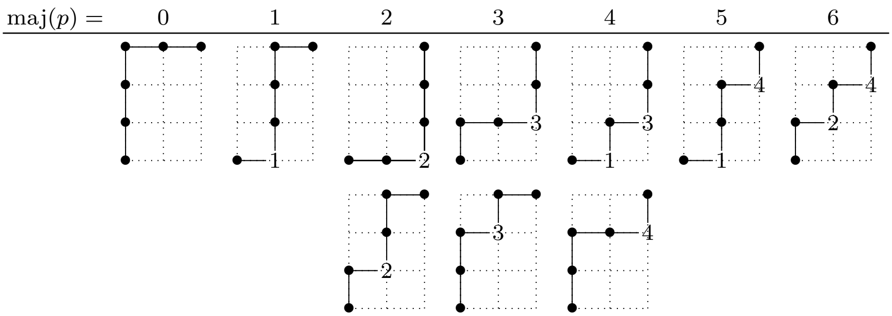


Table A.1 Lattice paths counted according to major index. Valleys are labeled with their position.

Define the generating function for paths in $L(k, n-k)$ by

$$L_{n,k}^{\mathrm{maj}}(q) = \sum_{p \in L(k,n-k)} q^{\mathrm{maj}(p)}.$$

a. Compute $L_{n,k}^{\mathrm{maj}}(q)$ for all $0 \le k \le n \le 4$ and arrange these polynomials in an array like Pascal's triangle.
b. Can you show that $L_{n,k}^{\mathrm{maj}}(q) = \begin{bmatrix} n \\ k \end{bmatrix}$? (Warning: this is not as straightforward as it is for area. If you think you have a simple explanation please let this textbook author know about it!)

11. The normalized area statistic for Dyck paths, area(p), counts the number of unit squares above the line $y = x$. (This is discussed in the essay at the end of Chapter 10.) Let $C_n^{\mathrm{area}}(q)$ denote the generating function for this area statistic on Dyck(n), the set of all Dyck paths. That is,

$$C_n^{\mathrm{area}}(q) = \sum_{p \in \mathrm{Dyck}(n)} q^{\mathrm{area}(p)}.$$

Show that

$$C_n^{\mathrm{area}}(q) = \sum_{i=0}^{n-1} q^i C_i^{\mathrm{area}}(q) C_{n-1-i}^{\mathrm{area}}(q).$$

12. In Problem 127, we studied the joint distribution of descents and major index for permutations, which we called the "Euler-Mahonian" distribution. This problem investigates a different Euler-Mahonian distribution, the joint distribution of descents and inversions. Let

$$S_n^{\mathrm{inv}}(q,t) = \sum_{w \in S_n} q^{\mathrm{inv}(w)} t^{\mathrm{des}(w)},$$

with $S_0^{\mathrm{inv}}(q,t) = 1$ for convenience.

a. Compute $S_n^{\mathrm{inv}}(q,t)$ for $n \le 4$, and group terms according to powers of t to see how these "q-Eulerian numbers" differ from those in Problem 127.
b. Show that

$$S_n^{\mathrm{inv}}(q,t) = S_{n-1}^{\mathrm{inv}}(q,t) + t \sum_{i=0}^{n-2} \begin{bmatrix} n-1 \\ i \end{bmatrix} S_i^{\mathrm{inv}}(q,t) q^{n-1-i} S_{n-1-i}^{\mathrm{inv}}(q,t).$$

c. Let us consider the following q-analogue of an exponential generating function as follows (compare with the generating function in Problem 101):

$$S(q,t,z) = 1 + \sum_{n \ge 1} S_n^{\mathrm{inv}}(q,t) \frac{z^n}{[n]!}$$

Show that

$$S(q,t,z) = \frac{(1-t)\exp(z(1-t);q)}{1-t\exp(z(1-t);q)},$$

where

$$\exp(z;q) = \sum_{n\geq 0} \frac{z^n}{[n]!}$$

is a q-analogue of the exponential function.

Exercises for Chapter 11

1. Suppose we have an experiment in which we toss a coin twice then roll a six-sided die.
   a. What is the sample space for this experiment?
   b. Let X denote the random variable that counts the number of heads plus the number on the die, e.g., if we toss heads, toss tails, then roll a 4, $X = 1 + 0 + 4 = 5$. What is the probability generating function
   $$p(t) = \sum_{k \geq 0} \Pr(X = k)t^k?$$
   c. What is $\Pr(X \leq 3)$?
   d. What is $E(X)$?
   e. What is $\mathrm{Var}(X)$?
2. What is the probability that, in a sequence of n coin tosses, you never see heads twice in a row?
3. Consider the experiment in which we toss a coin until we see heads twice in a row. Let X denote the random variable that records how many times we toss the coin before seeing two heads. For example, if our sequence of tosses is $HTTHTTHTHH$, then $X = 10$.
   a. Compute $\Pr(X = n)$ for $n \leq 7$.
   b. Find and prove a recurrence for $\Pr(X = n)$ in terms of $\Pr(X = n - 1)$ and $\Pr(x = n - 2)$.
   c. What is the generating function
   $$\sum_{n \geq 0} \Pr(X = n)z^n?$$
   d. What is $\sum_{n \geq 0} \Pr(X = n)$? Explain this with and without the generating function from part c.
4. What is the probability that, in a sequence of 100 coin tosses, you never see four heads in a row?
5. What is the probability that in a sequence of 100 coin tosses, you *do* have a run of five heads in a row?
6. In Figure A.12 we see two sets of data from an experiment in which a subject was asked to toss a coin one hundred times. (The tosses are ordered left to right, top to bottom.) One of the test subjects performed the experiment faithfully. The other got lazy and faked it after the first ten tosses or so. Which is the real data set? Explain your reasoning.
7. (**Essay opportunity**) We often imagine that a sequence of coin flips can come up heads indefinitely: $HHHHH\ldots$. Thus it is entirely *possible* that someone might toss a fair coin and have it come up heads one hundred

A	B
$HHTHTHTHTT$	$HHHTTTTHTT$
$THTHTHTHHT$	$HTTTHTTHTT$
$HTHTHTHTHH$	$HTHHHTHTTT$
$HTHTHTHTHT$	$THTTTHHTTH$
$HTHTHHHHTH$	$HTHHHHTTTH$
$HTHTHTHHTH$	$HHHTTHTHHT$
$THTHTHHTTT$	$THHTHHTHTT$
$THHTHTHTHT$	$HHHTHHHTTT$
$HTTTHHHTHT$	$THTTTHHHHH$
$HHTTHHTHTH$	$HTTHTTHTHT$

Fig. A.12 Two data sets. Which is real?

times in a row. However, the probability of that happening is $1/2^{100} \approx 7.9 \times 10^{-31}$. For "practical purposes" will this ever happen?
It seems like we are wading into some deep philosophical waters here. To help us distinguish between the *mathematically possible* and the *realistically possible*, consider the following thought experiments.

a. First, you give every person on the planet a fair coin to toss, with instructions to continue until the coin comes up tails. What is the probability that at least one person gets ten heads in a row? twenty heads in a row? thirty heads in a row? What is the expected maximum? Explain your reasoning with as much precision as possible.
b. As a variation on the previous experiment, we now instruct everyone to toss their coin a thousand times (this should take each person an hour or two), recording the length of the longest run of heads that they witness during those tosses. Now how likely is it that at least one person witnesses a hundred heads in a row?

8. In a standard deck of fifty-two cards, the cards come in thirteen ranks and four suits. Suppose you are dealt five cards at random as in a game of poker. See Figure 1.5 for the cards, and see Table A.2 for the definitions of the standard poker hands.

   a. What is the size of the sample space?
   b. Compute the probability of each of the ten hands, from *high card* to *royal flush*.
   c. What is the probability of getting a *pair* or better, i.e., a hand that is better than *high card*?
   d. Let X denote the random variable that assigns a number from 0 to 9 for each of the hands, with 0 corresponding to *high card*, 1 corresponding to a *pair*, and so on, with 9 corresponding to a *royal flush*. Compute

$$p(t) = \sum_{k=0}^{9} \Pr(X = k) t^k.$$

e. Use $p(t)$ to find $E(X)$ and $\mathrm{Var}(X)$. Does this information have any practical use? Explain.

Hand	Description	Example
Royal Flush	10 through A of one suit	$10\spadesuit, J\spadesuit, Q\spadesuit, K\spadesuit, A\spadesuit$
Straight Flush	five consecutive ranks of one suit (but not a royal flush)	$8\heartsuit, 9\heartsuit, 10\heartsuit, J\heartsuit, Q\heartsuit$
Four of a Kind	four cards of the same rank	$2\heartsuit, 2\clubsuit, 2\diamondsuit, 2\spadesuit, 6\heartsuit$
Full House	two cards of one rank and three cards of another rank	$4\heartsuit, 4\clubsuit, 8\heartsuit, 8\clubsuit, 8\spadesuit$
Flush	five cards of one suit (but not a straight)	$2\clubsuit, 4\clubsuit, 5\clubsuit, J\clubsuit, K\clubsuit$
Straight	five consecutive ranks (but not a flush)	$4\heartsuit, 5\clubsuit, 6\clubsuit, 7\spadesuit, 8\heartsuit$
Three of a Kind	three cards of the same rank (but not a full house or four of a kind)	$9\heartsuit, 9\clubsuit, 9\spadesuit, 7\spadesuit, A\heartsuit$
Two Pair	two pairs of cards of the same rank (but not four of a kind)	$9\heartsuit, 9\clubsuit, J\spadesuit, J\diamondsuit, 2\clubsuit$
Pair	one pair of cards of the same rank (but not three nor four of a kind, nor a full house)	$3\spadesuit, 3\clubsuit, 7\spadesuit, J\heartsuit, Q\clubsuit$
High Card	any other collection of cards	$4\spadesuit, 5\clubsuit, 10\spadesuit, J\heartsuit, Q\clubsuit$

Table A.2 There are ten standard poker hands, ordered from *high card* (lowest) to *royal flush* (highest).

9. You are designing a dice-rolling game called "Dice Wars." The game has three players, and each player has one special six-sided die (die A, die B, and die C). They play a game where they roll their dice two at a time

in "battle." Whoever rolls the largest number wins the battle. If the two combatants roll the same number they draw and nobody wins the battle. To make things interesting, the numbers on the faces are carefully chosen so that:

- Die A defeats die B more than half the time,
- Die B defeats die C more than half the time, and
- Die C defeats die A more than half the time.

How is this possible? Give an example of how to label the faces of the dice. To lend the game some elegance, try to label the faces so that the sum of the numbers is the same for each die. If N is this common number, what is the smallest positive whole number N that can be used?

10. There are 137 seats on a typical Boeing 737 airplane as it makes the short flight from Minneapolis to Chicago. It is a full flight and everyone is in line waiting to board.
The first person in line somehow loses his boarding pass on the way down the jetbridge, and rather than trying to remember his seat number he picks a purely random seat and sits down. From this point on, no one else loses their boarding pass. When a passenger boards the plane, either their seat is open, in which case they sit down in it, or someone is sitting in their seat. Being polite Minnesotans, a person in this situation won't cause a fuss but instead will pick one of the open seats purely at random.
You were in the bathroom when they started boarding the plane and so are stuck at the very back of the line. What is the probability you will get your own seat?
11. Suppose a sequence of participants in a drug trial enter a room to receive treatment. Each person will reach into a bag and pull out either a black ball or a white ball. If the ball is black they receive treatment A. If white, they receive treatment B.
Upon selecting a ball they return their ball to the bag *and add one more ball of the opposite color*. That is, if a person draws a black ball, they return their black ball and they add a new white ball. If a person draws a white ball, they return their white ball and they add a new black ball.
The purpose of this process is to make it more likely that person n and person $n+1$ draw different colors.
When the first patient enters the room, there is one black ball and one white ball in the bag.

    a. Let $p_{n,k}$ denote the probability that after n participants have received their treatment, precisely k of them received treatment A. Relate $p_{n,k}$ to a probability distribution you saw in Chapter 11. (Hint: look for a recurrence relation.)
    b. One of the goals of a random drug trial is to have about equal numbers of people receiving each treatment (without the doctors knowing or dictating who receives which treatment). With this in mind, a simpler

design for a random drug trial would be to have a coin toss for each patient, e.g., heads means treatment A and tails means treatment B. In what way is the black/white ball selection process described above superior to a coin-tossing assignment? It might help to compare the outcomes of each design with a population of $n = 1000$ participants.

12. We can extend Problems 128 and 139 to study the binomial distribution for a biased coin as follows. Suppose p is the probability that our coin comes up tails and $q = 1 - p$ is the probability that it comes up heads. Let $b_n(t)$ denote the biased binomial generating function. Then the probability generating function for just one coin toss is $b_1(t) = p + qt$, and with two tosses it is $b_2(t) = (p + qt)^2$, and so on. After n tosses, we get

$$b_n(t) = \sum_{k \geq 0} \Pr(X_n = k) t^k = (p + qt)^n,$$

where X_n is the random variable that counts the number of times heads appears in a sequence of n tosses.

a. Using the Binomial Theorem, what is $\Pr(X_n = k)$, the probability that after n tosses, our biased coin comes up heads exactly k times?

b. Mirroring Problem 128, find the ordinary and exponential generating functions for $b_n(t)$:

$$F(t, z) = \sum_{n \geq 0} b_n(t) z^n \quad \text{and} \quad G(t, z) = \sum_{n \geq 0} b_n(t) \frac{z^n}{n!}$$

c. Mirroring Problem 139, find generating functions for $E(X_n)$ and $\mathrm{Var}(X_n)$ for the biased binomial distributions, and find formulas for the expectation and variance in terms of p, q, and n.

Exercises for Chapter 12

1. Use computer software to find p_{100} (the number of partitions of $n = 100$) by expanding one of the generating functions from Problem 143.
2. In Problem 142, we investigated partitions whose parts were only ones and twos. We will investigate similar sets of restricted partitions here. For a set of positive integers S, let $p_n(S)$ denote the number of integer partitions of size n whose parts belong to S, i.e.,

$$p_n(S) = |\{\lambda : |\lambda| = n, \lambda_i \in S\}|$$

(Thus in Problem 142, we investigated $p_n(S)$ for $S = \{1, 2\}$.) For each set S below, compute the generating function

$$P_S(z) = 1 + \sum_{n \geq 1} p_n(S) z^n$$

and use it (possibly with software) to find $p_n(S)$ for $n \leq 10$.

   a. $S = \{1, 3\}$
   b. $S = \{2, 3\}$
   c. $S = \{1, 2, 3\}$
   d. $S = \{2, 4, 6, 8, \ldots\}$ (even numbers)
   e. $S = \{1, 2, 4, 8, 16, \ldots\}$ (powers of two)
   f. $S = \{1, 2, 3, 5, 8, 13, \ldots\}$ (Fibonacci numbers!)
   g. $S =$? (pick your own!)

3. Inspired by part f. of the previous exercise, let p'_n denote the number of partitions of n into *distinct* Fibonacci numbers. For example, $p'_3 = 2$, since (3) and $(2, 1)$ are partitions of 3 into Fibonacci numbers without any repeats.

   a. Compute p'_n for $n \leq 12$.
   b. For which values of n is $p'_n = 1$?

4. In Problem 145 we found the generating function for partitions (Young diagrams) that fit inside an $a \times b$ rectangle. In this problem you will do the same for partitions with distinct parts. (Problem 148 also looked at partitions with distinct parts.) Let $L'_{a,b}(t)$ denote the generating function for such partitions according to size:

$$L'_{a,b}(t) = 1 + \sum_{\substack{\lambda \subseteq a \times b \\ \text{(distinct parts)}}} t^{|\lambda|}$$

   a. Compute $L'_{a,b}(t)$ for $a + b \leq 5$.
   b. Find a simple factorization of $L'_{a,a}(t)$.

c. Is it easier to compute $L'_{a,b}(t)$ when $a < b$ (short and wide) or when $a > b$ (tall and narrow)? Explain.

5. For a partition $\lambda = (\lambda_1, \ldots, \lambda_k)$, define the *multiplicity* of i, denoted $m_i = m_i(\lambda)$, to be the number of parts of size i in λ. An alternate "multiplicity" notation for partitions is $\lambda = 1^{m_1}2^{m_2}\cdots$, with parts of multiplicity zero omitted. For example, with $\lambda = (5,5,3,3,3,3,2,1)$ we have $m_5 = 2$, $m_3 = 4$, and $m_2 = m_1 = 1$, while all other multiplicities are zero. Thus we would write $\lambda = 1^1 2^1 3^4 5^2$ in multiplicity notation. Note that by definition, $\sum m_i = \ell(\lambda)$ and $\sum i \cdot m_i = n$.
   With this notation in mind, show that for fixed n,
$$\prod_{\substack{\lambda = 1^{m_1}2^{m_2}\ldots \\ |\lambda|=n}} m_1(\lambda)!m_2(\lambda)!\cdots = \prod_{\substack{\lambda = 1^{m_1}2^{m_2}\ldots \\ |\lambda|=n}} 1^{m_1(\lambda)}2^{m_2(\lambda)}\cdots$$
   For example, for $n = 4$, we have five partitions, $4, 31, 2^2, 21^2, 1^4$, and the identity is:
$$(1!)(1!\cdot 1!)(2!)(1!\cdot 2!)(4!) = 96 = (4)(3\cdot 1)(2^2)(2\cdot 1^2)(1^4).$$
   Hint: let $m'_j(\lambda)$ denote the number of multiplicities greater than or equal to j, i.e., $m'_j(\lambda) = |\{i : m_i(\lambda) \geq j\}|$. The number of times j appears as a factor on the left-hand side of the identity is $\sum_{|\lambda|=n} m'_j(\lambda)$. The number of times j appears as a factor on the right-hand side is $\sum_{|\lambda|=n} m_j(\lambda)$. Show these two sums are equal.
6. Using the notation of the previous exercise, show that for fixed n,
$$\sum_{\substack{\lambda = 1^{m_1}2^{m_2}\ldots \\ |\lambda|=n}} \frac{1}{m_1(\lambda)!1^{m_1(\lambda)}m_2(\lambda)!2^{m_2(\lambda)}\cdots} = 1$$
   For example, for $n = 4$, we have five partitions, $4, 31, 2^2, 21^2, 1^4$, and the identity is:
$$\frac{1}{1!\cdot 4}+\frac{1}{1!\cdot 3\cdot 1!\cdot 1}+\frac{1}{2!\cdot 2^2}+\frac{1}{1!\cdot 2\cdot 2!\cdot 1^2}+\frac{1}{4!\cdot 1^4} = \frac{1}{4}+\frac{1}{3}+\frac{1}{8}+\frac{1}{4}+\frac{1}{24} = 1.$$
7. Find the generating function for self-conjugate partitions. (Hint: Problem 147 can make this easier to work out.)

Exercises for Chapter 13

1. What sequence has Dirichlet generating function

$$-\frac{d}{ds}\left[\zeta(s)\right]?$$

2. Generalize our approach to counting binary necklaces (Problems 163 and 164) to count necklaces with up to m colors of beads. Show that there are

$$\frac{1}{n}\sum_{d|n}\phi(d)m^{\frac{n}{d}}$$

such necklaces, for any fixed choice of m.

3. This and the next few exercises investigate *Farey sequences.* A Farey sequence of size n, denoted $\mathcal{F}_n$, consists of all reduced fractions between 0 and 1 whose denominators are less than or equal to n, ordered from smallest to biggest. For example, here is $\mathcal{F}_5$:

$$\left\{\frac{0}{1},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{2}{5},\frac{1}{2},\frac{1}{5},\frac{3}{5},\frac{2}{3},\frac{3}{4},\frac{4}{5},\frac{1}{1}\right\}.$$

   a. Make a table of the Farey sequences $\mathcal{F}_n$ for $n \leq 7$.
   b. Compute $|\,\mathcal{F}_n\,|$ for $n \leq 7$, and let $a_n = |\,\mathcal{F}_n\,|$. Show that

$$a_n = a_{n-1} + \phi(n),$$

   where $\phi(n)$ is the totient function. Hint: consider those fractions that have n as their smallest denominator.

4. Consider two reduced fractions:

$$0 \leq \frac{a}{b} < \frac{c}{d} \leq 1,$$

where $\gcd(a,b) = \gcd(c,d) = 1$. Show that if these two fractions are consecutive elements of a Farey sequence $\mathcal{F}_n$, then

$$\frac{c}{d} - \frac{a}{b} = \frac{1}{bd}.$$

Hint: use the fact that $\mathcal{F}_{n-1} \subset \mathcal{F}_n$ to reduce to the case where b or d equals n.

5. Define the *mediant* of two reduced fractions a/b and c/d to be $(a+c)/(b+d)$.

   a. Show that the mediant lies strictly between the other two fractions:

$$\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}.$$

b. Suppose a/b and c/d are reduced fractions with $a/b < c/d$. Show that if $bc - ad = 1$, then for $n = \max\{b, d\}$, a/b and c/d are consecutive elements of the Farey sequence $\mathcal{F}_n$.
c. Show $a/b < (a+c)/(b+d) < c/d$ form three consecutive elements of $\mathcal{F}_{b+d}$.

6. Show that if we have three consecutive members of a Farey sequence $\mathcal{F}_n$,

$$0 \le \frac{a}{b} < \frac{e}{f} < \frac{c}{d} \le 1,$$

then

$$\frac{e}{f} = \frac{a+c}{b+d}.$$

That is, elements between others in the Farey sequence *must* be mediants.

7. This exercise yields a fun result at the intersection of probability and number theory.

a. Use a computer program to search through all pairs of distinct integers, $1 \le a < b \le 100$, and compute the proportion of these pairs that are relatively prime.
b. Show that the probability that a random integer is divisible by p is $1/p$. That is, given prime p and any integer n, let $n = pk + r$ where $0 \le r < p$. Explain why k is the number of integers less than or equal to n that are multiples of p. Let $\Pr(n, p)$ denote the probability that a random integer between 1 and n is divisible by p, and conclude

$$\Pr(n,p) = \frac{k}{n} = \frac{1}{p} \cdot \frac{k}{k + \frac{r}{p}},$$

and hence

$$\lim_{n\to\infty} \Pr(n,p) = \frac{1}{p}.$$

c. Using similar reasoning, explain why the probability that two random numbers are both multiples of p approaches $(1/p)^2$. Conclude that the probability that two random numbers do *not* share a common factor of p approaches $1 - (1/p)^2$.
d. Explain why the probability that two random numbers are relatively prime is

$$\frac{1}{\zeta(2)} = \frac{6}{\pi^2} = 0.6079\ldots.$$