The Hierarchy of Integer Sequences

An Introduction to Combinatorics

Eric Rowland

Department of Mathematics, Hofstra University, Hempstead, NY

July 27, 2023 Draft version

Contents

Preface

Goals of this book

Sequences of integers appear throughout mathematics. In classical enumeration questions, one runs into the Fibonacci sequence, the Catalan sequence, and other sequences that satisfy recurrences of various types. This half-book provides an introduction to combinatorics by working up the hierarchy, from relatively simple sequences to more sophisticated sequences, exploring the relationship between the complexity of a sequence and the complexity of the objects it counts. A second half is planned. I have tried to carefully motivate topics and provide as much intuition as possible, especially in places where other sources do not include intuition that to me seems important and useful. Anyone from advanced high school students to senior researchers in other areas should find it informative.

This book emphasizes guessing, which as a skill is greatly undervalued. Although you'd never know it by reading most papers and textbooks, researchers typically discover theorems experimentally before proving them. In the context of sequences, usually we can compute the first several terms of a sequence, and then we want to know what the nth term is. A major goal of the book is determining how to make such conjectures. Indeed, guessing is often the hard part. Once you have a good conjecture, proving it is often routine or even automatic if you have the right framework.

Additionally, this book invites experimentation. I never liked exercises of the form "Prove X" when I had no reason to think that X is true. Exercises like this give away the punchline, so I have tried to keep them to a minimum. Instead there are questions. Some can be answered by performing a computation, but for many the form of the answer will not be clear at the outset. How do you answer a new question? You play. You explore, you look at examples, you do experiments — all to piece together intuition where you had no intuition before. This will sometimes take longer than working through a traditional exercise, but it is a lot more authentic and fun. And when you do see "Prove X", interpret it as "Use examples to first convince yourself that X is true, and then prove X".

Prerequisites

There are not many prerequisites for this book, since care has been taken to make the material as accessible as possible. We assume familiarity with basic linear algebra, although the reader can pick up the necessary material as needed. Solving systems of linear equations will be crucial in guessing. Starting in Chapter 7, vector spaces play a large role. However, all the vector spaces are fairly concrete; they are finite-dimensional vector spaces over $\mathbb{Q}$ of sequences of rational numbers. Most results hold more generally for sequences with entries in other fields, but we stick to $\mathbb{Q}$ except for a brief lapse into the complex numbers in Chapter 19.

Using mathematical software

You can certainly do mathematics without a computer, but you probably don't want to. Programming is a superpower. It lets you see and do things that ordinary humans can't, at speeds that were unimaginable a few decades ago. It gives you a higher-level perspective, allowing you to relegate details of computations and focus on the big picture. It allows you to wield powerful tools constructed by experts.

Fortunately, it's a superpower you can acquire relatively painlessly, while doing real work the whole time (instead of training exercises). The trick is to use software for every computation you come across, and whenever you don't know how to implement something yet, look it up. This way you'll constantly be getting better.

Several of the questions in this book assume you are able to use a calculator or mathematical software to carry out standard computations from calculus and linear algebra. You're encouraged to use software for experiments as well, since it will be less error-prone and faster in many cases.

Some of the questions ask you to write a program. My advice is to not try to write a program linearly from the beginning to the end. Don't start by choosing a function name and names for its arguments. Instead, start with an explicit example input that the program will eventually compute with. Ask yourself "What do I do with it first?", then implement the first step, and check that the step works as expected. Then implement the second step, and so on, until you complete the computation for that example. Finally, generalize the code so that it works with other examples.

The larger intention in this book is that everything we work with is computable. Existence theorems are nice to look at, but we assume we'll want to apply theorems to specific examples that arise. Accordingly, the theorems in this book provide algorithms and bounds in addition to existence results.

While not necessary for reading this book or answering the questions, you may find the Mathematica package IntegerSequences [**20, 21**] useful. It is a package, written by me, for identifying and computing with sequences from the classes discussed in this book.

Notes on nomenclature

Basic mathematical objects often have different names in different fields. One of the main classes we discuss consists of sequences that satisfy linear recurrences with constant coefficients. For example, the Fibonacci sequence belongs to this class. They are often called 'linear recurrence sequences'. However, a major theme of this book is that many nice classes of integer sequences satisfy some class of linear recurrences, so 'linear' is too ambiguous. We call them 'constant-recursive sequences'. They are also called 'C-finite sequences'.

Analogously, polynomial-recursive sequences satisfy recurrences with polynomial coefficients. They are also called 'holonomic sequences'.

And now

Enjoy the book!

Part 1

Combinatorial fundamentals

CHAPTER 1

Integer sequences and the OEIS

This book is about sequences of integers. Why are we interested in sequences of integers? Primarily because they count things! The usual story is that you're going about your day, when suddenly an integer sequence wanders into view. For example, you may be working on a problem whose solution comes down to the following question.

Example 1.1. Suppose you have a sum $a_1 + a_2 + \cdots + a_n$ of n numbers. In how many ways can parentheses be inserted so that each $+$ is a binary operation (only taking two summands)? Let's generate some data. For the sum $a + b$, there is only 1 way: $(a + b)$. The sum $a + b + c$ can be grouped either $(a + b) + c$ or $a + (b + c)$, in 2 ways. The sum $a + b + c + d$ can be grouped 5 ways:

$$((a + b) + c) + d, \qquad (a + (b + c)) + d, \qquad (a + b) + (c + d),$$
$$a + ((b + c) + d), \qquad a + (b + (c + d)).$$

The sequence of numbers so far is $1, 2, 5$. We can probably compute a couple more by hand. But what if they don't look familiar?

Fortunately there's a database, The On-Line Encyclopedia of Integer Sequences, which contains several hundred thousand sequences. Here are a few highlights; how many do you recognize?

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, \ldots \qquad (\mathsf{A000045})$$
$$1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, \ldots \qquad (\mathsf{A000079})$$
$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, \ldots \qquad (\mathsf{A000108})$$
$$0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, 120, \ldots \qquad (\mathsf{A000217})$$
$$0, 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169, 196, \ldots \qquad (\mathsf{A000290})$$
$$0, 1, 1, 3, 5, 11, 21, 43, 85, 171, 341, 683, 1365, 2731, \ldots \qquad (\mathsf{A001045})$$
$$1, 2, 2, 4, 2, 4, 4, 8, 2, 4, 4, 8, 4, 8, 8, 16, 2, 4, 4, 8, 4, 8, \ldots \qquad (\mathsf{A001316})$$
$$0, 1, 0, 2, 0, 1, 0, 3, 0, 1, 0, 2, 0, 1, 0, 4, 0, 1, 0, 2, 0, 1, \ldots \qquad (\mathsf{A007814})$$

The OEIS is an immensely valuable resource. It was created by Neil Sloane[1], who began collecting sequences in the 1960s. When you come across a sequence you don't recognize, you can look it up at `https://oeis.org` by typing in the first few numbers. If your sequence is there, then you're done. Great!

But frequently you'll come across sequences that haven't been seen before. The purpose of this book is to give you tools for obtaining information of the type that

[1]Neil Sloane was born in 1939 in Beaumaris, Wales.

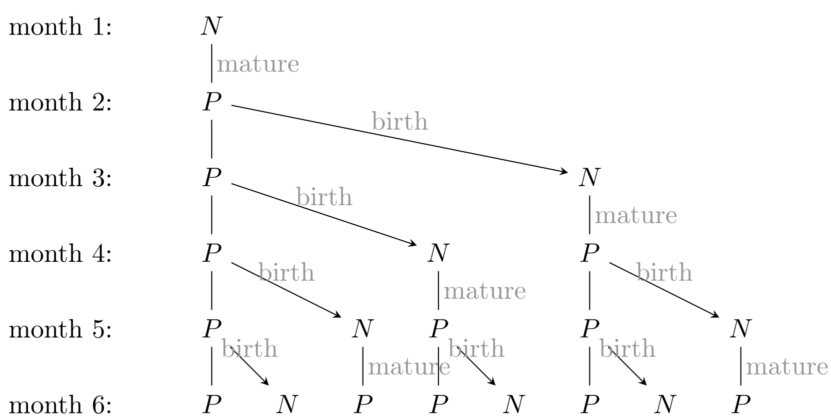
the OEIS would provide, even when your sequence isn't there. We'll explore several natural classes of sequences, each with a form that allows one to guess a sequence from its first several numbers. If you use the correct class and can compute enough numbers, then you can guess your sequence. From there you may even be able to verify it's the one you're interested in and determine many of its properties.

Example 1.2. Perhaps the most famous integer sequence is the Fibonacci sequence. It appears in the book *Liber Abaci* from the year 1202, written by Fibonacci[2], and it was known earlier in other parts of the world. Some mathematics goes out of fashion after a while, but the Fibonacci sequence is alive and well. It even has its own academic journal, *The Fibonacci Quarterly*. In Fibonacci's treatment:

- You start with a pair of newborn rabbits.
- Newborn rabbits take 1 month to reach maturity (not true in real life!).
- Each mature pair of rabbits that is not pregnant becomes immediately pregnant with 1 new pair of rabbits.
- The gestation period (that is, length of pregnancy) for rabbits is 1 month (true in real life!).
- Rabbits don't die.

Essentially, each mature pair produces 1 new pair per month. How many pairs of rabbits do you have each month?

In the first month you have 1 pair of newborn rabbits. A month later, you have 1 pair of pregnant rabbits. One month after that, you have 1 pair of (re-)pregnant rabbits and 1 pair of newborn rabbits, for a total of 2 pairs. We can track of all these rabbits in a tree diagram where each successive month is represented by a level. Each pair is either made up of newborns (N) or is pregnant (P).



The number of rabbit pairs you have each month is $1, 1, 2, 3, 5, 8, \ldots$. The number of pairs alive in a given month is the number of pairs that were already alive the previous month, plus the number of pairs that were pregnant the previous month. And how many were pregnant the previous month? As many as were alive the month before that. In other words, every subsequent number in the Fibonacci sequence is the sum of the two previous numbers. We can now extend the sequence

[2]Leonardo of Pisa was born around 1170 in Pisa (now in Italy) and died around 1250 in Pisa. He wasn't known as 'Fibonacci' until the 1800s.

without explicitly counting pairs of rabbits, by performing arithmetic instead:

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, \ldots.$$

To find out how many rabbits there are in a particular month, we extend this sequence appropriately.

The hierarchy of integer sequences we discuss in this book has its roots in combinatorics and computer science. There are plenty of sequences in the OEIS that we don't know how to fit into this hierarchy. But we won't ignore them, since it's important to understand why they don't fit and what would be necessary to extend the hierarchy to include them.

Indexing and shifts

Throughout the book we'll use $s(n)_{n\geq 0}$ to denote a sequence

$$s(0), s(1), s(2), s(3), \ldots.$$

Each number in a sequence is called a *term*, so $s(n)$ is the *nth term*. Note that we allow the possibility of negative terms. For example,

$$5, 3, 1, -1, -3, -5, -7, -9, -11, -13, -15, -17, \ldots$$

is a perfectly fine sequence. But often the sequences we'll be interested in will consist of non-negative integers, since such numbers have a *combinatorial interpretation*. That is, they arise in counting questions. For example, if we have sets $A_0, A_1, A_2, \ldots$, then letting $s(n) = |A_n|$ gives a sequence whose nth term is the number of elements in the nth set. This was the case for the Fibonacci sequence.

Convention. We will always index infinite sequences starting from $n = 0$.

However, when you find a sequence in the wild, it might not be clear which term should be assigned to the index $n = 0$, or, equivalently, what index n should be assigned to the first term. Notice we didn't explicitly say which term of the Fibonacci sequence corresponds to $n = 0$. We got away with this because the terms of the Fibonacci sequence aren't defined directly in terms of n; you can continue a portion of the sequence simply by knowing two terms, without knowing their indices:

$$\ldots, 1346269, 2178309, 3524578, \underline{\hspace{2cm}}, \ldots.$$

Maybe the initial terms 1 and 1 should be the $n = 0$ and $n = 1$ terms. On the other hand, that choice is somewhat arbitrary. Suppose we started with initial terms $1, 2$ instead:

$$1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, \ldots.$$

We get the *shift* of the previous sequence. Starting with $2, 3$ instead gives the shift of the shift. Or maybe we should start with $-1, 1$, which eventually gets us to $1, 1$:

$$-1, 1, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots.$$

Lots of initial conditions result in essentially the same sequence.

When adding a sequence in a database such as the OEIS, it's useful to pin down the initial conditions so we can write precise formulas for the nth term that depend on n or on other sequences. In this book (as in the OEIS and elsewhere) we use the following.

Definition 1.3. The *Fibonacci sequence* is the sequence $F(n)_{n\geq 0}$ defined by initial conditions $F(0) = 0$ and $F(1) = 1$ and the recurrence $F(n) = F(n-1) + F(n-2)$ for each $n \geq 2$.

Therefore the Fibonacci sequence is

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, \dots \qquad (\text{A000045}).$$

In terms of rabbit pairs, $F(n)$ is the number of pairs on month n, for each $n \geq 1$.

The OEIS uses a sequence's "offset" field to record the index of its first term. (The offset field usually also contains a second number; this is 1 plus the number of initial terms with absolute value ≤ 1.) In rare cases the OEIS contains multiple shifts of a single sequence. Shifts of $(2n)_{n\geq 0}$ (A005843) include $(2n+6)_{n\geq 0}$ (A020739) and $(2n+8)_{n\geq 0}$ (A020744).

In general, a *recurrence* for $s(n)_{n\geq 0}$ is a method of computing the nth term $s(n)$ recursively in terms of previous terms. For example, $s(n) = n \cdot s(n-1)$ is a recurrence satisfied by the sequence $1, 1, 2, 6, 24, 120, \dots$; it expresses $s(n)$ in terms of $s(n-1)$. Another example is $s(n) = s(n-2)+2$; along with the initial conditions $s(0) = 0$ and $s(1) = 1$, this recurrence generates the sequence $0, 1, 2, 3, 4, 5, \dots$ of non-negative integers. On the other hand, $s(n) = n$ is not a recurrence; instead we would call this a *formula*.

Questions

Computations.

(1) Use the Fibonacci recurrence to compute $F(20)$.

Experiments.

(2) Glance through the first 50 sequences in the OEIS (A000001 to A000050) to get a sense of what's there. Are they the 50 most important sequences in all of mathematics?

(3) What happens when you look up the terms $1, 1, 2, 3, 5, 8$ in the OEIS? What if you include more or fewer terms?

(4) Compute the first several terms of the sequence $s(n)_{n\geq 0}$ defined by the recurrence $s(n) = s(n-1) + s(n-2)$ and initial conditions $s(0) = -3$ and $s(1) = 2$. You'll see that eventually it contains the same numbers as the Fibonacci sequence. Find other initial conditions involving a negative number that have this same property. Are there infinitely many?

(5) In real life, newborn rabbits take closer to 6 months to reach maturity. Under this condition, what integer sequence counts the rabbit pairs on month n?

(6) Let $s(n)_{n\geq 0}$ be the sequence defined by initial conditions $s(0) = 1$ and $s(1) = 2$ and the recurrence $s(n) = s(n-1)s(n-2)$ for each $n \geq 2$. Is there a formula for the nth term $s(n)$?

Proofs.

(7) Prove that

$$2 \cdot 6 \cdot 10 \cdots (4n-2) = (n+1)(n+2)(n+3) \cdots (2n-1)(2n)$$

for all $n \geq 1$. Hint: Show that both sides satisfy the same recurrence and initial conditions.

(8) Are there countably many or uncountably many integer sequences?

Programs.

(9) Write a program that takes a list of integers and opens the web page containing the OEIS search results for those terms.

CHAPTER 2

Tuples, words, and bijections

Tuples and words

Combinatorics is populated by *combinatorial objects*. One of the simplest combinatorial objects is the finite list, which in mathematics is commonly referred to as a *tuple*. For example, we might be interested in all tuples of length 3 containing 0s and 1s:

$$(0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1),(1,1,0),(1,1,1).$$

Definition 2.1. Let Σ be a set[1] and let $\ell \geq 0$ be an integer. An *ℓ-tuple on* Σ is a list $(a_1, a_2, \ldots, a_\ell)$ of length ℓ, where each entry a_i belongs to Σ.

The most fundamental question about a combinatorial object is: How many are there? How many ℓ-tuples are there on a set of size $|\Sigma| = k$?

Theorem 2.2. *Let $k \geq 0$ and $\ell \geq 0$ be integers. The number of ℓ-tuples with entries from a set of size k is k^ℓ. (We interpret 0^0 as 1.)*

Proof. We use induction on the length ℓ. Let $s_k(\ell)$ be the number of ℓ-tuples with entries from a set of size k. If $\ell = 0$, there is only one tuple, namely (). Therefore $s_k(0) = 1 = k^0$. Let $\ell \geq 1$, and inductively assume $s_k(\ell-1) = k^{\ell-1}$. Each ℓ-tuple is obtained by choosing an $(\ell-1)$-tuple and appending a new entry. For each $(\ell-1)$-tuple, there are k possible new entries to append, so $s_k(\ell) = s_k(\ell-1) \cdot k = k^\ell$. $\square$

Closely related to tuples are *words*. Here are the words of length 3 containing 0s and 1s:

$$000, 001, 010, 011, 100, 101, 110, 111.$$

We have listed them in *lexicographic order* (a fancy way of saying *dictionary order*), where we consider the letter 0 to come before the letter 1 in the alphabet.

Definition 2.3. Let Σ be a set. A *word on* Σ is a finite sequence of elements from Σ.

Just as there is a single tuple of length 0, there is a single word of length 0, called the *empty word* and denoted by ε.

The *concatenation* of two words $v = a_1 a_2 \cdots a_\ell$ and $w = b_1 b_2 \cdots b_m$ is $vw := a_1 a_2 \cdots a_\ell b_1 b_2 \cdots b_m$. For example, concatenating the English words abs and orb produces absorb. Since we can build words from smaller words, we can also extract words.

Definition 2.4. The word y is a *factor* of w if $w = xyz$ for some (possibly empty) words x and z.

[1] Not to be confused with summation; here Σ is just a letter.

For example, 11 is a factor of 0110, but 00 is not.

When using the terminology of words, we refer to the set Σ as the *alphabet* and the entries of a word as *letters*. The set of words of length ℓ on the alphabet Σ is denoted by Σ^ℓ. We use $\Sigma^* := \bigcup_{\ell \geq 0} \Sigma^\ell$ to denote the set of all words on Σ. For example,

$$\{0,1\}^* = \{\varepsilon, 0, 1, 00, 01, 10, 11, 000, \dots\}.$$

Bijections

Words are very similar to tuples. On paper it's easy to think of them as the same objects; after all, they only differ in some parentheses and commas. But in many programming languages there is a clear distinction between words (strings) and tuples (lists). For instance, text processing is typically done on strings while numeric data is typically stored in lists or arrays.

However, a tuple can be transformed into a word, and vice versa, without losing any information.

$$\begin{array}{cccccccc} (0,0,0) & (0,0,1) & (0,1,0) & (0,1,1) & (1,0,0) & (1,0,1) & (1,1,0) & (1,1,1) \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \end{array}$$

We can formalize this correspondence as a bijection.

Definition 2.5. Let A and B be sets, and let $f\colon A \to B$ be a function.

- f is *surjective* if, for each $b \in B$, there exists $a \in A$ such that $f(a) = b$.
- f is *injective* if, for all $a, c \in A$, $f(a) = f(c)$ implies $a = c$.
- f is a *bijection* if f is surjective and injective. In this case, A and B are said to be *in bijection*.

Example 2.6. Let T be the set of ℓ-tuples with entries from Σ, and let $W = \Sigma^\ell$. Define $f\colon T \to W$ by $f((a_1, a_2, \dots, a_\ell)) = a_1 a_2 \cdots a_\ell$. We claim that f is a bijection. To show that f is surjective, let $a_1 a_2 \cdots a_\ell \in W$. Since f maps the tuple $(a_1, a_2, \dots, a_\ell)$ to $f((a_1, a_2, \dots, a_\ell)) = a_1 a_2 \cdots a_\ell$, surjectivity follows. To show that f is injective, assume the two tuples $(a_1, a_2, \dots, a_\ell)$ and $(b_1, b_2, \dots, b_\ell)$ satisfy $f((a_1, a_2, \dots, a_\ell)) = f((b_1, b_2, \dots, b_\ell))$. This implies $a_1 a_2 \cdots a_\ell = b_1 b_2 \cdots b_\ell$, so $a_1 = b_1$, $a_2 = b_2$, and so on. Therefore $(a_1, a_2, \dots, a_\ell) = (b_1, b_2, \dots, b_\ell)$, so f is injective.

It's helpful to think about surjectivity and injectivity as statements about existence and uniqueness. If $f\colon A \to B$ is surjective, then every element $b \in B$ "comes from" an element in A: There exists an $a \in A$ such that $f(a) = b$. If f is also injective, then there is a *unique* element a with this property, since if two elements a, c satisfy $f(a) = f(c)$ then in fact $a = c$. In other words, surjectivity is the property that there are enough elements in A to reach everything in B, and injectivity is the property that there aren't any extra.

We use this in the next theorem, which shows that when we transform an object (for example, a tuple) using a bijection, we can undo this transformation and recover the original object. This implies there is no loss of information when we perform the transformation, so the two objects represent the same information in two different ways.

Theorem 2.7. *If $f\colon A \to B$ is a bijection, then there is a function $g\colon B \to A$ such that $g(f(a)) = a$ for all $a \in A$ and $f(g(b)) = b$ for all $b \in B$.*

The function g is called the *inverse* of f, and we write $g = f^{-1}$.

PROOF OF THEOREM 2.7. Let $b \in B$. We claim there exists a unique $a \in A$ such that $f(a) = b$. Such an a exists, by surjectivity. Moreover, it is unique, by injectivity. Define $g(b) := a$; then $f(g(b)) = f(a) = b$.

Now let $a \in A$; we must show that $g(f(a)) = a$. Let $b := f(a)$. Since $g(b)$ is defined to be the unique element $x \in A$ such that $f(x) = b$, we have $g(b) = a$. Therefore $g(f(a)) = g(b) = a$. □

This brings us to one of the most important principles of enumeration.

Corollary 2.8. *If finite sets A and B are in bijection, then $|A| = |B|$.*

PROOF. Let $f : A \to B$ be a bijection. Every $a \in A$ corresponds to a unique $b \in B$, obtained by $b = f(a)$. Similarly, every $b \in B$ corresponds to a unique element $f^{-1}(b) \in A$. Since $f^{-1}(b) = f^{-1}(f(a)) = a$, we can pair a and b with each other. Therefore the elements of A and the elements of B are in one-to-one correspondence. It follows that A and B have the same number of elements. □

Since tuples and words are in bijection, we have an immediate corollary of Theorem 2.2.

Corollary 2.9. *Let $k \geq 0$ and $\ell \geq 0$ be integers. The number of words of length ℓ on an alphabet of size k is k^ℓ.*

This corollary justifies the notation Σ^ℓ, since $|\Sigma^\ell| = |\Sigma|^\ell$.

Example 2.10. The bijection we gave between tuples and words is quite simple. Another fairly simple bijection is the map $w \mapsto \text{reverse}(w)$, where $\text{reverse}(a_1 a_2 \cdots a_\ell) := a_\ell \cdots a_2 a_1$ is the word obtained by reading the letters of a word in reverse order. From $\text{reverse}(w)$, we can obtain w by reversing again, so $\text{reverse}^{-1} = \text{reverse}$.

Example 2.11. On the alphabet $\{0, 1\}$, there are 5 words of length 3 that don't contain 00 as a factor:

$$010, 011, 101, 110, 111.$$

We say these words *avoid* 00. There are also 5 words in $\{0, 1\}^3$ that avoid 11:

$$000, 001, 010, 100, 101.$$

There is a bijection between these two sets of words. Let f be the function that replaces each 0 with 1 and each 1 with 0. Then f transforms a word avoiding 00 into a word avoiding 11:

$$\begin{array}{ccccc} 010 & 011 & 101 & 110 & 111 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 101 & 100 & 010 & 001 & 000 \end{array}$$

You can check that, for each $\ell \geq 0$, f is a bijection from words of length ℓ avoiding 00 to words of length ℓ avoiding 11. Therefore the number of words avoiding 00 is equal to the number of words avoiding 11.

Questions

Experiments.

(1) What fraction of length-ℓ words on 3 letters contain all 3 letters?

(2) Let $f\colon \{0,1\}^* \to \{0\}^*$ be the function that deletes all 1s from a word. For example, $f(01010) = 000$. Is f surjective? Is f injective?

(3) Let $f\colon (\{0,1\}^* \setminus \{\varepsilon\}) \to \{0,1\}^*$ be the function that deletes the first letter of a word. For example, $f(01010) = 1010$. Is f surjective? Is f injective?

(4) What's an example of a function $f\colon \{0,1\}^* \to \{0,1\}^*$ that is injective but not surjective?

(5) Let w be a word of length ℓ, all of whose letters are distinct. How many factors does w have?

(6) A word w is a *palindrome* if $\text{reverse}(w) = w$. For example, 0102010 is a palindrome. How many length-ℓ palindromes are there on an alphabet of size k?

(7) What is the maximum possible number of distinct palindromic factors a word of length ℓ on an alphabet of size k can have?

(8) A word is a *permutation* if no letter occurs in it more than once. How many length-ℓ permutations are there on an alphabet of size k? How many length-ℓ permutations are there on an alphabet of size ℓ?

Proofs.

(9) Prove that if A and B are finite sets and $|A| = |B|$, then there is a bijection $f\colon A \to B$.

(10) As in Example 2.11, let $f\colon \{0,1\}^* \to \{0,1\}^*$ be the function that replaces each 0 with 1 and each 1 with 0. Prove that, for each $\ell \geq 0$, f is a bijection from words of length ℓ avoiding 00 to words of length ℓ avoiding 11.

(11) Prove the special case $\ell = 2$ of Theorem 2.2 using a bijection from the 2-tuples on Σ to points in a square. Does a similar proof work for all ℓ?

(12) The vertices of a hypercube in d dimensions, centered at the origin and with side length 2, are the points whose coordinates have entries -1 and 1. How many vertices does this hypercube have?

(13) Let $s(n)_{n\geq 0}$ be the sequence $0, 1, -1, 2, -2, 3, -3, \dots$.
   (a) What is the nth term $s(n)$ as a function of n?
   (b) Prove that s is a bijection from the set of non-negative integers to the set of integers.

(14) Prove that if $f\colon A \to B$ and $g\colon B \to C$ are both bijections then the function $h\colon A \to C$ defined by $h(a) = g(f(a))$ is a bijection.

(15) Prove that if f is a bijection then f^{-1} is a bijection.

Experiences.

(16) Look up a video of a child asked to perform the conservation tasks designed by psychologist Jean Piaget, specifically conservation of number.

CHAPTER 3

Binary words avoiding 00

A *binary word* is a word on the alphabet $\{0,1\}$. Thanks to Corollary 2.9, we know that the number of binary words of length ℓ is 2^ℓ. In this chapter we'll be interested in binary words that avoid 00.

Let $s(\ell)$ be the number of words in $\{0,1\}^\ell$ that avoid 00. What is the sequence $s(\ell)_{\ell\geq 0}$? From Example 2.11, we know $s(3)=5$. For $\ell=2$ there are 3 words that avoid 00 (all except 00). For $\ell=4$ there are 8:

$$0101, 0110, 0111, 1010, 1011, 1101, 1110, 1111.$$

The sequence $s(\ell)_{\ell\geq 0}$ begins $1,2,3,5,8,\dots$. This looks like (a shift of) the Fibonacci sequence! Is it? We conjecture $s(\ell)=F(\ell+2)$. If we are confident in this conjecture, there are two possible approaches toward proving it. We could try to show that $s(\ell)$ satisfies the Fibonacci recurrence, or we could try to find a bijection between length-ℓ words avoiding 00 and objects that we know are counted by $F(\ell+2)$ — pairs of rabbits, for example. In fact one approach will likely inform the other.

Let's start with the recurrence. Can we see combinatorially that $s(\ell)=s(\ell-1)+s(\ell-2)$? This would be true if each word of length ℓ avoiding 00 "comes from" either a word of length $\ell-1$ or a word of length $\ell-2$ avoiding 00. Here are the words of lengths 2 and 3 avoiding 00:

$$01, 10, 11, \qquad 010, 011, 101, 110, 111.$$

Can we see how to obtain the length-4 words

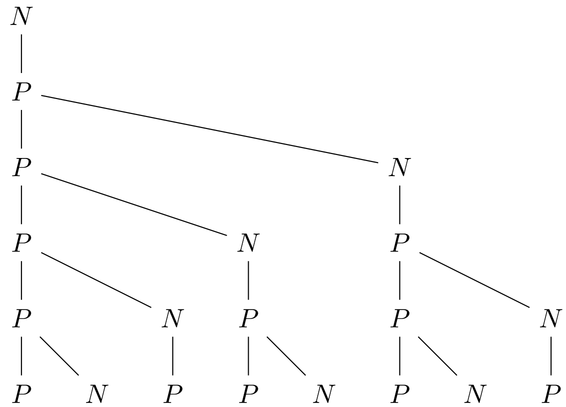
$$0101, 0110, 0111, 1010, 1011, 1101, 1110, 1111$$

from them? Yes, look at the first letter of each length-4 word. The first 3 words begin with 0, and the last 5 begin with 1. It looks like the first 3 words are obtained from the length-2 words by forming $01w$, and the last 5 are obtained from the length-3 words by forming $1w$.

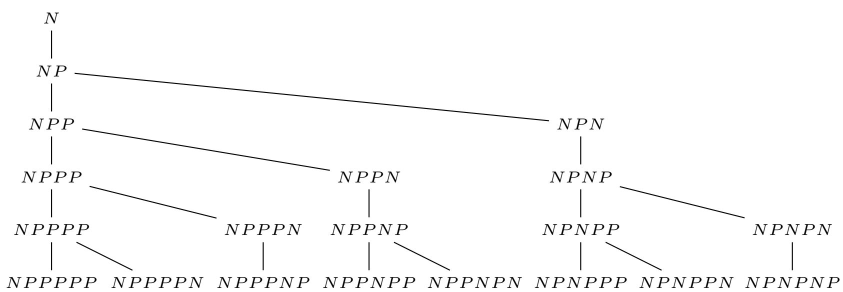
$$\begin{array}{cccccccc}
01 & 10 & 11 & 010 & 011 & 101 & 110 & 111 \\
\updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\
\color{red}{01}01 & \color{red}{01}10 & \color{red}{01}11 & \color{red}{1}010 & \color{red}{1}011 & \color{red}{1}101 & \color{red}{1}110 & \color{red}{1}111
\end{array}$$

Does this work for general ℓ? If a word w of length $\ell-2$ avoids 00, then $01w$ is a word of length ℓ that avoids 00. If a word w of length $\ell-1$ avoids 00, then $1w$ is a word of length ℓ that avoids 00. Moreover, every length-ℓ word that avoids 00 arises in exactly one of these two ways, so $s(\ell)=s(\ell-1)+s(\ell-2)$.

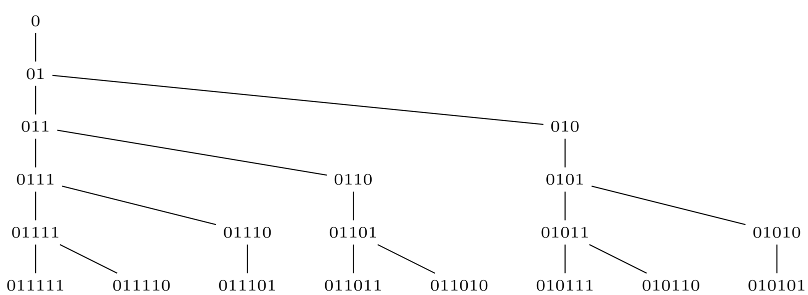
Does the combinatorial interpretation of the recurrence suggest a bijection to rabbit pairs? One issue is that in Chapter 1 we didn't distinguish between different pairs of rabbits. In order to give a bijection, we need to name each pair. Let's look at our tree diagram again.



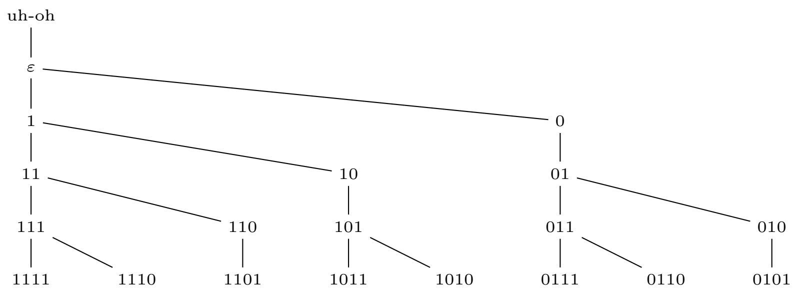
To distinguish the rabbit pairs on a given level from each other, we'll name each pair according to its ancestry in the tree — the path traversed from the top N as a word on the alphabet $\{N, P\}$. In the first month, the pair is called N. In month $n = 2$, that same pair is called NP. In month $n = 3$, we have two pairs, NPP and NPN, and so on.



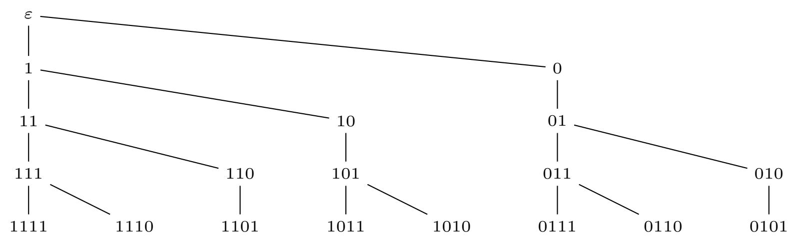
Now the length of each word tells us its level. Notice that these words avoid NN, since a newborn pair must mature and become pregnant before it can produce more newborns. If we write 0 instead of N and 1 instead of P, we get words avoiding 00.



But there's a problem. The rabbit pairs on level $n = 4$ are 0111, 0110, and 0101. These should correspond to the 3 words 01, 10, and 11 avoiding 00. The lengths are wrong. Apparently we need to drop the first two letters of each word in the tree to get the words we want. This is an invertible (bijective) operation, since the first two letters are always 01 (so we can add them back when going in reverse), except that there is an issue with the word 0 at month 1, because it doesn't start with 01.



This often happens when trying to find a bijection; something doesn't quite fit. Usually this suggests that we're not quite looking at the situation in the most natural way. Let's remember what objects we're trying to put into correspondence: words of length ℓ that avoid 00, and rabbit pairs on month $\ell + 2$. The smallest value of ℓ we can consider is $\ell = 0$, so that means the bijection only involves months $2, 3, \ldots$. Month 1 isn't involved at all, so we can chop it off the tree.



Let's write the result formally.

Theorem 3.1. *For each $\ell \geq 0$, the number of length-ℓ words on $\{0, 1\}$ that contain no consecutive 0s is $F(\ell + 2)$.*

Proof. Let $\ell \geq 0$. Let A be the set of rabbit pairs obtained on month $\ell + 2$ according to the five rules in Example 1.2. Let B be the set of words in $\{0, 1\}^\ell$ that avoid 00. We form a bijection $f \colon A \to B$ as a composition of three functions f_1, f_2, f_3 defined as follows.

Given a pair of rabbits $r \in A$, let $f_1(r)$ be the word on $\{N, P\}$ that encodes the ancestry of r, namely, the sequence of letters read along the shortest path from the top of the tree to r. The function f_1 is a bijection from A to the set A_1 of all words in $\{N, P\}^{\ell+2}$ that begin with NP and avoid NN.

Given a word $v \in A_1$, let $f_2(v)$ be the word on $\{0, 1\}$ obtained from v by replacing each N with 0 and each P with 1. The function f_2 is a bijection from A_1 to the set A_2 of all words in $\{0, 1\}^{\ell+2}$ that begin with 01 and avoid 00.

Given a word $w \in A_2$, let $f_3(w)$ be the word on $\{0, 1\}$ obtained from w by deleting its first two letters (which are 01). The function f_3 is a bijection from A_2 to B.

Since the three functions $f_1 \colon A \to A_1$, $f_2 \colon A_1 \to A_2$, and $f_3 \colon A_2 \to B$ are bijections, the function $f \colon A \to B$ defined by $f(r) = f_3(f_2(f_1(r)))$ is a bijection. Since the Fibonacci sequence satisfies $|A| = F(\ell + 2)$, it follows from Corollary 2.8 that $|B| = F(\ell + 2)$. □

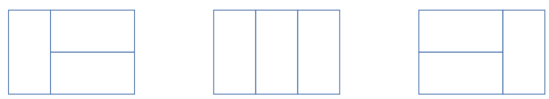
Questions

Computations.

(1) Use the recurrence to generate the 13 words in $\{0,1\}^5$ that avoid 00.
(2) Use the bijection with rabbit pairs to generate the 13 words in $\{0,1\}^5$ that avoid 00.

Proofs.

(3) Fill in the details of the proof of Theorem 3.1 by proving that the functions f_1, f_2, f_3 are surjective and injective.
(4) What is the inverse function f^{-1} of the function f in the proof of Theorem 3.1?
(5) What is the relationship between the order in which the rabbit pairs on month $\ell + 2$ appear in the tree and the lexicographic order of length-ℓ words that avoid 00? Why?
(6) A *domino* is a 1×2 rectangle or a 2×1 rectangle. A *domino tiling* of an $n \times m$ rectangle is an arrangement of non-overlapping dominos in the shape of an $n \times m$ rectangle. For example, here are the domino tilings of a 3×2 rectangle:



How many domino tilings does an $n \times 2$ rectangle have?

Experiments.

(7) Can you come up with a different set of rabbit reproduction rules that admits a bijection to words on $\{0,1\}$ avoiding 000 rather than words avoiding 00?

CHAPTER 4

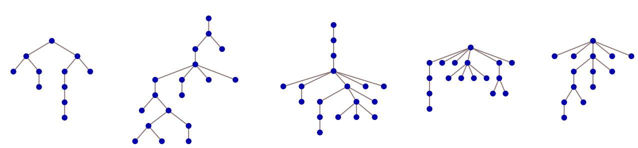
Catalan numbers

A major theme of this book is that an integer sequence $s(n)_{n\geq 0}$ often has multiple descriptions. In particular, there may be many different combinatorial objects with the property that $s(n)$ is the number of objects of size n. The poster child for sequences with multiple combinatorial interpretations is the sequence of *Catalan*[1] *numbers*, which are so common in combinatorics that their combinatorial interpretations fill an entire book [**30**]. In this chapter we'll see three combinatorial objects counted by the Catalan numbers, and several more will appear later in the book.

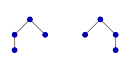
Two objects that share an enumeration sequence are necessarily closely related. The underlying structures of these objects must be the same, and there may even be natural bijections between them. Since the terms of the sequence can often be computed easily, usually recognizing a sequence is the first clue that two objects are in fact related.

Plane trees

The first objects we'll discuss are *plane trees*. We employed plane trees, without naming them as such, to keep track of rabbits in Chapter 3. Here are some more:



Plane trees are composed of points in the plane, connected by line segments. We'll refer to the points as *vertices* and the line segments as *edges*. Each plane tree has a special *root vertex*. Rather than drawing a tree with its root at the bottom, we put the root on the top, more like a family tree than a forest tree. Each vertex, including the root, has 0 or more *children vertices*, connected to their parent by an edge. The children of each vertex are ordered (say, left to right). This means, for example, that the trees



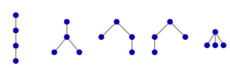
are different, because in one tree the root's first child has a child of its own, and in the other tree the root's second child has a child of its own. If we didn't care about child order, then you could think of a tree as a mobile hung by its root; rather than being embedded in the plane, the children could swap places with each other freely in space. For plane trees, we do care about order.

[1]Eugène Catalan was born in 1814 in Bruges (now in Belgium) and died in 1894 in Liège, Belgium.

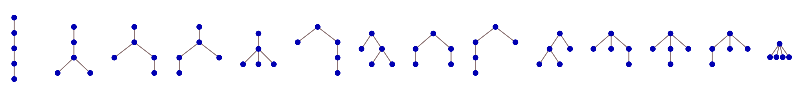
How many plane trees are there? First we should figure out what parameters we would like to enumerate them according to. The number of edges? The number of vertices? The depth (number of generations)?

To understand the structure of an object, it's often useful to understand how to build objects from a smaller objects. Words can be built from smaller words by appending letters. Trees can be built from smaller trees by adding children vertices along with edges connecting them to their parents. The smallest plane tree consists of a root vertex with no children. When we add a vertex to a general tree, the depth may or may not change, so the depth is not a fundamental measure of the size of a tree. When we add a vertex, certainly the number of vertices increases by 1. The number of edges also increases by 1. Since the smallest plane tree has 0 edges and 1 vertex, this implies that a plane tree with n edges has $n+1$ vertices. Therefore it doesn't matter much whether we enumerate plane trees according to the number of edges or number of vertices. We will choose edges.

Here are the plane trees with $n = 3$ edges:



Here are the plane trees with $n = 4$ edges:



Definition 4.1. For each $n \geq 0$, the nth *Catalan number* $C(n)$ is the number of plane trees with n edges (and $n+1$ vertices).

The sequence $C(n)_{n\geq 0}$ of Catalan numbers (A000108) is

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, 742900, 2674440, 9694845, \dots.$$

Balanced parentheses

How many valid ways are there to arrange n pairs of parentheses? For $n = 3$ there are 5 ways:

$$()()(), \quad ()(()), \quad (()()), \quad ((())), \quad (())().$$

These are valid because each open parenthesis has a unique matching close parenthesis to its right. On the other hand, the arrangement $)()()($ is not valid.

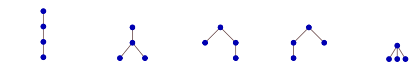
A valid arrangement of n pairs of parentheses is called a *Dyck*[2] *word* of length $2n$. For $n = 0$, the empty word ε is the only Dyck word. For $n = 1$, there is 1 Dyck word, namely $()$. For $n = 2$ there are 2: $()()$ and $(())$. For $n = 4$ there are 14:

$$\begin{array}{ccccccc} ()()()(), & ()()(()), & ()(()()), & ()((())), & ()(())(), & (()()()), & (()(())), \\ ((()())), & (((()))), & ((())()), & (())()(), & (())(()), & (()())(), & ((()))(). \end{array}$$

It seems the Catalan numbers are appearing.

[2]Walther von Dyck (pronounced 'dik') was born in 1856 in Munich, in the Kingdom of Bavaria (now in Germany) and died in 1934 in Munich.

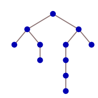
In fact there is a bijection from plane trees to Dyck words. Here is the correspondence for $n = 3$:



$\updownarrow$ $\updownarrow$ $\updownarrow$ $\updownarrow$ $\updownarrow$

((())) (()()) ()(()) (())() ()()()

In general, suppose you have a plane tree. Imagine you're standing on the plane at a point just above the root. Put your left hand on the root, and then walk counterclockwise around the tree, maintaining contact with the tree. Every time you walk down along an edge, record an open parenthesis, and every time you walk up along an edge, record a close parenthesis. Stop when you get back to where you started. You will have traversed each edge twice, once down and once up. For example, the tree



can be described by (()(()))((((())))()). Between the two parentheses corresponding to a given edge, you will have recorded the word corresponding to the subtree rooted at that edge's lower vertex. By induction, since each subtree corresponds to a balanced word, the word for the entire tree is also balanced.

Dyck paths

It's easy to see that the words ()())(and ())(() are unbalanced. It's a little harder to determine whether (())())((()) is balanced. How could you systematically determine whether a given word is balanced or not?

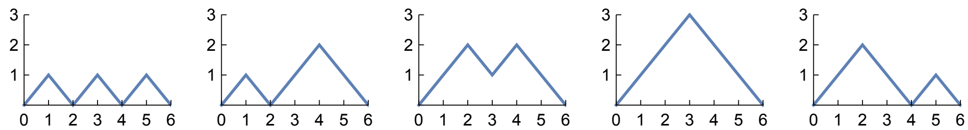
Start reading the word, left to right. The trick is to keep track of the number of unmatched open parentheses you've seen so far. That way, when you read a close parenthesis, you know whether there is an available open parenthesis to match it. When you read an open parenthesis, add 1 to your count of unmatched open parentheses. When you read a close parenthesis, subtract 1 from the count. For example, reading the letters of (()))(one at a time, we count 0 (before reading the first letter), 1, 2, 1, 0, -1, 0 unmatched open parentheses. If there are n open parentheses and n close parentheses, then we'll end up with a count of 0 at the end. But the -1 indicates that we found a close parenthesis without an open parenthesis to match. This gives us another characterization of valid arrangements of parentheses.

Theorem 4.2. *A word consisting of open and close parentheses is a Dyck word if and only if its running count of unmatched open parentheses consists of non-negative integers and ends with* 0.

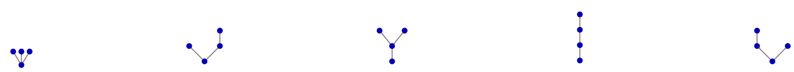
Theorem 4.2 gives yet another way to represent a plane tree. Namely, the function taking a Dyck word to its sequence of unmatched open parenthesis counts is also a bijection. For $n = 3$ here is the correspondence:

()()()	()(())	(()())	((()))	(())()
$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$
0101010	0101210	0121210	0123210	0121010

If we plot these sequences, we obtain *Dyck paths*:



A *Dyck path of length* $2n$ is a path from $(0, 0)$ to $(2n, 0)$ consisting of northeast and southeast steps, with the restriction that no point on the path lies below the x axis. There is a resemblance between these paths and the plane trees. Namely, horizontally collapsing a Dyck path by identifying each upward step with its corresponding downward step gives an upside down plane tree:



Equivalently, replace each open parenthesis with N and each close parenthesis with S. Then the word on the alphabet $\{N, S\}$ encodes the steps of the path.

Corollary 4.3. *The number of Dyck paths of length* $2n$ *is* $C(n)$*.*

Questions

Experiments.

(1) Match up the 14 plane trees that have 4 edges with the corresponding Dyck words of length 8.

(2) How many paths are there from $(0, 0)$ to $(2n, 0)$ consisting of northeast and southeast steps (now allowing the path to dip below the x axis)?

(3) Consider paths from $(0, 0)$ to (n, n) consisting of north and east steps, with the extra condition that no point on the path lies above the line $y = x$. Draw all such paths for $n = 1$, $n = 2$, and $n = 3$. How many such paths are there for each $n \geq 0$?

(4) Revisit Example 1.1 to compute, for more values of n, the number of ways of inserting parentheses in a sum of n numbers. Is there a relationship to any of the combinatorial objects in Chapter 4?

(5) Define the *rotation* of a nonempty word w to be the word $\rho(w)$ obtained by moving the first letter of w to the end. For example, $\rho(00110) = 01100$. Let $f(w)$ be the element of $\{w, \rho(w), \rho^2(w), \ldots, \rho^{|w|-1}(w)\}$ that occurs first in lexicographic order. For example, $f(00110) = 00011$. For each $n \geq 1$, let $S(n)$ be the set of words in $\{0, 1\}^{2n-1}$ consisting of n instances of the letter 0 and $n - 1$ instances of the letter 1. What is $|\{f(w) : w \in S(n)\}|$? In other words, how many words in $S(n)$ are lexicographically first among their rotations?

(6) In a Dyck path, a *valley* is a southeast step immediately followed by a northeast step. How many Dyck paths of length $2n$ have the property that each valley is strictly higher than all previous valleys (to its left)?

(7) How many Dyck paths of length $2n$ have the property that each valley is at least as high as all previous valleys? This was studied by Barcucci, Del Lungo, Fezzi, and Pinzani [**5**].

(8) Use the first 16 terms of $C(n)_{n\geq 0}$ to compute as many terms of $\left(\frac{C(n+1)}{C(n)}\right)_{n\geq 0}$ as possible. Stare at them, and try to guess a method for computing $C(n)$.

Proofs.

(9) Prove that the function we described for turning plane trees with n edges into Dyck words of length $2n$ is a bijection.

(10) Prove Theorem 4.2.

Part 2

Polynomial sequences

CHAPTER 5

Closure properties for polynomial sequences

The first class of sequences we will study are sequences whose terms are obtained by evaluating polynomials. For example, here's a polynomial in the symbol x:

$$6x^3 - 2x + 5.$$

Evaluating it at $x = 0, x = 1, \ldots$ gives the sequence $\left(6n^3 - 2n + 5\right)_{n \geq 0}$:

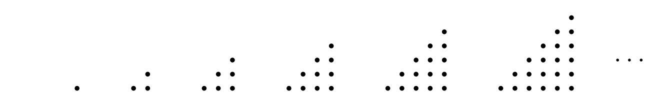
$$5, 9, 49, 161, 381, 745, 1289, 2049, 3061, 4361, 5985, 7969, \ldots.$$

As usual, the set of integers is denoted by $\mathbb{Z}$, and the set of rational numbers is denoted by $\mathbb{Q}$. The set of polynomials in x with rational coefficients is denoted by $\mathbb{Q}[x]$.

Definition 5.1. A sequence $s(n)_{n \geq 0}$ of rational numbers is a *polynomial sequence* if there exists a polynomial $f(x) \in \mathbb{Q}[x]$ such that $s(n) = f(n)$ for all $n \geq 0$.

Although we're primarily interested in sequences of integers in this book, we've allowed sequences of rational numbers in Definition 5.1 for an important reason that will become clear in Chapter 7. We could have allowed even more general sequences; if you want to consider a sequence of real numbers generated by a polynomial, go right ahead! The theorems in this book are stated for sequences of rational numbers, but most require little or no modification to apply to sequences of algebraic numbers, computable numbers, real numbers, or complex numbers.

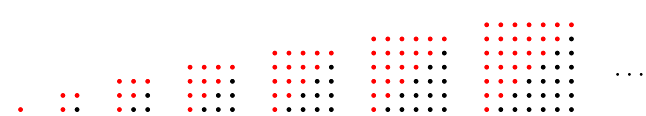
Example 5.2. How many points are in each of these triangular diagrams?



The sequence

$$0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, 120, \ldots \qquad \text{(A000217)}$$

counting these points is called the sequence of *triangular numbers* and denoted by $T(n)_{n \geq 0}$. (The 0 points corresponding to the initial term aren't visible in the diagram, but they're there!) The area of a triangle is $\frac{1}{2}bh$, but the number of points isn't exactly equal to the area. However, we can count the points in the same way we derive $\frac{1}{2}bh$, namely by forming a rectangle with a copy of the triangle:

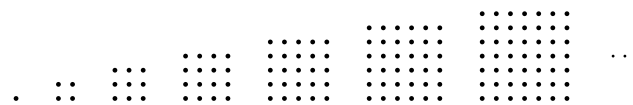


Now we see that the number of points in the original triangle is $T(n) = \frac{n(n+1)}{2}$. In particular, the sequence of triangular numbers is a polynomial sequence. Note that the polynomial $\frac{n(n+1)}{2} = \frac{1}{2}n^2 + \frac{1}{2}n$ has non-integer coefficients but evaluates to an integer for every $n \geq 0$; it may not have been obvious that this can happen.

Triangular numbers are examples of *figurate numbers* — numbers that count points in polygonal regions. Another famous sequence of figurate numbers is the sequence

$$0, 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169, 196, 225, \dots \qquad \text{(A000290)}.$$

of squares, which count points in these diagrams:



There are also figurate numbers in higher dimensions. For example, $(n^3)_{n\geq 0}$ counts points arranged in cubes, and so on.

One nice property of polynomials is that adding or multiplying polynomials gives another polynomial. These properties are passed on to polynomial sequences.

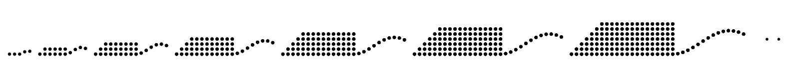
Theorem 5.3. *If $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ are polynomial sequences, then $(s(n) + t(n))_{n\geq 0}$ and $(s(n)t(n))_{n\geq 0}$ are polynomial sequences.*

PROOF. Let $f(x) \in \mathbb{Q}[x]$ such that $s(n) = f(n)$ for all $n \geq 0$. Let $g(x) \in \mathbb{Q}[x]$ such that $t(n) = g(n)$ for all $n \geq 0$. Then $f(x) + g(x)$ and $f(x)g(x)$ are polynomials in $\mathbb{Q}[x]$ with the property that $s(n) + t(n) = f(n) + g(n)$ and $s(n)t(n) = f(n)g(n)$ for all $n \geq 0$. $\square$

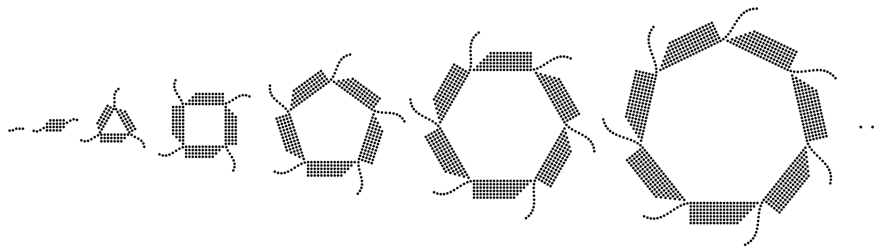
We summarize Theorem 5.3 by saying that polynomial sequences are *closed* under addition and multiplication. That is, performing these operations on polynomial sequences results in other polynomial sequences. Closure properties such as these are useful because they immediately imply that certain sequences built from other sequences belong to the same class.

In the case of addition and multiplication, closure properties are especially useful because addition and multiplication have combinatorial interpretations. Let A and B be finite sets. If A and B are disjoint, then $|A \cup B| = |A| + |B|$, so the sum of two numbers can be interpreted as the size of the union of two disjoint sets. Similarly, the $|A \times B| = |A| \cdot |B|$. (Recall that the Cartesian product $A \times B$ consists of all pairs (a, b) where $a \in A$ and $b \in B$.) Therefore the product of two numbers can be interpreted as the size of the Cartesian product of two sets. We proved a version of this in Theorem 2.2.

Example 5.4. For each $n \geq 0$, let's consider a triangular array of points with width n and height n. Next to it, place a $2n \times n$ rectangular array and attach a tail of length $2n$ to form a mouse:



Now arrange n copies of the nth mouse in a cyclic mouse train:



We just invented a sequence — the sequence of *cyclic mouse train numbers* — which counts the points in cycle mouse trains:

$$0, 5, 30, 90, 200, 375, 630, 980, 1440, 2025, 2750, 3630, \ldots.$$

From the closure properties of polynomial sequences, it follows immediately that this sequence is a polynomial sequence.

Sums and products

Let's introduce some additional notation and definitions that will be used throughout the book. If a and b are integers such that $a \leq b$, we denote

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b)$$

and

$$\prod_{i=a}^{b} f(i) = f(a)f(a+1)\cdots f(b).$$

The symbols $\sum$ and $\prod$ are the Greek letters sigma and pi (for "sum" and "product"). For example,

$$\sum_{i=0}^{3} i^2 = 0^2 + 1^2 + 2^2 + 3^2 = 14.$$

More generally, for a nonempty set $A = \{a_1, \ldots, a_n\}$ we will write $\sum_{a \in A} f(a) = f(a_1) + \cdots + f(a_n)$ for the sum over all elements in A.

Let A and B be nonempty, finite, disjoint sets. Then

$$\sum_{i \in A \cup B} f(i) = \sum_{a \in A} f(a) + \sum_{b \in B} f(b).$$

This identity suggests a natural way to interpret an empty sum $\sum_{i \in \{\}} f(i)$. Namely, we would like the identity to hold even if A or B is empty. For example, if $A = \{\}$, then the identity becomes

$$\sum_{b \in B} f(b) = \sum_{a \in \{\}} f(a) + \sum_{b \in B} f(b),$$

which implies

$$0 = \sum_{a \in \{\}} f(a).$$

Therefore, for the identity to hold when A is empty, an empty sum must be 0. Since this is quite useful, we define empty sums to be 0. For example, $\sum_{i \in \{\}} i^2 = 0$.

Similarly, if $b = a - 1$, we define $\sum_{i=a}^{b} f(i)$ to be 0. For example, $\sum_{i=2}^{1} i^2 = 0$. The definition of an empty sum has an implication for polynomials. Since we will write a general polynomial as $\sum_{i=0}^{r-1} c_i x^i$, if $r = 0$ then the polynomial $\sum_{i=0}^{r-1} c_i x^i$ is the zero polynomial 0.

While we're on the subject of 0... In calculus you were probably told that 0^0 cannot be defined, since x^y approaches different values as $x \to 0$ and $y \to 0$ depending on how x and y approach 0. However, when an exponent only takes integer values, we should always, absolutely, without a doubt, define $0^0 = 1$. The combinatorial reason for this is Theorem 2.2. The symbolic reason is that evaluating a polynomial, written as a sum, at 0 requires $0^0 = 1$. For example, write $f(x) = 6x^3 - 2x + 5 = \sum_{i=0}^{3} c_i x^i$ where $(c_0, c_1, c_2, c_3) = (5, -2, 0, 6)$. The constant term is $f(0) = 5$. This is equivalent to $\sum_{i=0}^{3} c_i 0^i = 5$, which simplifies to $5 \cdot 0^0 = 5$, implying $0^0 = 1$.

When working with sums and products, it is often useful to re-index. For example, we may want to reverse the order of $\sum_{i=2}^{13} (10 - i)^2$. The systematic way to think about this is as a change of variables $i = -j$. First replace 13 with $i = 13$ in the upper limit: $\sum_{i=2}^{i=13} (10 - i)^2$. Then replace every appearance of i with $-j$: $\sum_{-j=2}^{-j=13} (10 + j)^2$. Now clean this up: $\sum_{j=-2}^{j=-13} (10 + j)^2$. Finally, fix the order of summation: $\sum_{j=-13}^{-2} (10 + j)^2$. We have successfully reversed the order. Suppose we want the indices to start at 0 rather than -13. Another change of variables $j = k - 13$ produces $\sum_{k-13=-13}^{k-13=-2} (10 + k - 13)^2 = \sum_{k=0}^{11} (k - 3)^2$.

Ranking sequences

If we want to measure how complicated a polynomial sequence is, it makes sense to look at its underlying polynomial. A common measure of the complexity of a polynomial is the degree. If $c_{r-1} \neq 0$, the *degree* of the polynomial $f(x) = c_{r-1} x^{r-1} + \cdots + c_1 x + c_0$ is $r - 1$ and is denoted by $\deg f(x)$. But in some ways, the degree of a polynomial isn't an ideal measure. For one thing, what is the degree of the polynomial 0? For our purposes, the following measure is more natural (and avoids the silliness of defining $\deg 0 = -\infty$).

Definition 5.5. Let $s(n)_{n \geq 0}$ be a polynomial sequence, and let $f(x)$ be the polynomial such that $s(n) = f(n)$ for all $n \geq 0$. The *rank* of $s(n)_{n \geq 0}$ is

$$\begin{cases} 0 & \text{if } f(x) \text{ is the zero polynomial} \\ 1 + \deg f(x) & \text{otherwise.} \end{cases}$$

We denote it by $\operatorname{rank}(s)$.

The rank is a non-negative integer. Intuitively, the rank is the number of pieces of information needed to specify the polynomial that generates $s(n)_{n \geq 0}$. For example, the rank of $(6n^3 - 2n + 5)_{n \geq 0}$ is 4 because the tuple $(5, -2, 0, 6)$ of coefficients has length 4. This same sequence can also be specified by the 7-tuple $(5, -2, 0, 6, 0, 0, 0)$, but three 0s can be removed to obtain a shorter tuple.

If $s(x)$ is a polynomial that is not the zero polynomial 0, then the degree d of $s(x)$ and the rank r of $s(n)_{n \geq 0}$ are related by $r = 1 + d$.

Example 5.6. What is the rank of the zero sequence $0, 0, 0, \ldots$? By Definition 5.5, its rank is 0. This agrees with the intuitive interpretation of the rank as the length of a coefficient tuple, since the tuple () has length 0 and corresponds to the polynomial formed by the empty sum $\sum_{i=0}^{-1} a_i x^i = 0$. In fact the zero sequence is the only polynomial sequence with rank 0, since the empty sum is always 0. We define $\deg 0 = -1$ so that the equation $\operatorname{rank}(s) = 1 + \deg s(x)$ holds for all polynomials $s(x)$ and so that we can always write $s(x) = \sum_{i=0}^{\deg s(x)} c_i x^i$.

Given a polynomial sequence, there is a unique polynomial of minimum rank which generates it. It will be convenient to refer to this polynomial as *the* polynomial for $s(n)_{n\geq 0}$.

How does the rank behave under addition and multiplication? We can refine Theorem 5.3 as follows.

Theorem 5.7. *If $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ are polynomial sequences, then $(s(n)+t(n))_{n\geq 0}$ is a polynomial sequence with rank at most $\max(\operatorname{rank}(s), \operatorname{rank}(t))$, and $(s(n)t(n))_{n\geq 0}$ is a polynomial sequence with rank equal to*

$$\begin{cases} 0 & \text{if } s(n)_{n\geq 0} \text{ or } t(n)_{n\geq 0} \text{ is the zero sequence} \\ \operatorname{rank}(s) + \operatorname{rank}(t) - 1 & \text{otherwise.} \end{cases}$$

Proof. Let $\sigma = \operatorname{rank}(s)$ and $\tau = \operatorname{rank}(t)$. Let $f(x) = \sum_{i=0}^{\sigma-1} b_i x^i \in \mathbb{Q}[x]$ and $g(x) = \sum_{i=0}^{\tau-1} c_i x^i \in \mathbb{Q}[x]$ such that $s(n) = f(n)$ and $t(n) = g(n)$ for all $n \geq 0$. If $f(x) + g(x)$ is the zero polynomial, then the rank of $(s(n)+t(n))_{n\geq 0}$ is at most $\max(\sigma, \tau)$ since $\max(\sigma, \tau) \geq 0$. Otherwise, since $\deg(f(x)+g(x)) \leq \max(\sigma-1, \tau-1)$, the rank of $(s(n)+t(n))_{n\geq 0}$ is at most $\max(\sigma-1, \tau-1)+1 = \max(\sigma, \tau)$.

If $\sigma = 0$ or $\tau = 0$, then $(s(n)t(n))_{n\geq 0}$ is the zero sequence and has rank 0. Otherwise $\deg(f(x)g(x)) = (\sigma-1)+(\tau-1) = \sigma+\tau-2$, so the rank of $(s(n)t(n))_{n\geq 0}$ is $\sigma + \tau - 1$. □

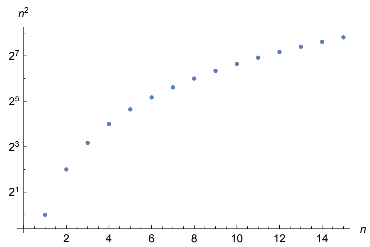
For example, consider the sequences $(n^3)_{n\geq 0}$ and $(n^4)_{n\geq 0}$. The sequence $(n^3)_{n\geq 0}$ has rank 4, and the sequence $(n^4)_{n\geq 0}$ has rank 5. Their product is $(n^7)_{n\geq 0}$, which has rank $4 + 5 - 1 = 8$.

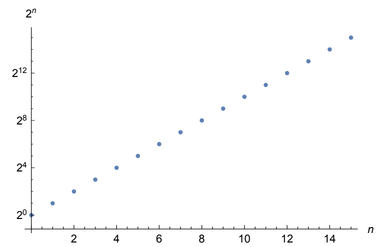
We will have a notion of rank for each class of sequences we consider in this book. Why call it the "rank"? In addition to being a measure of complexity or size, the rank of a sequence will turn out to coincide with the ranks of certain matrices associated with the sequence, as we will see starting in Chapter 15.

Log plots

If you've computed the first few terms of a sequence, it can be quite informative to plot them. By doing this you primarily get an idea of the growth rate, and this can tell you a lot about the sequence. Of course, with only finitely many initial terms you can't be sure that the behavior won't qualitatively change later on, but often the asymptotic behavior of a sequence is reflected early.

Logarithmic plots are particularly useful for studying combinatorial sequences. In a log plot, you're effectively plotting $\log s(n)$ rather than $s(n)$. Equivalently, you're using a logarithmic scale for the vertical axis. For example, here are log plots of n^2 and 2^n (with vertical axes marked with a logarithmic scale):





The height of n^2 in the first plot is $\log(n^2) = 2\log n$, whereas the height of 2^n in the second plot is $\log(2^n) = (\log 2)n$. These plots illustrate some general facts. The sequence $(n^2)_{n\geq 0}$ is a polynomial sequence, whereas $(2^n)_{n\geq 0}$ is not. Sequences that grow exponentially appear as lines in log plots, since $\log(ba^n) = (\log a)n + \log b$ is a line with slope $\log a$. On the other hand, sequences that grow sub-exponentially, including all polynomial sequences, appear as sub-linear in log plots.

Showing that a sequence grows too quickly is an easy way to show that it's not a polynomial sequence, and looking at its log plot is an easy way to guess that it grows too quickly. If you find yourself without a computer, you can substitute a log plot by counting digits in the first few terms of the sequence. Since $\log_{10}(10^\ell) = \ell$, the logarithm of an integer is proportional to its number of digits. If the number of digits in the nth term grows linearly, that means the sequence grows exponentially.

Questions

Computations.

(1) What is the rank of the sequence $(n^2 - (n+1)(n-1))_{n\geq 0}$?

(2) If $s(x) = 7x^6 + \cdots$ is a polynomial with degree 6 and $t(x) = -3x^4 + \cdots$ is a polynomial with degree 4, what is the rank of

$$\left(\left(s(n)^2 t(n) + n^3 s(n) + n^{10} t(n)\right)^3\right)_{n\geq 0}?$$

(3) What is the rank of the sequence of cyclic mouse train numbers?

(4) Let $s(n)_{n\geq 0}$ be the rank-2 sequence $3, 5, 7, \ldots$, and let $t(n)_{n\geq 0}$ be the rank-2 sequence $4, 7, 10, \ldots$.
   (a) Give formulas for $s(n)$ and $t(n)$.
   (b) Compute the rank of $(s(n) + t(n))_{n\geq 0}$ in two different ways — using closure properties and using the formulas for $s(n)$ and $t(n)$.
   (c) Compute the rank of $(s(n)t(n))_{n\geq 0}$ in two different ways — using closure properties and using the formulas for $s(n)$ and $t(n)$.

(5) Let A_n be the set of odd integers between 10 and $6n + 34$. Let B_n be the set of integers between 10 and $6n + 34$ that are not divisible by 3.
   (a) Give formulas for $|A_n|$ and $|B_n|$.
   (b) Give a formula for $|A_n \cap B_n|$.
   (c) Use the formulas for $|A_n|$, $|B_n|$, and $|A_n \cap B_n|$ to obtain a formula for $|A_n \cup B_n|$.
   (d) How many points (x, y, z) have coordinates with $x \in A_n$, $y \in B_n$, and $z \in A_n \cap B_n$?

Experiments.

(6) Plot each sequence on a log plot, using at least 10 or 20 terms so you start to see the long-term behavior. Does it seem to be a polynomial sequence?
  (a) the Fibonacci sequence
  (b) the Catalan sequence
  (c) the sequence of prime numbers

(7) The holiday of n-*Hanukah* lasts for n nights. On the first night, 2 candles are lit; on the second night, 3 candles are lit; and so on (increasing the number of candles by 1 each night).
  (a) How many candles are lit during 8-Hanukah?
  (b) How many candles are lit during n-Hanukah?
  (c) How many candles are lit during 365-Hanukah?

(8) Let p, q, r be distinct primes.
  (a) How many divisors does $p^3q^4r^2$ have?
  (b) How many divisors does $p^iq^jr^k$ have, where i, j, k are non-negative integers?

(9) Play the following game. Initially you only have the sequences $1, 1, 1, 1, \dots$ and $0, 1, 2, 3, \dots$ at your disposal. Whenever you have two sequences, you can add them and multiply them to get new sequences. You can use the new sequences to build even more sequences, and so on. If you continue in this way, what are all the sequences can you build?

(10) (a) Are polynomial sequences closed under subtraction? If so, what is the rank of the difference of two polynomial sequences?
  (b) Are polynomial sequences closed under division? If so, what is the rank of the quotient of two polynomial sequences?

(11) Are polynomial sequences closed under the following operations? (Let $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ be polynomial sequences, and let $a, b \in \mathbb{Z}$.)
  (a) Addition by a constant: $(s(n) + a)_{n\geq 0}$.
  (b) Multiplication by a constant: $(a\, s(n))_{n\geq 0}$.
  (c) Perturbation: Change a single term to a.
  (d) Shift (dropping the first term): $s(n+1)_{n\geq 0}$.
  (e) Prepending a term: $a, s(0), s(1), \dots$.
  (f) Stutter: $s(0), s(0), s(1), s(1), \dots$.
  (g) Subsequence of periodic indexing: $s(an + b)_{n\geq 0}$.
  (h) Difference sequence: $(s(n+1) - s(n))_{n\geq 0}$.
  (i) Linear combination: $(a\, s(n) + b\, t(n))_{n\geq 0}$.
  (j) Composition: $s(t(n))_{n\geq 0}$.
  (k) Riffle: $s(0), t(0), s(1), t(1), \dots$.

(12) Do any of the operations in the previous question have combinatorial interpretations?

(13) Definition 5.1 only allows polynomials with rational coefficients. Are there integer sequences we're missing by excluding more general coefficients?

(14) Is there a polynomial with integer coefficients that produces the nth triangular number?

Proofs.

(15) We showed that the only reasonable interpretation of an empty sum is 0. What is an empty product?

(16) Prove that $(2^n)_{n\geq 0}$ is not a polynomial sequence.

(17) For the Fibonacci rabbits of Chapter 1, prove that at least half of all rabbit pairs on month n are pregnant, for all $n \geq 2$. Conclude that $F(n+1) \geq \frac{3}{2}F(n)$ for all $n \geq 2$. Use this to determine whether $F(n)_{n\geq 0}$ is a polynomial sequence.

(18) Prove that if $s(n)_{n\geq 0}$ is a polynomial sequence then there is a unique polynomial $f(x)$ such that $s(n) = f(n)$ for all $n \geq 0$. Conclude that the rank in Definition 5.5 is well defined.

CHAPTER 6

Guessing a polynomial sequence

The method of undetermined coefficients

Formulas for sequences like the triangular numbers, whose ranks are small, can usually be guessed by experimenting with the first few terms. But what if you come across a sequence whose rank is large? How can you guess the polynomial? One way is to set up a polynomial whose coefficients you don't know and then solve for the coefficients. This is known as the *method of undetermined coefficients.*

Perhaps the rank of the sequence is $\leq r$. If this is the case, then there are coefficients $c_0, \ldots, c_{r-1}$ such that $s(n) = c_0 + c_1 n + \cdots + c_{r-1} n^{r-1}$ for all $n \geq 0$. In particular, we have the following system of linear equations in the unknown coefficients c_i.

$$\begin{aligned}
s(0) &= c_0 \\
s(1) &= c_0 + c_1 + \cdots + \phantom{2^{r-2}}c_{r-2} + \phantom{2^{r-1}}c_{r-1} \\
s(2) &= c_0 + 2c_1 + \cdots + 2^{r-2}c_{r-2} + 2^{r-1}c_{r-1} \\
s(3) &= c_0 + 3c_1 + \cdots + 3^{r-2}c_{r-2} + 3^{r-1}c_{r-1} \\
&\vdots
\end{aligned}$$

A *solution* of a system of equations in r unknowns is an r-tuple $(c_0, c_1, \ldots, c_{r-1})$ of rational numbers. Recall from linear algebra the following important fact.

Theorem 6.1. *An inhomogeneous system of linear equations with rational coefficients either has no solutions, exactly* 1 *solution, or infinitely many solutions.*

Since the polynomial for $s(n)_{n\geq 0}$ is unique, we are looking for a unique solution. Taking fewer than r equations will not allow us to solve uniquely for $c_0, \ldots, c_{r-1}$ since we would have more equations than unknowns. With exactly r equations there's a chance of having a unique solution, and in fact we will see that taking r equations always produces a unique solution. We can write the first r equations as the single matrix equation

$$\begin{bmatrix} s(0) \\ s(1) \\ s(2) \\ \vdots \\ s(r-1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{r-2} & 2^{r-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & r-1 & \cdots & (r-1)^{r-2} & (r-1)^{r-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{r-1} \end{bmatrix}. \tag{6.1}$$

The $r \times r$ matrix appearing here, whose (i, j) entry is $(i-1)^{j-1}$, is an example of a *Vandermonde*[1] *matrix*. It turns out that the determinant of a Vandermonde matrix has a simple product formula. The following proposition can be proved by performing well-chosen row and column operations.

Proposition 6.2. *Let $a_0, a_1, \ldots, a_{r-1}$ be complex numbers. Let M be the $r \times r$ matrix whose (i, j) entry is $(a_{i-1})^{j-1}$. Then*

$$\det M = \prod_{i=0}^{r-1} \prod_{j=i+1}^{r-1} (a_j - a_i).$$

Consequently, we can easily tell whether or not a Vandermonde matrix is invertible, by checking whether the determinant is nonzero. In particular, the $r \times r$ Vandermonde matrix in Equation (6.1) is obtained by setting $a_i = i$; therefore $a_j - a_i \neq 0$ for each pair of indices where $i \neq j$. By Proposition 6.2, this matrix is invertible. Therefore Equation (6.1) has a unique solution $(c_0, c_1, \ldots, c_{r-1})$. We have proved the following.

Theorem 6.3. *Every finite sequence $s(0), s(1), s(2), \ldots, s(r-1)$ of length r has a unique extension to a polynomial sequence $s(n)_{n \geq 0}$ with* $\operatorname{rank}(s) \leq r$.

In other words, the first r terms of a polynomial sequence with rank at most r determine that sequence uniquely.

Example 6.4. Suppose we've computed the initial terms of a sequence $s(n)_{n \geq 0}$ to be $0, 1, 3, 6$. We can guess what sequence this is by using the method of undetermined coefficients. The first $r = 4$ terms produces the system

$$\begin{bmatrix} 0 \\ 1 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

The unique solution of this system is $(0, \frac{1}{2}, \frac{1}{2}, 0)$. This solution corresponds to the formula $\frac{1}{2}n^2 + \frac{1}{2}n$, so our guess is that $s(n) = \frac{1}{2}n^2 + \frac{1}{2}n = T(n)$. In other words, the sequence of triangular numbers $T(n)_{n \geq 0} = 0, 1, 3, 6, 10, 15, 21, 28, \ldots$ is the unique polynomial sequence with rank ≤ 4 that extends $0, 1, 3, 6$.

The method of undetermined coefficients *always produces a formula* that reproduces the given terms of a sequence, so we can always use it to guess a formula for $s(n)$. But this guess can be wrong if we don't use enough terms (or if the sequence is not a polynomial sequence).

Example 6.5. Suppose we are trying to guess a formula for the sequence $T(n)_{n \geq 0}$ of triangular numbers. We compute the first 2 terms to be $0, 1$. Using only these terms gives

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}.$$

The solution of this system is $(0, 1)$, corresponding to the guess $s(n) = n$. Indeed $(n)_{n \geq 0}$ is the rank-2 sequence that agrees with sequence of triangular numbers on the first 2 terms, but $s(n) = n$ is not the correct formula for the sequence of triangular numbers.

[1]Alexandre-Théophile Vandermonde was born in 1735 in Paris, France and died in 1796 in Paris.

To take another example, the sequence $s(n)_{n\geq 0}$ in Example 6.4 could be given not by $s(n) = \frac{1}{2}n^2 + \frac{1}{2}n$ but instead by $s(n) = \frac{1}{12}n^4 - \frac{1}{2}n^3 + \frac{17}{12}n^2$, which produces the same first 4 terms $0, 1, 3, 6$.

To get the best guess, the safest approach is to compute as many terms as is feasible and let r be the number of terms we've computed. If this value of r is larger than the rank, then the coefficients will be 0 beyond some point.

Example 6.6. Suppose we have computed the first 8 terms of $T(n)_{n\geq 0}$ to be $0, 1, 3, 6, 10, 15, 21, 28$. Solving the system of linear equations for the coefficients of a degree-7 polynomial gives $(0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0)$. This corresponds to the guessed formula $s(n) = \frac{1}{2}n^2 + \frac{1}{2}n$. Additionally, the five leading 0s indicate that there is significant redundancy in the equations; we only needed a polynomial with degree 2 (and a sequence with rank 3) to capture 8 values of the sequence. The more leading 0s there are, the more confident we can be that the polynomial we guessed doesn't just fit the terms we computed but in fact fits all terms of the sequence.

Finally, let's talk a bit about the role of intuition in guessing. For naturally occurring sequences there is some expectation that the right formula will match the sequence in complexity. Simple sequences should, in some sense, have correspondingly simple formulas. Therefore a polynomial that is obtained by guessing and that is either too simple or too complex looks suspicious.

Example 6.7. Let's guess a polynomial for the sequence of prime numbers, using the first 8 primes $2, 3, 5, 7, 11, 13, 17, 19$. The method of undetermined coefficients gives

$$-\tfrac{53}{5040}n^7 + \tfrac{91}{360}n^6 - \tfrac{431}{180}n^5 + \tfrac{815}{72}n^4 - \tfrac{20021}{720}n^3 + \tfrac{6017}{180}n^2 - \tfrac{5791}{420}n + 2.$$

Plugging $n = 0$ into this polynomial gives 2, as expected. Plugging in $n = 1$ gives 3, and so on. But does this polynomial generate the entire sequence of prime numbers? If it does, then the rational numbers appearing as coefficients are incredibly special, given that the prime numbers play a central role in mathematics. The first indication that this polynomial probably doesn't generate the primes is its length. We input 8 pieces of information and got 8 coefficients out; the information wasn't compressed, just transformed. The lack of leading 0s means that if the sequence of primes is a polynomial sequence with rank ≤ 8 then we happened to choose exactly the minimum number of terms that finds it.

Of course, other information about a sequence might tell us something. In this example, we can rule out this polynomial as a formula for the primes since the leading coefficient is negative. Alternatively, we can plug in $n = 8$ and observe that the output is not 23. The sequence of values it produces is

$$2, 3, 5, 7, 11, 13, 17, 19, -92, -769, -3231, -10129, \dots.$$

Indeed, this is not the sequence of primes.

However, this sequence nonetheless appears to be a sequence of integers, despite the complicated rational coefficients. At this point it's not clear why this should happen, but in fact extending a sequence of integers using the method of undetermined coefficients always produces an integer sequence. We will develop the tools to understand why in Chapter 10.

Rigorous guessing

When we guess the nth term of a polynomial sequence, even if there are many leading 0s, it's still a guess. Confirming the guess with a proof requires an additional step. Sometimes this step is completely separate, perhaps requiring an algebraic or combinatorial argument that eventually makes the guess obsolete for the purposes of a proof. But if we have just a little information about the sequence, then the guess can be proved correct automatically, making the guess part of the actual proof. The information we need is the rank. If we know that the rank of $s(n)_{n\geq 0}$ is r, then using the first r terms guarantees that the guess will be correct, by Theorem 6.3. How might we know the rank? This is where the bounds in Theorem 5.7 come into play.

Example 6.8. The sequence $s(n)_{n\geq 0}$ of cyclic mouse train numbers from Chapter 5 is $0, 5, 30, 90, 200, \dots$. Guessing a polynomial from the first 5 terms gives $\frac{5}{2}n^3 + \frac{5}{2}n^2$. This guess is not hard to prove by writing out the definition of $s(n)$ and doing some algebra. But we can prove it with less hassle by noting that $s(n)_{n\geq 0}$ is the sum of polynomial sequences with rank ≤ 3, times a polynomial sequence with rank 2. By Theorem 5.7, $s(n)_{n\geq 0}$ is a polynomial sequence with rank ≤ 4. Since $s(n)_{n\geq 0}$ agrees with $\left(\frac{5}{2}n^3 + \frac{5}{2}n^2\right)_{n\geq 0}$ on the first 4 terms, Theorem 6.3 guarantees that $s(n) = \frac{5}{2}n^3 + \frac{5}{2}n^2$ for all $n \geq 0$. This is a rigorous proof, carried out by simply keeping track of the rank and then testing 4 terms! Moreover, if we had first determined that $\operatorname{rank}(s) \leq 4$, before guessing a polynomial, then this would have told us 4 terms are sufficient to guess the right polynomial.

Example 6.9. Suppose you remember that the sum $1^2 + 2^2 + 3^2 + \cdots + n^2$ of the first n squares is given by a polynomial in n with degree 3, but you don't remember the polynomial. The values for $n \in \{0, 1, 2, 3\}$ are $0, 1, 5, 14$. Fitting a polynomial to these values gives $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$. Therefore $\sum_{i=1}^{n} i^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ for all $n \geq 0$.

Another way of phrasing Theorem 6.3 is that if two rank-r polynomial sequences agree on their first r terms, then they are equal. Consequently, rigorous guessing can also be used to prove identities.

Example 6.10. Suppose we want to prove

$$(n+2)^4 + (n-2)^4 - (n+1)^4 - (n-1)^4 = 6\left(6n^2 + 5\right)$$

for all $n \geq 0$. There's nothing to guess here. Each side generates a polynomial sequence with rank ≤ 5. Therefore plugging in $0, 1, 2, 3, 4$ and checking that both sides are equal for each of these 5 values is a rigorous proof of this polynomial identity.

It might seem like overkill to prove polynomial identities like this, because expanding the two sides is also an automatic way to prove them. But it's good to have options. Expanding polynomials is a symbolic computation, and evaluating polynomials is a numeric computation, so in a given situation one may be more advantageous than the other.

Questions

Computations.

(1) Check that Proposition 6.2 gives the correct value for the determinant of each matrix.

(a) $\begin{bmatrix} 1 & 5 \\ 1 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 3 & 9 \\ 1 & -4 & 16 \\ 1 & 7 & 49 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 3 & 9 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}$

Programs.

(2) Write a program that takes a list of integers and performs the method of undetermined coefficients to guess a formula for the sequence.

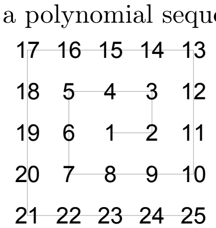
Experiments.

(3) Let $s(n)$ be the number of pairs (a, b) of integers such that $1 \leq a \leq b \leq n$. Compute the first several terms of $s(n)_{n \geq 0}$, and guess a formula for $s(n)$. Does $s(n)_{n \geq 0}$ seem to be a polynomial sequence?

(4) Let $s(n)$ be the number of 3-tuples (a, b, c) of integers such that $n \leq a \leq b \leq c \leq 2n$. Guess a formula for $s(n)$. Does $s(n)_{n \geq 0}$ seem to be a polynomial sequence?

(5) Guess a polynomial from the first several terms of each sequence. Does it seem to be a polynomial sequence?
  (a) the Fibonacci sequence
  (b) the Catalan sequence

(6) The *Ulam*[2] *spiral* is the arrangement of the positive integers in a counterclockwise spiral, with 1 at the origin, 2 at the point $(1, 0)$, 3 at the point $(1, 1)$, and so on. Does the sequence $1, 2, 11, \ldots$ of values along the nonnegative x axis seem to be a polynomial sequence?



(7) Let $s(n)$ be the number of points in the plane with integer coordinates that have distance at most n from the origin. For example, $s(1) = 5$. Does $s(n)_{n \geq 0}$ seem to be a polynomial sequence?

(8) Compute the first several powers of the matrix

$$M = \begin{bmatrix} 1 & 0 & 4 \\ -2 & 3 & 2 \\ -2 & 2 & -1 \end{bmatrix}.$$

  (a) Is there a formula for the top left entry of M^n?
  (b) Is there a formula for M^n?

(9) Compute the first several terms of the Taylor series, centered at $x = 0$, of each function. What is the coefficient of x^n in the series? (Hint: These are not polynomial sequences, but they are close.)
  (a) $-\frac{(x^2+2)(3x^2+1)}{(x-1)^5}$

---

[2]Stanisław Ulam was born in 1909 in Lemberg, Austria-Hungary (now Lviv, Ukraine) and died in 1984 in Santa Fe, New Mexico, USA.

(b) $\frac{30x^2+48x+20}{(x+1)^3}$

(c) $\frac{(x-1)(x+2)}{(x+1)^3}$

(10) A *Pythagorean*[3] *triple* is a 3-tuple (a, b, c) of positive integers such that $a^2 + b^2 = c^2$. A Pythagorean triple is *primitive* if $\gcd(a, b) = 1$. The primitive Pythagorean triples in which $c = b+1$ are listed in the following table. Extend this table by searching for additional examples. Which columns seem to be polynomial sequences?

a	b	c
3	4	5
5	12	13
7	24	25
$\vdots$	$\vdots$	$\vdots$

(11) The primitive Pythagorean triples in which $c = a + 2$ are listed in the following table. Extend this table by searching for additional examples. Which columns seem to be polynomial sequences?

a	b	c
3	4	5
15	8	17
35	12	37
$\vdots$	$\vdots$	$\vdots$

(12) Suppose you know r terms of a sequence, but they aren't necessarily the first r terms and aren't necessarily consecutive. For example, you know $s(3) = 0$, $s(7) = -8$, $s(8) = -5$, and $s(25) = 1474$. Can you still guess a polynomial? What's the rank of the resulting polynomial sequence?

(13) Suppose you know that $s(n)_{n\geq 0}$ is a polynomial sequence with rank ≤ 5. You only know 3 values; for example, $s(0) = -1$, $s(1) = -6$, and $s(2) = 5$. But you also know that the polynomial for the sequence is divisible by $x^2 + 1$. Can you still guess a polynomial?

(14) Proposition 6.2 gives a formula for the determinant of a general Vandermonde matrix. For the particular $r \times r$ Vandermonde matrix in Equation (6.1), can this formula be written more simply?

(15) How many numbers end up being multiplied together when evaluating the product formula for $\det M$ in Proposition 6.2?

(16) Let M_r be the $r\times r$ Vandermonde matrix in Equation (6.1). Does $(\det M_{r+1})_{r\geq 0}$ seem to be a polynomial sequence?

(17) Arrange the sequence $0, 1, 2, \ldots$ of non-negative integers along northwest diagonals in the first quadrant of the plane, with 0 at the origin, 1 at the point $(1, 0)$, 2 at the point $(0, 1)$, and so on.

[3]Pythagoras was born in Samos (now in Greece) around 570 BCE and probably died in Croton or Metapontum (now Crotone and Metaponto, Italy) around 495 BCE.

20					
14	19				
9	13	18			
5	8	12	17		
2	4	7	11	16	
0	1	3	6	10	15

For $x \geq 0$ and $y \geq 0$, let $s(x, y)$ be the number assigned to the point (x, y).

(a) Does $s(x, 0)_{x \geq 0}$ seem to be a polynomial sequence?
(b) Does $s(x, 1)_{x \geq 0}$ seem to be a polynomial sequence?
(c) Does $s(x, 2)_{x \geq 0}$ seem to be a polynomial sequence?
(d) The function $s(x, y)$ is known as the *Cantor*[4] *pairing function*. Guess polynomials for the coefficients of $s(x, 0), s(x, 1), s(x, 2), \ldots$, and use them to guess a general formula for $s(x, y)$.

(18) For each sequence, guess a polynomial for the first 2 terms, the first 3 terms, the first 4 terms, and so on. What happens to the coefficient of n in these polynomials as you use more and more terms?
(a) $(2^n)_{n \geq 0}$
(b) the periodic sequence $0, 1, 0, -1, 0, 1, 0, -1, \ldots$ consisting of repeated blocks of $(0, 1, 0, -1)$

Proofs.

(19) Let $T(n)$ be the nth triangular number.
(a) Give an upper bound on the rank of $(T(n + 3) - T(n))_{n \geq 0}$.
(b) Give an upper bound on the rank of $(T(n + 2) - T(n + 1))_{n \geq 0}$.
(c) Compute the first few terms of these two sequences. What is the relationship between them?

(20) Prove each identity by bounding the ranks of the sequences on each side and plugging in sufficiently many values of n.
(a) $n^2 - (n + 1)(n - 1) = 1$
(b) $(n + 2)^3 - (n - 1)^3 = 9(n^2 + n + 1)$
(c) $(n^2 - 1)^2 + (2n)^2 = (n^2 + 1)^2$

(21) Use closure properties to bound the rank of each sequence. Then rigorously guess the polynomial by computing sufficiently many terms and solving for the coefficients.
(a) $\big((n^2 - 2n - 4)(n^2 - n + 2)\big)_{n \geq 0}$
(b) $\big(T(n)^2 - T(n - 1)^2\big)_{n \geq 0}$, where $T(n)$ is the nth triangular number

(22) Prove Proposition 6.2.

[4]Georg Cantor was born in 1845 in Saint Petersburg, Russia and died in 1918 in Halle, Germany.

CHAPTER 7

The vector space of polynomial sequences

We now have two ways of representing a rank-r polynomial sequence $s(n)_{n\geq 0}$. The first is by an explicit formula, in the form of a polynomial $s(x)$ with degree $r-1$. The second is by its first r terms, which uniquely determine the sequence by Theorem 6.3. We will introduce a third representation of polynomial sequences in Chapter 10. So in this chapter we will develop the infrastructure to keep all of these representations straight and to understand the relationships between them.

To do this, we will use the terminology of vector spaces. Vector spaces are central objects in linear algebra. Since we already used linear algebra to guess a polynomial sequence in Chapter 6 by solving a system of linear equations, it is perhaps not terribly surprising that vector spaces will play an important role.

Let's recall the definition of a vector space (over $\mathbb{Q}$). For concreteness, you can think of $\mathbf{s}$, $\mathbf{t}$, and $\mathbf{u}$ as sequences $s(n)_{n\geq 0}$, $t(n)_{n\geq 0}$, and $u(n)_{n\geq 0}$. The operation $+$ is defined on sequences by

$$s(n)_{n\geq 0} + t(n)_{n\geq 0} := (s(n) + t(n))_{n\geq 0}. \tag{7.1}$$

If a is a rational number, the operation $\cdot$ is defined by

$$a \cdot (s(n)_{n\geq 0}) := (a\, s(n))_{n\geq 0}. \tag{7.2}$$

Definition 7.1. A set V with two operations $+\colon V \times V \to V$ and $\cdot\colon \mathbb{Q} \times V \to V$ is a *vector space*, and the elements of V are called *vectors*, if there exists an element $\mathbf{0} \in V$ such that, for all $\mathbf{s}, \mathbf{t}, \mathbf{u} \in V$ and for all $a, b \in \mathbb{Q}$, the operation $+$ satisfies

- $\mathbf{0} + \mathbf{s} = \mathbf{s}$,
- there exists an element $-\mathbf{s} \in V$ such that $\mathbf{s} + (-\mathbf{s}) = \mathbf{0}$,
- $(\mathbf{s} + \mathbf{t}) + \mathbf{u} = \mathbf{s} + (\mathbf{t} + \mathbf{u})$,
- $\mathbf{s} + \mathbf{t} = \mathbf{t} + \mathbf{s}$,

the operation $\cdot$ satisfies

- $1 \cdot \mathbf{s} = \mathbf{s}$,
- $(ab) \cdot \mathbf{s} = a \cdot (b \cdot \mathbf{s})$,

and the two operations $+$ and $\cdot$ together satisfy

- $a \cdot (\mathbf{s} + \mathbf{t}) = a \cdot \mathbf{s} + a \cdot \mathbf{t}$,
- $(a + b) \cdot \mathbf{s} = a \cdot \mathbf{s} + b \cdot \mathbf{s}$.

There are a lot of axioms there. However, what's *not* there is multiplication of a vector by another vector. As far as a vector space is concerned, two elements of V are never multiplied together; there need not be any such multiplication defined.

Example 7.2. The most familiar vector space is the set of points $\mathbb{R}^d$ in d-dimensional Euclidean[1] space. Each point is represented by a d-tuple $(x_1, x_2, \ldots, x_d)$ consisting

[1] Euclid (pronounced 'yooklid') lived around 300 BCE in Alexandria, Greece (now in Egypt).

of its Cartesian[2] coordinates. Vector addition and scalar multiplication are defined coordinatewise:

$$(x_1, x_2, \ldots, x_d) + (y_1, y_2, \ldots, y_d) := (x_1 + y_1, x_2 + y_2, \ldots, x_d + y_d)$$
$$a \cdot (x_1, x_2, \ldots, x_d) := (ax_1, ax_2, \ldots, ax_d).$$

The tuple $(0, 0, \ldots, 0)$ is the $\mathbf{0}$ element of this vector space. Check that the axioms of Definition 7.1 are satisfied!

A *basis* of a vector space V is a finite sequence $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_d$ of elements in V such that every $\mathbf{t} \in V$ has a unique representation $\mathbf{t} = a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + \cdots + a_d\mathbf{s}_d$ where $a_1, a_2, \ldots, a_d \in \mathbb{Q}$. In particular, the basis vectors $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_d$ are linearly independent; that is, if $\mathbf{0} = a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + \cdots + a_d\mathbf{s}_d$ then $0 = a_1 = a_2 = \cdots = a_d$. If V has a basis $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_d$, then one can show that all bases of V have exactly d elements, so we say that the *dimension* of V is d. The dimension of V is the answer to an enumeration question — what is the largest number of linearly independent vectors that can be chosen from V? The dimension is also the number of pieces of information required to specify an element of V, given a particular basis.

Example 7.3. For d-dimensional Euclidean space, the *standard basis* consists of the vectors $(1, 0, 0, \ldots, 0)$, $(0, 1, 0, \ldots, 0)$, $\ldots$, $(0, 0, 0, \ldots, 1)$. These vectors are linearly independent, and every vector can be written as a linear combination of them.

Throughout the book we will make use of various vector spaces of sequences. The first of these is the following.

Notation. We denote the set of all polynomial sequences (of rational numbers) with rank $\leq r$ by $\mathsf{Poly}(r)$.

For example, $\mathsf{Poly}(1)$ contains the sequences $(0)_{n\geq 0}$, $(1)_{n\geq 0}$, $(2)_{n\geq 0}$, $(-1)_{n\geq 0}$, $(\frac{4}{3})_{n\geq 0}$, and so on. The set $\mathsf{Poly}(2)$ consists of all sequences of the form $(an + b)_{n\geq 0}$ where $a, b \in \mathbb{Q}$.

The operations $+$ and $\cdot$ for sequences are defined as in Equations (7.1) and (7.2).

Theorem 7.4. *For every $r \geq 0$, $\mathsf{Poly}(r)$ is a vector space.*

Proof. We show that $\mathsf{Poly}(r)$ satisfies the conditions in Definition 7.1. The sum of two elements of $\mathsf{Poly}(r)$ has rank $\leq r$, as established in Theorem 5.7. If $a \in \mathbb{Q}$ and $s(n)_{n\geq 0} \in \mathsf{Poly}(r)$ then $(a\, s(n))_{n\geq 0}$ has rank $\leq r$. Therefore $+$ and $\cdot$ map into $\mathsf{Poly}(r)$. The zero sequence $(0)_{n\geq 0}$ is an element of $\mathsf{Poly}(r)$ and satisfies the conditions on the element $\mathbf{0}$. The remaining vector space axioms are true for all sequences $\mathbf{s}, \mathbf{t}, \mathbf{u}$, not just sequences in $\mathsf{Poly}(r)$, and follow from basic properties of addition and multiplication of rational numbers. $\square$

Theorem 7.4 is the main reason we included sequences of rational numbers in Definition 5.1. In order to have a vector space, we need $(a\, s(n))_{n\geq 0} \in \mathsf{Poly}(r)$ for every rational number a.

Definition 7.5. Let V be a vector space. A set $W \subseteq V$ is a *subspace* of V if W is itself a vector space.

[2]René Descartes (pronounced 'daykart') was born in 1596 in La Haye en Touraine (which was renamed Descartes in 1967), France and died in 1650 in Stockholm, Sweden.

Example 7.6. The vector space $\mathsf{Poly}(1)$ is a subspace of $\mathsf{Poly}(2)$.

Is there a natural basis of $\mathsf{Poly}(r)$? Yes, in fact there are several. Perhaps the first that comes to mind is the *monomial basis*

$$\left((1)_{n\geq 0}, (n)_{n\geq 0}, (n^2)_{n\geq 0}, \dots, (n^{r-1})_{n\geq 0}\right).$$

Since the monomial basis of $\mathsf{Poly}(r)$ has size r, we have the following refinement of Theorem 7.4.

Theorem 7.7. *For every $r \geq 0$, $\mathsf{Poly}(r)$ is a vector space with dimension r.*

Proof. It suffices to show that the monomial basis is in fact a basis. Every sequence in $\mathsf{Poly}(r)$ is a linear combination of the sequences $(1)_{n\geq 0}$, $(n)_{n\geq 0}$, $(n^2)_{n\geq 0}$, $\dots$, $(n^{r-1})_{n\geq 0}$. Namely,

$$\left(c_0 + c_1 n + \cdots + c_{r-1} n^{r-1}\right)_{n\geq 0} = c_0\,(1)_{n\geq 0} + c_1\,(n)_{n\geq 0} + \cdots + c_{r-1}\,(n^{r-1})_{n\geq 0}.$$

To show that the sequences $(1)_{n\geq 0}$, $(n)_{n\geq 0}$, $(n^2)_{n\geq 0}$, $\dots$, $(n^{r-1})_{n\geq 0}$ are linearly independent, suppose that

$$c_0\,(1)_{n\geq 0} + c_1\,(n)_{n\geq 0} + \cdots + c_{r-1}\,(n^{r-1})_{n\geq 0}$$

is the zero sequence. Then its first r terms are 0, and Equation (6.1) becomes

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r-1 & \cdots & (r-1)^{r-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{r-1} \end{bmatrix}.$$

As in the proof of Theorem 6.3, this Vandermonde matrix is invertible by Proposition 6.2, so the only solution is $c_0 = c_1 = \cdots = c_{r-1} = 0$. $\square$

The coordinates of a sequence in the monomial basis are simply the coefficients of its polynomial formula.

Example 7.8. Let $T(n)_{n\geq 0}$ be the sequence of triangular numbers, and let $r = 5$. What are the coordinates of $T(n)_{n\geq 0}$ in the monomial basis of $\mathsf{Poly}(5)$? The nth triangular number can be written

$$T(n) = \tfrac{n(n+1)}{2} = \tfrac{1}{2}n^2 + \tfrac{1}{2}n,$$

so $T(n)_{n\geq 0} = \frac{1}{2}(n^2)_{n\geq 0} + \frac{1}{2}(n)_{n\geq 0}$. That is, the coordinates of $T(n)_{n\geq 0}$ in the monomial basis are $(0, \frac{1}{2}, \frac{1}{2}, 0, 0)$.

The Lagrange basis

As we have seen, the coefficients of a polynomial $s(x)$ give the coordinates of $s(n)_{n\geq 0}$ in the monomial basis. Theorem 6.3 implies that we can also specify a polynomial sequence $s(n)_{n\geq 0}$ by its first r terms $(s(0), s(1), \dots, s(r-1))$; we refer to the basis in which these are the coordinates of $s(n)_{n\geq 0}$ as the *Lagrange*[3] *basis*. For example, the coordinates of $(n^3)_{n\geq 0} \in \mathsf{Poly}(4)$ in the Lagrange basis are $(0, 1, 8, 27)$.

What do the basis elements look like? By definition, the coordinates of a Lagrange basis element in the Lagrange basis are of the form $(0, \dots, 0, 1, 0, \dots, 0)$.

---

[3] Joseph-Louis Lagrange was born in 1736 in Turin, Kingdom of Sardinia (now in Italy) and died in 1813 in Paris, France.

There's nothing special about the Lagrange basis in this respect; coordinates of monomial basis elements in the monomial basis also look like $(0, \dots, 0, 1, 0, \dots, 0)$, as do standard Euclidean basis vectors in the standard basis of Euclidean space. Since the coordinates of a sequence in the Lagrange basis tell us its first r terms, the basis elements are as follows. Fix $r \geq 0$. For each i in the range $0 \leq i \leq r-1$, let $e_i(n)_{n\geq 0} \in \mathsf{Poly}(r)$ be the sequence whose first r terms $0, \dots, 0, 1, 0, \dots, 0$ satisfy

$$e_i(n) = \begin{cases} 1 & \text{if } n = i \\ 0 & \text{if } 0 \leq n \leq r-1 \text{ and } n \neq i. \end{cases} \tag{7.3}$$

The first r terms of $e_i(n)_{n\geq 0}$ determine the entire sequence, by Theorem 6.3. Note that $e_i(n)_{n\geq 0}$ depends on r, even though our notation doesn't reflect this. For example, if $r = 1$, then the sequence $e_0(n)_{n\geq 0}$ is the constant sequence $1, 1, 1, \dots$, so $e_0(n) = 1$. However, if $r = 2$, then $e_0(n)_{n\geq 0}$ is the unique sequence in $\mathsf{Poly}(2)$ beginning with $1, 0$; therefore $e_0(n) = 1 - n$.

Definition 7.9. Let $r \geq 0$. The *Lagrange basis* of $\mathsf{Poly}(r)$ is

$$(e_0(n)_{n\geq 0}, e_1(n)_{n\geq 0}, \dots, e_{r-1}(n)_{n\geq 0}).$$

Example 7.10. Again let $T(n)_{n\geq 0}$ be the sequence of triangular numbers, and let $r = 5$. The first 5 terms of $T(n)_{n\geq 0}$ are $0, 1, 3, 6, 10$. The definition of $e_i(n)$ implies that

$$T(n) = 0e_0(n) + 1e_1(n) + 3e_2(n) + 6e_3(n) + 10e_4(n)$$

for each $n \in \{0, 1, 2, 3, 4\}$, since each $e_i(n)$ was defined to be 1 at $n = i$ and 0 elsewhere in the first r terms. For example,

$$T(3) = 6 = 0 + 0 + 0 + 6 \cdot 1 + 0.$$

By Theorem 6.3, this identity holds not just for $n \in \{0, 1, 2, 3, 4\}$ but for all $n \geq 0$. Therefore the coordinates of $T(n)_{n\geq 0}$ in the Lagrange basis of $\mathsf{Poly}(5)$ are $(0, 1, 3, 6, 10)$.

We can perform vector space operations on sequences using either the monomial basis or the Lagrange basis.

Example 7.11. Let $s(n) = 5n^2 - 2n + 2$ and $t(n) = 2n^2 + 4n$. The sequences $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ are elements of $\mathsf{Poly}(3)$. The linear combination $s(n)_{n\geq 0} - 2\,t(n)_{n\geq 0}$, in the monomial basis, has coordinates

$$(2, -2, 5) - 2\,(0, 4, 2) = (2, -10, 1),$$

which corresponds to the polynomial $n^2 - 10n + 2$. In the Lagrange basis, the coordinates of $(n^2 - 10n + 2)_{n\geq 0}$ are $(2, -7, -14)$, its first 3 terms. Alternatively, we could have computed these coordinates by first writing $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ in the Lagrange basis as $(2, 5, 18)$ and $(0, 6, 16)$. Then

$$(2, 5, 18) - 2\,(0, 6, 16) = (2, -7, -14).$$

Given two bases of a vector space, one can convert the coordinates of a sequence in one basis to its coordinates in the other basis by multiplying by an invertible *change-of-basis matrix*. In fact, looking back at Equation (6.1) we see that the Vandermonde matrix is the change-of-basis matrix for converting from the monomial basis to the Lagrange basis. This is most transparent for a sequence such as

$(n^3)_{n\geq 0} \in \mathsf{Poly}(5)$, for which we have

$$\begin{bmatrix} 0 \\ 1 \\ 8 \\ 27 \\ 64 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 3 & 9 & 27 & 81 \\ 1 & 4 & 16 & 64 & 256 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

The vector on the left contains the coordinates of $(n^3)_{n\geq 0}$ in the Lagrange basis, and the vector on the right contains the coordinates of $(n^3)_{n\geq 0}$ in the monomial basis. The matrix multiplication clearly picks out the appropriate column of the Vandermonde matrix. Multiplication by the inverse of the Vandermonde matrix goes the other direction; it converts coordinates in the Lagrange basis to coordinates in the monomial basis.

Lagrange interpolation

An element $(n^i)_{n\geq 0}$ of the monomial basis has coordinates $(0^i, 1^i, \dots, (r-1)^i)$ in the Lagrange basis. But how do the elements $e_i(n)_{n\geq 0}$ of the Lagrange basis look in the monomial basis?

Example 7.12. Consider $e_3(n)_{n\geq 0} \in \mathsf{Poly}(5)$. Its coordinates in the Lagrange basis are $(0,0,0,1,0)$. What are its coordinates in the monomial basis? In other words, what are the coefficients in $e_3(n) = c_0 + c_1 n + c_2 n^2 + c_3 n^3 + c_4 n^4$? Solving the system

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 3 & 9 & 27 & 81 \\ 1 & 4 & 16 & 64 & 256 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

gives the coordinates in the monomial basis, namely

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 0 \\ 8 \\ -14 \\ 7 \\ -1 \end{bmatrix}.$$

In other words, we have $e_3(n) = \frac{1}{6}(-n^4 + 7n^3 - 14n^2 + 8n)$, and indeed the sequence $\left(\frac{1}{6}(-n^4 + 7n^3 - 14n^2 + 8n)\right)_{n\geq 0}$ has the expected initial terms:

$$0, 0, 0, 1, 0, -10, -40, -105, -224, -420, -720, -1155, \dots.$$

The simple structure of the polynomial $\frac{1}{6}(-n^4 + 7n^3 - 14n^2 + 8n)$ in the previous example is not apparent by looking at its coefficients, but it turns out to factor nicely:

$$-\tfrac{1}{6} n(n-1)(n-2)(n-4).$$

This is because the factorization of a polynomial is tightly entwined with the values for which it evaluates to 0. Let $f(x) \in \mathbb{Q}[x]$. If $f(a) = 0$ for some integer (or rational number) a, then dividing $f(x)$ by $x - a$ leaves remainder 0. That is, $f(x)$ factors as $(x-a)g(x)$ for some new $g(x) \in \mathbb{Q}[x]$. In Example 7.12, the initial terms $0, 0, 0, 1, 0$ determine a polynomial sequence with rank ≤ 5. The four 0 terms at $n = 0$, $n = 1$, $n = 2$, and $n = 4$ imply that $f(x) = x(x-1)(x-2)(x-4)g(x)$ for some $g(x)$.

Since the rank is at most 5, this implies that $g(x) = c$ is a constant. Further, the nonzero term $s(3) = 1$ determines c; namely, $1 = s(3) = 3(3-1)(3-2)(3-4)c$, so $c = -\frac{1}{6}$.

The same argument works more generally.

Theorem 7.13. *Let $r \geq 0$ and $0 \leq i \leq r-1$. Let $e_i(n)_{n\geq 0} \in \mathsf{Poly}(r)$ be the sequence defined by Equation (7.3). For all $n \geq 0$,*

$$e_i(n) = \frac{n-0}{i-0} \cdot \frac{n-1}{i-1} \cdots \frac{n-(i-1)}{i-(i-1)} \cdot \frac{n-(i+1)}{i-(i+1)} \cdots \frac{n-(r-1)}{i-(r-1)}. \tag{7.4}$$

Proof. Let $e_i(x)$ be the polynomial for $e_i(n)_{n\geq 0}$. Since $e_i(n) = 0$ for each n in the range $0 \leq n \leq r-1$ for which $n \neq i$, by polynomial division we have

$$e_i(x) = (x-0)\cdot(x-1)\cdots(x-(i-1))\cdot(x-(i+1))\cdots(x-(r-1))\cdot g(x)$$

where the term $x - i$ is absent from the product and where $g(x) \in \mathbb{Q}[x]$ is some polynomial. Since $\operatorname{rank}(e_i) \leq r$, we have $g(x) = c \in \mathbb{Q}$. Letting $x = i$ gives

$$1 = (i-0)\cdot(i-1)\cdots(i-(i-1))\cdot(i-(i+1))\cdots(i-(r-1))\cdot c,$$

which allows us to solve for c and leads to Equation (7.4). $\square$

The beauty of Equation (7.4) is that for $n = i$ each numerator cancels with a denominator, and for all other n in the range $0 \leq n \leq r-1$ one numerator is 0. Therefore

$$e_i(n) = \begin{cases} 1 & \text{if } n = i \\ 0 & \text{if } 0 \leq n \leq r-1 \text{ and } n \neq i \end{cases}$$

as desired. Expanding the polynomial in Equation (7.4) gives the coordinates of $e_i(n)_{n\geq 0}$ in the monomial basis.

Now that we know the explicit polynomial for $e_i(n)_{n\geq 0}$, we have a new method for constructing the polynomial for a sequence in $\mathsf{Poly}(r)$, given its first few terms. The method of undetermined coefficients in Chapter 6 used the monomial basis. The method using Equation (7.4) is known as *Lagrange interpolation*.

Example 7.14. Let's compute the polynomial $s(n)$ for the sequence in $0, 1, 3, 6, \ldots \in \mathsf{Poly}(4)$. Lagrange interpolation (using the Lagrange basis) gives

$$\begin{aligned}
s(n) &= s(0)e_0(n) + s(1)e_1(n) + s(2)e_2(n) + s(3)e_3(n) \\
&= 0 \cdot \frac{n-1}{0-1} \cdot \frac{n-2}{0-2} \cdot \frac{n-3}{0-3} + 1 \cdot \frac{n-0}{1-0} \cdot \frac{n-2}{1-2} \cdot \frac{n-3}{1-3} \\
&\qquad + 3 \cdot \frac{n-0}{2-0} \cdot \frac{n-1}{2-1} \cdot \frac{n-3}{2-3} + 6 \cdot \frac{n-0}{3-0} \cdot \frac{n-1}{3-1} \cdot \frac{n-2}{3-2} \\
&= 0 + 1 \cdot \frac{n(n-2)(n-3)}{2} + 3 \cdot \frac{n(n-1)(n-3)}{-2} + 6 \cdot \frac{n(n-1)(n-2)}{6} \\
&= \tfrac{1}{2}n^2 + \tfrac{1}{2}n.
\end{aligned}$$

Questions

Computations.

(1) Find the coordinates of each sequence in the Lagrange basis.
   (a) $(n^2)_{n\geq 0} \in \mathsf{Poly}(3)$
   (b) $(-n^3 + n^2 - n + 1)_{n\geq 0} \in \mathsf{Poly}(5)$

(c) $(n(n-1)(n-3)(n-4))_{n\geq 0} \in \mathsf{Poly}(6)$
(d) $(0)_{n\geq 0} \in \mathsf{Poly}(r)$

(2) The following sequences are specified by their coordinates in the Lagrange basis. Convert them to their coordinates in the monomial basis.
(a) $(3, 8)$
(b) $(6, 6, 6)$
(c) $(1, 0, 1, 0)$
(d) $(0, 0, 4, 18, 48)$
(e) $(3, 15, 67, 213, 531)$

(3) Use Lagrange interpolation to compute the polynomial for each sequence.
(a) $16, 25, \ldots \in \mathsf{Poly}(2)$
(b) $2, 6, 9, \ldots \in \mathsf{Poly}(3)$
(c) $0, 54, 0, 0, 0, \ldots \in \mathsf{Poly}(5)$
(d) $4, 3, -1, 4, 3, -1, \ldots \in \mathsf{Poly}(6)$

Programs.

(4) Write a program that takes a list of integers and uses Lagrange interpolation to guess a formula for the sequence. How does its speed compare to the method of undetermined coefficients?

Experiments.

(5) What sequences are contained in $\mathsf{Poly}(0)$?

(6) Theorem 7.4 states that $\mathsf{Poly}(r)$ is a vector space. Is the set of polynomial sequences with rank exactly r also a vector space?

(7) Are there any ways in which the method of undetermined coefficients is better than Lagrange interpolation, or vice versa?

(8) There are other bases of $\mathsf{Poly}(r)$ as well. For example,

$$B = \big((1)_{n\geq 0}, (n+1)_{n\geq 0}, (n^2+n)_{n\geq 0}\big)$$

is a basis of $\mathsf{Poly}(3)$.
(a) What are the coordinates of the sequence $(n^2)_{n\geq 0}$ in the basis B?
(b) What is the change-of-basis matrix from the monomial basis of $\mathsf{Poly}(3)$ to B?
(c) What is the change-of-basis matrix from B to the Lagrange basis of $\mathsf{Poly}(3)$?

Proofs.

(9) Write out the details of the proof of Theorem 7.4.

(10) (a) Prove that the polynomials $f(x) \in \mathbb{Q}[x]$ with degree at most 4, along with the zero polynomial, form a vector space.
(b) Is the set of all polynomials a vector space?
(c) Is the set $\{ax^2 + b : a, b \in \mathbb{Q}\}$ a vector space?

(11) Use the Lagrange basis to give a proof of Theorem 7.7 that doesn't involve Vandermonde matrices. Then use Theorem 7.7 to give a proof of Theorem 6.3 that doesn't involve Vandermonde matrices.

CHAPTER 8

Permutations and subsets

In this chapter we pause our study of polynomial sequences to introduce some fundamental combinatorial objects.

Permutations

In words, letters can be repeated. Words are good models for phone numbers (made up of digits), text (made up of letters, punctuation, spaces, and line breaks), melodies (made up of notes), passwords, license plates, and so on.

But in other settings, you don't want repetitions. For example, if you're seating people in a row at a theater, you don't have duplicate people. If you're making a playlist, you may not want repeat songs. And if you're running errands, you don't want repeat destinations. These situations are better modeled by *permutations*.

Definition 8.1. Let Σ be a set. A *permutation* on Σ is word on Σ in which no letter appears more than once.

When Σ is finite, the longest permutations on Σ have length $|\Sigma|$. For example, here are the length-3 permutations on $\{1, 2, 3\}$:

$$123, \quad 132, \quad 213, \quad 231, \quad 312, \quad 321.$$

As usual when introducing a combinatorial object, the first question we will ask is an enumeration question. How many length-n permutations on n letters are there? There are n choices for the first letter, $n-1$ choices for the second letter (since the first letter cannot be used again), $n-2$ choices for the third, and so on. Therefore there are $n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$ length-n permutations on n letters. This product arises frequently enough that it has a standard notation,

$$n! := n(n-1)(n-2)\cdots 3\cdot 2\cdot 1,$$

and is read "n factorial". For $n = 3$, the number of permutations is $3! = 3\cdot 2\cdot 1 = 6$. For $n = 0$, there is exactly 1 permutation on 0 letters, namely the empty word. The sequence $(n!)_{n\geq 0}$ of factorials is

$$1, 1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, 39916800, \ldots \qquad \text{(A000142)}.$$

The nth term can be obtained by multiplying the previous term by n.

A similar argument allows us to count length-m permutations.

Theorem 8.2. *Let $0 \leq m \leq n$. The number of length-m permutations on a set of n elements is $\frac{n!}{(n-m)!}$.*

Proof. There are n choices for the first letter, $n-1$ choices for the second letter, $\ldots$, and $n-(m-1)$ choices for the mth letter. Therefore the number of length-m

permutations is

$$n(n-1)(n-2)\cdots(n-(m-2))(n-(m-1)) = \frac{n!}{(n-m)!}. \qquad \square$$

We'll often speak of "the permutations on Σ". When the length is not specified, this means the permutations with length $|\Sigma|$. This convention is useful is because permutations (with length $|\Sigma|$) can be interpreted as functions. For example, each permutation *abc* on $\{1, 2, 3\}$ can be interpreted as encoding the image of each letter, namely $1 \mapsto a, 2 \mapsto b, 3 \mapsto c$:

$$\begin{array}{cccccc} 1\,2\,3 & 1\,2\,3 & 1\,2\,3 & 1\,2\,3 & 1\,2\,3 & 1\,2\,3 \\ \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow \\ 1\,2\,3 & 1\,3\,2 & 2\,1\,3 & 2\,3\,1 & 3\,1\,2 & 3\,2\,1 \end{array}$$

Formally, permutations on Σ are in bijection with bijections from Σ to itself. Since we know how many permutations there are, we have also enumerated such bijections.

Corollary 8.3. *Let Σ be a set of size n. The number of bijections $f\colon \Sigma \to \Sigma$ is $n!$.*

Another consequence of this bijection is that every permutation has an inverse permutation. Namely, if π is a permutation and $f\colon \Sigma \to \Sigma$ is its corresponding bijection, then f has an inverse bijection f^{-1}, and f^{-1} corresponds to some permutation π^{-1}. The inverse permutation can be obtained by reversing the arrows in the previous table:

$$\begin{array}{ll} (123)^{-1} = 123 & \qquad (231)^{-1} = 312 \\ (132)^{-1} = 132 & \qquad (312)^{-1} = 231 \\ (213)^{-1} = 213 & \qquad (321)^{-1} = 321. \end{array}$$

Moreover, the map $\pi \mapsto \pi^{-1}$ is *itself* a bijection on the set of permutations. (There are lots of bijections!)

Subsets

A different combinatorial object models situations like the following. You have a group of people (say, the members of a team) and you need to form a 4-person committee. Or you have many shirts but can only fit 6 in your suitcase. Or you can choose 3 side dishes to go with your main course. How many ways are there to do each of these?

Like a permutation, repetitions in these situations are not possible (unless maybe you *really* like collard greens). Unlike a permutation, the order of the selections doesn't matter. These objects are *subsets*.

Example 8.4. Let $\Sigma = \{1, 2, 3\}$. There are 8 subsets of Σ:

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

Theorem 8.5. *If $|\Sigma| = n$, then the number of subsets of Σ is 2^n.*

The proof uses the following function.

Notation. Let A be a set, and let $a \in A$. For each word $w \in A^*$, define $\mathrm{pos}_a(w)$ to be the set consisting of the positions in w where the letter a occurs.

For example, $\mathrm{pos}_Y(NYNNY) = \{2, 5\}$.

Proof of Theorem 8.5. Without loss of generality, let $\Sigma = \{1, 2, \dots, n\}$. Consider the alphabet $\{N, Y\}$. We claim that pos_Y is a bijection from the set of words $\{N, Y\}^n$ to the set of all subsets of Σ. The result will then follow from Corollary 2.9, which implies that $|\{N, Y\}^n| = 2^n$.

First we show surjectivity. Given a subset $S \subseteq \Sigma$, we exhibit a word w such that $\mathrm{pos}_Y(w) = S$. Let $w = w_1 w_2 \cdots w_n \in \{N, Y\}^n$ be the word whose ith letter is

$$w_i = \begin{cases} Y & \text{if } i \in S \\ N & \text{if } i \notin S. \end{cases}$$

Then $\mathrm{pos}_Y(w) = S$, so pos_Y is surjective.

Next we show injectivity. Let v and w be length-n words on $\{N, Y\}$ such that $\mathrm{pos}_Y(v) = \mathrm{pos}_Y(w)$. Then v and w have Ys in the same positions. Since there are only two letters, v and w also have Ns in the same positions. Therefore $v = w$. $\square$

The appearance of the number 2 in Theorem 8.5 is made clear by the bijection pos_Y in the proof. Here is the correspondence for $n = 3$:

$$\begin{array}{cccccccc} NNN & NNY & NYN & NYY & YNN & YNY & YYN & YYY \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ \{\} & \{3\} & \{2\} & \{2,3\} & \{1\} & \{1,3\} & \{1,2\} & \{1,2,3\} \end{array}$$

We'll also be interesting in refining Theorem 8.5 to count subsets of Σ with exactly m elements.

Notation. Let $n \geq 0$ and $m \geq 0$. The number of subsets of $\{1, 2, \dots, n\}$ of size m is denoted by $\binom{n}{m}$ and is read "n choose m".

Example 8.6. Here are the subsets of $\{1, 2, 3, 4\}$ of size 2:

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

Therefore $\binom{4}{2} = 6$. What formula gives this value? We can enumerate subsets by first counting permutations and then disregarding the order of elements. Namely, the permutations on $\{1, 2, 3, 4\}$ of length 2 are

$$12, 13, 14, 21, 23, 24, 31, 32, 34, 41, 42, 43.$$

There are $\frac{4!}{2!} = 12$ of them, by Theorem 8.2. But if we don't care about the order of letters, then they collapse in pairs: $\{1, 2\} = \{2, 1\}$, $\{1, 3\} = \{3, 1\}$, and so on. Therefore the number of subsets of $\{1, 2, 3, 4\}$ of size 2 is $\frac{12}{2} = 6$.

Theorem 8.7. *Let $0 \leq m \leq n$. Then $\binom{n}{m} = \frac{n!}{m!(n-m)!}$.*

Proof. The number of permutations on $\{1, 2, \dots, n\}$ of length m is $\frac{n!}{(n-m)!}$ by Theorem 8.2. Each of these permutations can be permuted into $m!$ permutations on the same set of letters. Therefore, since we don't care about the order of letters, there are $\frac{n!}{m!(n-m)!}$ distinct subsets. $\square$

For $m > n$, we have $\binom{n}{m} = 0$ since there are no subsets of $\{1, 2, \dots, n\}$ of size m. Since $\binom{n}{m}$ has two parameters n and m, we can consider the 2-dimensional sequence $\binom{n}{m}_{n \geq 0, m \geq 0}$, as in the following grid. The nonzero region is *Pascal's*[1]

[1]Blaise Pascal was born in 1623 in Clermont-Ferrand, France and died in 1662 in Paris, France.

triangle, although it had been studied at least a thousand years earlier.

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

The entries of this 2-dimensional sequence count subsets, but they also seem to be the coefficients in expansions of $(x+y)^n$:

$$\begin{aligned}
(x+y)^0 &= 1\\
(x+y)^1 &= x+y\\
(x+y)^2 &= x^2+2xy+y^2\\
(x+y)^3 &= x^3+3x^2y+3xy^2+y^3\\
(x+y)^4 &= x^4+4x^3y+6x^2y^2+4xy^3+y^4.
\end{aligned}$$

Theorem 8.8. *The coefficient of $x^{n-m}y^m$ in $(x+y)^n$ is $\binom{n}{m}$. Consequently,*

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

Example 8.9. We will give a proof shortly, but to get a sense of why Theorem 8.8 is true, consider

$$(x+y)^3 = (x+y)(x+y)(x+y).$$

If we expand this product without using $xy = yx$, we obtain

$$\begin{aligned}
(x+y)^3 &= (x+y)^2(x+y)\\
&= (xx+xy+yx+yy)(x+y)\\
&= xxx+xxy+xyx+xyy+yxx+yxy+yyx+yyy.
\end{aligned}$$

These 8 monomials are the 8 words of length 3 on the alphabet $\{x,y\}$. When we let x and y commute again, the words that become $x^{3-m}y^m$ are the words with exactly m instances of y (and $3-m$ instances of x):

$$(x+y)^3 = \begin{bmatrix}\text{\# words with}\\ 3\,x\text{s and }0\,y\text{s}\end{bmatrix}x^3+\begin{bmatrix}\text{\# words with}\\ 2\,x\text{s and }1\,y\end{bmatrix}x^2y+\begin{bmatrix}\text{\# words with}\\ 1\,x\text{ and }2\,y\text{s}\end{bmatrix}xy^2+\begin{bmatrix}\text{\# words with}\\ 0\,x\text{s and }3\,y\text{s}\end{bmatrix}y^3.$$

However, we would like to interpret the coefficients as counting subsets, not words. The bijection pos_y gives us this interpretation. The correspondence

$$\begin{array}{cccccccc}
xxx & xxy & xyx & xyy & yxx & yxy & yyx & yyy\\
\updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow\\
\{\} & \{3\} & \{2\} & \{2,3\} & \{1\} & \{1,3\} & \{1,2\} & \{1,2,3\}
\end{array}$$

under pos_y allows us to convert from words to subsets, since words with exactly m instances of y correspond to subsets with exactly m elements. The number of such subsets is $\binom{3}{m}$, so

$$(x+y)^3 = \binom{3}{0}x^3 + \binom{3}{1}x^2y + \binom{3}{2}xy^2 + \binom{3}{3}y^3.$$

Proof of Theorem 8.8. Without using $xy = yx$, the expansion of $(x+y)^n = (x+y)(x+y)\cdots(x+y)$ contains every word in $\{x,y\}^n$ precisely once, since the ith letter of a word in this expansion is the letter (x or y) contributed to that word by the ith factor $x+y$. Therefore

$$(x+y)^n = \sum_{w \in \{x,y\}^n} w.$$

Now allow $xy = yx$; words with exactly m instances of y become $x^{n-m}y^m$. For each $m \in \{0,1,\dots,n\}$, the coefficient of $x^{n-m}y^m$ in $(x+y)^n$ is the number of words in $\{x,y\}^n$ with exactly m instances of y. As proved in the proof of Theorem 8.5, the map pos_y is a bijection from $\{x,y\}^n$ to the set of subsets of $\{1,2,\dots,n\}$. Under this bijection, words with exactly m instances of y correspond to subsets of size m, of which there are $\binom{n}{m}$. Therefore the coefficient of $x^{n-m}y^m$ in $(x+y)^n$ is $\binom{n}{m}$. $\square$

In light of Theorem 8.8, we refer to $\binom{n}{m}$ as a *binomial coefficient*.

Questions

Experiments.

(1) Does $(n!)_{n\geq 0}$ seem to be a polynomial sequence?
(2) Does $\binom{2n}{n}_{n\geq 0}$ seem to be a polynomial sequence?
(3) What is the determinant of the 0×0 empty matrix?
(4) How many points with integer coordinates have distance 1 from the origin in d-dimensional Euclidean space?
(5) Consider an n-sided polygon with all diagonals drawn in. How many diagonals are there?
(6) Suppose n people attend a conference and all shake hands with each other.
 (a) What combinatorial object models a handshake?
 (b) How many handshakes occur?
(7) Let w be a length-n word on the alphabet $\{0,1\}$. How many ways are there to insert 3 hyphens into w (resulting in a word of length $n+3$) such that each hyphen is immediately preceded by an element of $\{0,1\}$ and immediately followed by an element of $\{0,1\}$?
(8) How many subsets of $\{1,2,\dots,n\}$ contain no pairs of consecutive numbers?
(9) How many 2-element sets $\{A,B\}$ are there, where A and B are subsets of $\{1,2,\dots,n\}$ and $A\cap B = \{\}$?
(10) Let $f(x)$ and $g(x)$ be differentiable functions. What is the formula for the nth derivative $\frac{d^n}{dx^n}(f(x)g(x))$ in terms of derivatives of $f(x)$ and $g(x)$?
(11) Can the formula $\binom{n}{m} = 0$ for $0 \leq n < m$ be seen to agree with the formula $\binom{n}{m} = \frac{n!}{m!(n-m)!}$? (We haven't defined $(n-m)!$ when $n-m<0$.)
(12) A permutation $w_1w_2\cdots w_n$ on $\{1,2,\dots,n\}$ *contains* 123 if there exist indices i,j,k satisfying $1 \leq i < j < k \leq n$ and $w_i < w_j < w_k$. That is, the letters w_i, w_j, w_k are in the same relative order as $1,2,3$. Let $s(n)$ be the

number of permutations on $\{1, 2, \ldots, n\}$ that do not contain 123. Does $s(n)_{n\geq 0}$ seem to be a polynomial sequence?

(13) A permutation $w_1 w_2 \cdots w_n$ on $\{1, 2, \ldots, n\}$ *contains* 231 if there exist indices i, j, k satisfying $1 \leq i < j < k \leq n$ and $w_k < w_i < w_j$. Let $s(n)$ be the number of permutations on $\{1, 2, \ldots, n\}$ that do not contain 123 and do not contain 231. Does $s(n)_{n\geq 0}$ seem to be a polynomial sequence?

Proofs.

(14) Prove that your answer to "Does $(n!)_{n\geq 0}$ seem to be a polynomial sequence?" in Question (1) is correct.

(15) (a) Give an alternate proof of Theorem 8.5 by setting $x = 1$ and $y = 1$ in Theorem 8.8.
(b) What is the sum of the nth row of Pascal's triangle?

(16) Prove that each row of Pascal's triangle is left–right symmetric in three ways.
(a) Use Theorem 8.7.
(b) Use a bijection on the set $\{N, Y\}^n$.
(c) Use a bijection on the set of subsets of $\{1, 2, \ldots, n\}$.

Experiences.

(17) Listen to minimalist composer Tom Johnson's 1986 piece *The Chord Catalogue*, which is subtitled "all the 8178 chords possible in one octave".

CHAPTER 9

The ubiquity of binomial coefficients

Binomial coefficients play a central role in mathematics. In this chapter we'll explore several of their combinatorial interpretations and start fitting them into our study of polynomial sequences.

Column sequences

One way to get insight into a 2-dimensional sequence such as $\binom{n}{m}_{n\geq 0, m\geq 0}$ is to look at 1-dimensional cross sections. The row sequences in Pascal's triangle aren't very promising, since each row is eventually 0. So let's look at column sequences. For $m = 0$ we obtain the constant sequence $1, 1, 1, \ldots$. (What's the combinatorial reason for this?) The sequence $\binom{n}{1}_{n\geq 0}$ is evidently $0, 1, 2, 3, \ldots$, whose nth term is simply n. The sequence $\binom{n}{2}_{n\geq 0}$ is

$$0, 0, 1, 3, 6, 10, 15, 21, \ldots,$$

which appears to be the sequence of triangular numbers, with an extra 0 on the front. If this is true, then the nth term is $\binom{n}{2} = T(n-1) = \frac{1}{2}n^2 - \frac{1}{2}n$. In fact this is not hard to prove using Theorem 8.7, since $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)(n-2)\cdots 2\cdot 1}{2(n-2)\cdots 2\cdot 1} = \frac{n(n-1)}{2}$ after canceling the common factors of the factorials. Similarly, Theorem 8.7 shows that the first several column sequences are all polynomial sequences:

$$\begin{aligned}
\binom{n}{0} &= \frac{n!}{0!(n-0)!} = 1 \\
\binom{n}{1} &= \frac{n!}{1!(n-1)!} = n \\
\binom{n}{2} &= \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2} \\
\binom{n}{3} &= \frac{n!}{3!(n-3)!} = \frac{n(n-1)(n-2)}{6} \\
\binom{n}{4} &= \frac{n!}{4!(n-4)!} = \frac{n(n-1)(n-2)(n-3)}{24}.
\end{aligned}$$

Since all but m factors of $n!$ are canceled by the denominator, this trend continues.

Proposition 9.1. *Let $m \geq 0$. The sequence $\binom{n}{m}_{n\geq 0}$ is a polynomial sequence with rank $m+1$.*

These polynomials allow us to extend the definition of $\binom{n}{m}$ in a natural way. Let $m \geq 0$. For any complex number x, define $\binom{x}{m} := \frac{x(x-1)(x-2)\cdots(x-(m-1))}{m!}$.

In particular, we can extend Pascal's triangle up the page by evaluating $\binom{n}{m}$ at negative integers n:

	$m = 0$	1	2	3	4	5	6	7
-5	1	-5	15	-35	70	-126	210	-330
-4	1	-4	10	-20	35	-56	84	-120
-3	1	-3	6	-10	15	-21	28	-36
-2	1	-2	3	-4	5	-6	7	-8
-1	1	-1	1	-1	1	-1	1	-1
$n = 0$	1	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0
2	1	2	1	0	0	0	0	0
3	1	3	3	1	0	0	0	0
4	1	4	6	4	1	0	0	0
5	1	5	10	10	5	1	0	0

Each column is still given by a polynomial.

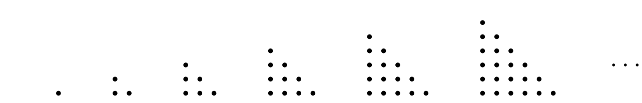
For combinatorial reasons, we also define $\binom{n}{m} = 0$ if $n \geq 0$ and $m \leq -1$. (There are $\binom{5}{-3} = 0$ ways to choose -3 objects from a set of 5.) We leave $\binom{n}{m}$ undefined if $n \leq -1$ and $m \leq -1$.

Simplices and the Pascal relation

The column sequences suggest another combinatorial interpretation of $\binom{n}{m}$. The sequence $\binom{n}{0}_{n \geq 0}$ trivially counts points in points:

The sequence $\binom{n}{1}_{n \geq 0}$ counts points in line segments, which are stacks of points:

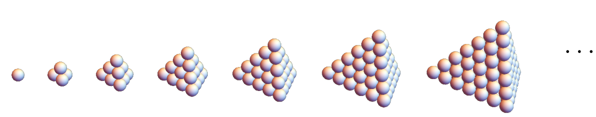
The sequence $\binom{n}{2}_{n \geq 0}$ counts points in triangles, which are stacks of line segments:



The sequence $\binom{n}{3}_{n \geq 0}$,

$$0, 0, 0, 1, 4, 10, 20, 35, 56, 84, 120, 165, \ldots,$$

is the sequence of *tetrahedral numbers* (A000292), with two extra 0s on the front. Tetrahedral numbers count points in tetrahedral diagrams obtained by stacking triangles:



What do we get when we stack tetrahedra? We've run out of dimensions to visualize them, but we get *pentatopes* in 4-dimensional space. In general d-dimensional space, the analogue of a triangle is called a *d-simplex*. Pentatopes are 4-simplices, tetrahedra are 3-simplices, triangles are 2-simplices, line segments are 1-simplices, and points are 0-simplices. A d-simplex is obtained by stacking $(d-1)$-simplices. To see why points in d-simplices are counted by binomial coefficients, we will need following recurrence, known as the *Pascal relation*.

Theorem 9.2 (Pascal relation)**.** *For all $n \geq 0$ and $m \geq 0$,*

$$\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}.$$

Proof. We give a combinatorial proof. By definition, $\binom{n+1}{m+1}$ is the number of subsets of $\{1, 2, \ldots, n+1\}$ of size $m+1$. These subsets come in two kinds — those that contain $n+1$ and those that do not. The number of subsets of $\{1, 2, \ldots, n+1\}$ of size $m+1$ that contain $n+1$ is $\binom{n}{m}$, since removing $n+1$ gives a subset of $\{1, 2, \ldots, n\}$ of size m. The number of subsets of $\{1, 2, \ldots, n+1\}$ of size $m+1$ that do not contain $n+1$ is $\binom{n}{m+1}$. □

The Pascal relation allows you to compute the nth row of Pascal's triangle from the previous row. The leftmost entry is 1, and each remaining entry is the sum of the entries above it and to its upper left.

Corollary 9.3. *For all $n \geq -1$ and $m \geq 0$,*

$$\sum_{i=0}^{n} \binom{i}{m} = \binom{n+1}{m+1}.$$

Proof. We use induction on n. For $n = -1$, both sides are 0. Inductively,

$$\begin{aligned} \sum_{i=0}^{n} \binom{i}{m} &= \binom{n}{m} + \sum_{i=0}^{n-1} \binom{i}{m} \\ &= \binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}. \end{aligned}$$ □

Now we can show that binomial coefficients count points in simplices. When $m = 1$, Corollary 9.3 simplifies to $\sum_{i=0}^{n} i = T(n)$, which is essentially the original definition of the nth triangular number. For each $n \geq 0$ and $d \geq 1$, let $s(n, d)$ be the number of points in the d-simplex constructed by stacking the first $n+1$ nonempty $(d-1)$-simplices. Let $s(n, 0)$ be the number of points in the nth 0-simplex. Then

$$s(n,d) = \begin{cases} 1 & \text{if } d = 0 \\ \sum_{i=0}^{n} s(i, d-1) & \text{if } d \geq 1. \end{cases} \tag{9.1}$$

The table of values of $s(n, d)$ appears to be a vertically-sheared Pascal's triangle, where the dth column is shifted up by d entries:

	$d=0$	1	2	3	4
$n=0$	1	1	1	1	1
1	1	2	3	4	5
2	1	3	6	10	15
3	1	4	10	20	35
4	1	5	15	35	70

Therefore we conjecture $s(n, d) = \binom{n+d}{d}$. Indeed, we can see this by rewriting

$$\sum_{i=0}^{n} s(i, d-1) = \sum_{i=0}^{n-1} s(i, d-1) + s(n, d-1) = s(n-1, d) + s(n, d-1).$$

Equation (9.1) now becomes

$$s(n,d) = \begin{cases} 1 & \text{if } n = 0 \text{ or } d = 0 \\ s(n-1,d) + s(n,d-1) & \text{if } n \geq 1 \text{ and } d \geq 1. \end{cases}$$

In particular, if $n \geq 1$ and $d \geq 1$ then $s(n,d)$ is the sum of the entries above it and to its left. Since the Pascal relation for binomial coefficients is similar, and since binomial coefficients satisfy $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$, the recurrence for $s(n,d)$ emulates the recurrence for binomial coefficients, and it follows that $s(n,d) = \binom{n+d}{d}$ for all $n \geq 0$ and $d \geq 0$.

By observing that points in simplices can be indexed by tuples (namely, their coordinates in d-dimensional space), we also have the following.

Theorem 9.4. *Let $n \geq 0$ and $d \geq 0$. The number of d-tuples $(a_1, a_2, \ldots, a_d)$ of non-negative integers such that $a_1 + a_2 + \cdots + a_d \leq n$ is $\binom{n+d}{d}$.*

For example, there are $\binom{2+2}{2} = 6$ different 2-tuples with sum at most 2, namely

$$(0,0), \quad (0,1), \quad (0,2), \quad (1,0), \quad (1,1), \quad (2,0).$$

Proof. It suffices to show that the d-tuples $(a_1, a_2, \ldots, a_d)$ satisfying $a_1 + a_2 + \cdots + a_d \leq n$ are the coordinates of points of a d-simplex obtained by stacking the first $n+1$ nonempty $(d-1)$-simplices. We use an induction on d. For $d = 0$, the only tuple is the empty tuple (), and these are the coordinates of the only point in 0-dimensional space. For $d \geq 1$, the tuples satisfying $a_1 + a_2 + \cdots + a_d \leq n$ can be partitioned according to their last entries:

$$\begin{aligned} a_d = 0: &\qquad a_1 + a_2 + \cdots + a_{d-1} \leq n, \\ a_d = 1: &\qquad a_1 + a_2 + \cdots + a_{d-1} \leq n-1, \\ \vdots\quad &\qquad\qquad\qquad \vdots \\ a_d = n: &\qquad a_1 + a_2 + \cdots + a_{d-1} \leq 0. \end{aligned}$$

By the inductive assumption, each of these is a $(d-1)$-simplex. In fact they are the first $n+1$ nonempty $(d-1)$-simplices, so stacking them produces the desired d-simplex. $\square$

Multisets and lattice paths

In Chapter 2 we introduced words, in which the order of letters matters. Words and permutations differ in that words allow repeated letters but permutations do not.

Analogously, we can consider a version of subsets where repeated elements are allowed, but order still doesn't matter. These are called *multisets*. For example, here are the multisets of size 2 on $\{1,2,3\}$:

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

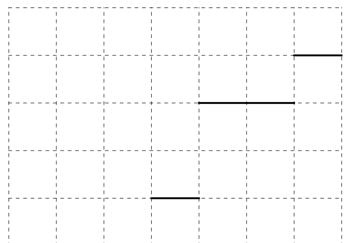
Multisets are often more natural than sets. For example, the zeros of a polynomial $f(x)$ should be counted with multiplicity. The zeros of $(x-1)^3$ make up the multiset $\{1,1,1\}$. This way, the number of zeros of $f(x)$ in the complex numbers equals the degree of $f(x)$.

Multisets complete the following table of basic combinatorial objects.

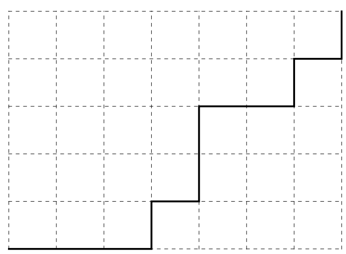
	ordered	unordered
repeats allowed	word/tuple	multiset
repeats disallowed	permutation	set

How many multisets of size m can be formed from a set of n distinct objects? To enumerate them, we'll convert multisets into lattice paths. For this purpose, it is convenient to first consider multisets on $\{0, 1, \ldots, n-1\}$ rather than $\{1, 2, \ldots, n\}$.

Example 9.5. Let $n = 6$, and consider the multiset $\{0, 0, 0, 1, 3, 3, 4\}$ of size $m = 7$ on $\{0, 1, 2, 3, 4, 5\}$. Since reordering the elements in a multiset doesn't change the multiset, we can assume they are sorted in non-decreasing order, as we have written them. Then we place a horizontal bar at height a for each element a in the multiset:



We can produce a *lattice path* from $(0, 0)$ to $(m, n - 1) = (7, 5)$ by connecting the bars with vertical lines:



(We ended this path at $(7, 5)$ rather than $(7, 4)$ for uniformity; multisets that contain the element 5 need that extra height.)

Definition 9.6. A *northeast lattice path* is a sequence of north and east steps, where each step has length 1.

Theorem 9.7. *Multisets of size m on $\{0, 1, \ldots, n - 1\}$ are in bijection with northeast lattice paths from $(0, 0)$ to $(m, n - 1)$.*

To enumerate multisets, it therefore suffices to enumerate northeast lattice paths. We use another bijection for this.

Example 9.8. We can represent the path from Example 9.5 as a word on $\{N, E\}$ specifying the sequence of steps it takes from $(0, 0)$ to $(m, n-1)$, where N represents a north step and E represents an east step. For the multiset $\{0, 0, 0, 1, 3, 3, 4\}$, the corresponding word is $EEENENNEENEN$.

Theorem 9.9. *Northeast lattice paths from $(0, 0)$ to $(m, n-1)$ are in bijection with words in $\{N, E\}^{m+n-1}$ containing exactly m instances of the letter E.*

Proof. The word representing a northeast lattice path from $(0, 0)$ to $(m, n - 1)$ is a word on $\{N, E\}$ of length $m + n - 1$, since the path consists of m east steps and $n - 1$ north steps. Moreover, every word containing m instances of the letter E and

$n-1$ instances of the letter N corresponds to a unique northeast lattice path from $(0,0)$ to $(m, n-1)$. □

Corollary 9.10. *The number of multisets of size m on $\{1,2,\dots,n\}$ is $\binom{m+n-1}{m}$.*

PROOF. By subtracting 1 from each element, we see that multisets of size m on $\{1,2,\dots,n\}$ are in bijection with multisets of size m on $\{0,1,\dots,n-1\}$. From there we use a chain of bijections

$$\text{multisets} \leftrightarrow \text{northeast lattice paths} \leftrightarrow \text{words} \leftrightarrow \text{subsets}$$

through these objects. Let M be the set of multisets of size m on $\{0,1,\dots,n-1\}$. Let W be the set of words in $\{N,E\}^{m+n-1}$ containing exactly m instances of the letter E. Theorems 9.7 and 9.9 imply that M is in bijection with W.

Recall the notation $\mathrm{pos}_a(w)$ from Chapter 8. Let S be the set of subsets of $\{1,2,\dots,m+n-1\}$ with exactly m elements. The function $\mathrm{pos}_E\colon W \to S$ is a bijection, since a word on $\{N,E\}$ is uniquely determined by the positions of the letter E. It follows that M is in bijection with S. By definition, $\binom{m+n-1}{m}$ is the number of elements in S. □

Using pos_N rather than pos_E for the bijection from words to subsets has the effect of interchanging the roles of the two variables m and n, and we obtain the following equivalent result.

Corollary 9.11. *The number of multisets of size $n-1$ on $\{0,1,\dots,m\}$ is $\binom{m+n-1}{m}$.*

Questions

Computations.

(1) Illustrate the bijection in the proof of Theorem 9.2 for $n=3$ and $m=2$.

(2) For each multiset, construct the corresponding northeast lattice path, word on $\{N,E\}$, and subset of $\{1,2,\dots,m+n-1\}$.
  (a) $\{1,2,2,4,5,5,5\}$ as a multiset on $\{0,1,2,3,4,5\}$
  (b) $\{1,3,6,7\}$ as a multiset on $\{0,1,\dots,9\}$
  (c) $\{\}$ as a multiset on $\{0,1,2\}$
  (d) $\{0,0,0,1,1,1,2,2,2\}$ as a multiset on $\{0,1,2\}$

(3) Assuming the streets of Manhattan form a grid, how many minimal routes are there from the intersection of 5th Ave. & 42nd St. to the intersection of 8th Ave. & 59th St.?

(4) Consider all tuples (a,b,c) of non-negative integers satisfying $a+b+c \le 3$.
  (a) How many such tuples are there?
  (b) What shape is formed by the points with their coordinates? (A line segment? Triangle? Tetrahedron? Pentatope? Higher-dimensional simplex?)
  (c) List the tuples.
  (d) Define $f((a,b,c)) = \{a, a+b, a+b+c\}$. List $f((a,b,c))$ for each tuple satisfying $a+b+c \le 3$. What objects are these?
  (e) Draw all northeast lattice paths from $(0,0)$ to $(3,3)$.

Experiments.

(5) Let $s_m(n)$ be the coefficient of x^m in the expansion of $(1+x+x^2)^n$.

(a) Guess a formula for $s_2(n)$. Does the formula look simpler in expanded form or when factored?
(b) Guess a formula for $s_3(n)$.
(c) Guess a formula for $s_4(n)$.
(d) Guess a formula for $s_5(n)$.
(e) Make a conjecture identifying a binomial coefficient as part of a formula for $s_m(n)$.

(6) When n is negative, the value of $\binom{n}{m}$ is ± 1 times an entry in Pascal's triangle. What is the exact relationship?

(7) Let $s(n)$ be the number of words $a_0 a_1 \cdots a_{n-1}$ where each a_i is an integer satisfying $a_i \leq i$ and $0 \leq a_0 \leq a_1 \leq \cdots \leq a_{n-1}$. For example, the length-3 words are $000, 001, 002, 011, 012$. What is $s(n)$?

(8) How many multisets of size $\leq 2n$ on $\{1, 2\}$ contain at least as many 1s as 2s? For $n = 2$, there are 9, namely $\{\}$, $\{1\}$, $\{1, 1\}$, $\{1, 2\}$, $\{1, 1, 1\}$, $\{1, 1, 2\}$, $\{1, 1, 1, 1\}$, $\{1, 1, 1, 2\}$, $\{1, 1, 2, 2\}$.

(9) Let $P = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}$ be the set of the first 10 prime numbers.
(a) How many distinct integers can be obtained as a product $p_1 p_2 \cdots p_7$, where $p_i \in P$ for each $i \in \{1, 2, \dots, 7\}$?
(b) How many distinct integers can be obtained as a product $p_1 p_2 \cdots p_7$, where $p_i \in P$ for each $i \in \{1, 2, \dots, 7\}$ and the primes $p_1, p_2, \dots, p_7$ are all distinct?

(10) A university is moving their Greek Literature department to a new building. There are 5 faculty members in the department, and the new building has 7 empty offices. In how many ways can the department choose which faculty member moves into which office?

(11) Let M be an $m \times n$ matrix, all of whose entries are distinct. A *submatrix* of M is obtained by choosing a subset of rows and a subset of columns. How many square submatrices of M are there?

(12) Let p, q, r be distinct primes. How many integers can be written as $p^i q^j r^k$ where i, j, k are integers such that $i + j + k \leq n$?

(13) How many m-tuples $(a_1, a_2, \dots, a_m)$ of *positive* integers satisfy $a_1 + a_2 + \cdots + a_m \leq n$? For example, the 2-tuples whose sums are at most 4 are $(1, 1)$, $(1, 2)$, $(1, 3)$, $(2, 1)$, $(2, 2)$, $(3, 1)$.

(14) How many m-tuples $(a_1, a_2, \dots, a_m)$ with entries from $\{1, 2, \dots, n\}$ satisfy $a_1 < a_2 < \cdots < a_m$? For example, the increasing 2-tuples with entries at most 4 are $(1, 2)$, $(1, 3)$, $(1, 4)$, $(2, 3)$, $(2, 4)$, $(3, 4)$.

(15) An *integer composition* of n is a tuple $(a_1, a_2, \dots, a_m)$ of positive integers such that $a_1 + a_2 + \cdots + a_m = n$. For example, the length-2 integer compositions of 4 are $(1, 3)$, $(2, 2)$, $(3, 1)$.
(a) How many integer compositions of n have length m?
(b) How many integer compositions of n are there in total?

(16) How many m-tuples $(a_1, a_2, \dots, a_m)$ of non-negative integers satisfy $a_1 + a_2 + \cdots + a_m = n$? For example, the 2-tuples with sum 4 are $(0, 4)$, $(1, 3)$, $(2, 2)$, $(3, 1)$, $(4, 0)$.

(17) (a) For which tuples (i, j, k) does a monomial of the form $cx^i y^j z^k$ appear in the expansion of $(x + y + z)^n$?
(b) What is the coefficient of $x^i y^j z^k$ in the expansion of $(x + y + z)^n$?

(18) Is it faster to compute $\binom{n}{m}$ using the formula $\frac{n!}{m!(n-m)!}$ or recursively using the Pascal relation?

(19) As a subset of $\mathbb{R}^d$, the edges and vertices of a d-simplex can be defined as follows. Choose $d+1$ points in general position (that is, no 3 points are collinear). Connect each pair of points with a line segment.
  (a) How many edges (line segments) does a d-simplex contain?
  (b) How many triangles does a d-simplex contain?
  (c) How many m-simplices does a d-simplex contain?

(20) Construct a *d-cube* in $\mathbb{R}^d$ as follows. Consider the 2^d points $(x_1, x_2, \ldots, x_d)$ where each $x_i \in \{-1, 1\}$. For each pair of points that differ in exactly one coordinate, connect them with a line segment.
  (a) How many edges does a d-cube contain?
  (b) How many squares does a d-cube contain?
  (c) How many m-cubes does a d-cube contain?

(21) What is

$$\sum_{i_1=1}^{n} \sum_{i_2=1}^{n-i_1} \sum_{i_3=1}^{n-i_1-i_2} \cdots \sum_{i_m=1}^{n-i_1-i_2-\cdots-i_{m-1}} 1?$$

(22) Let $r \geq 0$, and let $e_i(n)$ be the polynomial that generates the ith Lagrange basis sequence $e_i(n)_{n\geq 0} \in \mathsf{Poly}(r)$ from Chapter 7. Put the monomials of $e_i(n)$ under their least common denominator $d_{r,i}$, so that $e_i(n) = \frac{1}{d_{r,i}} f_{r,i}(n)$ for some polynomial $f_{r,i}(x) \in \mathbb{Z}[x]$. What is $d_{r,i}$?

Proofs.

(23) Prove Theorem 9.2 using the formula $\binom{n}{m} = \frac{n!}{m!(n-m)!}$.

(24) Prove that the Pascal relation in Theorem 9.2 holds more generally for $n \in \mathbb{Z}$ and $m \geq 0$.

(25) Define $f((a, b, c)) = \{a, a+b, a+b+c\}$, as in Question (4).
  (a) Prove that f is a bijection from the set of 3-tuples (a, b, c) of nonnegative integers satisfying $a + b + c \leq 3$ to the set of multisets on $\{0, 1, 2, 3\}$ of size 3.
  (b) What is the inverse function f^{-1}?
  (c) Does the proof generalize to give a bijection from the set of d-tuples whose sum is at most n to some set of multisets?

CHAPTER 10

Discrete calculus

An integer sequence is a discrete function, so there doesn't seem to be much hope of applying any tools from calculus to the study of integer sequences. But in fact there are analogues of differentiation and integration in discrete mathematics whose properties directly parallel those of differentiation and integration in continuous mathematics.

The difference operator

Consider the sequence $T(n)_{n\geq 0}$ of triangular numbers:

$$0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, \dots.$$

If we take the difference of each pair of consecutive terms, we obtain

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \dots.$$

This sequence is the *difference sequence* of $T(n)_{n\geq 0}$. The operator that turns a sequence into its difference sequence is the discrete analogue of the derivative.

Definition 10.1. The *difference operator* Δ_n with respect to the symbol n is defined by $\Delta_n f(n) := f(n+1) - f(n)$. The *difference sequence* of $s(n)_{n\geq 0}$ is $(s(n+1) - s(n))_{n\geq 0}$. We denote the difference sequence of $s(n)_{n\geq 0}$ by $((\Delta s)(n))_{n\geq 0}$ (rather than "$(\Delta_n s(n))_{n\geq 0}$" to avoid expressions such as $\Delta_0 s(0)$).

For example, $\Delta_n T(n) = \frac{(n+1)(n+2)}{2} - \frac{n(n+1)}{2} = n + 1$. More generally, if $s(n)$ is a nonzero polynomial in n, then the leading term cancels:

$$\begin{aligned}\Delta_n \left(cn^{r-1} + \cdots\right) &= \left(c(n+1)^{r-1} + \cdots\right) - \left(cn^{r-1} + \cdots\right) \\ &= (c - c)n^{r-1} + \cdots.\end{aligned}$$

Therefore we have the following.

Proposition 10.2. *If $s(n)_{n\geq 0}$ is a polynomial sequence with rank $r \geq 1$, then $((\Delta s)(n))_{n\geq 0}$ is a polynomial sequence with rank at most $r - 1$.*

Just as iteratively differentiating a polynomial function eventually produces the 0 function, iteratively applying the difference operator to a polynomial sequence eventually produces the zero sequence.

Example 10.3. Let $s(n) = n^3$.

$$\begin{array}{rccccccc}
n^3: & 0, & 1, & 8, & 27, & 64, & 125, & \dots \\
\Delta_n(n^3): & & 1, & 7, & 19, & 37, & 61, & \dots \\
\Delta_n(\Delta_n(n^3)): & & & 6, & 12, & 18, & 24, & \dots \\
\Delta_n(\Delta_n(\Delta_n(n^3))): & & & & 6, & 6, & 6, & \dots \\
\Delta_n(\Delta_n(\Delta_n(\Delta_n(n^3)))): & & & & & 0, & 0, & \dots
\end{array}$$

We can compute formulas for each of these sequences directly:

$$\begin{aligned}
\Delta_n(n^3) &= (n+1)^3 - n^3 = 3n^2 + 3n + 1 \\
\Delta_n^2(n^3) &= \big(3(n+1)^2 + 3(n+1) + 1\big) - \big(3n^2 + 3n + 1\big) = 6n + 6 \\
\Delta_n^3(n^3) &= (6(n+1) + 6) - (6n + 6) = 6 \\
\Delta_n^4(n^3) &= 6 - 6 = 0.
\end{aligned}$$

The antidifference operator

Differentiation has an inverse operation — integration. A natural question is whether we can also invert the difference operator Δ_n.

Definition 10.4. A function $F(n)$ is an *antidifference* of $f(n)$ if $\Delta_n F(n) = f(n)$.

(In this chapter we'll stop using $F(n)$ to exclusively denote the nth Fibonacci number.)

If $F(n)$ is an antidifference of $f(n)$, then so is $F(n) + C$ for any constant C, since $\Delta_n C = 0$. Conversely, we have the following.

Proposition 10.5. *If $F_1(n)$ and $F_2(n)$ are both antidifferences of $f(n)$, then there exists a constant C such that $F_2(n) - F_1(n) = C$ for all $n \in \mathbb{Z}$.*

Proof. Since $F_2(n+1) - F_2(n) = f(n) = F_1(n+1) - F_1(n)$, we have $F_2(n+1) - F_1(n+1) = F_2(n) - F_1(n)$. Let $C = F_2(0) - F_1(0)$. It follows by induction that $F_2(n) - F_1(n) = C$ for all $n \geq 0$. A second induction shows that $F_2(n) - F_1(n) = C$ for all $n \leq 0$. □

Notation. If $F(n)$ is an antidifference of $f(n)$, we write $\Sigma_n f(n) = F(n) + C$.

By definition, $\Sigma_n(\Delta_n F(n)) = F(n) + C$, which is the discrete analogue of $\int F'(x)\,dx = F(x) + C$. For example, we computed $\Delta_n T(n) = n + 1$ above. Therefore $\Sigma_n(n+1) = T(n) + C$.

The reason for the notation Σ_n is that antidifferences are closely related to sums. Indeed, we have seen (from Corollary 9.3, if not before) that $\sum_{i=0}^{n-1}(i+1) = \sum_{i=1}^{n} i = T(n)$. Just as $F(x) := \int_a^x f(t)\,dt$ is an antiderivative of $f(x)$, we can construct an antidifference by forming a sum.

Theorem 10.6. *Let $a \in \mathbb{Z}$. The function $F(n) := \sum_{i=a}^{n-1} f(i)$ is an antidifference of $f(n)$, provided that $n \geq a$.*

Proof. We have

$$\Delta_n\left(\sum_{i=a}^{n-1} f(i)\right) = \sum_{i=a}^{n} f(i) - \sum_{i=a}^{n-1} f(i) = f(n). \qquad \square$$

In light of Theorem 10.6, we refer to Σ_n as the *indefinite summation* operator.

Since two antidifferences of $f(n)$ differ by a constant, every antidifference of $f(n)$ can be obtained by adding a constant C to the definite sum $\sum_{i=0}^{n-1} f(i)$. (More generally, if $f(x)$ is a function of a real or complex variable, then every antidifference of $f(x)$ can be obtained by adding a periodic function $C(x)$ satisfying $C(x+1) = C(x)$.)

Example 10.7. What are the antidifferences of n? By Theorem 10.6 with $a = 0$, one antidifference of n is $\sum_{i=0}^{n-1} i$. We happen to know a formula for this sum, namely

$$\sum_{i=0}^{n-1} i = T(n-1) = \frac{(n-1)n}{2}.$$

If we choose a different value, say $a = 10$, then the antidifference

$$\sum_{i=10}^{n-1} i$$

we obtain is only defined for $n \geq 10$. However, if we write

$$\sum_{i=10}^{n-1} i = \sum_{i=0}^{n-1} i - \sum_{i=0}^{9} i = T(n-1) - T(9)$$

then this gives an interpretation of this antidifference for $n \leq 9$. This antidifference differs from the previous antidifference $T(n-1)$ by a constant.

In the proof of Theorem 10.6 we saw that the difference of a sum is essentially the summand. The sum of a difference also exhibits cancellation, analogous to $\int_a^b F'(t)\,dt = F(b) - F(a)$.

Theorem 10.8. *Let $F(n)$ be an antidifference of $f(n)$, and let $a, b \in \mathbb{Z}$ such that $a \leq b$. Then*

$$\sum_{i=a}^{b-1} f(i) = F(b) - F(a).$$

Note the upper limit is $b-1$ here rather than b.

Proof. Since $f(i) = \Delta_i F(i) = F(i+1) - F(i)$,

$$\sum_{i=a}^{b-1} f(i) = \sum_{i=a}^{b-1} \big(F(i+1) - F(i)\big).$$

This sum is an example of a *telescoping sum*; the inner instances of the summand cancel with pieces from other instances. We can see this by writing the sum in reverse order (starting from $i = b-1$ and going down to $i = a$):

$$\begin{aligned}\sum_{i=a}^{b-1} f(i) &= \big(F(b) - F(b-1)\big) + \big(F(b-1) - F(b-2)\big) + \cdots + \big(F(a+1) - F(a)\big) \\ &= F(b) - F(a). \end{aligned}$$

□

To summarize, here is a table showing the corresponding operations in continuous versus discrete calculus:

continuous calculus	discrete calculus
derivative $\frac{d}{dx}$	difference Δ_n
antiderivative $\int\, dx$	antidifference Σ_n
definite integral $\int_a^b\, dt$	definite sum $\sum_{i=a}^{b-1}$

Example 10.9. Let us determine an antidifference $\Sigma_n n^2$ and use it to compute $\sum_{i=0}^{n-1} i^2$. One way to compute an antidifference $F(n)$ is by using undetermined coefficients. The polynomial n^2 has degree 2, so if its antidifferences are also polynomials then their degrees are at least 3. Let's suppose that the degree is 3 and see if we find one. Let $F(n) = c_3n^3 + c_2n^2 + c_1n + C$. The statement $\Sigma_n n^2 = F(n)$ is equivalent to $\Delta_n F(n) = n^2$. In other words, $F(n+1) - F(n) = n^2$. Expanding, we find

$$3c_3n^2 + (2c_2 + 3c_3)n + (c_1 + c_2 + c_3) = n^2.$$

If this equation holds for all n, then

$$\begin{aligned} 3c_3 &= 1 \\ 2c_2 + 3c_3 &= 0 \\ c_1 + c_2 + c_3 &= 0. \end{aligned}$$

The solution of this system is $c_1 = \frac{1}{6}, c_2 = -\frac{1}{2}, c_3 = \frac{1}{3}$, so $F(n) = \frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{6}n + C = \frac{(2n-1)n(n-1)}{6} + C$ is an antidifference of n^2 for every C. By Theorem 10.8,

$$\sum_{i=0}^{n-1} i^2 = F(n) - F(0) = \frac{(2n-1)n(n-1)}{6}.$$

The antidifference of n^2 may seem unsatisfactorily complicated when compared to the antiderivative of x^2. This suggests that the monomial basis

$$\big((1)_{n\geq 0}, (n)_{n\geq 0}, (n^2)_{n\geq 0}, \dots, (n^{r-1})_{n\geq 0}\big)$$

is not particularly easy to work with when computing differences and antidifferences. We would also like to show that the antidifference of a polynomial is in fact a polynomial. Next we'll introduce a basis that conveniently solves both of these problems.

The binomial coefficient basis

Unlike the monomials n^m, the columns of Pascal's triangle do behave simply under differences and antidifferences.

Theorem 10.10. *Let $m \geq 0$. Then $\binom{n}{m+1}$ is an antidifference of $\binom{n}{m}$ with respect to n.*

Proof. The Pascal relation (Theorem 9.2)

$$\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}$$

implies

$$\Delta_n \binom{n}{m+1} = \binom{n+1}{m+1} - \binom{n}{m+1} = \binom{n}{m}. \qquad \square$$

Theorem 10.10 gives the analogue for Δ_n of the "power rule" $\frac{d}{dx}x^{m+1} = (m+1)x^m$. For example, an antidifference of $\binom{n}{2}$ is simply $\binom{n}{3}$, and $\sum_{i=0}^{n-1}\binom{i}{2} = \binom{n}{3} - \binom{0}{3} = \binom{n}{3}$.

Since $\binom{n}{m}_{n\geq 0}$ is a polynomial sequence with rank $m+1$,

$$\left(\binom{n}{0}_{n\geq 0}, \binom{n}{1}_{n\geq 0}, \binom{n}{2}_{n\geq 0}, \ldots, \binom{n}{r-1}_{n\geq 0}\right)$$

is a basis of $\mathsf{Poly}(r)$. We refer to this basis as the *binomial coefficient basis.*

Example 10.11. Let's write n^2 in the binomial coefficient basis. We can use undetermined coefficients. Write $n^2 = b_2\binom{n}{2} + b_1\binom{n}{1} + b_0\binom{n}{0}$. Expanding and solving for the coefficients gives $n^2 = 2\binom{n}{2} + 1\binom{n}{1} + 0\binom{n}{0}$. Now we can easily compute an antidifference of n^2:

$$\begin{aligned}\Sigma_n n^2 &= \Sigma_n\left(2\binom{n}{2} + \binom{n}{1}\right)\\ &= 2\binom{n}{3} + \binom{n}{2} + C.\end{aligned}$$

This antidifference gives an alternative to Example 10.9 for summing i^2:

$$\begin{aligned}\sum_{i=0}^{n-1} i^2 &= \left(2\binom{n}{3} + \binom{n}{2} + C\right) - \left(2\binom{0}{3} + \binom{0}{2} + C\right)\\ &= 2\binom{n}{3} + \binom{n}{2}.\end{aligned}$$

Some polynomials that arise in combinatorics have much simpler structure when written in the binomial coefficient basis. We will see one family of examples in Chapter 11.

Corollary 10.12. *If $s(n)_{n\geq 0}$ is a polynomial sequence with rank $r \geq 1$, then every antidifference of $s(n)_{n\geq 0}$ is a polynomial sequence with rank $r+1$.*

In particular, polynomial sequences are closed under partial summation.

Proof. Let $F(n) = \sum_{i=0}^{n-1} s(i)$. By Theorem 10.6, $F(n)$ is an antidifference of $s(n)$. Writing $s(n)_{n\geq 0}$ in the binomial coefficient basis and applying Theorem 10.10 shows that $F(n)_{n\geq 0}$ is a polynomial sequence with rank $r+1$. By Proposition 10.5, every antidifference of $s(n)_{n\geq 0}$ is a polynomial sequence with rank $r+1$. ☐

If $s(n)_{n\geq 0}$ is the zero sequence $(0)_{n\geq 0}$, then it is an antidifference of itself. All other antidifferences of the zero sequence have rank 1.

The binomial coefficient basis can also be used to extrapolate from finitely many terms.

Example 10.13. Let us extend the four terms $1, 3, 2, 6$ to a polynomial sequence $s(n)_{n\geq 0}$ with rank ≤ 4. The successive differences of $s(n)_{n\geq 0}$ are as follows.

$$\begin{array}{rccccccccc} s(n): & 1, & & 3, & & 2, & & 6, & & \ldots\\ \Delta_n s(n): & & 2, & & -1, & & 4, & & \ldots & \\ \Delta_n^2 s(n): & & & -3, & & 5, & & \ldots & & \\ \Delta_n^3 s(n): & & & & 8, & & \ldots & & & \end{array}$$

We start with the final sequence and work our way back up to $s(n)_{n\geq 0}$. Since $s(n)_{n\geq 0}$ has rank ≤ 4, the sequence $\big((\Delta^3 s)(n)\big)_{n\geq 0}$ is constant. Since its first term is 8, we have $\Delta_n^3 s(n) = 8\binom{n}{0}$ for all $n \geq 0$. The next sequence up is the antidifference of $8\binom{n}{0}$, so

$$\Delta_n^2 s(n) = 8\binom{n}{1} + C\binom{n}{0}$$

for some C. Since $\binom{0}{1} = 0$, we have $C = -3$. Continuing up, we have

$$\Delta_n s(n) = 8\binom{n}{2} - 3\binom{n}{1} + C\binom{n}{0},$$

where now $C = 2$. Finally,

$$s(n) = 8\binom{n}{3} - 3\binom{n}{2} + 2\binom{n}{1} + \binom{n}{0}.$$

The previous example shows that we can read the coordinates of a polynomial sequence in the binomial coefficient basis directly from the first terms of the successive difference sequences.

Theorem 10.14. *Let $r \geq 0$ and $s(n)_{n\geq 0} \in \mathsf{Poly}(r)$. For all $n \geq 0$, we have $s(n) = \sum_{i=0}^{r-1} b_i\binom{n}{i}$, where $b_i = (\Delta^i s)(0)$.*

PROOF. The idea is that the difference sequence of an integer-valued polynomial is also an integer-valued polynomial, so we can get every integer-valued polynomial by iteratively taking antidifferences from a constant polynomial. Specifically, $\operatorname{rank}(s) \leq r$ implies that $\Delta_n^{r-1} s(n)$ is a rational number b_{r-1} that does not depend on n. In particular, b_{r-1} is the first term of $\big((\Delta^{r-1} s)(n)\big)_{n\geq 0}$. Writing $b_{r-1} = b_{r-1}\binom{n}{0}$ and applying Σ_n gives

$$\Delta_n^{r-2} s(n) = b_{r-1}\binom{n}{1} + b_{r-2}\binom{n}{0},$$

where $b_{r-2} \in \mathbb{Q}$ such that this equality holds at $n = 0$; namely, b_{r-2} is the first term of $\big((\Delta^{r-2} s)(n)\big)_{n\geq 0}$. Applying Σ_n additionally $r-2$ times gives

$$s(n) = b_{r-1}\binom{n}{r-1} + b_{r-2}\binom{n}{r-2} + \cdots + b_0\binom{n}{0},$$

where at each step the constant b_i we introduce is the first term of the corresponding sequence $\big((\Delta^i s)(n)\big)_{n\geq 0}$. □

By taking successive differences until we obtain a sequence that appears to be constant, we are able to guess a polynomial sequence without prescribing the size of the polynomial in advance. When we know many initial terms of a sequence and the rank turns out to be small, this method has a huge speed advantage over both the method of undetermined coefficients and Lagrange interpolation.

Example 10.15. Suppose the first few terms of $s(n)_{n\geq 0}$ are $7, 6, 15, 70, 231, 582, 1231, 2310$. We obtain a sequence that appears to be constant after 4 successive differences:

$$\begin{array}{rl} s(n): & 7, \quad 6, \quad 15, \quad 70, \quad 231, \quad 582, \quad 1231, \quad 2310, \quad \dots \\ \Delta_n s(n): & -1, \quad 9, \quad 55, \quad 161, \quad 351, \quad 649, \quad 1079, \quad \dots \\ \Delta_n^2 s(n): & 10, \quad 46, \quad 106, \quad 190, \quad 298, \quad 430, \quad \dots \\ \Delta_n^3 s(n): & 36, \quad 60, \quad 84, \quad 108, \quad 132, \quad \dots \\ \Delta_n^4 s(n): & 24, \quad 24, \quad 24, \quad 24, \quad \dots \end{array}$$

Therefore our guess is

$$\begin{aligned} s(n) &= 24\binom{n}{4} + 36\binom{n}{3} + 10\binom{n}{2} - \binom{n}{1} + 7\binom{n}{0} \\ &= n^4 - 2n^2 + 7. \end{aligned}$$

Integer-valued polynomials

We've seen that a polynomial can have non-integer coefficients and yet generate an integer sequence. For example, the nth triangular number $T(n) = \frac{1}{2}n^2 + \frac{1}{2}n$ is an integer. Another example is $s(n) = \frac{5}{6}n^3 - \frac{1}{2}n^2 + \frac{8}{3}n$, which produces the integer sequence $0, 3, 10, 26, 56, 105, 178, 280, \dots$. Similarly, the polynomial we computed in Example 6.7 seems to only output integers despite its complicated rational coefficients. We conclude this chapter with an explanation of this phenomenon.

Definition 10.16. A polynomial $f(x)$ is *integer-valued* if $f(n) \in \mathbb{Z}$ for all $n \in \mathbb{Z}$.

Example 10.17. We can see that $T(x)$ is integer-valued by writing it in the binomial coefficient basis. The first several difference sequences of $T(n)_{n\geq 0}$ are as follows.

$$\begin{array}{rl} T(n): & 0, \quad 1, \quad 3, \quad 6, \quad 10, \quad \dots \\ \Delta_n T(n): & 1, \quad 2, \quad 3, \quad 4, \quad \dots \\ \Delta_n^2 T(n): & 1, \quad 1, \quad 1, \quad \dots \end{array}$$

By Theorem 10.14,

$$T(n) = 1\binom{n}{2} + 1\binom{n}{1} + 0\binom{n}{0}.$$

Since the polynomial $\binom{x}{m}$ is integer-valued for each $m \geq 0$, so is $T(x)$.

Example 10.18. Let $s(n) = \frac{5}{6}n^3 - \frac{1}{2}n^2 + \frac{8}{3}n$. The differences of $s(n)_{n\geq 0}$ are as follows.

$$\begin{array}{rl} s(n): & 0, \quad 3, \quad 10, \quad 26, \quad 56, \quad \dots \\ \Delta_n s(n): & 3, \quad 7, \quad 16, \quad 30, \quad \dots \\ \Delta_n^2 s(n): & 4, \quad 9, \quad 14, \quad \dots \\ \Delta_n^3 s(n): & 5, \quad 5, \quad \dots \end{array}$$

Therefore

$$s(n) = 5\binom{n}{3} + 4\binom{n}{2} + 3\binom{n}{1} + 0\binom{n}{0},$$

and it follows that $s(n)_{n\geq 0}$ is an integer sequence.

Example 10.19. On the other hand, expressing $\frac{1}{2}n^2$ in the binomial coefficient basis indicates a possible obstruction to being integer-valued:

$$\frac{1}{2}n^2 = 1\binom{n}{2} + \frac{1}{2}\binom{n}{1} + 0\binom{n}{0}.$$

It turns out that having integer coordinates in the binomial coefficient basis is necessary and sufficient for a polynomial to be integer-valued. In other words, writing a polynomial in the binomial coefficient basis "squeezes out" any non-integer values so they are visible.

Theorem 10.20. *A polynomial $f(x) \in \mathbb{Q}[x]$ is integer-valued if and only if it has integer coordinates when expressed in the binomial coefficient basis.*

Proof. First suppose $f(x)$ has integer coordinates $(b_0, b_1, \ldots, b_{r-1})$ in the binomial coefficient basis. Then $f(n) = b_{r-1}\binom{n}{r-1} + \cdots + b_1\binom{n}{1} + b_0\binom{n}{0}$ is a sum of products of integers and is therefore an integer for every integer n.

Now assume $f(x)$ is integer-valued. Then each difference sequence $\left((\Delta^i f)(n)\right)_{n\geq 0}$ is an integer sequence. By Theorem 10.14, the coordinates of $f(n)_{n\geq 0}$ in the binomial coefficient basis are integers. □

Corollary 10.21. *The extension of a finite sequence of r integers to a sequence in* $\mathsf{Poly}(r)$ *is an integer sequence.*

Proof. By Theorem 10.14, the coefficient of $\binom{n}{i}$ when we represent $s(n)_{n\geq 0}$ in the binomial coefficient basis is $b_i = (\Delta^i s)(0)$. Since this term is an integer,

$$s(n) = b_{r-1}\binom{n}{r-1} + b_{r-2}\binom{n}{r-2} + \cdots + b_0\binom{n}{0}$$

where each b_i is an integer. □

Example 10.22. Let $s(n)_{n\geq 0}$ be the sequence in $\mathsf{Poly}(8)$ whose first 8 terms are the first 8 primes $2, 3, 5, 7, 11, 13, 17, 19$. We computed a polynomial for this sequence in Example 6.7. In the binomial coefficient basis, it is much easier to see why $s(n)_{n\geq 0}$ is an integer sequence, since all the coefficients are necessarily integers:

$$s(n) = -53\binom{n}{7} + 23\binom{n}{6} - 9\binom{n}{5} + 3\binom{n}{4} - \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + 2\binom{n}{0}.$$

Questions

Computations.

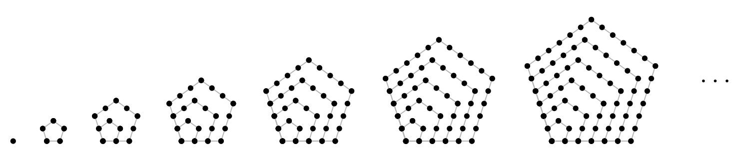
(1) Find an antidifference of the polynomial $3n^2 - 2n + 1$.

(2) (a) Find an antidifference of the polynomial n^3.
  (b) Use the antidifference to evaluate the sum $\sum_{i=1}^{100} i^3$.

(3) The sequence $P(n)_{n\geq 0}$ of *pentagonal numbers*

$$0, 1, 5, 12, 22, 35, 51, 70, \ldots$$

counts points in pentagonal diagrams:



The $(i+1)$th pentagonal diagram has $3(i+1)-2=3i+1$ more points than the ith pentagonal diagram, so $P(n)=\sum_{i=0}^{n-1}(3i+1)$. What is the polynomial formula for $P(n)$?

(4) The nth hexagonal number is defined analogously.
  (a) Compute the first several hexagonal numbers.
  (b) What is the polynomial formula for the nth hexagonal number?

(5) What is $\sum_{i=0}^{n} T(i)$, where $T(i)$ is the ith triangular number?

(6) Compute the terms of the first several iterated difference sequences to write each sequence in the binomial coefficient basis.
  (a) $16, 25, \ldots \in \mathsf{Poly}(2)$
  (b) $2, 6, 9, \ldots \in \mathsf{Poly}(3)$
  (c) $0, 54, 0, 0, 0, \ldots \in \mathsf{Poly}(5)$
  (d) $4, 3, -1, 4, 3, -1, \ldots \in \mathsf{Poly}(6)$

(7) Write each polynomial in the binomial coefficient basis. Is it integer-valued?
  (a) $\frac{1}{4}(n+1)(n-2)$
  (b) $\frac{1}{6}(2n+1)(n+3)(n-2)$
  (c) $\frac{1}{6}(2n^3+3n^2+6)$
  (d) $\frac{1}{24}(-n^4+6n^3+n^2+42n+24)$

Experiments.

(8) Some sequences show their structure not by taking *differences* of consecutive terms but by taking *ratios* of consecutive terms. Compute the first several ratios of each sequence, guess a formula for the nth ratio, and use this formula to obtain a product formula for the nth term.
  (a) $(2^n)_{n\geq 0}$
  (b) $(n!)_{n\geq 0}$
  (c) the sequence of triangular numbers $T(n)_{n\geq 0}$

(9) What is the difference sequence of each sequence?
  (a) $(\log 2^n)_{n\geq 0}$
  (b) $(\log n!)_{n\geq 0}$
  (c) $(\log T(n+1))_{n\geq 0}$

(10) Carrying out the method of undetermined coefficients in the binomial coefficient basis (rather than in the monomial basis, as in Chapter 6) results in an $r\times r$ Vandermonde-like matrix whose (i,j) entry is $\binom{i-1}{j-1}$. What is the determinant of this matrix? When is it invertible?

(11) Is there a product rule for the difference operator that gives $\Delta_n(s(n)t(n))$?

(12) Is there a quotient rule for the difference operator that gives $\Delta_n\left(\frac{s(n)}{t(n)}\right)$?

(13) Let b be a positive integer.
  (a) What is the difference $\Delta_n b^n$?
  (b) What is the antidifference $\sum_n b^n$?
  (c) The real number e has the special property that e^x is its own derivative. Is there an analogue of e for the difference operator?

(14) What is the formula for the ith iterated difference $\Delta_n^i f(n)$ in terms of $f(n), f(n+1), f(n+2), \ldots$?

(15) In Theorem 10.6 we required $n \geq a$. What happens if $n = a - 1$? Is there some way to interpret $\sum_{i=a}^{b-1} f(i)$ when $a > b$ that allows us to remove the condition $n \geq a$?

(16) What are the coordinates of $\frac{d}{dx}\binom{x}{m}$ in the binomial coefficient basis?

(17) Does Descartes's rule of signs (for the number of positive zeros) work when a polynomial is written in the binomial basis?

Proofs.

(18) Prove that if $s(n)_{n\geq 0}$ is a polynomial sequence such that the first term of the ith difference sequence $\big((\Delta^i s)(n)\big)_{n\geq 0}$ is an integer for all $i \geq 0$ then $s(n)_{n\geq 0}$ is an integer sequence.

Programs.

(19) Write a program that takes a list of integers and uses successive differences to guess a formula for the sequence. How does its speed compare to the method of undetermined coefficients and Lagrange interpolation when the rank is small compared to the number of given terms? How does its speed compare when the rank is roughly equal to the number of given terms?

(20) Write a program that determines whether a given polynomial is integer-valued.

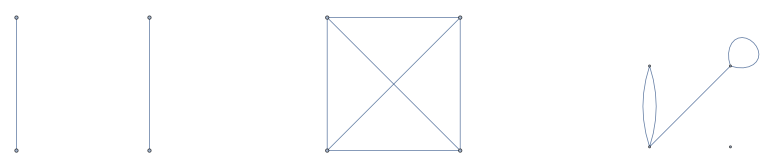
CHAPTER 11

Graphs and their chromatic polynomials

Before moving on from polynomial sequences, we discuss one last combinatorial object and an important connection to polynomial sequences.

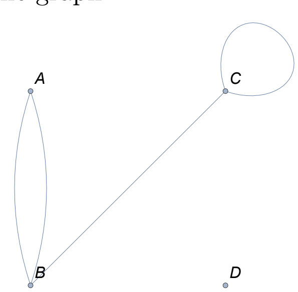
Graphs

A graph is a network. It consists of vertices (points) and edges (connections between points). This is the same terminology we used to define plane trees in Chapter 4, but for a general graph we don't require any relationships among vertices. For example, here are three graphs on 4 vertices (two of which contain two disconnected pieces):



Plane trees are graphs with extra structure imposed on them. This structure has a significant implication: There is a unique way to describe a plane tree, for example as a Dyck word.

However, a general graph does not have any distinguished "root" vertex, there is no order given to vertices connected to a given vertex, and there is no hierarchy of vertices. To work with a graph (for example, to apply a function to it or represent it in a computer) we will need a more concrete description. Let's start by naming the vertices. Then we can specify each edge by the pair of vertices it connects. For example, we can represent the graph



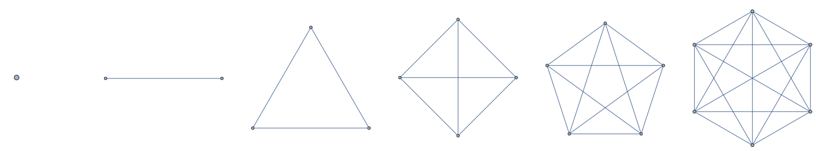
by its vertices $V = \{A, B, C, D\}$ and edges $E = \{\{A, B\}, \{A, B\}, \{B, C\}, \{C, C\}\}$. Since the vertices have no order, V is a set. The edge $\{B, C\}$ is the same as the edge $\{C, B\}$, so there is no order on the vertices in an edge. Since an edge can be a *loop* connecting a vertex to itself, as in the case of $\{C, C\}$, each edge is a multiset of size 2. Moreover, E itself is a multiset, since there can be multiple edges connecting a pair of vertices. Therefore the full definition is as follows.

Definition 11.1. A *graph* is a 2-tuple (V, E), where V is a set (of "vertices") and E is a multiset of 2-element multisets on V.

If we don't want to bother with all the multiset business, we can restrict to simple graphs.

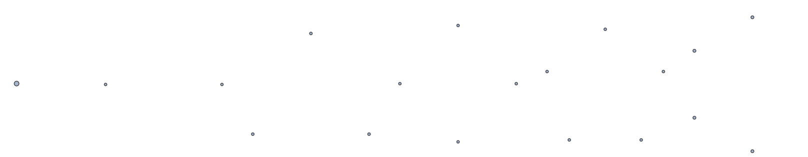
Definition 11.2. A graph (V, E) is *simple* if it contains no loops or multiple edges; that is, E is a set of 2-element subsets of the vertex set V.

Example 11.3. A *complete graph* is a simple graph whose edge set consists of all 2-element subsets of V. Here are complete graphs on $1, 2, \ldots, 6$ vertices:

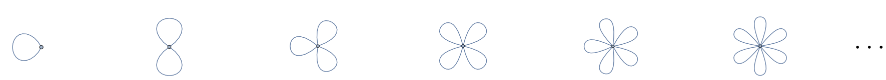


There's a nice connection between complete graphs and simplices. The "frame" of a d-simplex, consisting of just its vertices and edges, is a complete graph on $d + 1$ vertices.

Example 11.4. A complete graph contains all possible edges; an *empty graph* contains no edges. That is, $E = \{\}$. Here are empty graphs on $1, 2, \ldots, 6$ vertices:



Now that we've seen some examples, cue the standard enumeration question: How many graphs are there on n vertices? Well, even on 1 vertex there are infinitely many:



So that's the wrong question.

Enumerating *simple* graphs will solve the problem of loops. How many simple graphs are there on n vertices? Still infinitely many! For each $i \geq 0$ we can form a simple graph with $V = \{i, i + 1\}$ and $E = \{\{i, i + 1\}\}$. These graphs don't have the same vertex sets, so they aren't the same:

And yet, they all *look* the same when we ignore the vertex names. So there are two different counting questions here.

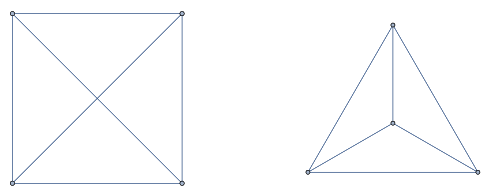
If we care about the names of vertices (and therefore the graph with $V = \{0, 1\}$ and $E = \{\{0, 1\}\}$ is not the same as the graph with $V = \{1, 2\}$ and $E = \{\{1, 2\}\}$) then we are counting *labeled graphs*. We could restrict the alphabet size, as we did for words and permutations and other objects, and ask how many simple graphs there are on the vertex set $V = \{0, 1, \ldots, n\}$. We'll return to this in the Questions at the end of the chapter.

On the other hand, if (as was the case for plane trees) we only care about the structure of the graph and don't care about the names of vertices, then we are counting *unlabeled graphs*, which are equivalence classes of labeled graphs. Two labeled graphs are equivalent if they can be relabeled into each other. The following definition makes this precise.

Definition 11.5. Two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are *isomorphic* if there is a bijection $f\colon V_1 \to V_2$ such that $\{\{f(v), f(w)\}, \ldots, \{f(y), f(z)\}\} = E_2$, where $E_1 = \{\{v, w\}, \ldots, \{y, z\}\}$.

That is, two graphs are isomorphic if they differ only by a renaming of the vertices. Isomorphism also allows us to capture sameness among graph diagrams that we haven't labeled:

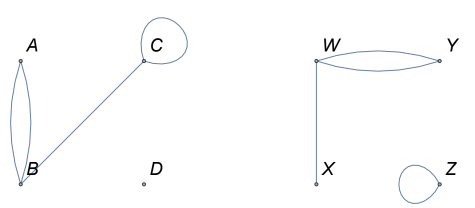
(11.1)



Those two graphs look different but are isomorphic; if we name the left graph's vertices to be $V_1 = \{A, B, C, D\}$ and the right graph's vertices to be $V_2 = \{W, X, Y, Z\}$, then the function $f\colon V_1 \to V_2$ defined by $f(A) = W, f(B) = X, f(C) = Y, f(D) = Z$ is a bijection that maps the edges $\{\{A, B\}, \{A, C\}, \{A, D\}, \{B, C\}, \{B, D\}, \{C, D\}\}$ to $\{\{W, X\}, \{W, Y\}, \{W, Z\}, \{X, Y\}, \{X, Z\}, \{Y, Z\}\}$.

In fact the two graphs (11.1) above are both complete graphs on 4 vertices. Since all complete graphs on 4 vertices are isomorphic to each other, we can speak of *the* complete graph on 4 vertices, with the understanding that if we need to refer to vertices by name then we will have to make some naming decisions. But often it's convenient to put off making those decisions. We use K_n to denote the complete graph on n vertices.

As an example of two graphs that are not isomorphic to each other, consider these:



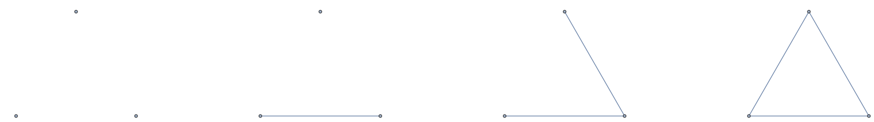
These graphs are not isomorphic, since every bijection $f\colon \{A, B, C, D\} \to \{W, X, Y, Z\}$ sends the vertex D, which has no incident edges, to a vertex with at least one incident edge.

We now returning to the question of enumerating unlabeled graphs. The right question is this: How many simple graphs are there on n vertices, counted up to isomorphism? That is, if two graphs are isomorphic then we only include one of them in the count.

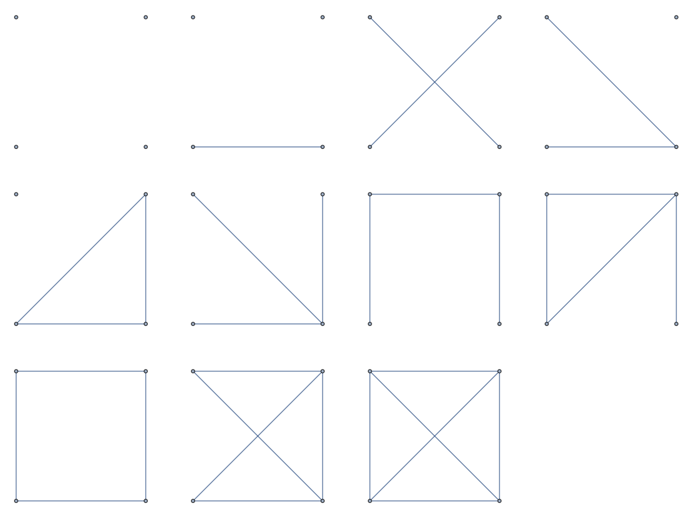
For $n = 0$ there is only 1 graph, namely the graph with $V = \{\}$ and $E = \{\}$. For $n = 1$, the graph consisting of one vertex and no edges is the only simple graph.

For $n = 2$ there are 2 simple graphs up to isomorphism:

For $n = 3$ there are 4:



For $n = 4$ there are 11:



The number of non-isomorphic simple graphs on n vertices for $n = 0, 1, \ldots$ is

$$1, 1, 2, 4, 11, 34, 156, 1044, 12346, 274668, 12005168, 1018997864, \ldots \qquad (\text{A000088}).$$

The growth rate of this sequences suggests it is not a polynomial sequence. In fact there is no simple "formula" for the nth term. The lack of a formula implies that the objects we're counting — simple graphs, in this case — are correspondingly complex. Counting objects is hard when the information we forget about causes objects to merge into equivalence classes in a way that is not easy to track. This is what happens when identify isomorphic graphs with each other.

Coloring planar graphs

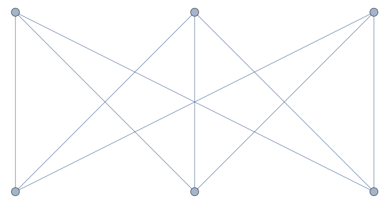
The study of graphs comprises an entire field, known as *graph theory*. Much of graph theory was historically motivated by a single conjecture concerning planar graphs. We'll content ourselves with a survey of the main results in this area without proofs.

A graph is *planar* if it can be drawn in the plane without any edges crossing. Edges are allowed to be curved. For example, each plane tree is a planar graph. The complete graph K_4 is also planar; even though the first drawing in (11.1) has a pair of crossing edges, the second drawing shows that it can be drawn without crossing edges. However, K_5 is not.

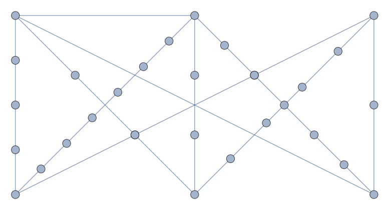
Theorem 11.6. *The complete graph on* 5 *vertices is not planar.*

Since K_5 is not planar, it follows that K_n is not planar for each $n \geq 5$ since any planar drawing of K_n would necessarily contain a planar drawing of K_5.

The graph



is known as $K_{3,3}$, since it consists of two sets of 3 vertices (drawn here in two rows) and all possible edges connecting the two sets. The graph $K_{3,3}$ is also not planar. If a graph G contains a copy of $K_{3,3}$, then a planar drawing of G would necessarily contain a planar drawing of $K_{3,3}$; therefore G is not planar. Here the precise meaning of "copy of H" allows each edge of H to be subdivided into several edges; for example, the graph



contains a copy of $K_{3,3}$ and is therefore not planar.

There's a remarkable characterization of planar graphs due to Kuratowski[1], who in 1930 proved that K_5 or $K_{3,3}$ are the *only* obstructions to planarity. In other words, a graph is planar if and only if it contains no copy of K_5 or $K_{3,3}$.

Planar graphs arise naturally from geographic maps, since maps depict regions of Earth's 2-dimensional surface. Take a map. For each country on the map, create a vertex. Then connect two vertices with an edge precisely when the two corresponding countries share a border. Assuming that each country is one contiguous region of land, the resulting graph is planar.

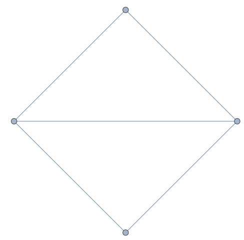
When making a map, it's natural to color countries in such a way that neighboring countries receive different colors. In 1852, Guthrie[2] noticed that he could color a map of the counties in England using only 4 colors. Coloring a map is equivalent to coloring its associated planar graph. A *coloring* of a graph is a way of assigning a color to each vertex in such a way that no edge connects two vertices with the same color. We can make this more precise as follows.

Definition 11.7. Let $G = (V, E)$ be a graph. A *coloring of G when there are n available colors* is a function $f\colon V \to \{1, 2, \dots, n\}$ such that $f(v) \neq f(w)$ for all $\{v, w\} \in E$.

[1]Kazimierz Kuratowski was born in 1896 in Warsaw, Russia (now in Poland) and died in 1980 in Warsaw.

[2]Francis Guthrie was born in 1831 in London, UK and died in 1899 in Cape Town, South Africa.

For example, the *diamond graph*



can be colored with 3 colors by coloring the top and bottom vertices with the same color. Some planar graphs require 4 colors; for example, K_4 cannot be colored with 3 colors. For more than 100 years, it was an open problem whether 4 colors suffice to color every planar graph. A huge number of people worked on it, developing a large tool set in the process. A proof was finally obtained in 1976 by Appel and Haken[3] [**3, 4**] by improving upon previous proof attempts.

Theorem 11.8 (Four color theorem)**.** *Every loopless planar graph has a coloring when there are* 4 *available colors.*

The four color theorem is probably the most famous theorem in discrete mathematics. Partly this is because it answered a famous open question. But its fame — and notoriety — also derives from the central role that computers played in its proof.

The proof consists of two parts. In the first part, humans painstakingly identified 1834 configurations with the property that if there is a counterexample to the four color theorem (that is, a planar graph requiring at least 5 colors) then this counterexample arises from one of these 1834 configurations. This is already a strong result, since it's far from obvious that a finite set of configurations like this exists. In the second part, computers checked that each of the 1834 configurations is *reducible*; that is, if a graph G contains one of these configurations, then G can be reduced to a smaller graph that can be colored using the same number of colors. It follows that every minimal counterexample to the four color theorem can be reduced to a smaller counterexample. Therefore there are no minimal counterexamples, hence the four color theorem is true.

Allegedly the computer check took over a month of computer time in 1976. The four color theorem was the first major theorem to be proved using non-human computation, so it challenged older notions of what is acceptable as a mathematical proof. What if there is an error in the code or the compiler or the hardware? Even if there are no errors, how can we *understand* a proof that relies on a computer program outputting True after a month of churning away? Some members of the mathematical community were of the opinion that the theoremhood of such a result was extremely dubious. Up until 1976, every proof of every theorem that had ever been written had been verified by humans; this proof could not be.

Fortunately, no serious errors have been found in Appel and Haken's approach. Other researchers have improved upon it and have developed other proofs, and this has given additional credence to the four color theorem. Today the community widely accepts it as a theorem.

[3]Kenneth Appel was born in 1932 in Brooklyn, New York, USA and died in 2013 in Dover, New Hampshire, USA. Wolfgang Haken was born in 1928 in Berlin, Germany.

The chromatic polynomial

One of the unsuccessful approaches toward proving the four color theorem introduced the following question, which has become worthy of attention in its own right. Given a graph (not necessarily planar), how many colorings does it have when there are n available colors?

Example 11.9. Let G be the following graph.

Given n available colors, there are n choices for the first vertex and $n-1$ choices for the second, so there are $n(n-1)$ colorings.

In general, the number of colorings of a graph is given by a polynomial.

Theorem 11.10. *Let $G = (V, E)$ be a simple graph whose vertex set V is finite. Let $\chi_G(n)$ be the number of colorings of G when there are n available colors. Then $\chi_G(n)_{n\geq 0}$ is a polynomial sequence with rank $|V| + 1$.*

Definition 11.11. The polynomial $\chi_G(n)$ is the *chromatic polynomial* of G.

Proof of Theorem 11.10. We break up the colorings of G according to how many of the n available colors actually appear in each coloring:

$$\chi_G(n) = \begin{bmatrix}\text{\# colorings using}\\ \text{exactly 0 colors}\end{bmatrix} + \begin{bmatrix}\text{\# colorings using}\\ \text{exactly 1 color}\end{bmatrix} + \cdots + \begin{bmatrix}\text{\# colorings using}\\ \text{exactly } |V| \text{ colors}\end{bmatrix}.$$

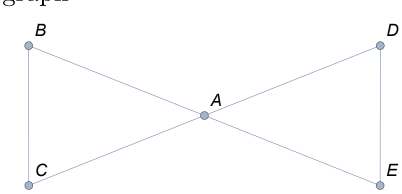
We can stop that sum at $|V|$ colors since no coloring of G uses more than $|V|$ colors. It remains to determine the number of colorings of G that use exactly i colors, for each i. Let b_i be the number of colorings of G that use each color $1, 2, \ldots, i$ at least once and use no other colors. Since there are $\binom{n}{i}$ subsets of $\{1, 2, \ldots, n\}$ of size i, the number of colorings of G that use exactly i of the n available colors is $b_i \cdot \binom{n}{i}$. Therefore

$$\chi_G(n) = b_0\binom{n}{0} + b_1\binom{n}{1} + \cdots + b_{|V|}\binom{n}{|V|}.$$

Since each $\binom{n}{i}$ is a polynomial in n, so is $\chi_G(n)$. Moreover, the rank of $\chi_G(n)_{n\geq 0}$ is $|V| + 1$ by Proposition 9.1, since $b_{|V|} \neq 0$. $\square$

Notice that the proof wrote the polynomial $\chi_G(n)$ in the binomial coefficient basis.

Example 11.12. The graph



is known as the *butterfly graph*. Let's compute its chromatic polynomial. By the proof of Theorem 11.10 it suffices to compute $b_0, b_1, \ldots, b_5$. Since G contains a vertex, this implies $b_0 = 0$. Since G contains K_3 as a subgraph, we also have $b_1 = 0$ and $b_2 = 0$.

For larger i, there is an art to computing b_i by hand. It is helpful to focus on vertices with a large number of incident edges. Let $i = 3$, so that we must use all three colors $1, 2, 3$. Color the vertex A first; there are 3 choices. Now there are 2 choices for the color of B; once we've chosen its color, the color of C is determined. Similarly, there are 2 choices for the color of D and 1 choice for E. Therefore $b_3 = 3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 = 12$.

Let $i = 4$; we must use all four colors $1, 2, 3, 4$. Therefore exactly two of the five vertices will receive the same color. Vertex A receives a different color than all other vertices, so it cannot participate in the pair of vertices that receive the same color. There are four pairs of vertices that are not connected to each other by an edge, so we examine the four cases where each of these pairs receives the same color. In each case there are $4!$ colorings, so $b_4 = 4! + 4! + 4! + 4! = 96$.

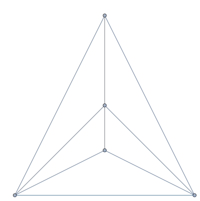
For $i = 5$, each vertex receives a different color, so $b_5 = 5! = 120$. Therefore

$$\chi_G(n) = 12\binom{n}{3} + 96\binom{n}{4} + 120\binom{n}{5}.$$

Questions

Computations.

(1) Let G be the following graph.

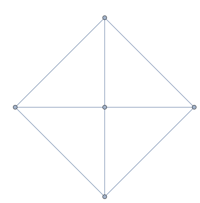


(a) Determine the chromatic polynomial of G.

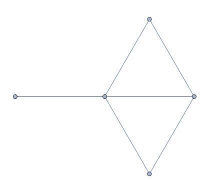
(b) How many colorings of G are there, if you have 100 available colors?

(2) Determine the chromatic polynomial of each graph.

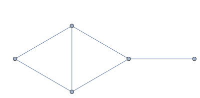
(a)



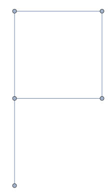
(b)



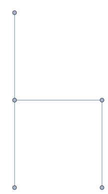
(c)



(d)



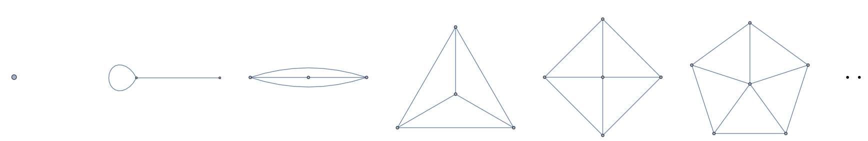
(e)



(3) Determine the chromatic polynomial of each complete graph.
 (a) K_3
 (b) K_4
 (c) K_5
 (d) K_m. What is $\chi_{K_m}(4)$?
(4) Let E_m denote the empty graph on m vertices. Determine the chromatic polynomial of each graph. (Simplify as much as possible.)
 (a) E_2
 (b) E_3
 (c) E_m
(5) Suppose
 - G is a graph with 4 vertices,
 - G has no colorings when there are 0, 1, or 2 available colors,
 - G has 12 colorings when there are 3 available colors (not necessarily using all 3 colors!), and
 - G has 72 colorings when there are 4 available colors (again, not necessarily using all 4).
 (a) What is the chromatic polynomial of G?
 (b) Find all graphs that have the same chromatic polynomial as G.

Experiments.

(6) How many edges does the complete graph on n vertices have?
(7) For each $n \geq 1$, the *wheel graph* W_n is the graph on n vertices such that $n-1$ vertices form a cycle and the remaining vertex is connected to all other vertices.



 (a) How many edges does the wheel graph W_n have?
 (b) How many regions (including the exterior region) do the edges of W_n cut the plane into when drawn without any edges crossing?

(c) Find constants a, b, c such that $a|V_n| + b|E_n| + c|R_n| = 1$ holds for all $n \geq 1$, where V_n, E_n, R_n are the sets of vertices, edges, and regions of W_n.

(8) Consider an n-sided polygon in *general position*, with all the diagonals drawn in; that is, place the vertices in the plane so that no three diagonals intersect in a single point.

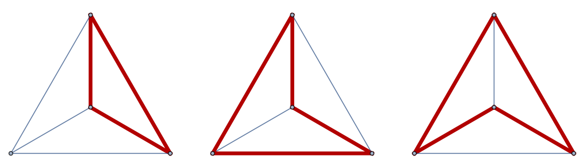
(a) How many intersection points are there?

(b) How many line segments do the diagonals cut themselves into? (In other words, if we consider each intersection point to be a new vertex, how many edges are there?)

(c) How many regions do the segments cut the plane into?

(d) Find constants a, b, c such that $a|V_n| + b|E_n| + c|R_n| = 1$ holds for all $n \geq 1$, where V_n, E_n, R_n are the sets of vertices, edges, and regions.

(9) A *cycle* in a graph $G = (V, E)$ is a subset $S \subseteq E$ of edges such that no vertex in V belongs to more than 2 edges in S and the edges of S form a loop. For example, 3 of the cycles in K_4 are as follows.

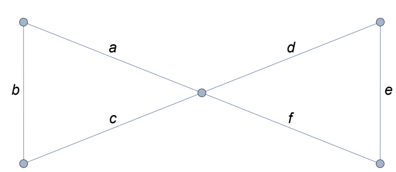


For each $n \in \{3, 4, 5, 6\}$, compute the number of cycles in W_n, where W_n is the wheel graph introduced in Question (7). Use this to guess a formula for the number of cycles in W_n.

(10) The coordinates of the chromatic polynomial in the monomial basis also have a combinatorial interpretation, as found by Whitney[4] in 1932 [**31**]. Given a graph G, choose a permutation on the set E of edges. A set $S \subseteq E$ of edges is a *broken cycle* if it can be obtained from a cycle of G by removing the edge that appears earliest in the chosen permutation on E. Let $b_G(i)$ be the number of subsets of E of size i that do not contain a broken cycle as a subset. Then

$$\chi_G(n) = \sum_{i=0}^{|V|} (-1)^i b_G(i) n^{|V|-i}.$$

(In particular, $b_G(i)$ does not depend on the permutation on E.) For example, consider the butterfly graph from Example 11.12. Name the edges as follows, and choose the permutation (a, b, c, d, e, f).

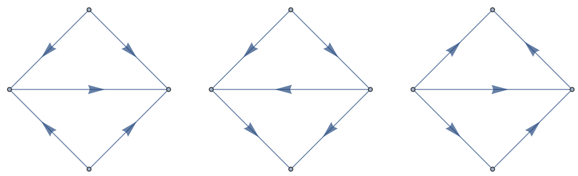


There are two cycles — $\{a, b, c\}$ and $\{d, e, f\}$. The broken cycles are $\{b, c\}$ and $\{e, f\}$. Therefore we are interested in subsets of $E = \{a, b, c, d, e, f\}$

[4]Hassler Whitney was born in 1907 in New York City, New York, USA and died in 1989 in Princeton, New Jersey, USA.

that do not contain $\{b, c\}$ or $\{e, f\}$ as a subset. The number of such subsets for each size $0, 1, 2, 3, 4, 5, 6$ is $1, 6, 13, 12, 4, 0, 0$, so $\chi_G(n) = n^5 - 6n^4 + 13n^3 - 12n^2 + 4n$. Verify Whitney's identity for the graph in Question (2), part (b).

(11) The chromatic polynomial gives information about a graph even for negative values of n. An *acyclic orientation* of a graph G is an assignment of a direction to each edge in G such that no cycle in G is oriented completely clockwise or completely counterclockwise. For example,



are some acyclic orientations of the diamond graph.
(a) Find all acyclic orientations of the diamond graph.
(b) Compute $\chi_G(-1)$ for the diamond graph by first computing $\chi_G(n)$ and then letting $n = -1$.

(12) (a) Determine the number of acyclic orientations of the wheel graph W_m for each $m \in \{3, 4, 5, 6\}$.
(b) For each $m \in \{3, 4, 5, 6\}$, compute $\chi_{W_m}(-1)$ by first computing $\chi_{W_m}(n)$ and then letting $n = -1$.

Proofs.

(13) The relationship between the number of vertices, edges, and regions that you found in Questions (7)–(8) is known as Euler's[5] formula. Prove that Euler's formula holds for all connected, planar graphs with finitely many vertices and edges. A graph is *connected* if every vertex can be reached from every other vertex by walking along edges. (Hint: Use induction.)

(14) Find a relationship between the number of edges and the number of regions in a simple, connected, planar graph drawn in such a way that no additional edges can be added without introducing a pair of crossing edges.

(15) Prove Theorem 11.6 that K_5 is not planar. (Hint: Use the previous question.)

(16) Prove that $K_{3,3}$ is not planar.

(17) Let V be a set of n elements. How many simple graphs have vertex set V? (These are called *labeled graphs*; we're not counting them up to isomorphism.)

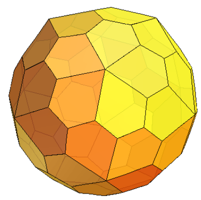
More computations.

(18) Euler's formula also applies to polyhedra that can be flattened into the plane without introducing edge crossings, since each face of the polyhedron becomes a region in the plane.

[5]Leonhard Euler (pronounced 'oiler') was born in 1707 in Basel, Switzerland and died in 1783 in Saint Petersburg, Russia.

(a) Verify that Euler's formula gives a valid relationship between the number of vertices, edges, and faces for the tetrahedron, cube, and octahedron.

(b) A *pentagonal hexecontahedron* has 60 pentagonal faces. Use Euler's formula to determine how many vertices it has.



Part 3

Constant-recursive sequences

CHAPTER 12

Friends of Fibonacci

Recurrences

A polynomial sequence has an explicit formula for the nth term. But as we begin to see in this chapter, formulas are the exception rather than the rule. For most sequences that arise in mathematics, it is more natural to define them indirectly by a recurrence. In fact we've met several already. The most famous is the Fibonacci sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, \ldots \qquad (\text{A000045}),$$

which counted rabbit pairs in Chapter 1. The Fibonacci sequence $F(n)_{n\geq 0}$ is defined by

$$\begin{aligned} F(0) &= 0 \\ F(1) &= 1 \\ F(n+2) &= F(n+1) + F(n) \text{ for } n \geq 0. \end{aligned}$$

As before, we refer to the values $F(0) = 0$ and $F(1) = 0$ as *initial conditions* and $F(n+2) = F(n+1) + F(n)$ as the *recurrence*.

The following defines a general class of Fibonacci-like sequences.

Definition 12.1. A sequence $s(n)_{n\geq 0}$ is *constant-recursive* if there is an integer $r \geq 0$ and rational numbers $c_0, c_1, \ldots, c_{r-1}$ such that

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n)$$

for all $n \geq 0$. The minimal such r is the *rank* of $s(n)_{n\geq 0}$ as a constant-recursive sequence.

Just as the rank of a polynomial sequence measures its size, the rank of a constant-recursive gives us a notion of size for constant-recursive sequences. The recurrence for the Fibonacci sequence shows that its rank is at most 2. Is the rank in fact less than 2? For the rank to be 1 there would need to be a rational number c_0 such that $F(n+1) = c_0 F(n)$ for all $n \geq 0$. However, this would imply $1 = \frac{F(2)}{F(1)} = c_0 = \frac{F(3)}{F(2)} = 2$, which is false. Therefore the rank of $F(n)_{n\geq 0}$ is 2.

A sequence can satisfy more than one recurrence.

Example 12.2. Let $s(n) = (-1)^n$. The sequence $s(n)_{n\geq 0}$ is $1, -1, 1, -1, \ldots$. Since $(-1)^{n+2} = (-1)^n$, the sequence $s(n)_{n\geq 0}$ satisfies the recurrence $s(n+2) = s(n)$. It also satisfies the smaller recurrence $s(n+1) = -s(n)$.

However, the minimal recurrence is unique.

Proposition 12.3. *Let $s(n)_{n\geq 0}$ be a constant-recursive sequence with rank r. There is exactly one recurrence of the form*

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n)$$

that holds for all $n \geq 0$.

Proof. Assume that $c_0, c_1, \ldots, c_{r-1}, d_0, d_1, \ldots, d_{r-1}$ are rational numbers such that both

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n)$$

and

$$s(n+r) = d_{r-1}s(n+r-1) + \cdots + d_1 s(n+1) + d_0 s(n)$$

hold for all $n \geq 0$. Subtracting the second equation from the first, we get

$$0 = (c_{r-1} - d_{r-1})s(n+r-1) + \cdots + (c_1 - d_1)s(n+1) + (c_0 - d_0)s(n)$$

for all $n \geq 0$. Since $\operatorname{rank}(s) = r$, this third equation cannot be used to define $s(n)_{n\geq 0}$, because otherwise $\operatorname{rank}(s)$ would be less than r. Therefore each coefficient is 0, so $c_i = d_i$ for each $i \in \{0, 1, \ldots, r-1\}$. □

Well-known sequences

There are two variants of the Fibonacci sequence that have their own names.

Example 12.4. First we'll keep the Fibonacci recurrence but change the initial conditions. Let $L(0) = 2$ and $L(1) = 1$. For $n \geq 0$, let $L(n+2) = L(n+1) + L(n)$. The sequence $L(n)_{n\geq 0}$ is known as the *Lucas*[1] *sequence*. Its terms are

$$2, 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, 322, 521, 843, 1364, \ldots \qquad (\text{A000032}).$$

Example 12.5. The *Tribonacci sequence* is defined by adding 3 previous terms rather than 2. Since we've already used $T(n)$ to denote the nth triangular number, we'll use $E(n)$ to denote the nth Tribonacci number. (The letter F has two prongs, and E has three!) Let $E(0) = 0$, $E(1) = 0$, and $E(2) = 1$. For $n \geq 0$, let $E(n+3) = E(n+2) + E(n+1) + E(n)$. The terms of $E(n)_{n\geq 0}$ are

$$0, 0, 1, 1, 2, 4, 7, 13, 24, 44, 81, 149, 274, 504, 927, 1705, \ldots \qquad (\text{A000073}).$$

Like the Fibonacci sequence, the Lucas and Tribonacci sequences also arise when counting certain words.

Example 12.6. How many length-n words on $\{0, 1\}$ avoid 00 as *cyclic words*? That is, when looking for 00 in a word $w = w_1 w_2 \cdots w_n$ we consider the last letter w_n to be the left neighbor of the first letter of w_1. For $n = 3$ there are 4 such words:

011, 101, 110, 111.

The word 010 avoids 00 in the usual sense but not as a cyclic word. For $n = 4$ there are 7:

0101, 0111, 1010, 1011, 1101, 1110, 1111.

For $n = 5$ there are 11:

01011, 01101, 01111, 10101, 10110, 10111, 11010, 11011, 11101, 11110, 11111.

Can you see how to build these words from smaller words?

[1]Édouard Lucas (pronounced 'lookah') was born in 1842 in Amiens, France and died in 1891 in Paris, France.

Theorem 12.7. *For each $n \geq 2$, the number of length-n words on $\{0,1\}$ that avoid* 00 *as cyclic words is the nth Lucas number $L(n)$.*

PROOF. Let $S(n)$ be the set of length-n words on $\{0,1\}$ that avoid 00 as cyclic words. For $n = 2$, we have $S(2) = \{01, 10, 11\}$, so indeed $|S(2)| = 3 = L(2)$. We have already verified that $|S(3)| = L(3)$, so it suffices to show that $|S(n+2)| = |S(n+1)| + |S(n)|$. We claim that the function $f\colon S(n+1) \cup S(n) \to S(n+2)$ defined by

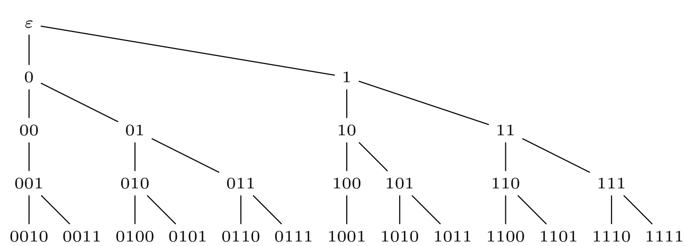
$$f(w) = \begin{cases} w1 & \text{if } |w| = n+1 \\ w01 & \text{if } |w| = n \text{ and } w \text{ starts with } 0 \\ w10 & \text{if } |w| = n \text{ and } w \text{ starts with } 1 \end{cases}$$

is a bijection.

To show surjectivity, let $v \in S(n+2)$. There are three cases. If v ends with 0, then v starts with 1 and ends with 10, so dropping the last two letters of v produces a word in $S(n)$. If v ends with 01, then v ends with 101, so dropping the last two letters of v produces a word in $S(n)$. Finally, if v ends with 11, then dropping the last letter of v produces a word in $S(n+1)$.

To show injectivity, suppose $f(w) = f(v)$ for some words $w, v \in S(n+1) \cup S(n)$. There are two cases. If $|w| = |v|$, then $w1 = v1$, so $w = v$. If $|w| \neq |v|$, then without loss of generality we can assume $|w| < |v|$. Then $w01 = v1$ and w starts with 0. This implies v starts with 0 and ends with 0, but this contradicts $v \in S(n+1)$. $\square$

Example 12.8. How many length-n words on $\{0,1\}$ avoid 000? Let's build a tree organizing such words according to their prefixes.



The number of words on each level is $1, 2, 4, 7, 13, \ldots$.

Theorem 12.9. *For each $n \geq 0$, the number of length-n words on $\{0,1\}$ that avoid* 000 *is the Tribonacci number $E(n+3)$.*

PROOF. Let $S(n)$ be the set of length-n words on $\{0,1\}$ that avoid 000. Define the function $f\colon S(n+2) \cup S(n+1) \cup S(n) \to S(n+3)$ by

$$f(w) = \begin{cases} w1 & \text{if } |w| = n+2 \\ w10 & \text{if } |w| = n+1 \\ w100 & \text{if } |w| = n. \end{cases}$$

One checks that f is a bijection. Therefore $|S(n+3)| = |S(n+2)| + |S(n+1)| + |S(n)|$. This is the same recurrence satisfied by the Tribonacci sequence, so it suffices to check the initial conditions $|S(0)| = 1 = E(3)$, $|S(1)| = 2 = E(4)$, and $|S(2)| = 4 = E(5)$. $\square$

Difference equations

Since the discrete analogue of the derivative is the difference operator Δ_n from Chapter 10, it is natural to be interested in difference equations, which involve $s(n), \Delta_n s(n), \dots, \Delta_n^{r-1} s(n)$, in addition to recurrences, which involve $s(n), s(n+1), \dots, s(n+r-1)$.

Example 12.10. Does the Fibonacci sequence satisfy an equation of the form

$$\Delta_n^r F(n) + d_{r-1} \Delta_n^{r-1} F(n) + \cdots + d_1 \Delta_n F(n+1) + d_0 F(n) = 0$$

for some r, where $d_i \in \mathbb{Q}$? Beginning with the Fibonacci recurrence $0 = F(n+2) - F(n+1) - F(n)$, use the identity $\Delta_n^2 F(n) = F(n+2) - 2F(n+1) + F(n)$ to replace $F(n+2)$ with $\Delta_n^2 F(n) + 2F(n+1) - F(n)$:

$$\begin{aligned} 0 &= \left(\Delta_n^2 F(n) + 2F(n+1) - F(n)\right) - F(n+1) - F(n) \\ &= \Delta_n^2 F(n) + F(n+1) - 2F(n). \end{aligned}$$

Then use $\Delta_n F(n) = F(n+1) - F(n)$ to replace $F(n+1)$ with $\Delta_n F(n) + F(n)$. We obtain the difference equation

$$\Delta_n^2 F(n) + \Delta_n F(n) - F(n) = 0.$$

In general, iteratively applying Δ_n to $s(n)$ produces linear combinations of shifts of $s(n)$:

$$\begin{aligned} s(n) &= s(n) \\ \Delta_n s(n) &= s(n+1) - s(n) \\ \Delta_n^2 s(n) &= s(n+2) - 2s(n+1) + s(n) \\ \Delta_n^3 s(n) &= s(n+3) - 3s(n+2) + 3s(n+1) - s(n). \end{aligned}$$

The coefficients are as follows.

Theorem 12.11. *For each $i \geq 0$,*

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

PROOF. The statement is true for $i = 0$, since $s(n) = s(n)$. Inductively, assume it holds for i. The term $s(n+m)$ arises in $\Delta_n^{i+1} s(n)$ in two ways when we apply Δ_n to $\Delta_n^i s(n)$, namely from $\Delta_n s(n+m-1) = s(n+m) - s(n+m-1)$ and from $\Delta_n s(n+m) = s(n+m+1) - s(n+m)$. Therefore the coefficient of $s(n+m)$ in $\Delta_n^{i+1} s(n)$ is the difference of the coefficients of $s(n+m-1)$ and $s(n+m)$ in $\Delta_n^i s(n)$, which is

$$\begin{aligned} (-1)^{i-(m-1)} \binom{i}{m-1} - (-1)^{i-m} \binom{i}{m} &= (-1)^{i+1-m} \left(\binom{i}{m-1} + \binom{i}{m} \right) \\ &= (-1)^{i+1-m} \binom{i+1}{m} \end{aligned}$$

by Theorem 9.2. This completes the induction. ☐

The result of iteratively applying Δ_n in the previous proof can also be seen by representing each successive iteration $\Delta_n^i s(n)$ by its list of coefficients. For example, we see that we obtain $(1, -3, 3, -1)$ by subtracting $(0, 1, -2, 1)$ from the shifted copy $(1, -2, 1, 0)$:

$$(1, -2, 1, 0) - (0, 1, -2, 1) = (1, -3, 3, -1).$$

This is an emulation of the Pascal relation, so the coefficients are the entries of a row of Pascal's triangle, with alternating signs.

Now we can show that difference equations are equivalent to recurrences.

Theorem 12.12. *The sequence $s(n)_{n\geq 0}$ is constant-recursive with rank $\leq r$ if and only if there exist rational numbers $d_0, d_1, \dots, d_{r-1}$ such that*

$$\Delta_n^r s(n) + d_{r-1}\Delta_n^{r-1} s(n) + \cdots + d_1 \Delta_n s(n) + d_0 s(n) = 0 \tag{12.1}$$

for all $n \geq 0$.

Proof. First assume that $s(n)_{n\geq 0}$ satisfies Equation (12.1). Use Theorem 12.11 to rewrite each $\Delta_n^i s(n)$ in terms of $s(n), s(n+1), \dots, s(n+r)$. This produces a recurrence for $s(n)_{n\geq 0}$ in which the coefficient of $s(n+r)$ is 1 since the only difference $\Delta_n^i s(n)$ that involves $s(n+r)$ is $\Delta_n^r s(n)$. Therefore $s(n)_{n\geq 0}$ is constant-recursive with rank $\leq r$.

In the other direction, assume that $s(n)_{n\geq 0}$ is constant-recursive, and let

$$s(n + r) = c_{r-1}s(n + r - 1) + \cdots + c_1 s(n + 1) + c_0 s(n)$$

be its minimal recurrence. For each $i \geq 0$, Theorem 12.11 implies that

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

so successively replacing $s(n+r), s(n+r-1), \dots, s(n+1)$ in the recurrence produces a difference equation in the form of Equation (12.1). The coefficient of $\Delta_n^r s(n)$ in this equation is 1 since this term only arises from $s(n + r)$. □

Computation of the nth term

For a polynomial sequence, the nth term can be computed quickly by simply evaluating a polynomial at an integer n. On the other hand, to compute the nth term of a constant-recursive sequence $s(n)_{n\geq 0}$, the definition suggests that we need to first compute $s(0), s(1), \dots, s(n-1)$. Fortunately, there is faster way if you're not interested in previous terms. It works by emulating the recurrence with a matrix product.

Example 12.13. Let

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

The matrix M has the property that

$$\begin{bmatrix} F(n+1) \\ F(n+2) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} F(n) \\ F(n+1) \end{bmatrix}$$

since the second row of this matrix equation states $F(n + 2) = F(n) + F(n + 1)$ and the first row states $F(n+1) = F(n+1)$. The two 2×1 matrices each contain two consecutive Fibonacci numbers and therefore contain enough information to compute additional terms. Repeated matrix multiplication corresponds to applying the recurrence repeatedly; for example,

$$\begin{bmatrix} F(n+2) \\ F(n+3) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^2 \begin{bmatrix} F(n) \\ F(n+1) \end{bmatrix}.$$

Starting with the initial conditions and iterating the recurrence n times gives

$$\begin{bmatrix} F(n) \\ F(n+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

In general, to emulate the recurrence

$$s(n+r) = c_0 s(n) + c_1 s(n+1) + \cdots + c_{r-1} s(n+r-1),$$

we form the $r \times r$ matrix

$$M = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ c_0 & c_1 & c_2 & \cdots & c_{r-1} \end{bmatrix}.$$

This matrix is known as the *companion matrix* of the recurrence. When we multiply M by

$$\begin{bmatrix} s(n) \\ s(n+1) \\ s(n+2) \\ \vdots \\ s(n+r-1) \end{bmatrix},$$

the last row of M computes $s(n+r)$ from the previous r terms, and the other rows shift the previous $r-1$ terms up one entry and discard $s(n)$. Therefore, for all $n \geq 1$,

$$\begin{bmatrix} s(n) \\ s(n+1) \\ s(n+2) \\ \vdots \\ s(n+r-1) \end{bmatrix} = M^n \begin{bmatrix} s(0) \\ s(1) \\ s(2) \\ \vdots \\ s(r-1) \end{bmatrix}.$$

We define M^0 to be the $r \times r$ identity matrix so that this identity also holds for $n = 0$.

Computing the nth power of a matrix M can be done by multiplying n copies of M, but it is faster to use *repeated squaring*.

Example 12.14. What is $F(50)$? Let

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

To compute M^{50}, we write $50 = 32 + 16 + 2$ as a sum of distinct powers of 2. (This is equivalent to the base-2 representation of 50.) Then $M^{50} = M^{32} M^{16} M^2$, and the matrices $M^1, M^2, M^4, M^8, M^{16}, M^{32}$ can be computed successively by squaring:

$$M^1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$M^2 = \left(M^1\right)^2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$M^4 = \left(M^2\right)^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^2 = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

$$M^8 = \left(M^4\right)^2 = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}^2 = \begin{bmatrix} 13 & 21 \\ 21 & 34 \end{bmatrix}$$

$$M^{16} = \left(M^8\right)^2 = \begin{bmatrix} 13 & 21 \\ 21 & 34 \end{bmatrix}^2 = \begin{bmatrix} 610 & 987 \\ 987 & 1597 \end{bmatrix}$$

$$M^{32} = \left(M^{16}\right)^2 = \begin{bmatrix} 610 & 987 \\ 987 & 1597 \end{bmatrix}^2 = \begin{bmatrix} 1346269 & 2178309 \\ 2178309 & 3524578 \end{bmatrix}.$$

Finally, we multiply the appropriate powers together to compute

$$\begin{aligned} M^{50} = M^{32}M^{16}M^2 &= \begin{bmatrix} 1346269 & 2178309 \\ 2178309 & 3524578 \end{bmatrix} \begin{bmatrix} 610 & 987 \\ 987 & 1597 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 7778742049 & 12586269025 \\ 12586269025 & 20365011074 \end{bmatrix}. \end{aligned}$$

Rather than performing 49 matrix multiplications to compute M^{50}, we only performed 7. (It's a coincidence that $49 = 7^2$ here.) Therefore

$$\begin{bmatrix} F(50) \\ F(51) \end{bmatrix} = \begin{bmatrix} 7778742049 & 12586269025 \\ 12586269025 & 20365011074 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 12586269025 \\ 20365011074 \end{bmatrix}.$$

To compare the method of repeated squaring to direct application of the recurrence, notice that multiplying two 2×2 matrices involves 4 dot products on vectors with length 2. Overall we performed $7 \cdot 4 + 2 = 30$ dot products to compute $F(50)$, whereas using the recurrence to compute $F(50)$ requires 49 dot products. We can do even better by backing up a few steps and skipping the computation of M^{50}. Instead, multiply

$$\begin{bmatrix} F(50) \\ F(51) \end{bmatrix} = \begin{bmatrix} 1346269 & 2178309 \\ 2178309 & 3524578 \end{bmatrix} \begin{bmatrix} 610 & 987 \\ 987 & 1597 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

from right to left. Then we only need $5 \cdot 4 + 3 \cdot 2 = 26$ dot products.

Questions

Computations.

(1) Compute the first several terms of the sequence $s(n)_{n \geq 0}$ defined by the initial conditions $s(0) = 0, s(1) = 1, s(2) = 2$ and the recurrence $s(n+3) = s(n+2) + s(n+1) - s(n)$ for all $n \geq 0$.
(2) What difference equation is satisfied by the Tribonacci sequence $E(n)_{n \geq 0}$?
(3) Use repeated squaring to compute $E(22)$.

Experiments.

(4) Does 2^n satisfy any recurrence?
(5) Does $(10^n - 1)/9$ satisfy any recurrence?
(6) What is the rank of the sequence defined by the recurrence $s(n+4) = s(n+3) + s(n+1) + s(n)$ and initial conditions $s(0) = 0, s(1) = 1, s(2) = 1, s(3) = 2$?
(7) What happens in Example 12.6 for $n = 0$ and $n = 1$?
(8) Draw a tree to organize the words in Example 12.6 according to their prefixes. What goes wrong?
(9) How many constant-recursive sequences have rank 0? What are they?
(10) What do constant-recursive sequences with rank 1 look like?
(11) How many length-n words on $\{0, 1\}$ avoid
   (a) 11?
   (b) 01?
   (c) 10?
(12) How many length-n words on $\{0, 1\}$ avoid 0000?
(13) How many length-n words on $\{0, 1\}$ avoid
   (a) 0001?
   (b) 0011?
(14) How many length-n words on $\{0, 1\}$ avoid
   (a) 0010?
   (b) 0110?
(15) How many length-n words on $\{0, 1, 2\}$ avoid
   (a) 00?
   (b) 01?
(16) How many length-n words on $\{0, 1, 2\}$ avoid
   (a) both 01 and 10?
   (b) both 01 and 00?
   (c) both 01 and 22?
(17) A *Pythagorean triple* is a 3-tuple (a, b, c) of positive integers such that $a^2 + b^2 = c^2$. The Pythagorean triples in which a is odd and $|a - b| = 1$ are listed in the following table, along with the corresponding values of $\sqrt{c-b}$ and $\frac{a}{\sqrt{c-b}}$. Guess recurrences for the last two columns, and use these recurrences to extend the table by several rows.

a	b	c	$\sqrt{c-b}$	$\frac{a}{\sqrt{c-b}}$
3	4	5	1	3
21	20	29	3	7
119	120	169	7	17
697	696	985	17	41
4059	4060	5741	41	99
			99	239
	$\vdots$		$\vdots$	$\vdots$

(18) For each sequence $s(n)_{n\geq 0}$, compute several values of $\frac{s(n+1)}{s(n)}$ numerically. What happens as $n \to \infty$?
   (a) the Fibonacci sequence
   (b) the Lucas sequence
   (c) the Tribonacci sequence

(d) the sequence $(2^n)_{n\geq 0}$
(e) the sequence $(n^3)_{n\geq 0}$

(19) Suppose the recurrence satisfied by $s(n)_{n\geq 0}$ is large, so that repeated squaring involves the multiplication of large matrices. For sufficiently large n, is repeated squaring still faster than using the recurrence to compute $s(n)$? For which n should one use the recurrence, and for which n should one use repeated squaring?

Proofs.

(20) Write out the details of the proof that f is a bijection in Theorem 12.9.

(21) Modify the proof of Theorem 12.9 to prove Theorem 3.1 without using rabbit pairs.

CHAPTER 13

The base of the hierarchy

In this chapter we discuss several special cases of constant-recursive sequences. This will allow us to start exploring the hierarchy promised by the title of this book.

Periodic sequences

Our first special case consists of sequences that repeat after some fixed number of terms.

Definition 13.1. A sequence $s(n)_{n\geq 0}$ is *periodic* if there is an integer $\ell \geq 1$ such that $s(n+\ell) = s(n)$ for all $n \geq 0$. The minimal such ℓ is the *period length*. The ℓ-tuple $(s(0), s(1), \dots, s(\ell-1))$ is the *period*.

For example, the sequence $1, 1, 0, 1, 1, 0, \dots$ with period $(1, 1, 0)$ satisfies $s(n+3) = s(n)$. This sequence also satisfies $s(n+6) = s(n)$, but its period length is 3 since 3 is minimal. Periodic sequences are constant-recursive, since the recurrence $s(n+\ell) = s(n)$ satisfies the conditions of Definition 12.1.

Definition 13.2. A sequence $s(n)_{n\geq 0}$ is *eventually periodic* if there are integers $\ell \geq 1$ and $N \geq 0$ such that $s(n+\ell) = s(n)$ for all $n \geq N$.

For example, $5, 4, 3, 2, 1, 0, 1, 1, 0, 1, 1, 0, \dots$ satisfies $s(n+3) = s(n)$ for all $n \geq 4$.

Eventually periodic sequences arise in decimal expansions of rational numbers. For example, $\frac{1}{110} = .0090909\cdots$ and $\frac{3}{8} = .3750000\cdots = .3749999\cdots$. A real number x is rational if and only if the sequence of digits following the decimal point in its decimal expansion(s) is eventually periodic.

Eventually periodic sequences also arise in the following famous open problem.

Example 13.3. The *Collatz*[1] *function* is the function

$$C(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ 3n+1 & \text{if } n \text{ is odd.} \end{cases}$$

Iterating the usually produces an eventually periodic sequence. For example, starting with 6, we compute $C(6) = 3$, $C(3) = 10$, $C(10) = 5$, and so on to obtain the sequence

$$6, 3, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, \dots.$$

Try starting with different values. Eventually you'll reach 1 and enter the cycle $1, 4, 2, 1, 4, 2, \dots$, but no one has been able to prove that this happens for every starting value.

Are eventually periodic sequences constant-recursive?

[1] Lothar Collatz was born in 1910 in Arnsberg, Germany and died in 1990 in Varna, Bulgaria.

Example 13.4. Let $s(n)_{n\geq 0}$ be the sequence $3,7,5,2,7,2,7,2,7,\ldots$ satisfying $s(n+2)=s(n)$ for all $n\geq 3$. We'd like a recurrence that holds for all $n\geq 0$. Replacing n with $n+3$ gives

$$s(n+5)=0\cdot s(n+4)+s(n+3)+0\cdot s(n+2)+0\cdot s(n+1)+0\cdot s(n).$$

This recurrence holds for all $n\geq 0$, so $s(n)_{n\geq 0}$ is constant-recursive with rank 5.

This trick of delaying the recurrence at a cost of increasing the rank shows that every eventually periodic sequence is constant-recursive. In fact we can use the same trick to show that if a sequence is "eventually constant-recursive" then it is constant-recursive.

Proposition 13.5. *If there exists $r\geq 0$ and $N\geq 0$ such that $s(n)_{n\geq 0}$ satisfies $s(n+r)=\sum_{i=0}^{r-1}c_i s(n+i)$ for all $n\geq N$, then $s(n)_{n\geq 0}$ is constant-recursive.*

PROOF. Replacing n with $n+N$ and i with $i-N$ gives

$$\begin{aligned}s(n+N+r)&=\sum_{i=N}^{N+r-1}c_{i-N}s(n+i)\\&=\sum_{i=N}^{N+r-1}c_{i-N}s(n+i)+\sum_{i=0}^{N-1}0\cdot s(n+i)\end{aligned}$$

for all $n\geq 0$. □

Example 13.6. Let $s(n)_{n\geq 0}$ be a sequence satisfying $s(n+3)=2s(n+2)-s(n)$ for all $n\geq 2$. Shifting by 2 shows that

$$s(n+5)=2s(n+4)+0\cdot s(n+3)-s(n+2)+0\cdot s(n+1)+0\cdot s(n)$$

for all $n\geq 0$. Therefore $s(n)_{n\geq 0}$ is constant-recursive with rank ≤ 5.

Polynomial sequences

We already spent several chapters discussing polynomial sequences, but there's more to say! A natural question is this: Which polynomial sequences are also constant-recursive?

Example 13.7. Let $s(n)=3n+2$. Is $s(n)_{n\geq 0}$ constant-recursive? Its terms are

$$2,5,8,11,14,17,20,23,\ldots.$$

We can get each term from the previous term by adding 3, so $s(n+1)=s(n)+3$. But this recurrence does not fit Definition 12.1. It's fairly easy to check that no recurrence of the form $s(n+1)=c_0 s(n)$ will do the job. Suppose instead that there are some constants c_0 and c_1 such that $s(n+2)=c_0 s(n)+c_1 s(n+1)$. The system

$$\begin{aligned}8&=5c_1+2c_0\\11&=8c_1+5c_0\end{aligned}$$

has the solution $c_1=2, c_0=-1$, which corresponds to the recurrence

$$s(n+2)=2s(n+1)-s(n).$$

Moreover, this recurrence seems to hold for larger values of n as well. We can prove that this guess is correct by expanding $2s(n+1)-s(n)=2(3(n+1)+2)-(3n+2)=3n+8=3(n+2)+2=s(n+2)$. Therefore $(3n+2)_{n\geq 0}$ is constant-recursive.

We now find ourselves in the following awkward position. The sequence $(3n+2)_{n\geq 0}$ is simultaneously a polynomial sequence and a constant-recursive sequence. This means it has both a rank as a polynomial sequence and a rank as a constant-recursive sequence, so it might be ambiguous to refer its rank without additional context. However, we will show in Theorems 13.9 and 13.13 that the two ranks are equal when this happens, so it is safe to refer to *the* rank.

Example 13.8. Let's choose another linear polynomial. The sequence

$$3, 13, 23, 33, 43, 53, 63, 73, \ldots$$

is given by $s(n) = 10n + 3$. Is it constant-recursive? How can we write 23 as a linear combination of 13 and 3? How can we write 33 as a linear combination of 23 and 13? After some trial and error, we guess that

$$s(n+2) = 2s(n+1) - s(n),$$

which again can be proved by expanding. Curiously, this is the same recurrence as in Example 13.7!

In fact all polynomial sequences with rank r satisfy the *same recurrence*. Moreover, the coefficients in this recurrence are signed binomial coefficients.

Theorem 13.9. *If $s(n)_{n\geq 0}$ is a polynomial sequence with rank r, then $s(n)$ satisfies*

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

for all $n \geq 0$. In particular, $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq r$.

Proof. Since $s(n)_{n\geq 0}$ is a polynomial sequence with rank r, we have $\Delta_n^r s(n) = 0$ by Proposition 10.2. The claimed recurrence now follows from Theorem 12.11, so $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq r$. $\square$

Theorem 13.9 implies that every polynomial $s(n) = an + b$ satisfies

$$s(n+2) - 2s(n+1) + s(n) = 0.$$

Similarly, every polynomial $s(n) = an^2 + bn + c$ satisfies

$$s(n+3) - 3s(n+2) + 3s(n+1) - s(n) = 0,$$

every polynomial $s(n) = an^3 + bn^2 + cn + d$ satisfies

$$s(n+4) - 4s(n+3) + 6s(n+2) - 4s(n+1) + s(n) = 0,$$

and so on.

Example 13.10. It is now easy to write down a recurrence satisfied by $s(n) = n^3$. Its rank is 4, so

$$s(n+4) = 4s(n+3) - 6s(n+2) + 4s(n+1) - s(n).$$

As a check, let's use this recurrence to compute $s(4) = 4^3$. We have $s(0) = 0$, $s(1) = 1$, $s(2) = 8$, and $s(3) = 27$, so

$$s(4) = 4 \cdot 27 - 6 \cdot 8 + 4 \cdot 1 - 0 = 64.$$

In fact, not only does every rank-r polynomial sequence satisfy the same recurrence, but every sequence in $\mathsf{Poly}(r)$ satisfies the same recurrence — whether the sequence has rank r or less than r.

Corollary 13.11. *Every sequence in* $\mathsf{Poly}(r)$ *satisfies the recurrence in Theorem 13.9.*

PROOF. This follows from the proof of Theorem 13.9, since the only property of $s(n)_{n\geq 0}$ we used is $\Delta_n^r s(n) = 0$. □

We would like to know whether the recurrence in Theorem 13.9 is in fact minimal. This turns out to be surprisingly tricky.

Example 13.12. Let $s(n) = a_0 + a_1 n + a_2 n^2 + a_3 n^3$ with $a_3 \neq 0$. If the rank of $s(n)_{n\geq 0}$ as a constant-recursive sequence is ≤ 3, then there are some rational numbers c_0, c_1, c_2, c_3, not all 0, such that

$$c_0 s(n) + c_1 s(n+1) + c_2 s(n+2) + c_3 s(n+3) = 0$$

for all $n \geq 0$. Suppose a relation of this form does hold, and we will try to solve for c_0, c_1, c_2, c_3. Expand the left side as a polynomial in n; then each coefficient must be equal to 0. For the coefficient of n^0, this implies

$$\begin{aligned}a_0(c_0 + c_1 + c_2 + c_3) &+ a_1(c_1 + 2c_2 + 3c_3)\\ &+ a_2(c_1 + 4c_2 + 9c_3) + a_3(c_1 + 8c_2 + 27c_3) = 0.\end{aligned}$$

Similarly, for the coefficients of n^1, n^2, and n^3 we have

$$\begin{aligned}a_1(c_0 + c_1 + c_2 + c_3) + 2a_2(c_1 + 2c_2 + 3c_3) + 3a_3(c_1 + 4c_2 + 9c_3) &= 0\\ a_2(c_0 + c_1 + c_2 + c_3) + 3a_3(c_1 + 2c_2 + 3c_3) &= 0\\ a_3(c_0 + c_1 + c_2 + c_3) &= 0.\end{aligned}$$

This is a large system of equations, but it suggests a strategy. Define

$$\begin{aligned}C_0 &= c_0 + c_1 + c_2 + c_3\\ C_1 &= c_1 + 2c_2 + 3c_3\\ C_2 &= c_1 + 4c_2 + 9c_3\\ C_3 &= c_1 + 8c_2 + 27c_3\end{aligned}$$

since these quantities occur several times. Then we can write the system as

$$\begin{aligned}a_0C_0 + a_1C_1 + a_2C_2 + a_3C_3 &= 0\\ a_1C_0 + 2a_2C_1 + 3a_3C_2 &= 0\\ a_2C_0 + 3a_3C_1 &= 0\\ a_3C_0 &= 0.\end{aligned}$$

Working from the bottom to the top, these equations successively imply $C_0 = 0$, $C_1 = 0$, $C_2 = 0$, and $C_3 = 0$. This in turn implies

$$\begin{bmatrix}0\\0\\0\\0\end{bmatrix} = \begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 2 & 3\\ 0 & 1 & 4 & 9\\ 0 & 1 & 8 & 27\end{bmatrix}\begin{bmatrix}c_0\\c_1\\c_2\\c_3\end{bmatrix}.$$

The coefficient matrix is the transpose of a Vandermonde matrix, and it is invertible by Proposition 6.2. Therefore $c_0 = c_1 = c_2 = c_3 = 0$, so $\operatorname{rank}(s) \geq 4$. Theorem 13.9 now implies $\operatorname{rank}(s) = 4$.

The previous example suggests that the coefficients c_0, c_1, c_2, c_3 of the recurrence are not the most natural parameters for the question of minimality. We'll use a change of variables to prove the general case, using a difference equation rather than a recurrence. Since Δ_n behaves nicely when working in the binomial coefficient basis, we use can further simplify the argument by writing $s(n)$ using binomial coefficients.

Theorem 13.13. *If $s(n)_{n\geq 0}$ is a polynomial sequence with rank r, then the rank of $s(n)_{n\geq 0}$ as a constant-recursive sequence is also r.*

Proof. By Theorem 13.9, $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq r$. To show that this rank is exactly r, we show that $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, \dots, s(n+r-1)_{n\geq 0}$ are linearly independent. Write $s(n) = b_0\binom{n}{0} + b_1\binom{n}{1} + \cdots + b_{r-1}\binom{n}{r-1}$ with $b_{r-1} \neq 0$, and suppose $c_0, c_1, \dots, c_{r-1}$ are rational numbers such that $0 = \sum_{m=0}^{r-1} c_m s(n+m)$ for all $n \geq 0$. By Theorem 12.12, there are rational numbers $d_0, d_1, \dots, d_{r-1}$ such that $0 = \sum_{m=0}^{r-1} d_m (\Delta^m s)(n)$ for all $n \geq 0$. Rewrite this as

$$\begin{aligned} 0 &= \sum_{m=0}^{r-1} d_m \left(b_m \binom{n}{0} + b_{m+1}\binom{n}{1} + \cdots + b_{r-1}\binom{n}{r-1-m}\right) \\ &= \sum\sum_{0\leq m\leq i\leq r-1} b_i d_m \binom{n}{i-m}; \end{aligned}$$

writing the double sum like this lets us avoid changing the order of summation twice in the next step. We would like to collect all the terms involving $\binom{n}{j}$ together. Performing the change of variables $m = i - j$ shows that

$$\begin{aligned} 0 &= \sum\sum_{0\leq i-j\leq i\leq r-1} b_i d_{i-j}\binom{n}{j} \\ &= \sum_{j=0}^{r-1}\left(\sum_{i=j}^{r-1} b_i d_{i-j}\right)\binom{n}{j}. \end{aligned}$$

Since this equation holds for all $n \geq 0$, the coefficient of each $\binom{n}{j}$ is equal to 0. Therefore we have the system

$$\begin{aligned} b_0 d_0 + b_1 d_1 + \cdots + b_{r-2}d_{r-2} + b_{r-1}d_{r-1} &= 0 \\ &\vdots \\ b_{r-3}d_0 + b_{r-2}d_1 + b_{r-1}d_2 &= 0 \\ b_{r-2}d_0 + b_{r-1}d_1 &= 0 \\ b_{r-1}d_0 &= 0. \end{aligned}$$

Since $b_{r-1} \neq 0$, we have $d_0 = 0, d_1 = 0, \dots, d_{r-1} = 0$. This implies $c_0 = 0, c_1 = 0, \dots, c_{r-1} = 0$. Therefore $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, \dots, s(n+r-1)_{n\geq 0}$ are linearly independent. $\square$

In light of Theorem 13.9, a final consequence of Proposition 13.5 is that "eventually polynomial" sequences are constant-recursive. That is, let $f(x) \in \mathbb{Q}[x]$, let $s(0), s(1), \dots, s(N-1)$ be arbitrary rational numbers, and define $s(n) = f(n)$ for all $n \geq N$; then $s(n)_{n\geq 0}$ is constant-recursive.

Quasi-polynomial sequences

In Theorem 5.3 we showed that polynomial sequences are closed under addition and multiplication. It's not too difficult to convince yourself that periodic sequences (and also eventually periodic sequences) have this property as well. But what if we start combining polynomial sequences with periodic (or eventually periodic) sequences? Then we may end up with sequences that are neither polynomial nor periodic.

Example 13.14. Let $t(n) = 3n + 2$, and let

$$u(n) = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd.} \end{cases}$$

Then $(t(n) + u(n))_{n \geq 0}$ is

$$2, 6, 8, 12, 14, 18, 20, 24, 26, 30, 32, 36, 38, 42, 44, 48, \ldots \qquad \text{(shift of A047238).}$$

Clearly this sequence is not periodic. Since its difference sequences never become the 0 sequence, it's also not polynomial. It has a simple formula, however:

$$t(n) + u(n) = \begin{cases} 3n + 2 & \text{if } n \text{ is even} \\ 3n + 3 & \text{if } n \text{ is odd.} \end{cases}$$

Sequences like this consisting of polynomials patched together in a periodic way are called *quasi-polynomial sequences*.

Definition 13.15. A sequence $s(n)_{n\geq0}$ is *quasi-polynomial* if there is an integer $\ell \geq 1$ such that $s(\ell n)_{n\geq0}$, $s(\ell n + 1)_{n\geq0}$, $\ldots$, $s(\ell n + \ell - 1)_{n\geq0}$ are all polynomial sequences.

Example 13.16. Let

$$s(n) = \begin{cases} 3n + 2 & \text{if } n \text{ is even} \\ 3n + 3 & \text{if } n \text{ is odd.} \end{cases}$$

The sequence $s(2n)_{n\geq0}$ of even-indexed terms is $2, 8, 14, 20, \ldots$ and is given by the formula $s(2n) = 6n + 2$. The sequence $s(2n + 1)_{n\geq0}$ of odd-indexed terms $6, 12, 18, 24, \ldots$ and is given by $s(2n + 1) = 6n + 6$. Therefore $s(n)_{n\geq0}$ satisfies Definition 13.15 with $\ell = 2$.

Example 13.17. Let $a(n)_{n\geq0}$ be the periodic sequence with period $(1, 1, 0)$, let $b(n)_{n\geq0}$ be the (constant) periodic sequence with period (3), and let $c(n)_{n\geq0}$ be the periodic sequence with period $(2, 1)$. Define $s(n) = a(n) \cdot n^2 + b(n) \cdot n + c(n)$. The sequence $s(n)_{n\geq0}$ is

$$2, 5, 8, 19, 30, 16, 56, 71, 26, 109, 132, 34, 182, 209, 44, 271, 306, 52, \ldots.$$

It is a quasi-polynomial sequence with $\ell = 6$.

Example 13.18. We saw another quasi-polynomial function earlier in this chapter. The Collatz function from Example 13.3 is

$$C(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ 3n + 1 & \text{if } n \text{ is odd.} \end{cases}$$

The sequence $C(n)_{n\geq 0}$ obtained by evaluating $C(n)$ at each $n \geq 0$ (rather than iterating) is quasi-polynomial:

$$0, 4, 1, 10, 2, 16, 3, 22, 4, 28, 5, 34, 6, 40, 7, 46, \ldots \qquad (\mathsf{A006370}).$$

Quasi-polynomial sequences simultaneously generalize periodic sequences and sequences: Every periodic sequence is a quasi-polynomial sequence, and every polynomial sequence is a quasi-polynomial sequence. They also possess closure properties.

Theorem 13.19. *Quasi-polynomial sequences are closed under addition and multiplication.*

Proof. Let $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ be quasi-polynomial sequences. Let ℓ_1 and ℓ_2 be integers such that $s(\ell_1 n+i)_{n\geq 0}$ and $t(\ell_2 n+j)_{n\geq 0}$ are polynomial sequences for all $i \in \{0,1,\ldots,\ell_1-1\}$ and $j \in \{0,1,\ldots,\ell_2-1\}$. Then in fact $s(\ell_1 n+i)_{n\geq 0}$ and $t(\ell_2 n+j)_{n\geq 0}$ are polynomial sequences for all $i \geq 0$ and $j \geq 0$. Let $\ell = \operatorname{lcm}(\ell_1, \ell_2)$. Then $(s(\ell n+i)+t(\ell n+i))_{n\geq 0}$ and $(s(\ell n+i)t(\ell n+i))_{n\geq 0}$ are polynomial sequences for all $i \in \{0,1,\ldots,\ell-1\}$. Therefore $(s(n)+t(n))_{n\geq 0}$ and $(s(n)t(n))_{n\geq 0}$ are quasi-polynomial sequences. $\square$

Like periodic sequences and polynomial sequences, quasi-polynomial sequences are constant-recursive.

Theorem 13.20. *If $s(n)_{n\geq 0}$ is a quasi-polynomial sequence, then $s(n)_{n\geq 0}$ is a constant-recursive sequence.*

Proof. Let $\ell \geq 1$ be an integer such that $s(\ell n+i)_{n\geq 0}$ is a polynomial sequence for each $i \in \{0,1,\ldots,\ell-1\}$. Let

$$r = \max_{0\leq i\leq \ell-1} \operatorname{rank}(s(\ell n+i)_{n\geq 0})$$

be the maximum rank among the polynomial sequences $s(\ell n+i)_{n\geq 0}$. By Corollary 13.11, for each $i \in \{0,1,\ldots,\ell-1\}$ the sequence $s(\ell n+i)_{n\geq 0}$ satisfies the recurrence

$$\sum_{j=0}^{r} (-1)^{r-j}\binom{r}{j} s(\ell(n+j)+i) = 0$$

for all $n \geq 0$. We rewrite this recurrence as

$$\sum_{j=0}^{r} (-1)^{r-j}\binom{r}{j} s(\ell n+i+\ell j) = 0.$$

Since every non-negative integer has a representation as $\ell n+i$ for some $n \geq 0$ and $i \in \{0,1,\ldots,\ell-1\}$, we have

$$\sum_{m=0}^{r} (-1)^{r-m}\binom{r}{m} s(n+\ell m) = 0$$

for all $n \geq 0$. Substituting $m = \frac{j}{\ell}$ "inflates" this recurrence, giving

$$\sum_{j=0}^{\ell r} c_j s(n+j) = 0$$

where

$$c_j = \begin{cases} (-1)^{r-j/\ell}\binom{r}{j/\ell} & \text{if } j \text{ is divisible by } \ell \\ 0 & \text{if } j \text{ is not divisible by } \ell. \end{cases}$$

Therefore $s(n)_{n\geq 0}$ is constant-recursive. □

Notice that the proof gives an upper bound of ℓr on the rank of the quasi-polynomial sequence.

We conclude with some additional examples of quasi-polynomial sequences.

Example 13.21. The *floor function* $\lfloor x \rfloor$ is defined to be the greatest integer m such that $m \leq x$. The sequence $\lfloor n/2 \rfloor_{n\geq 0}$ is

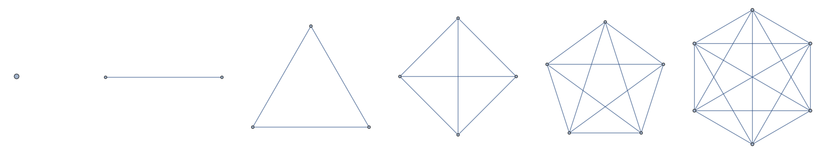
$$0, 0, 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7, \dots \qquad (\text{A004526}).$$

Similarly, the *ceiling function* $\lceil x \rceil$ is the least integer m such that $x \leq m$. The sequence $\lceil n/2 \rceil_{n\geq 0}$ is

$$0, 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7, 8, \dots \qquad (\text{A110654}).$$

Both of these sequences are quasi-polynomial sequences with $\ell = 2$.

A much more complicated quasi-polynomial sequence arises from regular polygons. In how many points do the diagonals of a regular n-sided polygon intersect?



The diagonals of a square intersect in 5 points (including the vertices). The diagonals of a regular pentagon intersect in 10 points. The general sequence is

$$0, 1, 2, 3, 5, 10, 19, 42, 57, 135, 171, 341, 313, 728, 771, 1380, \dots \qquad (\text{A007569}).$$

If we adjust the first few terms, we obtain a quasi-polynomial sequence.

Theorem 13.22 (Poonen–Rubinstein 1998 [**18**])**.** *Define $s(0) = 1$, $s(1) = 1$, and $s(2) = 3$. For each $n \geq 3$, let $s(n)$ be the number of intersection points made by the diagonals of a regular n-sided polygon. The sequence $s(n)_{n\geq 0}$ is a quasi-polynomial sequence comprised of $\ell = 2520$ polynomials.*

Questions

Computations.

(1) Let $s(n) = n^2$ for all $n \neq 5$, and let $s(5) = 20$. Compute a recurrence for $s(n)_{n\geq 0}$ that holds for all $n \geq 0$.

Experiments.

(2) (a) Find a formula of the form $a+b\,(-1)^n$ for the nth term of the periodic sequence $3, 8, 3, 8, 3, 8, \dots$.
    (b) Let $s(n)_{n\geq 0}$ be a periodic sequence with period length 2. Is there always a formula for $s(n)$ of the form $a + b\,(-1)^n$?

(3) For which real numbers ω is the sequence $(\sin(\omega n))_{n\geq 0}$ periodic?

(4) If $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ are two periodic sequences of non-negative integers, what can you say about the composition $s(t(n))_{n\geq 0}$? For example, is it periodic?
(5) Fix an integer $m \geq 1$. Consider the class of periodic sequences with period length dividing m. What closure properties does this class have?
(6) Let $s(n)_{n\geq 0}$ be periodic with period length ℓ and $t(n)_{n\geq 0}$ be periodic with period length k. Are the following sequences periodic? If so, what's the period length? (Let $a, b \in \mathbb{Z}$.)
   (a) Addition by a constant: $(s(n) + a)_{n\geq 0}$.
   (b) Difference sequence: $(s(n+1) - s(n))_{n\geq 0}$.
   (c) Subsequence of periodic indexing: $s(an+b)_{n\geq 0}$.
   (d) Composition: $s(t(n))_{n\geq 0}$.
(7) Let $s(n)_{n\geq 0}$ be periodic with period length ℓ. How many (distinct) sequences are in the set
$$\{s(n+i)_{n\geq 0} : i \geq 0\}?$$
(8) To compute the nth term of a polynomial sequence, which is faster — evaluating the polynomial or using the recurrence guaranteed by Theorem 13.9, as in Example 13.10?
(9) (a) The terms of a sequence $s(n)_{n\geq 0}$ are
$$14, 18, 20, 22, 26, 34, 38, 40, 42, 46, 54, 58, 60, 62, 66, 74, \dots.$$
   Compute terms of its difference sequence $\Delta_n s(n)$, and use them to guess a formula for $s(n)$.
   (b) Under what conditions is the difference sequence of a quasi-polynomial sequence periodic?
(10) (a) Let $C(n)$ be the Collatz function from Example 13.18. Find formulas for $\Delta_n C(n)$, $\Delta_n^2 C(n)$, and $\Delta_n^3 C(n)$.
   (b) If $s(n)_{n\geq 0}$ is a polynomial sequence with rank r, then $\Delta_n^r s(n) = 0$. Is there an analogous statement for quasi-polynomial sequences?
(11) Let $m \geq 2$ be an integer. Is there a formula for the nth positive integer that is not divisible by m?
(12) Is $\lfloor f(n) \rfloor_{n\geq 0}$ a quasi-polynomial sequence for every polynomial $f(x) \in \mathbb{Q}[x]$?
(13) Is the composition $s(t(n))$ of two quasi-polynomials a quasi-polynomial?
(14) What about compositions of polynomials, floor, and ceiling? Is the resulting function always quasi-polynomial?
(15) What is the set of polynomial sequences $s(n)_{n\geq 0}$ that satisfy the Fibonacci recurrence $s(n+2) = s(n+1) + s(n)$ for all $n \geq 0$?

Proofs.

(16) Let S be a finite set, let $a \in S$, and let $f\colon S \to S$. Prove that $(f^n(a))_{n\geq 0}$ is eventually periodic.
(17) Prove that if $s(n)_{n\geq 0}$ is periodic then $s(n)_{n\geq 0}$ is quasi-polynomial.
(18) Prove that $s(n)_{n\geq 0}$ is eventually periodic if and only if $s(n)_{n\geq 0}$ is constant-recursive and $\{s(n) : n \geq 0\}$ is finite.
(19) Prove that if a polynomial sequence $s(n)_{n\geq 0}$ is eventually periodic then $s(n) = c$ is constant.

(20) Prove that periodic sequences are closed under addition and multiplication.

(21) Prove that eventually periodic sequences are closed under addition and multiplication.

(22) A sequence $s(n)_{n\geq 0}$ is *quasi-quasi-polynomial* if there is an integer $\ell \geq 1$ such the sequence $s(\ell n+i)_{n\geq 0}$ is quasi-polynomial for each $i \in \{0, 1, \ldots, \ell-1\}$. Is the class of quasi-quasi-polynomial sequences strictly larger than the class of quasi-polynomial sequences?

Artwork.

(23) Draw an Euler diagram showing the relationships between the classes of constant sequences, polynomial sequences, periodic sequences, quasi-polynomial sequences, and constant-recursive sequences.

CHAPTER 14

A sequence's siblings

To fully understand a mathematical object, we should understand it in the context of related objects. For the Fibonacci sequence, a natural context is the set of all sequences satisfying $s(n+2) = s(n+1) + s(n)$.

Definition 14.1. Let $s(n)_{n\geq 0}$ be a constant-recursive sequence with rank r. Let

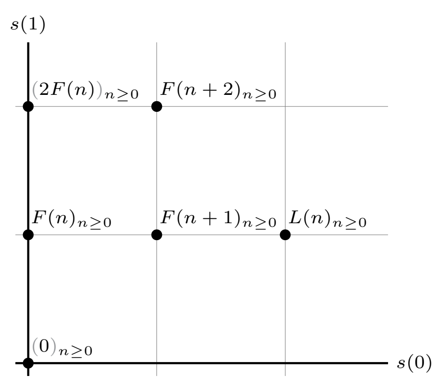
$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n) \tag{14.1}$$

be the minimal recurrence satisfied by $s(n)_{n\geq 0}$. The *set of siblings* of $s(n)_{n\geq 0}$ is the set of all sequences of rational numbers satisfying Equation (14.1) for all $n \geq 0$. We denote it by $\mathsf{Siblings}(s)$.

For the Fibonacci sequence $F(n)_{n\geq 0}$, every pair of initial conditions generates a sequence in $\mathsf{Siblings}(F)$. Here are some examples:

$s(0)$	$s(1)$	sequence	
0	0	$0, 0, 0, 0, 0, 0, 0, 0, \dots$	
0	1	$0, 1, 1, 2, 3, 5, 8, 13, \dots$	Fibonacci
0	2	$0, 2, 2, 4, 6, 10, 16, 26, \dots$	
1	0	$1, 0, 1, 1, 2, 3, 5, 8, \dots$	
1	1	$1, 1, 2, 3, 5, 8, 13, 21, \dots$	
1	2	$1, 2, 3, 5, 8, 13, 21, 34, \dots$	
2	0	$2, 0, 2, 2, 4, 6, 10, 16, \dots$	
2	1	$2, 1, 3, 4, 7, 11, 18, 29, \dots$	Lucas
2	2	$2, 2, 4, 6, 10, 16, 26, 42, \dots$	

The initial conditions $2, 1$ generate the Lucas sequence, defined in Example 12.4. We can identify each sequence $s(n)_{n\geq 0} \in \mathsf{Siblings}(F)$ with the point in the plane whose coordinates are $(s(0), s(1))$. For example, $F(n)_{n\geq 0}$ is identified with the point $(0, 1)$. Here are the locations of several sequences in the plane:



The previous table and graphic suggest several directions we will pursue in this chapter. One observation is that the initial conditions $0, 2$ seem to generate the sequence $(2F(n))_{n\geq 0}$. This makes sense, since $0 = 2F(0)$ and $2 = 2F(1)$, and the Fibonacci recurrence implies $2F(n+2) = 2F(n+1) + 2F(n)$. More generally, we have the following.

Proposition 14.2. *If $t(n)_{n\geq 0}$ is a constant-recursive sequence and $a \in \mathbb{Q}$, then $(a\,t(n))_{n\geq 0}$ belongs to $\mathsf{Siblings}(t)$.*

PROOF. Multiply both sides of the recurrence for $t(n)_{n\geq 0}$ by a. □

Another observation is that the sequence with initial conditions $2, 2$ is the sum of the sequences with initial conditions $0, 2$ and $2, 0$. This is a special case of the following result.

Proposition 14.3. *If $t(n)_{n\geq 0}$ and $u(n)_{n\geq 0}$ are constant-recursive sequences satisfying the same recurrence, then $(t(n) + u(n))_{n\geq 0}$ is constant-recursive and also satisfies that recurrence.*

PROOF. Let $s(n) = t(n) + u(n)$. Suppose

$$t(n+r) = c_{r-1}t(n+r-1) + \cdots + c_1 t(n+1) + c_0 t(n)$$

and

$$u(n+r) = c_{r-1}u(n+r-1) + \cdots + c_1 u(n+1) + c_0 u(n)$$

for all $n \geq 0$. Adding these equations gives

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n). \qquad \square$$

Those propositions aren't impressive on their own, but they have an implication about the difference operator Δ_n from Chapter 10. If $s(n)_{n\geq 0}$ is constant-recursive, then $(\Delta_n s(n))_{n\geq 0} \in \mathsf{Siblings}(s)$. More generally, they give us the following important result.

Theorem 14.4. *If $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank r, then $\mathsf{Siblings}(s)$ is a vector space with dimension r.*

PROOF. Let $c_0, c_1, \ldots, c_{r-1}$ be rational numbers such that

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n) \tag{14.2}$$

for all $n \geq 0$. Let $V = \mathsf{Siblings}(s)$. We check the conditions in Definition 7.1. By Proposition 14.3, V is closed under addition. By Proposition 14.2, V is closed under multiplication by a rational number. The zero sequence $(0)_{n\geq 0}$ is an element of V since it satisfies the defining recurrence (14.2) for all $n \geq 0$. If $t(n)_{n\geq 0} \in V$ then $-(t(n)_{n\geq 0}) = (-t(n))_{n\geq 0} \in V$ by Proposition 14.2. As in the proof of Theorem 7.4, the remaining vector space axioms follow from properties of addition and multiplication of rational numbers.

To show that $\dim V = r$, we establish a basis of V. For each i in the range $0 \leq i \leq r-1$, define $e_i(n)_{n\geq 0}$ to be the sequence in V whose initial conditions are given by

$$e_i(n) = \begin{cases} 1 & \text{if } n = i \\ 0 & \text{if } 0 \leq n \leq r-1 \text{ and } n \neq i. \end{cases} \tag{14.3}$$

These are analogous to the sequences comprising the Lagrange basis of $\mathsf{Poly}(r)$ in Definition 7.9. We claim that

$$B = (e_0(n)_{n\geq 0}, e_1(n)_{n\geq 0}, \ldots, e_{r-1}(n)_{n\geq 0})$$

is a basis of V. These r sequences are linearly independent, since their initial conditions imply that if the ith term of

$$a_0 \cdot \left(e_0(n)_{n\geq 0}\right) + a_1 \cdot \left(e_1(n)_{n\geq 0}\right) + \cdots + a_{r-1} \cdot \left(e_{r-1}(n)_{n\geq 0}\right)$$

is 0 for some i in the range $0 \leq i \leq r-1$ then $a_i = 0$. Moreover, every sequence in V can be written as a linear combination of the sequences in B; namely, if $t(n)_{n\geq 0} \in V$ then

$$t(n)_{n\geq 0} = t(0) \cdot \left(e_0(n)_{n\geq 0}\right) + t(1) \cdot \left(e_1(n)_{n\geq 0}\right) + \cdots + t(r-1) \cdot \left(e_{r-1}(n)_{n\geq 0}\right)$$

since the right side satisfies Recurrence (14.2) by Propositions 14.2 and 14.3 and has the same initial conditions as $t(n)_{n\geq 0}$. Therefore B is a basis of V, so $\dim V = r$. $\square$

The proof of Theorem 14.4 establishes a standard basis of $\mathsf{Siblings}(s)$ consisting of the r sequences $e_i(n)_{n\geq 0}$, defined by Equation (14.3).

Example 14.5. For the Fibonacci sequence, let

$$V = \mathsf{Siblings}(F) = \{s(n)_{n\geq 0} : s(n+2) = s(n+1) + s(n) \text{ for all } n \geq 0\},$$

which has dimension 2. The standard basis of V consists of $e_0(n)_{n\geq 0} = 1, 0, 1, 1, 2, 3, 5, 8, \ldots$, which is the Fibonacci sequence with an extra 1 term at the beginning, and $e_1(n)_{n\geq 0} = 0, 1, 1, 2, 3, 5, 8, 13, \ldots$, which is the Fibonacci sequence itself. Every sequence $s(n)_{n\geq 0} \in V$ is can be written as a linear combination of these basis sequences in a unique way. Moreover, the coordinates of $s(n)_{n\geq 0}$ in the basis $(e_0(n)_{n\geq 0}, e_1(n)_{n\geq 0})$ are $(s(0), s(1))$. For example, let $s(n)_{n\geq 0} \in V$ be the sequence with initial conditions $s(0) = 3$ and $s(1) = 1$. Its terms are $3, 1, 4, 5, 9, 14, 23, 37, \ldots$ (A104449). Its coordinates of $s(n)_{n\geq 0}$ are given by its initial conditions, namely $(3, 1)$.

Example 14.6. Another basis of $\mathsf{Siblings}(F)$ is $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$. Let $s(n)_{n\geq 0}$ be the sequence in Example 14.5 with initial conditions $s(0) = 3$ and $s(1) = 1$. To write $s(n)_{n\geq 0}$ in the basis $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$, it suffices to solve the system

$$\begin{aligned} 3 &= aF(0) + bL(0) \\ 1 &= aF(1) + bL(1). \end{aligned}$$

The result is that $s(n) = -\frac{1}{2}F(n) + \frac{3}{2}L(n)$ for all $n \geq 0$. Therefore the coordinates of $s(n)_{n\geq 0}$ in the basis $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$ are $(-\frac{1}{2}, \frac{3}{2})$. Let us revisit the table from the beginning of the chapter and include the coordinates of some of its sequences in the basis $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$.

sequence	coordinates in $(e_0(n)_{n\geq 0}, e_1(n)_{n\geq 0})$	coordinates in $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$
$0, 0, 0, 0, 0, \ldots$	$(0, 0)$	$(0, 0)$
$0, 1, 1, 2, 3, \ldots$	$(0, 1)$	$(1, 0)$
$1, 0, 1, 1, 2, \ldots$	$(1, 0)$	$(-\frac{1}{2}, \frac{1}{2})$
$1, 1, 2, 3, 5, \ldots$	$(1, 1)$	$(\frac{1}{2}, \frac{1}{2})$
$2, 0, 2, 2, 4, \ldots$	$(2, 0)$	$(-1, 1)$
$2, 1, 3, 4, 7, \ldots$	$(2, 1)$	$(0, 1)$

Shifts of the Fibonacci sequence

The final observation about the table at the beginning of this chapter is that several of the sequences in $\mathsf{Siblings}(F)$ are shifts of each other. For example, the Fibonacci sequence with its first term removed appears with initial conditions $1, 1$. The first few shifts of the Fibonacci sequence are as follows.

$$\begin{aligned} F(n)_{n\geq 0}: &\quad 0, 1, 1, 2, 3, 5, 8, 13, \ldots \\ F(n+1)_{n\geq 0}: &\quad 1, 1, 2, 3, 5, 8, 13, 21, \ldots \\ F(n+2)_{n\geq 0}: &\quad 1, 2, 3, 5, 8, 13, 21, 34, \ldots \\ F(n+3)_{n\geq 0}: &\quad 2, 3, 5, 8, 13, 21, 34, 55, \ldots \end{aligned}$$

Every shift of the Fibonacci sequence satisfies the Fibonacci recurrence.

Proposition 14.7. *If $s(n)_{n\geq 0}$ is a constant-recursive sequence and $i \geq 0$, then $s(n+i)_{n\geq 0} \in \mathsf{Siblings}(s)$.*

Proof. Let

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n)$$

be a recurrence satisfied by $s(n)_{n\geq 0}$. Since this recurrence holds for all $n \geq 0$, replacing n with $n+i$ shows that

$$s(n+i+r) = c_{r-1}s(n+i+r-1) + \cdots + c_1 s(n+i+1) + c_0 s(n+i)$$

holds for all integers n and i satisfying $n+i \geq 0$. In particular, for every $i \geq 0$ this recurrence holds for all $n \geq 0$. Therefore $s(n+i)_{n\geq 0}$ is an element of $\mathsf{Siblings}(s)$. $\square$

Not only do shifts of $F(n)_{n\geq 0}$ belong to $\mathsf{Siblings}(F)$, but we can from a basis from them, so that every sequence in $\mathsf{Siblings}(F)$ can be written as a linear combination of shifts of $F(n)_{n\geq 0}$. Namely, $F(n)_{n\geq 0}$ and $F(n+1)_{n\geq 0}$ form a basis since they are linearly independent.

Example 14.8. Let $s(n)_{n\geq 0} \in \mathsf{Siblings}(F)$ be the sequence with $s(0) = 3$ and $s(1) = 1$; we compute the coordinates of $s(n)_{n\geq 0}$ in the basis $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$ to be $(-2, 3)$. More generally, we can compute the coordinates of every sequence in the table from the beginning of the chapter in this basis.

The next example is mostly for fun but will turn out to be quite useful.

Example 14.9. By Proposition 14.7, each shift of $F(n)_{n\geq 0}$ belongs to $\mathsf{Siblings}(F)$. Since $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$ is a basis of $\mathsf{Siblings}(F)$, this implies that every shift of $F(n)_{n\geq 0}$ is a linear combination of $F(n)_{n\geq 0}$ and $F(n+1)_{n\geq 0}$. We can compute the coefficients in this linear combination using the Fibonacci recurrence itself. Namely, $F(n+2) = F(n+1) + F(n)$ says that the 2nd shift $F(n+2)_{n\geq 0}$ is the sum of the 0th and 1st shifts. Terms of the 3rd shift can be written

$$\begin{aligned} F(n+3) &= F(n+2) + F(n+1) \\ &= (F(n) + F(n+1)) + F(n+1) \\ &= 2F(n+1) + F(n). \end{aligned}$$

This implies that the 4th shift is given by

$$\begin{aligned}F(n+4) &= F(n+3) + F(n+2)\\ &= \big(2F(n+1) + F(n)\big) + \big(F(n+1) + F(n)\big)\\ &= 3F(n+1) + 2F(n).\end{aligned}$$

The 5th shift is

$$\begin{aligned}F(n+5) &= F(n+4) + F(n+3)\\ &= \big(3F(n+1) + 2F(n)\big) + \big(2F(n+1) + F(n)\big)\\ &= 5F(n+1) + 3F(n).\end{aligned}$$

One can prove the addition formula

$$F(n+i) = F(i)F(n+1) + F(i-1)F(n)$$

for each $i \geq 1$ by induction. Therefore the coordinates of $F(n+i)_{n\geq 0}$ in the basis $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$ are $(F(i-1), F(i))$. If we define $F(-1) = 1$ then this also holds for $i = 0$.

A natural basis

In a sense we are lucky that the initial conditions of the Fibonacci sequence are 0 and 1, since this causes the Fibonacci sequence to be one of the standard basis sequences of $\mathsf{Siblings}(F)$. But other sequences do not have initial conditions consisting of 0s along with a single 1, so their coordinates in the standard basis are not particularly simple. For example, $L(n)_{n\geq 0}$ is not one of the standard basis elements of $\mathsf{Siblings}(L) = \mathsf{Siblings}(F)$. Fortunately, we will see that the shifts of a sequence provide a natural basis in general.

Example 14.10. Consider the first two shifts of $L(n)_{n\geq 0}$.

$$\begin{aligned}L(n)_{n\geq 0}: &\quad 2, 1, 3, 4, 7, 11, 18, 29, \dots\\ L(n+1)_{n\geq 0}: &\quad 1, 3, 4, 7, 11, 18, 29, 47, \dots\end{aligned}$$

These sequences are linearly independent, so they form a basis of $\mathsf{Siblings}(L)$.

Example 14.11. Define $s(n)_{n\geq 0}$ by $s(0) = 2$, $s(1) = -1$, $s(2) = 3$, and $s(n+3) = s(n+2) + s(n+1) + s(n)$ for all $n \geq 0$. The first three shifts of $s(n)_{n\geq 0}$ are as follows.

$$\begin{aligned}s(n)_{n\geq 0}: &\quad 2, -1, 3, 4, 6, 13, 23, 42, \dots\\ s(n+1)_{n\geq 0}: &\quad -1, 3, 4, 6, 13, 23, 42, 78, \dots\\ s(n+2)_{n\geq 0}: &\quad 3, 4, 6, 13, 23, 42, 78, 143, \dots\end{aligned}$$

They are linearly independent and therefore form a basis of $\mathsf{Siblings}(s)$.

To show that the first few shifts of a constant-recursive sequence $s(n)_{n\geq 0}$ comprise a basis of $\mathsf{Siblings}(s)$, we introduce the following concept.

Definition 14.12. The *shift space* of a sequence $s(n)_{n\geq 0}$, denoted by $\mathsf{ShiftSpace}(s)$, is the vector space generated by the shifts of $s(n)_{n\geq 0}$. That is, $\mathsf{ShiftSpace}(s)$ is the smallest vector space containing $s(n+i)_{n\geq 0}$ for all $i \geq 0$.

For the Fibonacci sequence, $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$ is a basis of $\mathsf{ShiftSpace}(F)$, since we saw in Example 14.9 that each shift $F(n+i)_{n\geq 0}$ can be written as a linear combination of $F(n)_{n\geq 0}$ and $F(n+1)_{n\geq 0}$. Since $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$ is also a basis of $\mathsf{Siblings}(F)$, this means that $\mathsf{Siblings}(F)$ and $\mathsf{ShiftSpace}(F)$ are in fact the same vector space. This is true in general.

Theorem 14.13. *If $s(n)_{n\geq 0}$ is a constant-recursive sequence, then* $\mathsf{Siblings}(s) = \mathsf{ShiftSpace}(s)$.

Proof. Let $r = \operatorname{rank}(s)$, and let

$$B = (s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, \dots, s(n+r-1)_{n\geq 0}).$$

We show that B is a basis of both $\mathsf{Siblings}(s)$ and $\mathsf{ShiftSpace}(s)$; this will imply $\mathsf{Siblings}(s) = \mathsf{ShiftSpace}(s)$.

First we show that the r sequences comprising B are linearly independent. Suppose there is a linear relation

$$c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n) = 0$$

for all $n \geq 0$. If at least one of the coefficients c_i is not 0, then this relation is a nonzero recurrence satisfied by $s(n)_{n\geq 0}$, which contradicts $\operatorname{rank}(s) = r$. Therefore $c_i = 0$ for all $i \in \{0, 1, \dots, r-1\}$, and the sequences in B are linearly independent.

Next we show that B is a basis of $\mathsf{Siblings}(s)$. By Theorem 14.4, $\mathsf{Siblings}(s)$ is a vector space with dimension r. Since each sequence in B belongs to $\mathsf{Siblings}(s)$ (by Proposition 14.7) and these r sequences are linearly independent, it follows that B is a basis of $\mathsf{Siblings}(s)$.

Finally we show that B is a basis of $\mathsf{ShiftSpace}(s)$. An induction as in Example 14.9 shows that, for every $i \geq 0$, the shift $s(n+i)_{n\geq 0}$ can be written as a linear combination of the sequences in B. It follows that every sequence in $\mathsf{ShiftSpace}(s)$ is a linear combination of the sequences in B. Since the sequences in B are linearly independent, this implies that B is a basis of $\mathsf{ShiftSpace}(s)$. $\square$

Theorems 14.13 and 14.4 imply that if $s(n)_{n\geq 0}$ is constant-recursive then

$$\dim \mathsf{ShiftSpace}(s) = \operatorname{rank}(s).$$

Note that a constant-recursive sequence with rank r may also satisfy larger recurrences, as in the following example, but this doesn't mean that more than r shifts are linearly independent.

Example 14.14. Define $s(n)_{n\geq 0}$ to have initial conditions $s(0) = 3$, $s(1) = 1$, and $s(2) = 4$ and satisfy $s(n+3) = 2s(n+1) + s(n)$ for all $n \geq 0$. From this recurrence, one might assume that $\dim \mathsf{ShiftSpace}(s) = 3$. However, the terms of $s(n)_{n\geq 0}$ are $3, 1, 4, 5, 9, 14, 23, 37, \dots$, and in fact $s(n)_{n\geq 0}$ appears to satisfy the Fibonacci recurrence. We can prove this by showing that the sequence $t \in \mathsf{ShiftSpace}(F)$ defined by $t(0) = 3$, $t(1) = 1$, $t(2) = 4$ belongs to $\mathsf{ShiftSpace}(s)$. Example 14.9 shows that $t(n)_{n\geq 0}$ satisfies $t(n+3) = 2t(n+1) + t(n)$. It follows that both s and t are elements of $\mathsf{ShiftSpace}(s)$; since they agree on their first 3 terms, $s(n)_{n\geq 0} = t(n)_{n\geq 0}$. Therefore $\dim \mathsf{ShiftSpace}(s) = \dim \mathsf{ShiftSpace}(t) = 2$, and only the zeroth and first shifts of $s(n)_{n\geq 0}$ are linearly independent.

Since the space of siblings and the space of shifts are the same, we don't need both notations. The previous example shows that the definition of $\mathsf{ShiftSpace}(s)$ is more intrinsic to $s(n)_{n\geq 0}$ than the definition of $\mathsf{Siblings}(s)$, which requires us to know the minimal recurrence of $s(n)_{n\geq 0}$. Going forward, we will dispense with the notation $\mathsf{Siblings}(s)$ and use $\mathsf{ShiftSpace}(s)$ to denote this space.

Theorem 14.13 implies that the shifts of a constant-recursive sequence are contained in a finite-dimensional vector space. The converse is also true, so this is one characterization of constant-recursive sequences.

Theorem 14.15. *A sequence $s(n)_{n\geq 0}$ of rational numbers is constant-recursive if and only if the set of sequences $\{s(n+i)_{n\geq 0} : i \geq 0\}$ is contained in a finite-dimensional vector space. Moreover, if $s(n)_{n\geq 0}$ is constant-recursive then its rank is the smallest dimension of such a vector space.*

Proof. One direction follows from Proposition 14.7: If $s(n)_{n\geq 0}$ is constant-recursive, then each shift $s(n+i)_{n\geq 0}$ belongs to the vector space $\mathsf{ShiftSpace}(s)$. Moreover, $\mathsf{ShiftSpace}(s)$ is the vector space generated by the shifts of $s(n)_{n\geq 0}$, so it is the smallest vector space containing all shifts of $s(n)_{n\geq 0}$; by Theorems 14.4 and 14.13, $\operatorname{rank}(s) = \dim \mathsf{ShiftSpace}(s)$.

For the other direction, let V be a finite-dimensional vector space such that $s(n+i)_{n\geq 0} \in V$ for all $i \geq 0$. Let $r = \dim V$. Every set of $r+1$ sequences from V is linearly dependent. In particular, the $r+1$ sequences $s(n)_{n\geq 0}, \dots, s(n+r-1)_{n\geq 0}$ satisfy some relation

$$c_0 s(n) + c_1 s(n+1) + \cdots + c_{r-1} s(n+r-1) = 0$$

for all $n \geq 0$, where not all the coefficients are 0. Let $j \in \{0, 1, \dots, r-1\}$ be maximal such that $c_j \neq 0$; dividing this relation by c_j shows that $s(n)_{n\geq 0}$ is constant-recursive with rank at most $j+1$. □

Theorem 14.15 gives a quick way to determine that many sequences are constant-recursive.

Example 14.16. Let $s(n) = \frac{10^n - 1}{9}$. The sequence $s(n)_{n\geq 0}$ is $0, 1, 11, 111, 1111, \dots$. Each shift is a linear combination of $(10^n)_{n\geq 0}$ and $(1)_{n\geq 0}$, since $s(n+i) = \frac{10^i}{9} \cdot 10^n + \frac{1}{9} \cdot 1$. Therefore $s(n)_{n\geq 0}$ is constant-recursive with rank at most 2. In fact $s(n+2) = 11 s(n+1) - 10 s(n)$.

Example 14.17. Let $s(n)_{n\geq 0}$ be the periodic sequence $0, 1, 2, 1, 0, 1, 2, 1, \dots$. Since there are only 4 distinct shifts of $s(n)_{n\geq 0}$, it is constant-recursive with rank at most 4.

Example 14.18. Let $s(n) = n^3$. For every $i \geq 0$, expanding $(n+i)^3$ shows that it is a linear combination of $n^3, n^2, n, 1$. Therefore $s(n)_{n\geq 0}$ is constant-recursive with rank at most 4.

More generally, Theorem 14.15 provides a new, short proof that if $s(n)_{n\geq 0}$ is a polynomial sequence with rank r, then $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq r$. This was part of Theorem 13.9. Theorem 14.15 also shows that the shift space of a polynomial sequence is a vector space we are already familiar with.

Theorem 14.19. *If $s(n)_{n\geq 0}$ is a polynomial sequence with rank r, then* $\mathsf{ShiftSpace}(s) = \mathsf{Poly}(r)$.

Proof. Since every sequence in $\mathsf{Poly}(r)$ also belongs to $\mathsf{ShiftSpace}(s)$ by Theorem 13.9, we have $\mathsf{Poly}(r) \subseteq \mathsf{ShiftSpace}(s)$. However, $\mathsf{ShiftSpace}(s)$ and $\mathsf{Poly}(r)$ both have dimension r, by Theorems 13.13 and 7.7. Therefore $\mathsf{ShiftSpace}(s) = \mathsf{Poly}(r)$. □

We conclude with one more application of Theorem 14.15.

Example 14.20. Suppose $s(n)_{n\geq 0}$ satisfies the recurrence

$$s(n+3) = 5s(n+2) + 2s(n+1) + 4s(n) + F(n).$$

Due to the $F(n)$ term, we may not expect $s(n)_{n\geq 0}$ to be constant-recursive. However, every shift $s(n+i)_{n\geq 0}$ is a linear combination of the 5 sequences $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, s(n+2)_{n\geq 0}, F(n)_{n\geq 0}, F(n+1)_{n\geq 0}$, so in fact $s(n)_{n\geq 0}$ is constant-recursive. This is a *robustness result* — perturbing recurrences by adding constant-recursive sequences to one side does not change the class of sequences that we obtain. If $t(n)_{n\geq 0}$ is constant-recursive and $s(n+r) = t(n) + \sum_{i=0}^{r-1} c_i s(n+i)$ for all $n \geq 0$, then $s(n)_{n\geq 0}$ is constant-recursive with $\operatorname{rank}(s) \leq r + \operatorname{rank}(t)$.

We will continue to benefit from Theorem 14.15 in the next chapter.

Questions

Computations.

(1) What are the coordinates of the sequence $(3F(n+1) - F(n))_{n\geq 0}$ in the basis $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$?

(2) Write $L(n)_{n\geq 0}$ in the basis $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$, and use this to compute $L(10)$.

(3) Let $s \in \mathsf{ShiftSpace}(F)$ with $s(0) = 4$ and $s(1) = 7$. What are the coordinates of $s(n)_{n\geq 0}$ in each basis?
   (a) $(e_0(n)_{n\geq 0}, e_1(n)_{n\geq 0})$
   (b) $(F(n)_{n\geq 0}, L(n)_{n\geq 0})$
   (c) $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$

(4) Determine the coordinates in the basis $(F(n)_{n\geq 0}, F(n+1)_{n\geq 0})$ of each sequence in the table from the beginning of the chapter.

(5) Let $s(0) = 0$, $s(1) = 2$, $s(2) = 1$, $s(3) = 2$, and let

$$s(n+4) = 4s(n+2) - s(n+1) - 2s(n)$$

for all $n \geq 0$.
   (a) Compute the first several terms of $s(n)_{n\geq 0}$ and $(s(n+3) - 2s(n+2) + s(n))_{n\geq 0}$.
   (b) What is the dimension of $\mathsf{ShiftSpace}(s)$?
   (c) What is the minimal recurrence that $(6s(n+5) - 2s(n+3) + 5s(n))_{n\geq 0}$ satisfies?

(6) Let $E(n)_{n\geq 0}$ be the Tribonacci sequence.
   (a) Let $s(n)_{n\geq 0}$ be the sequence in $\mathsf{ShiftSpace}(E)$ with initial conditions $s(0) = a, s(1) = b, s(2) = c$. What are the coordinates of $s(n)_{n\geq 0}$ in the basis $(E(n)_{n\geq 0}, E(n+1)_{n\geq 0}, E(n+2)_{n\geq 0})$?
   (b) For each $i \geq 0$, what are the coordinates of $E(n+i)_{n\geq 0}$ in the standard basis $(e_0(n)_{n\geq 0}, e_1(n)_{n\geq 0}, e_2(n)_{n\geq 0})$?

(7) Let $E(n)_{n\geq 0}$ be the Tribonacci sequence. Use the Tribonacci recurrence to compute the coordinates of $E(n+i)_{n\geq 0}$ in the basis $(E(n)_{n\geq 0}, E(n+1)_{n\geq 0}, E(n+2)_{n\geq 0})$ for each $i \in \{0, 1, \ldots, 8\}$. Are the coordinates all Tribonacci numbers?

Experiments.

(8) Is the sibling relationship symmetric, or are there sequences $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ such that $s(n)_{n\geq 0} \in \mathsf{Siblings}(t)$ but $t(n)_{n\geq 0} \notin \mathsf{Siblings}(s)$?

Proofs.

(9) Prove that $F(n+i) = F(i)F(n+1) + F(i-1)F(n)$ for each $n \geq 0$ and $i \geq 1$.

(10) What are the coordinates of $L(n+i)_{n\geq 0}$ in the basis $(L(n)_{n\geq 0}, L(n+1)_{n\geq 0})$ of $\mathsf{ShiftSpace}(L)$?

(11) For a constant-recursive $s(n)_{n\geq 0}$ with rank r, what are the coordinates of $s(n+i)_{n\geq 0}$ in the basis $(s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, \ldots, s(n+r-1)_{n\geq 0})$ of $\mathsf{ShiftSpace}(s)$?

CHAPTER 15

Closure properties for constant-recursive sequences

Addition and multiplication

We saw in Proposition 14.3 that adding two sequences satisfying the same recurrence produces another sequence satisfying that recurrence. What if we add two sequences that satisfy different recurrences?

Example 15.1. Is the sum $(F(n) + E(n))_{n\geq 0}$ of the Fibonacci and Tribonacci sequences constant-recursive? Let $s(n) = F(n) + E(n)$. The sequence $s(n)_{n\geq 0}$ is

$$0, 1, 2, 3, 5, 9, 15, 26, 45, 78, 136, 238, \ldots.$$

If all the shifts of $s(n)_{n\geq 0}$ belong to a finite-dimensional vector space, then Theorem 14.15 will imply that $s(n)_{n\geq 0}$ is constant-recursive. Let us write out formulas for the first several shifts of $s(n)_{n\geq 0}$ using the Fibonacci and Tribonacci recurrences to rewrite $F(n + i)$ and $E(n + i)$ where possible:

$$\begin{aligned}
s(n) &= F(n) && && + E(n) && && \\
s(n+1) &= && F(n+1) && && + E(n+1) && \\
s(n+2) &= F(n) + && F(n+1) && && && + E(n+2) \\
s(n+3) &= F(n) + 2&&F(n+1) + && E(n) + && E(n+1) + && E(n+2) \\
s(n+4) &= 2F(n) + 3&&F(n+1) + && E(n) + 2&&E(n+1) + 2&&E(n+2) \\
s(n+5) &= 3F(n) + 5&&F(n+1) + 2&&E(n) + 3&&E(n+1) + 4&&E(n+2).
\end{aligned}$$

(The order and alignment will be useful shortly.) Let

$$\begin{aligned}
B_1 &= (F(n)_{n\geq 0}, F(n+1)_{n\geq 0}) \\
B_2 &= (E(n)_{n\geq 0}, E(n+1)_{n\geq 0}, E(n+2)_{n\geq 0}).
\end{aligned}$$

Since B_1 and B_2 are bases of $\mathsf{ShiftSpace}(F)$ and $\mathsf{ShiftSpace}(E)$, respectively, each shift $s(n + i)_{n\geq 0}$ can be written as a linear combination of sequences from

$$B_1 \cup B_2 = (F(n)_{n\geq 0}, F(n+1)_{n\geq 0}, E(n)_{n\geq 0}, E(n+1)_{n\geq 0}, E(n+2)_{n\geq 0}).$$

Since $B_1 \cup B_2$ generates a vector space with dimension ≤ 5, Theorem 14.15 implies that $s(n)_{n\geq 0}$ is constant-recursive with rank ≤ 5. In particular, the 6 sequences $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, \ldots, s(n+5)_{n\geq 0}$ are linearly dependent, so they satisfy a relation

$$c_0 s(n) + c_1 s(n+1) + \cdots + c_5 s(n+5) = 0. \tag{15.1}$$

This is a recurrence for $s(n)_{n\geq 0}$. To solve for the coefficients $c_0, c_1, \ldots, c_5$, substitute the formulas we computed above for $s(n), \ldots, s(n + 5)$ into the recurrence. This produces an equation involving the 5 sequences from $B_1 \cup B_2$. Collect all terms involving $F(n)$ together, and similarly for the other sequences in $B_1 \cup B_2$. It suffices for the coefficients of these 5 sequences to be 0. This lets us set up a system of equations.

$$\begin{array}{l} \text{coefficients of } F(n): \\ \text{coefficients of } F(n+1): \\ \text{coefficients of } E(n): \\ \text{coefficients of } E(n+1): \\ \text{coefficients of } E(n+2): \end{array} \qquad \begin{bmatrix} 1 & 0 & 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 2 & 3 & 5 \\ 1 & 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The first row stipulates that the coefficient of $F(n)$ on both sides of Equation (15.1) is 0, and the remaining rows put the same condition on the other sequences of $B_1 \cup B_2$. The entries of the 5×6 matrix can be read off directly from the formulas above. Row-reducing gives the augmented matrix

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{array}\right].$$

There are infinitely many solutions to the system. If we choose $c_5 = -1$ (so that $s(n+5)$ appears on the other side of the recurrence with coefficient 1), this determines the remaining coefficients uniquely. They can be read off from the last nonzero column, and we obtain the recurrence

$$-s(n) - 2s(n+1) - s(n+2) + s(n+3) + 2s(n+4) = s(n+5).$$

In Example 17.18, we will see a less intensive way to compute this recurrence.

A similar procedure works for the product of two constant-recursive sequences.

Example 15.2. Let $s(n) = F(n)E(n)$. The terms of $s(n)_{n \geq 0}$ are

$$0, 0, 1, 2, 6, 20, 56, 169, 504, 1496, 4455, 13261, \dots.$$

As in the previous example, the Fibonacci and Tribonacci recurrences allow us to write

$$\begin{aligned} s(n) &= F(n)E(n) \\ s(n+1) &= F(n+1)E(n+1) \\ s(n+2) &= \big(F(n) + F(n+1)\big)E(n+2) \\ s(n+3) &= \big(F(n) + 2F(n+1)\big)\big(E(n) + E(n+1) + E(n+2)\big) \end{aligned}$$

and so on. Therefore every shift $s(n+i)_{n \geq 0}$ belongs to the vector space generated by the 6 sequences $\big(F(n+i)E(n+j)\big)_{n \geq 0}$ where $i \in \{0, 1\}$ and $j \in \{0, 1, 2\}$. In particular, there is a linear relation among the first 7 shifts. We collect like terms and set the coefficients equal to 0.

$$\begin{array}{l} \text{coefficients of } F(n)E(n): \\ \text{coefficients of } F(n)E(n+1): \\ \text{coefficients of } F(n)E(n+2): \\ \text{coefficients of } F(n+1)E(n): \\ \text{coefficients of } F(n+1)E(n+1): \\ \text{coefficients of } F(n+1)E(n+2): \end{array} \qquad \begin{bmatrix} 1 & 0 & 0 & 1 & 2 & 6 & 20 \\ 0 & 0 & 0 & 1 & 4 & 9 & 30 \\ 0 & 0 & 1 & 1 & 4 & 12 & 35 \\ 0 & 0 & 0 & 2 & 3 & 10 & 32 \\ 0 & 1 & 0 & 2 & 6 & 15 & 48 \\ 0 & 0 & 1 & 2 & 6 & 20 & 56 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Row-reducing gives

$$\left[\begin{array}{ccccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right],$$

which corresponds to the recurrence

$$s(n+6) = s(n+5) + 4s(n+4) + 5s(n+3) + 2s(n+2) - s(n+1) + s(n).$$

These examples suggest the following analogue of Theorem 5.7.

Theorem 15.3. *Let $s_1(n)_{n\geq 0}$ and $s_2(n)_{n\geq 0}$ be constant-recursive sequences with respective ranks r_1 and r_2. Then $(s_1(n) + s_2(n))_{n\geq 0}$ is constant-recursive with rank $\leq r_1 + r_2$, and $(s_1(n) \cdot s_2(n))_{n\geq 0}$ is constant-recursive with rank $\leq r_1 r_2$.*

Proof. Let B_1 be a basis of $\mathsf{ShiftSpace}(s_1)$, and let B_2 be a basis of $\mathsf{ShiftSpace}(s_2)$. For every $i \geq 0$, the shift $(s_1(n+i) + s_2(n+i))_{n\geq 0}$ belongs to the vector space generated by $B_1 \cup B_2$. By Theorem 14.15, $(s_1(n) + s_2(n))_{n\geq 0}$ is constant-recursive with rank $\leq r_1 + r_2$. Similarly, for every $i \geq 0$, the shift $(s_1(n+i)s_2(n+i))_{n\geq 0}$ belongs to the vector space generated by $B_1 \times B_2$. Therefore $(s_1(n)s_2(n))_{n\geq 0}$ is constant-recursive with rank $\leq r_1 r_2$. $\square$

In Example 15.1, the sequences in $B_1 \cup B_2$ are linearly independent, so $B_1 \cup B_2$ is a basis of the vector space it generates. Similarly, in Example 15.2, the 6 sequences $(F(n+i)E(n+j))_{n\geq 0}$, where $i \in \{0,1\}$ and $j \in \{0,1,2\}$, form a basis of the vector space they generate. If this is not the case, then the procedure we used produces multiple recurrences, as in the following example, since the rank of the sum or product will not reach the upper bound of Theorem 15.3.

Example 15.4. What recurrence does $F(n)^2$ satisfy? Theorem 15.3 implies that the rank of this sequence is at most 4. We compute a recurrence for the product $F(n) \cdot F(n)$ as in Example 15.2, by writing $F(n)^2, F(n+1)^2, \ldots, F(n+4)^2$ as linear combinations of $F(n)^2$, $F(n)F(n+1)$, $F(n+1)^2$.

$$\begin{array}{l} \text{coefficients of } F(n)^2\text{:} \\ \text{coefficients of } F(n)F(n+1)\text{:} \\ \text{coefficients of } F(n+1)^2\text{:} \end{array} \quad \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 1 & 4 & 0 \\ 0 & 0 & 2 & 4 & 12 & 0 \\ 0 & 1 & 1 & 4 & 9 & 0 \end{array}\right]$$

When we row-reduce, we obtain

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -1 & -2 & 0 \\ 0 & 1 & 0 & 2 & 3 & 0 \\ 0 & 0 & 1 & 2 & 6 & 0 \end{array}\right].$$

The solution space is 2-dimensional. Choosing $c_4 = -1$ and $c_3 = 0$ gives the recurrence

$$F(n+4)^2 = 6F(n+2)^2 + 3F(n+1)^2 - 2F(n)^2.$$

But we can do better by choosing $c_4 = 0$ and $c_3 = -1$; then

$$F(n+3)^2 = 2F(n+2)^2 + 2F(n+1)^2 - F(n)^2,$$

which shows that the rank is at most 3 and less than the upper bound 4.

Powers

By iterating the multiplication closure property, one can compute powers of a sequence, as in Example 15.4. How does the rank grow as the exponent increases?

Example 15.5. For $F(n)^5$, Theorem 15.3 gives an upper bound of 2^5 on the rank. Our first observation is that, rather than performing 4 row reductions (one for each multiplication in $F(n) \cdot F(n) \cdot F(n) \cdot F(n) \cdot F(n)$), we can use repeated squaring, as discussed at the end of Chapter 12, to get by with only 3, since $F(n)^5 = (F(n)^2)^2 F(n)$. Having computed a recurrence for $F(n)^2$ in Example 15.4, we first multiply $F(n)^2 \cdot F(n)^2$, obtaining (among others) the recurrence

$$F(n+5)^4 = 5F(n+4)^4 + 15F(n+3)^4 - 15F(n+2)^4 - 5F(n+1)^4 + F(n)^4.$$

Next we compute recurrences for $F(n)^4 \cdot F(n)$ and find

$$F(n+6)^5 = 8F(n+5)^5 + 40F(n+4)^5 - 60F(n+3)^5 - 40F(n+2)^5 + 8F(n+1)^5 + F(n)^5.$$

Examples 15.4 and 15.5 suggest that the rank of $(F(n)^m)_{n\geq 0}$ is $m+1$. This is significantly less than the bound 2^m implied by Theorem 15.3. By taking a look at the proof of Theorem 15.3, we can see why. The generators we construct for the shift space of $(F(n)^2)_{n\geq 0}$ are $(F(n)F(n))_{n\geq 0}$, $(F(n)F(n+1))_{n\geq 0}$, $(F(n+1)F(n))_{n\geq 0}$, and $(F(n+1)F(n+1))_{n\geq 0}$. However, two of these sequences are the same, so we only need 3 of the 4. Additional duplication occurs for higher powers.

Theorem 15.6. *If $m \geq 0$ and $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank r, then $(s(n)^m)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq \binom{m+r-1}{m}$.*

Proof. The proof of Theorem 15.3 implies that the shift space of $(s(n)^m)_{n\geq 0}$ is generated by the r^m sequences $(s(n+i_1)s(n+i_2)\cdots s(n+i_m))_{n\geq 0}$ where each $i_j \in \{0, 1, \ldots, r-1\}$. The order of the sequences in each such product does not matter, since multiplication is commutative, and a sequence $s(n+i_j)$ can appear multiple times. Therefore the distinct generators are indexed by multisets with size m on $\{0, 1, \ldots, r-1\}$. By Corollary 9.10, there are $\binom{m+r-1}{m}$ distinct generators. $\square$

Example 15.7. The rank of the Fibonacci sequence is $r = 2$, so the rank of $(F(n)^m)_{n\geq 0}$ is at most $\binom{m+2-1}{m} = m+1$.

Subsequences

Another common operation on a sequence is the extraction of terms indexed by an arithmetic progression.

Example 15.8. Let $s(n)_{n\geq 0}$ be the subsequence $1, 3, 13, 55, 233, 987, \ldots$ of the Fibonacci sequence obtained by taking every third time, starting with $F(1)$, so that $s(n) = F(3n+1)$. Is $s(n)_{n\geq 0}$ constant-recursive? Theorem 14.15 suggests we consider the shifts $s(n+i)_{n\geq 0}$, but the trick is to write $s(n+i)$ using $F(3n)$ and $F(3n+1)$ rather than $F(n)$ and $F(n+1)$. The Fibonacci recurrence gives

$$\begin{aligned} s(n+1) = F(3n+4) &= F(3n+2) + F(3n+3) \\ &= F(3n+1) + 2F(3n+2) \\ &= 2F(3n) + 3F(3n+1). \end{aligned}$$

The addition formula $F(n+i) = F(i-1)F(n) + F(i)F(n+1)$ from Example 14.9 is extremely useful for obtaining such formulas without the intervening steps. For example, it gives

$$s(n+2) = F(3n+7) = 8F(3n) + 13F(3n+1).$$

Now we have the 3 sequences $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, s(n+2)_{n\geq 0}$ in the 2-dimensional vector space generated by $F(3n)_{n\geq 0}$ and $F(3n+1)_{n\geq 0}$, so $s(n)_{n\geq 0}$ is constant-recursive with rank at most 2. To compute a recurrence, we find a relation among the 3 shifts.

$$\begin{array}{ll} \text{coefficients of } F(3n): & \multirow{2}{*}{$\left[\begin{array}{ccc|c} 0 & 2 & 8 & 0 \\ 1 & 3 & 13 & 0 \end{array}\right]$} \\ \text{coefficients of } F(3n+1): & \end{array}$$

Row-reduce:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 4 & 0 \end{array}\right].$$

Therefore $s(n+2) = 4s(n+1) + s(n)$. Interestingly, the rank of the subsequence is the same as the original.

Theorem 15.9. *Let $s(n)_{n\geq 0}$ be a constant-recursive sequence with rank r. For each $a \geq 0$ and $b \geq 0$, the sequence $s(an+b)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq r$.*

Proof. For each $i \in \{0, 1, \dots, r\}$, we can use the recurrence for $s(n)_{n\geq 0}$ to write $s(a(n+i)+b)_{n\geq 0}$ as a linear combination of $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, \dots, s(n+r-1)_{n\geq 0}$. Therefore $s(a(n+i)+b)_{n\geq 0} \in \mathsf{ShiftSpace}(s)$. By Theorem 14.15, $s(an+b)_{n\geq 0}$ is constant-recursive with $\operatorname{rank}(s) \leq \dim \mathsf{ShiftSpace}(s) = r$. $\square$

The rank of the subsequence can be less than r. For example, if $s(2n) = 0$ and $s(2n+1) = F(n)$, then $s(n)_{n\geq 0} = 0, 0, 0, 1, 0, 1, 0, 2, 0, 3, 0, 5, \dots$ satisfies $s(n+4) = s(n+2) + s(n)$ and has rank 4, while $s(2n)_{n\geq 0}$ has rank 0 and $s(2n+1)_{n\geq 0}$ has rank 2.

Additional operations

Proposition 13.5 established another closure property: Modifying finitely many terms of a constant-recursive sequence produces a constant-recursive sequence. There are other operations we are interested in as well. For example, if $s(n)_{n\geq 0}$ is constant-recursive, is $\left(\sum_{i=0}^{n-1} s(i)\right)_{n\geq 0}$ constant-recursive? As a converse of Theorem 15.9, if $a \geq 1$ and $s(an+b)_{n\geq 0}$ is constant-recursive for each $b \in \{0, 1, \dots, a-1\}$, is $s(n)_{n\geq 0}$ necessarily constant-recursive? These will be easier to settle in Chapter 17 when we develop an additional tool for working with constant-recursive sequences.

Questions

Computations.

(1) As discussed in Chapter 3, the number of length-n binary words avoiding 00 is $F(n+2)$. Use closure properties to compute a recurrence for the number of binary words *containing* 00.

(2) Prove the Cassini identity $F(n-1)F(n+1) = F(n)^2 + (-1)^n$ for all $n \geq 0$ by computing a recurrence for $F(n-1)F(n+1) - F(n)^2 - (-1)^n$. (Here we interpret $F(-1) = 1$.)

(3) Let $s(0) = 1$, $s(1) = 3$, and $s(n+2) = 2s(n+1) + s(n)$ for all $n \geq 0$. Let $t(0) = 1$, $t(1) = 2$, and $t(n+2) = 2t(n+1) + t(n)$ for all $n \geq 0$.
   (a) Use closure properties to show that $2t(n)^2 - s(n)^2 = (-1)^n$.
   (b) Compute several values of $\frac{s(n)}{t(n)}$ numerically. Does $\lim_{n\to\infty} \frac{s(n)}{t(n)}$ exist?

(4) (a) Compute a recurrence for $L(n)^2 - 5F(n)^2$.
   (b) Compute the first several values of $L(n)^2 - 5F(n)^2$, make a conjecture for the nth term, and prove it.

Experiments.

(5) Compute a recurrence for $(F(n)^m)_{n\geq 0}$ for several values of m. Does the rank achieve the upper bound in Theorem 15.6?

(6) Compute a recurrence for $(F(n)^m + L(n)^m)_{n\geq 0}$ for several values of m. What is the rank?

(7) Compute a recurrence for $(E(n)^m)_{n\geq 0}$ for several values of m. Does the rank achieve the upper bound in Theorem 15.6?

(8) Let $s(n) = n^2$. Compute a recurrence for $(s(n)^m)_{n\geq 0}$ for several values of m. Does the rank achieve the upper bound in Theorem 15.6?

(9) Define $s(n)_{n\geq 0}$ by $s(0) = 0$, $s(1) = 0$, $s(2) = 1$, and $s(n+3) = s(n+2) + s(n+1) + 2s(n)$ for all $n \geq 0$.
   (a) Compute a recurrence for $(s(n)^3)_{n\geq 0}$. Does its rank achieve the upper bound in Theorem 15.6?
   (b) Use the first several terms of $s(n)_{n\geq 0}$ to guess a linear relation among the 10 sequences that the proof of Theorem 15.6 identifies as generators of the shift space of $(s(n)^3)_{n\geq 0}$.
   (c) Factor the linear relation, and use this to prove that the relation holds and to explain the rank of $(s(n)^3)_{n\geq 0}$.
   (d) If $t(n)_{n\geq 0}$ satisfies the same recurrence as $s(n)_{n\geq 0}$ but has different initial conditions, can the rank of $(t(n)^3)_{n\geq 0}$ be 10?

(10) Define $s(n)_{n\geq 0}$ by $s(0) = 0$, $s(1) = 0$, $s(2) = 1$, and $s(n+3) = -9s(n+2) - 6s(n+1) + 16s(n)$ for all $n \geq 0$.
   (a) Compute a recurrence for $(s(n)^3)_{n\geq 0}$. Does its rank achieve the upper bound in Theorem 15.6?
   (b) Use the first several terms of $s(n)_{n\geq 0}$ to guess a linear relation among the 10 sequences that the proof of Theorem 15.6 identifies as generators of the shift space of $(s(n)^3)_{n\geq 0}$.
   (c) Use induction to prove that the relation holds, and use this to explain the rank of $(s(n)^3)_{n\geq 0}$.
   (d) If $t(n)_{n\geq 0}$ satisfies the same recurrence as $s(n)_{n\geq 0}$ but has different initial conditions, can the rank of $(t(n)^3)_{n\geq 0}$ be 10?

Programs.

(11) Write a program that implements each closure property for constant-recursive sequences (given a recurrence and initial conditions for $s_1(n)_{n\geq 0}$ and $s_2(n)_{n\geq 0}$).

(a) sum $(s_1(n) + s_2(n))_{n\geq 0}$
(b) product $(s_1(n)s_2(n))_{n\geq 0}$
(c) power $(s_1(n)^m)_{n\geq 0}$ for a given m
(d) subsequence $s_1(an + b)_{n\geq 0}$ for given values of a, b

CHAPTER 16

Guessing a constant-recursive sequence

In this chapter, we discuss two methods of guessing a recurrence for a sequence. This will lead us to a way to test whether two constant-recursive sequences are equal, so that we can turn a guess into a rigorous proof.

Undetermined coefficients

In Chapter 6 we used the method of undetermined coefficients to guess a formula for the nth term of a polynomial sequence when given the first few terms. We can use a similar approach to guess a recurrence for a constant-recursive sequence. The idea is that we will use the biggest recurrence we can, to have the biggest chance of success. So first we should figure out how big this recurrence can be if we only have a fixed number of terms.

Example 16.1. Suppose the first 6 terms of $s(n)_{n\geq 0}$ are $1, 3, 5, 21, 95, 373$. Can we guess the nth term? If $s(n)_{n\geq 0}$ is a *polynomial* sequence with rank ≤ 6, then these 6 terms are enough to identify it. By solving a system of 6 equations in 6 unknowns, the method of undetermined coefficients produces the guess $\frac{3}{5}n^5 - \frac{19}{4}n^4 + \frac{95}{6}n^3 - \frac{93}{4}n^2 + \frac{407}{30}n + 1$ for $s(n)$. Since this polynomial generates a sequence with rank exactly 6, there is no redundancy, so it may not be the description of $s(n)_{n\geq 0}$ we are looking for.

Instead, suppose $s(n)_{n\geq 0}$ is constant-recursive. Do its first 6 terms determine a constant-recursive sequence with rank ≤ 6? We can set up a recurrence

$$s(n+6) = c_0 s(n) + c_1 s(n+1) + \cdots + c_5 s(n+5)$$

in 6 unknowns $c_0, \dots, c_5$, but how many equations in these unknowns can we get from this recurrence? Even plugging in $n = 0$ doesn't produce a useful equation, since we don't know $s(6)$, so we get 0 equations. We must lower the rank. With rank ≤ 4, the general recurrence

$$s(n+4) = c_0 s(n) + c_1 s(n+1) + c_2 s(n+2) + c_3 s(n+3)$$

produces equations for $n = 0$ and $n = 1$:

$$\begin{aligned} 95 &= 1c_0 + 3c_1 + 5c_2 + 21c_3 \\ 373 &= 3c_0 + 5c_1 + 21c_2 + 95c_3. \end{aligned}$$

But 2 equations aren't enough to determine 4 coefficients. With rank ≤ 3 instead, the recurrence

$$s(n+3) = c_0 s(n) + c_1 s(n+1) + c_2 s(n+2)$$

produces equations for $n = 0$, $n = 1$, and $n = 2$:

$$\begin{aligned} 21 &= 1c_0 + 3c_1 + 5c_2 \\ 95 &= 3c_0 + 5c_1 + 21c_2 \\ 373 &= 5c_0 + 21c_1 + 95c_2. \end{aligned}$$

This system of 3 equations in 3 unknowns corresponds to the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 3 & 5 & 21 \\ 3 & 5 & 21 & 95 \\ 5 & 21 & 95 & 373 \end{array}\right].$$

Row-reducing, we find that it has a unique solution $c_0 = 7$, $c_1 = -2$, $c_2 = 4$, so our guess is that

$$s(n+3) = 7s(n) - 2s(n+1) + 4s(n+2) \tag{16.1}$$

for all $n \geq 0$. There is no redundancy in this guess; we assumed the rank was ≤ 3 and guessed a sequence with rank exactly 3, so it also may not describe the sequence we are actually interested in.

Example 16.2. Let's extend the known terms by one. Suppose the first 7 terms of $s(n)_{n\geq 0}$ are $1, 3, 5, 21, 95, 373, 1449$. (Maybe this sequence counts some combinatorial object, and to compute $s(6) = 1449$ we needed to painstakingly generate all such objects with size 6.) If we set up the same recurrence

$$s(n+3) = c_0 s(n) + c_1 s(n+1) + c_2 s(n+2),$$

the extra term allows us to extract one more equation than in Example 16.1:

$$\left[\begin{array}{ccc|c} 1 & 3 & 5 & 21 \\ 3 & 5 & 21 & 95 \\ 5 & 21 & 95 & 373 \\ 21 & 95 & 373 & 1449 \end{array}\right].$$

Solving this system gives the same solution $c_0 = 7$, $c_1 = -2$, $c_2 = 4$ as before. Since a generic system of 4 equations in 3 unknowns has no solution, we are more confident than in Example 16.1 that the recurrence in Equation (16.1) describes our sequence.

In general, to guess a recurrence

$$s(n+r) = c_0 s(n) + c_1 s(n+1) + \cdots + c_{r-1} s(n+r-1)$$

with r coefficients, we need r equations, so we need to evaluate it at r values of n. Assuming we use the values $n \in \{0, 1, \ldots, r-1\}$, we therefore need to know the $2r$ initial terms $s(0), s(1), \ldots, s(2r-1)$ to determine the recurrence. This makes sense intuitively, since specifying a rank-r constant-recursive sequence requires $2r$ pieces of information, namely r coefficients in the recurrence and r initial conditions. This implies that, if we have the first k terms of a sequence but do not know what the rank is, we can set up a recurrence with size $\lfloor \frac{k}{2} \rfloor$ and solve a system of $\lceil \frac{k}{2} \rceil$ equations.

When guessing a recurrence, however, the resulting system of equations is not guaranteed to have a solution.

Example 16.3. Suppose the first 2 terms of $s(n)_{n\geq 0}$ are $0, 1$. Since we have only 2 terms, we set up a recurrence with $r = 1$: $s(n+1) = c_0 s(n)$. But $n = 0$ results in $1 = 0$, an inconsistent system. Therefore these initial terms cannot be extended to a constant-recursive sequence with rank ≤ 1.

Example 16.4. To take a larger example, suppose the first 4 terms of $s(n)_{n\geq 0}$ are $1, 2, 4, 6$. These 4 terms give rise to a system of 2 equations from the recurrence $s(n+2) = c_0 s(n) + c_1 s(n+1)$:

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 2 & 4 & 6 \end{array}\right].$$

This system has no solutions, so there is no constant-recursive sequence $1, 2, 4, 6, \ldots$ with rank ≤ 2.

Guessing a recurrence differs from guessing a polynomial in this way; there is no analogue of Proposition 6.2, so the system of equations does not always have a solution. If there is no solution, then r was too low and we need more terms to guess a recurrence. On the other hand, if there are infinitely many solutions, then r was higher than necessary.

Example 16.5. Suppose the first 4 terms of $s(n)_{n\geq 0}$ are $1, 2, 4, 8$. Now when we set up the recurrence $s(n+2) = c_0 s(n) + c_1 s(n+1)$ we obtain the system

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 2 & 4 & 8 \end{array}\right].$$

Solving, we find that $c_0 = 4 - 2c_1$ with no other restrictions. Namely, every value of c_1 gives a solution. For example, we can choose $c_1 = 2$ so that $c_0 = 0$; we obtain the recurrence $s(n+2) = 2s(n+1)$, which implies $s(n) = 2^n$ for all $n \geq 0$. What about other values for c_1? If we set $c_1 = 1$, then $c_0 = 2$, and the recurrence becomes $s(n+2) = 2s(n) + s(n+1)$; this also implies $s(n) = 2^n$, since $2^{n+2} = 2 \cdot 2^n + 2^{n+1}$. In fact every value of c_1 produces the same sequence; we can see this by checking that $2^{n+2} = (4 - 2c_1)2^n + c_1 2^{n+1}$.

In general, if there are infinitely many solutions, then all solutions describe the same sequence. This is the analogue of Theorem 6.3.

Theorem 16.6. *Every finite sequence $s(0), s(1), s(2), \ldots, s(2r-1)$ of length $2r$ has at most one extension to a constant-recursive sequence $s(n)_{n\geq 0}$ with $\operatorname{rank}(s) \leq r$.*

Proof. Suppose it has two extensions $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$. That is, $s(n) = t(n)$ for all $n \in \{0, 1, \ldots, 2r-1\}$ and the ranks satisfy $\operatorname{rank}(s) \leq r$ and $\operatorname{rank}(t) \leq r$. By Theorem 15.3 and Proposition 14.2, the sequence $(s(n) - t(n))_{n\geq 0}$ is constant-recursive with rank $\leq 2r$. Its first $2r$ terms are 0, so its recurrence implies that $s(n) - t(n) = 0$ for all $n \geq 0$. $\square$

In particular, if we know that $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank at most r, then its first $2r$ terms determine it uniquely.

Incremental guessing

If we know the first k terms of a sequence, it may be inefficient to solve a system of $\lceil \frac{k}{2} \rceil$ equations to guess a recurrence, particularly if k is large and the rank of the sequence turns out to be small. An alternative approach is to first try a small recurrence, since the corresponding system of equations has fewer unknowns;

if there is no solution, increase the size of the recurrence by 1 and try again. We refer to this as *incremental guessing*, for lack of a better term.

We will make one conceptual change, however. Rather than solving systems of equations, we will instead (but equivalently) look for linear dependence among the first several shifts of $s(n)_{n\geq 0}$. When we have eliminated the possibility of a nonzero linear relation among the r shifts $s(n)_{n\geq 0}, \dots, s(n+r-1)_{n\geq 0}$, then any nonzero relation among the $r+1$ shifts $s(n)_{n\geq 0}, \dots, s(n+r)_{n\geq 0}$ necessarily involves $s(n+r)_{n\geq 0}$ with a nonzero coefficient. Therefore we don't need to place $s(n+r)$ on the other side of the equation with coefficient 1, and we can treat all the shifts uniformly. Accordingly, we'll remove the vertical line from our augmented matrices to turn them into ordinary matrices.

Example 16.7. Suppose the first 8 terms of $s(n)_{n\geq 0}$ are

$$2, 6, 1, -5, 28, 49, -65, -44.$$

We look for linear relations among the first $r+1$ shifts of $s(n)_{n\geq 0}$, starting with $r = 0$ and subsequently incrementing. For $r = 0$, the sequence $s(n)_{n\geq 0}$ itself comprises a linearly independent set, since it is not the 0 sequence. For $r = 1$, we place the shifts $s(n)_{n\geq 0}$ and $s(n+1)_{n\geq 0}$ as column sequences:

$$\begin{bmatrix} 2 & 6 \\ 6 & 1 \\ 1 & -5 \\ -5 & 28 \\ 28 & 49 \\ 49 & -65 \\ -65 & -44 \end{bmatrix}.$$

Here we could not include the term -44 in the left column because we only know 7 terms of the right column. We are looking for linear relations among the columns. That is, we would like to compute the null space of this matrix. Row reduction produces

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which indicates that $s(n)_{n\geq 0}$ and $s(n+1)_{n\geq 0}$ are linearly independent. This implies that, if $s(n)_{n\geq 0}$ is constant-recursive, then its rank is not 1. Therefore we start over with the shifts $s(n)_{n\geq 0}$, $s(n+1)_{n\geq 0}$, and $s(n+2)_{n\geq 0}$ as column sequences in a matrix with 6 rows (since we only know 6 terms of $s(n+2)_{n\geq 0}$):

$$\begin{bmatrix} 2 & 6 & 1 \\ 6 & 1 & -5 \\ 1 & -5 & 28 \\ -5 & 28 & 49 \\ 28 & 49 & -65 \\ 49 & -65 & -44 \end{bmatrix}, \quad \text{which reduces to} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

indicating that $s(n)_{n\geq 0}$, $s(n+1)_{n\geq 0}$, and $s(n+2)_{n\geq 0}$ are linearly independent. Next we construct a matrix containing terms of the first 4 shifts:

$$\begin{bmatrix} 2 & 6 & 1 & -5 \\ 6 & 1 & -5 & 28 \\ 1 & -5 & 28 & 49 \\ -5 & 28 & 49 & -65 \\ 28 & 49 & -65 & -44 \end{bmatrix}, \quad \text{which reduces to} \quad \begin{bmatrix} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have found linear dependence among the first 5 terms of $s(n)_{n\geq 0}$, $s(n+1)_{n\geq 0}$, $s(n+2)_{n\geq 0}$, and $s(n+3)_{n\geq 0}$. Each row encodes a relation

$$c_0 s(n) + c_1 s(n+1) + c_2 s(n+2) + c_3 s(n+3) = 0,$$

so the row-reduced matrix corresponds to the system

$$\begin{aligned} c_0 + 6c_3 &= 0 \\ c_1 - 3c_3 &= 0 \\ c_2 + c_3 &= 0. \end{aligned}$$

We choose $c_3 = -1$, and this determines $c_0 = 6$, $c_1 = -3$, and $c_2 = 1$. Therefore

$$s(n+3) = 6s(n) - 3s(n+1) + s(n+2)$$

for all $n \in \{0, 1, 2, 3, 4\}$. We might conjecture that this recurrence holds for all $n \geq 0$.

Every time we create a column, we must delete a row, effectively deleting information from the bottom of several columns. Consequently, it is possible for a subset of columns to become linearly dependent that previously were known to be linearly independent. If this happens, then there will be at least one relation found that does not hold for all $n \geq 0$. To obtain a reliable guess, we would need to compute more terms of the sequence and start the guessing process over.

Example 16.8. Let's revisit the initial terms $1, 2, 4, 6$ from Example 16.5. We begin with the matrix

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 4 & 6 \end{bmatrix}, \quad \text{which reduces to} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Since the columns are linearly independent, next we add a column to obtain

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 6 \end{bmatrix}, \quad \text{which reduces to} \quad \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Deleting the third row in this step caused the first two columns to become linearly dependent, which we know does not reflect the relationship between the sequences that these columns represent.

In the previous examples, we row-reduced from scratch at every step. This is redundant because the first several row operations when using the first $k+1$ shifts have the same effect on the first several columns as they did when using the first k shifts. We can be more efficient by assigning sequences to rows rather than columns, as in the following example. However, when we do this, we must somehow keep track of the row operations we perform, since in general a row-reduced matrix does not contain information about the relations among the rows of the original matrix. Appending an identity matrix will do the trick.

Example 16.9. As in Example 16.7, suppose the first 8 terms of $s(n)_{n\geq 0}$ are

$$2, 6, 1, -5, 28, 49, -65, -44.$$

Start with a matrix whose rows contain terms of $s(n)_{n\geq 0}$ and $s(n+1)_{n\geq 0}$:

$$\left[\begin{array}{ccccccc|cc} 2 & 6 & 1 & -5 & 28 & 49 & -65 & 1 & 0 \\ 6 & 1 & -5 & 28 & 49 & -65 & -44 & 0 & 1 \end{array}\right].$$

Row reduction produces

$$\left[\begin{array}{ccccccc|cc} 1 & 0 & -\frac{31}{34} & \frac{173}{34} & \frac{133}{17} & -\frac{439}{34} & -\frac{199}{34} & -\frac{1}{34} & \frac{3}{17} \\ 0 & 1 & \frac{8}{17} & -\frac{43}{17} & \frac{35}{17} & \frac{212}{17} & -\frac{151}{17} & \frac{3}{17} & -\frac{1}{17} \end{array}\right].$$

Now create a new row containing terms of $s(n+2)_{n\geq 0}$. This requires deleting a column from the left of the matrix, since we only know 8 terms of $s(n)_{n\geq 0}$. We also create a new column on the right of the matrix:

$$\left[\begin{array}{cccccc|ccc} 1 & 0 & -\frac{31}{34} & \frac{173}{34} & \frac{133}{17} & -\frac{439}{34} & -\frac{1}{34} & \frac{3}{17} & 0 \\ 0 & 1 & \frac{8}{17} & -\frac{43}{17} & \frac{35}{17} & \frac{212}{17} & \frac{3}{17} & -\frac{1}{17} & 0 \\ 1 & -5 & 28 & 49 & -65 & -44 & 0 & 0 & 1 \end{array}\right].$$

We do not need to row-reduce this matrix from scratch, since the top left 2×2 submatrix is already reduced. Completing the row reduction, we get

$$\left[\begin{array}{cccccc|ccc} 1 & 0 & 0 & 6 & 6 & -12 & -\frac{3}{1063} & \frac{173}{1063} & \frac{31}{1063} \\ 0 & 1 & 0 & -3 & 3 & 12 & \frac{173}{1063} & -\frac{55}{1063} & -\frac{16}{1063} \\ 0 & 0 & 1 & 1 & -2 & 1 & \frac{31}{1063} & -\frac{16}{1063} & \frac{34}{1063} \end{array}\right].$$

Next we create a row containing terms of $s(n+3)_{n\geq 0}$:

$$\left[\begin{array}{ccccc|cccc} 1 & 0 & 0 & 6 & 6 & -\frac{3}{1063} & \frac{173}{1063} & \frac{31}{1063} & 0 \\ 0 & 1 & 0 & -3 & 3 & \frac{173}{1063} & -\frac{55}{1063} & -\frac{16}{1063} & 0 \\ 0 & 0 & 1 & 1 & -2 & \frac{31}{1063} & -\frac{16}{1063} & \frac{34}{1063} & 0 \\ -5 & 28 & 49 & -65 & -44 & 0 & 0 & 0 & 1 \end{array}\right],$$

which reduces to

$$\left[\begin{array}{ccccc|cccc} 1 & 0 & 0 & 6 & 6 & 0 & \frac{343}{2126} & \frac{63}{2126} & -\frac{1}{2126} \\ 0 & 1 & 0 & -3 & 3 & 0 & \frac{63}{2126} & -\frac{269}{6378} & \frac{173}{6378} \\ 0 & 0 & 1 & 1 & -2 & 0 & -\frac{1}{2126} & \frac{173}{6378} & \frac{31}{6378} \\ 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{6} & -\frac{1}{6} \end{array}\right].$$

By construction, the 4×4 matrix on the right is the product of the matrices that perform the row operations we applied to obtain the 4×5 matrix on the left. In particular, the last row implies

$$\begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{6} & -\frac{1}{6} \end{bmatrix} \cdot \begin{bmatrix} 2 & 6 & 1 & -5 & 28 \\ 6 & 1 & -5 & 28 & 49 \\ 1 & -5 & 28 & 49 & -65 \\ -5 & 28 & 49 & -65 & -44 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Scaling $\begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{6} & -\frac{1}{6} \end{bmatrix}$ so that its last entry is -1 shows that

$$s(n+3) = 6s(n) - 3s(n+1) + s(n+2)$$

for all $n \in \{0, 1, 2, 3, 4\}$. This is the same recurrence we guessed in Example 16.7.

You may have noticed that the top left 4×4 submatrices of the final row-reduced matrices in Examples 16.7 and 16.9 are identical, suggesting that we did not actually need to append an identity matrix in order to compute a linear relation in Example 16.9. However, this is a quirk of constant-recursive sequences and is explained by the fact that placing the first 4 terms of the shifts $s(n)_{n\geq 0}, s(n+1)_{n\geq 0}, s(n+2)_{n\geq 0}, s(n+3)_{n\geq 0}$ in columns produces the same 4×4 symmetric matrix as placing them in rows:

$$\begin{bmatrix} 2 & 6 & 1 & -5 \\ 6 & 1 & -5 & 28 \\ 1 & -5 & 28 & 49 \\ -5 & 28 & 49 & -65 \end{bmatrix}.$$

If we do not append an identity matrix when using row sequences, then in general we will not obtain a relation. For example, recomputing Example 16.9 with only the initial terms $2, 6, 1, -5, 28, 49$ produces

$$\left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 0 & \frac{343}{2126} & \frac{63}{2126} & -\frac{1}{2126} \\ 0 & 1 & 0 & 0 & \frac{63}{2126} & -\frac{269}{6378} & \frac{173}{6378} \\ 0 & 0 & 1 & 0 & -\frac{1}{2126} & \frac{173}{6378} & \frac{31}{6378} \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{6} & -\frac{1}{6} \end{array}\right],$$

where the matrix on the left gives no relation among the original rows.

Rigorous guessing

In Chapter 6, we discussed rigorous guessing for a polynomial sequence. We can also use guessing as a way to compute or rigorously prove a recurrence for a constant-recursive sequence.

Example 16.10. In Example 15.5 we computed a recurrence for $F(n)^5$. This required row-reducing 3 matrices — one for the product $F(n) \cdot F(n)$, a second for the product $F(n)^2 \cdot F(n)^2$, and a third for the product $F(n)^4 \cdot F(n)$. However, we can get by with only one row reduction by bounding the rank of $F(n)^5$ and guessing a recurrence using sufficiently many terms. Theorem 15.6 guarantees that $(F(n)^5)_{n\geq 0}$ is a constant-recursive sequence with rank at most $\binom{5+2-1}{5} = 6$. Therefore there is a linear relation among $(F(n)^5)_{n\geq 0}, (F(n+1)^5)_{n\geq 0}, \ldots, (F(n+6)^5)_{n\geq 0}$. We can compute this relation from the first 12 terms of $(F(n)^5)_{n\geq 0}$ by constructing the matrix

$$\begin{bmatrix} 0 & 1 & 1 & 32 & 243 & 3125 & 32768 \\ 1 & 1 & 32 & 243 & 3125 & 32768 & 371293 \\ 1 & 32 & 243 & 3125 & 32768 & 371293 & 4084101 \\ 32 & 243 & 3125 & 32768 & 371293 & 4084101 & 45435424 \\ 243 & 3125 & 32768 & 371293 & 4084101 & 45435424 & 503284375 \\ 3125 & 32768 & 371293 & 4084101 & 45435424 & 503284375 & 5584059449 \end{bmatrix}$$

containing the first 6 terms of the 7 shifts as column sequences. Row-reducing, we obtain

$$F(n+6)^5 = 8F(n+5)^5 + 40F(n+4)^5 - 60F(n+3)^5 - 40F(n+2)^5 + 8F(n+1)^5 + F(n)^5$$

as in Example 15.5. Since the rank of $(F(n)^5)_{n\geq 0}$ is at most 6 and the recurrence satisfied by the first 12 terms is unique (up to scaling the coefficients by a constant), this is a proof that the guessed recurrence holds for all $n \geq 0$.

The beauty of Theorems 15.3 and 15.6 is that their bounds provide a lot of information. Rather than carrying out the algorithms to explicitly compute closure properties, we can guess a recurrence for the result instead, and the recurrence is guaranteed to be correct if we use enough terms. This turns guessing into a method of proof.

Example 16.11. The Cassini[1] identity is

$$F(n-1)F(n+1) = F(n)^2 + (-1)^n.$$

Rather than computing recurrences for the two sides of the equation, we prove that this identity holds for all $n \geq 0$ (where $F(-1) = 1$) by bounding the rank of each side and checking the identity for sufficiently many values of n. There is no need to compute any recurrences or row-reduce any matrices. It is enough to observe that, by Theorems 15.3 and 15.6, $(F(n-1)F(n+1))_{n\geq 0}$ is constant-recursive with rank $\leq 2 \cdot 2 = 4$, and $(F(n)^2 + (-1)^n)_{n\geq 0}$ is constant-recursive with rank $\leq 3 + 1 = 4$. Therefore both sides generate sequences with rank ≤ 4, so if their first 8 terms agree, then are equal by Theorem 16.6. Indeed, the first 8 terms of both sequences are the same, namely $1, 0, 2, 3, 10, 24, 65, 168$.

Questions

Computations.

(1) Guess a recurrence for the sequence $1, 1, 3, 13, 59, 269, 1227, 5597, \dots$.

(2) Guess a recurrence for the sequence

$$0, 1, 1, 0, -2, -4, -4, 0, 8, 16, 16, 0, -32, -64, -64, 0, \dots.$$

(3) Guess a recurrence for the sequence

$$1, 1, 3, 5, 7, 11, 17, 27, 43, 67, 105, 165, 259, 407, 639, 1003, \dots.$$

Experiments.

(4) Pick some polynomial $s(n)$ with rank 7 (degree 6). Compute the first several terms of $s(n)_{n\geq 0}$, and guess a recurrence for this sequence. Is it the recurrence we expect a rank-7 polynomial sequence to satisfy?

(5) The first 4 terms of $s(n)_{n\geq 0}$ are a, b, c, d.
   (a) Guess a recurrence for $s(n)_{n\geq 0}$.
   (b) Which length-4 sequences a, b, c, d cannot be extended to a constant-recursive sequence with rank 2?

(6) Consider the matrix

$$M = \begin{bmatrix} 1 & 3 \\ -2 & 0 \end{bmatrix},$$

and let $s(n)$ be the top left entry of M^n. Does $s(n)_{n\geq 0}$ seem to be a constant-recursive sequence? What about the other entries?

(7) Compute the first few terms in the power series expansion, centered at $x = 0$, of each of the following functions. Does the sequence of coefficients seem to be constant-recursive?

[1] Giovanni Cassini was born in 1625 in Perinaldo (now in Italy) and died in 1712 in Paris, France.

(a) $\frac{2-x}{1-x-x^2}$
(b) $\frac{1}{1-2x-3x^2-4x^3}$

(8) Does the sequence $(n!)_{n\geq 0}$ seem to be constant-recursive?

(9) Does the sequence of Catalan numbers seem to be constant-recursive?

(10) A *Pythagorean triple* is a 3-tuple (a, b, c) of positive integers such that $a^2+b^2 = c^2$. The Pythagorean triples in which a is odd and $|a-b| = 1$ are listed in the following table. Extend this table by searching for additional examples. Which columns seem to be constant-recursive sequences?

a	b	c
3	4	5
21	20	29
119	120	169
$\vdots$	$\vdots$	$\vdots$

(11) A word on $\{0, 1, 2\}$ is *alternating* if no two consecutive letters are equal.
(a) Let $s(n)$ be the number of length-n alternating words on $\{0, 1, 2\}$. Does it seem to be a constant-recursive sequence? Is there a formula for $s(n)$?
(b) What if we only count alternating words that don't start with 2 and don't end with 2?
(c) What if we only count alternating words that don't start with 2 and don't end with 1?

(12) If $s(n)_{n\geq 0}$ is a periodic sequence, is it possible for $\dim \mathsf{ShiftSpace}(s)$ to be less than the period length?

Computation proofs.

(13) Let $F(n)$ be the nth Fibonacci number, and let $T(n)$ be the nth triangular number. Guess a recurrence for each sequence, and prove that this recurrence is correct by bounding the rank and testing sufficiently many values of n.
(a) $(F(n) + T(n))_{n\geq 0}$
(b) $(F(n)T(n))_{n\geq 0}$

(14) Prove each identity by bounding the rank of each side and testing sufficiently many values of n.
(a) $L(n)^2 = 5F(n)^2 + (-1)^n 4$
(b) $a(n)^2 + b(n)^2 = c(n)^2$ where $(a(n), b(n), c(n))$ is the nth Pythagorean triple in Question (10)

Programs.

(15) Write a program that implements the method of undetermined coefficients.

(16) Write a program that implements Incremental guessing, placing shifts in columns. How does its speed compare to the method of undetermined coefficients when the rank is small compared to the number of given terms? How does its speed compare when the rank is roughly equal to the number of given terms?

(17) Write a program that implements Incremental guessing, placing shifts in rows. How does its speed compare to the other methods when the rank is small and when the rank is large?

CHAPTER 17

Generating series

There is an important way to work with sequences that we have not explored yet — as power series. Translating between sequences and series will give us a natural way to perform several operations on sequences, lead to new characterizations of constant-recursive sequences, establish several additional closure properties, and provide a new technique for enumerating combinatorial objects.

Sequences as series

Definition 17.1. A *series in x* is an expression $c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots$ where $c_n \in \mathbb{Q}$ for all $n \geq 0$. The *generating series* of the sequence $s(n)_{n\geq 0}$ is the series

$$\sum_{n\geq 0} s(n)x^n = s(0) + s(1)x + s(2)x^2 + s(3)x^3 + \cdots .$$

For example, the generating series of the sequence of triangular numbers is $\sum_{n\geq 0} T(n)x^n = 0 + x + 3x^2 + 6x^3 + 10x^4 + \cdots$. Just as a polynomial in x is a mathematical object in its own right without ever substituting a number for x, so is a series. We won't be evaluating series at a point $x = a$, so we should think of x as a symbol, not a variable. In particular, we are not concerned with convergence of series.

We define addition and multiplication of series in the natural way.

Definition 17.2. Let $\sum_{n\geq 0} s(n)x^n$ and $\sum_{n\geq 0} t(n)x^n$ be series. Their sum is defined to be

$$\sum_{n\geq 0} (s(n) + t(n))x^n,$$

and their product is defined to be

$$\sum_{n\geq 0} \left(\sum_{i+j=n} s(i)t(j) \right) x^n = s(0)t(0) + (s(0)t(1) + s(1)t(0))x + \cdots , \tag{17.1}$$

where the inner sum is over all pairs (i, j) of non-negative integers that sum to n.

The first n terms of the product of two series can be computed by truncating the two series after n terms, multiplying the two resulting polynomials, and then truncating the product after n terms. Therefore multiplication of series is associative and commutative, and multiplication by the series $1 = 1 + 0x + 0x^2 + 0x^3 + \cdots$ leaves a series unchanged. Next we ask when a series has a multiplicative inverse.

Example 17.3. Is the series $x = 0 + 1x + 0x^2 + 0x^3 + \cdots$ invertible? Suppose the series $g(x) = \sum_{n\geq 0} s(n)x^n$ satisfies is $xg(x) = 1$. The left side has constant term 0, whereas the right side has constant term 1, so there is no such $g(x)$ and x is not invertible.

The series $\sum_{n\geq 0} T(n)x^n$ is also not invertible, for the same reason. Fortunately, a 0 constant term is the only obstruction to a series being invertible.

Proposition 17.4. *If $f(x)$ is a series whose constant term is not 0, then there exists a unique series $g(x)$ such that $f(x)g(x) = 1$.*

PROOF. We use induction on n to show that there exists a unique sequence $t(n)_{n\geq 0}$ such that $f(x) \cdot \sum_{n\geq 0} t(n)x^n = 1$. Let $s(n)$ be the coefficient of x^n in $f(x)$, and let $g(x) = \sum_{n\geq 0} t(n)x^n$. If $s(0) \neq 0$, then Equation (17.1) shows that $\frac{1}{s(0)}$ is the unique value of $t(0)$ that results in the constant term of $f(x)g(x)$ being 1. Inductively, assume that $n \geq 1$ and that $t(0), t(1), \ldots, t(n)$ are uniquely determined by the equation $f(x)g(x) = 1$. The coefficient of x^{n+1} in $f(x)g(x)$ is $\sum_{i=0}^{n+1} s(i)t(n+1-i)$. Since the coefficient of x^{n+1} in the right side of $f(x)g(x) = 1$ is 0, we have

$$\left(\sum_{i=1}^{n+1} s(i)t(n+1-i)\right) + s(0)t(n+1) = 0,$$

which uniquely determines $t(n+1)$, again since $s(0) \neq 0$. □

If $f(x)g(x) = 1$, then $f(x)$ is invertible, and we will write $\frac{1}{f(x)}$ for the series $g(x)$. More generally, if $f(x), g(x), h(x)$ are series such that $f(x)g(x) = h(x)$ and $f(x)$ is invertible, we write $g(x) = \frac{h(x)}{f(x)}$. In certain special cases, we will even extend this notation to non-invertible series. For example, the series x is not invertible, so "$\frac{1}{x}$" is not a series. However, we should still make sense of $\frac{0+x+x^2+x^3+\cdots}{x}$ as $1 + x + x^2 + \cdots$. Cancelling a power of x from the numerator and denominator is justified since $x^m f(x)g(x) = x^m h(x)$ implies $f(x)g(x) = h(x)$.

The proof of Proposition 17.4 gives an algorithm to compute the first n terms of the inverse of an invertible series.

Example 17.5. Let $f(x) = \sum_{n\geq 0} x^n = 1+x+x^2+x^3+\cdots$ be the generating series of the constant sequence whose nth term is $s(n) = 1$. Let $t(n)$ be the coefficient of x^n in $\frac{1}{f(x)}$. From Equation (17.1), the equations

$$\begin{aligned} s(0)t(0) &= 1 \\ s(0)t(1) + s(1)t(0) &= 0 \\ s(0)t(2) + s(1)t(1) + s(2)t(0) &= 0 \\ &\vdots \end{aligned}$$

uniquely determine $t(n)_{n\geq 0}$, namely $t(0) = 1$, $t(1) = -1$, $t(2) = 0$, $t(3) = 0$, and so on. In fact you can prove that $t(n) = 0$ for all $n \geq 2$, so we have

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

This implies the *geometric series formula*

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots. \tag{17.2}$$

In continuous mathematics, if x is a real number satisfying $|x| < 1$, then the geometric series formula is an equation relating two real numbers. However, for us x is not a number, so the condition "$|x| < 1$" doesn't make sense, and instead the geometric series formula is an equation relating two series.

In light of Equation (17.2), we say that the generating series of $1, 1, 1, \ldots$ is $\frac{1}{1-x}$. We will be able to represent many generating series in a finite way like this. This is the main feature of generating series — they provide a third way to represent sequences. Rather than specify $s(n)$ by a formula or with a recurrence, we can specify an expression whose nth series coefficient is $s(n)$.

We will be particularly interested in the correspondence between certain classes of sequences and certain classes of series. Throughout the remainder of the book, we'll see characterizations of several properties of sequences in terms of their generating series. Here is a simple one: A sequence has only finitely many nonzero terms if and only if its generating series is a polynomial. For example, the generating series of the 3rd row of Pascal's triangle $1, 3, 3, 1, 0, 0, 0, \ldots$ is $1+3x+3x^2+x^3 = (1+x)^3$.

Some simple series

We'll want to be able to convert in both directions — from sequences to series and from series to sequences. Let's start with some basic sequences to better understand the correspondence. We will use several tricks, but, to paraphrase Pólya[1], a trick that works more than once becomes a method.

Example 17.6. Let $f(x) = \sum_{n\geq 0} 2^n x^n = 1+2x+4x^2+8x^3+\cdots$ be the generating series of $(2^n)_{n\geq 0}$. Can we write $f(x)$ in some more compact way? The trick is that $(2^n)_{n\geq 0}$ can be transformed into $(2^{n+1})_{n\geq 0}$ in two different ways — by shifting or by multiplying by 2. Shifting a sequence is equivalent to subtracting the constant term from its generating series and dividing by x. In the language of series, we have

$$\begin{aligned}\frac{(1+2x+4x^2+8x^3+\cdots)-1}{x} &= 2+4x+8x^2+\cdots \\ &= 2(1+2x+4x^2+\cdots).\end{aligned}$$

Therefore $\frac{f(x)-1}{x} = 2f(x)$, so the generating series of $(2^n)_{n\geq 0}$ is $f(x) = \frac{1}{1-2x}$.

The same trick establishes the generating series of the sequence of powers of a.

Proposition 17.7. *Let $a \in \mathbb{Q}$. The generating series of $(a^n)_{n\geq 0}$ is $\frac{1}{1-ax}$.*

Proof. Let $f(x) = \sum_{n\geq 0} a^n x^n$. We have

$$f(x) - 1 = \sum_{n\geq 1} a^n x^n = \sum_{n\geq 0} a^{n+1} x^{n+1} = ax \sum_{n\geq 0} a^n x^n = axf(x).$$

Therefore $f(x) = \frac{1}{1-ax}$. □

Generating series behave nicely under dilations of a sequence.

Example 17.8. What is the generating series of the sequence $1, 0, 2, 0, 4, 0, 8, 0 \ldots$ obtained from $(2^n)_{n\geq 0}$ by inserting a 0 between each pair of consecutive terms? The generating series $1+0x+2x^2+0x^3+4x^4+0x^5+8x^6+0x^7+\cdots$ can be obtained from $1+2x+4x^2+8x^3+\cdots$ by replacing each x with x^2, so

$$1+0x+2x^2+0x^3+4x^4+0x^5+8x^6+0x^7+\cdots = \frac{1}{1-2x^2}.$$

[1] George Pólya was born in 1887 in Budapest, Austria-Hungary and died in 1985 in Palo Alto, California, U.S.

Next let's look at a simple polynomial sequence.

Example 17.9. What is the generating series of $(n)_{n\geq 0}$? It may not be clear how to write

$$\sum_{n\geq 0} nx^n = 0 + 1x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + 6x^6 + 7x^7 + \cdots$$

in a different form. We will develop a general technique to do this, but for now let's use another trick. Equation (17.1) describes the product of two series, so if we can think of two sequences $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ such that $\sum_{i=0}^{n} s(i)t(n-i) = n$ for all $n \geq 0$, then

$$\sum_{n\geq 0} nx^n = \left(\sum_{n\geq 0} s(n)x^n\right)\left(\sum_{n\geq 0} t(n)x^n\right).$$

A first guess might be $s(n) = 1$ and $t(n) = 1$; this is close, but $\sum_{i=0}^{n} s(i)t(n-i) = \sum_{i=0}^{n} 1 = n+1$. Instead, let $s(n) = 1$ and

$$t(n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n \geq 1; \end{cases}$$

then $\sum_{i=0}^{n} s(i)t(n-i) = \sum_{i=0}^{n-1} 1 + 0 = n$. It remains to determine the generating series of $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$. Example 17.5 shows that the generating series of $s(n)_{n\geq 0}$ is $\frac{1}{1-x}$. The generating series of $t(n)_{n\geq 0}$ is $\frac{1}{1-x} - 1 = \frac{x}{1-x}$. Therefore

$$\sum_{n\geq 0} nx^n = \frac{1}{1-x} \cdot \frac{x}{1-x} = \frac{x}{(1-x)^2}.$$

Generating series of constant-recursive sequences

Let us now turn our attention to obtaining generating series from a recurrence.

Example 17.10. What is the generating series of the Fibonacci sequence $F(n)_{n\geq 0}$? Let

$$f(x) = \sum_{n\geq 0} F(n)x^n = 0 + 1x + 1x^2 + 2x^3 + 3x^4 + 5x^5 + 8x^6 + 13x^7 + \cdots.$$

To attempt to write $f(x)$ in a finite form, we will use the Fibonacci recurrence $F(n+2) = F(n+1) + F(n)$. Accordingly, we should also consider the generating series of the shifts $F(n+1)_{n\geq 0}$ and $F(n+2)_{n\geq 0}$. Since $F(n+1)_{n\geq 0}$ is obtained from $F(n)_{n\geq 0}$ by dropping the first term, we have

$$\begin{aligned}\sum_{n\geq 0} F(n+1)x^n &= 1 + 1x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + 13x^6 + 21x^7 + \cdots \\ &= \frac{f(x) - 0}{x}.\end{aligned}$$

We can get the second shift $F(n+2)_{n\geq 0}$ by dropping the first two terms:

$$\begin{aligned}\sum_{n\geq 0} F(n+2)x^n &= 1 + 2x + 3x^2 + 5x^3 + 8x^4 + 13x^5 + 21x^6 + 34x^7 + \cdots \\ &= \frac{f(x) - 0 - x}{x^2}.\end{aligned}$$

Now use the Fibonacci recurrence:

$$\frac{f(x)-x}{x^2} = \sum_{n\geq 0} F(n+2)x^n = \sum_{n\geq 0} \big(F(n+1)+F(n)\big)x^n = \frac{f(x)}{x} + f(x).$$

We can rewrite this equation as $f(x) - x = xf(x) + x^2 f(x)$. Solving for $f(x)$ gives $f(x) = \frac{x}{1-x-x^2}$. This is the generating series of the Fibonacci sequence.

Despite the Fibonacci sequence being more sophisticated in some ways than $(2^n)_{n\geq 0}$ and $(n)_{n\geq 0}$, its generating series is nonetheless quite simple. It is a rational expression.

Definition 17.11. A *rational expression* is an expression $\frac{a(x)}{b(x)}$ where $a(x), b(x) \in \mathbb{Q}[x]$ and $b(x)$ is not the 0 polynomial. A series $\sum_{n\geq 0} s(n)x^n$ is *rational* if there are polynomials $a(x), b(x) \in \mathbb{Q}[x]$ such that $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)}$.

If general, given a constant-recursive sequence, we can set up the generating series of the shifts appearing in its recurrence and then play the same game as in Example 17.10.

Theorem 17.12. *If $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank r, then $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)}$ for some $a(x), b(x) \in \mathbb{Q}[x]$ such that $\deg a(x) \leq r-1$, $\deg b(x) \leq r$, and the constant term of $b(x)$ is 1. In particular, $\sum_{n\geq 0} s(n)x^n$ is rational.*

Proof. Let $f(x) = \sum_{n\geq 0} s(n)x^n$ be the generating series of $s(n)_{n\geq 0}$, and let

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n)$$

be the recurrence satisfied by $s(n)_{n\geq 0}$ for all $n \geq 0$. Multiply both sides by x^{n+r} and sum over $n \geq 0$ to obtain

$$\sum_{n\geq 0} s(n+r)x^{n+r} = c_{r-1}x \sum_{n\geq 0} s(n+r-1)x^{n+r-1} + \cdots + c_1 x^{r-1} \sum_{n\geq 0} s(n+1)x^{n+1} + c_0 x^r \sum_{n\geq 0} s(n)x^n.$$

Next we relate each of these $r+1$ sums to $f(x)$ by filling in the terms that are missing. The result is

$$f(x) - \sum_{n=0}^{r-1} s(n)x^n = c_{r-1}x\left(f(x) - \sum_{n=0}^{r-2} s(n)x^n\right) + \cdots + c_1 x^{r-1}\left(f(x) - \sum_{n=0}^{0} s(n)x^n\right) + c_0 x^r f(x).$$

Collecting the terms involving $f(x)$, we find

$$\left(1 - c_{r-1}x - \cdots - c_1 x^{r-1} - c_0 x^r\right) f(x) = \sum_{n=0}^{r-1} s(n)x^n - \sum_{i=1}^{r-1} c_i x^{r-i} \sum_{n=0}^{i-1} s(n)x^n.$$

At this point we solve for $f(x)$ and see that it is rational with a numerator whose degree is at most $r-1$ and a denominator whose degree is at most r and whose constant term is 1, which completes the proof. However, we perform one last manipulation on the numerator to show how each coefficient depends on the terms $s(0), \dots, s(r-1)$ and on $c_0, \dots, c_{r-1}$. We obtain the explicit rational expression

$$f(x) = \frac{\sum_{n=0}^{r-1} \big(s(n) - c_{r-1}s(n-1) - \cdots - c_{r-(n-1)}s(1) - c_{r-n}s(0)\big)x^n}{1 - c_{r-1}x - \cdots - c_1x^{r-1} - c_0x^r}$$

for the series $f(x)$. □

Now let's ask about the converse: Is the sequence of coefficients in a rational series necessarily constant-recursive?

Example 17.13. Let $s(n)$ be the coefficient of x^n in the series $\frac{x^2}{1-x-x^2-x^3}$. Is $s(n)_{n\geq 0}$ a constant-recursive sequence? We have

$$\big(1 - x - x^2 - x^3\big) \sum_{n\geq 0} s(n)x^n = x^2.$$

Let us attempt to reverse the steps of Example 17.10. First we expand:

$$\sum_{n\geq 0} s(n)x^n - \sum_{n\geq 0} s(n)x^{n+1} - \sum_{n\geq 0} s(n)x^{n+2} - \sum_{n\geq 0} s(n)x^{n+3} = x^2.$$

To get a recurrence, we should gather like powers of x. The problem is that the exponents n, $n+1$, $n+2$, and $n+3$ are all different. We can fix this by re-indexing the sums:

$$\sum_{m\geq -3} s(m+3)x^{m+3} - \sum_{m\geq -2} s(m+2)x^{m+3} - \sum_{m\geq -1} s(m+1)x^{m+3} - \sum_{m\geq 0} s(m)x^{m+3} = x^2.$$

Each of these sums includes the range $m \geq 0$, but the first three sums contain extra terms. Let's pull these terms out, since then we can combine the sums:

$$\text{polynomial} + \sum_{m\geq 0} \big(s(m+3) - s(m+2) - s(m+1) - s(m)\big)x^{m+3} = x^2,$$

where the omitted polynomial is $s(0) + \big(s(1) - s(0)\big)x + \big(s(2) - s(1) - s(0)\big)x^2$. Since this equation is an equality of two series, the coefficient of x^{m+3} in the series on the left equals the coefficient of x^{m+3} in the series on the right. In particular, the recurrence

$$s(m+3) - s(m+2) - s(m+1) - s(m) = 0$$

holds for all $m \geq 0$. The coefficients of x^0, x^1, and x^2 give us equations involving the initial conditions:

$$\begin{aligned} s(0) &= 0 \\ s(1) - s(0) &= 0 \\ s(2) - s(1) - s(0) &= 1. \end{aligned}$$

Therefore $s(0) = 0$, $s(1) = 0$, and $s(2) = 1$. These are the recurrence and initial conditions satisfied by the Tribonacci sequence $E(n)_{n\geq 0}$. So the generating series of the Tribonacci sequence is $\frac{x^2}{1-x-x^2-x^3}$.

The same procedure works for arbitrary constant-recursive sequences.

Theorem 17.14. *If* $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)}$ *where* $a(x), b(x) \in \mathbb{Q}[x]$, *then* $s(n)_{n\geq 0}$ *is constant-recursive with rank* $\leq \max(1 + \deg a(x), \deg b(x))$.

PROOF. We derive a recurrence for $s(n)_{n\geq 0}$. If $b(x)$ is divisible by x, then the coefficient of x^0 in $b(x)\sum_{n\geq 0} s(n)x^n = a(x)$ is 0, so $a(x)$ is also divisible by x and we can cancel an x from $b(x)$ and $a(x)$. Therefore, without loss of generality, assume that $b(x)$ is not divisible by x. Let $r = \deg b(x)$, and write

$$b(x) = c\left(1 - c_{r-1}x - \cdots - c_1 x^{r-1} - c_0 x^r\right)$$

where $c, c_{r-1}, \ldots, c_1, c_0 \in \mathbb{Q}$ and $c \neq 0$. We have

$$\left(1 - c_{r-1}x - \cdots - c_1 x^{r-1} - c_0 x^r\right) \sum_{n\geq 0} s(n)x^n = \frac{1}{c}a(x).$$

Define $c_r = -1$, so that we can write

$$\begin{aligned}\frac{1}{c}a(x) &= \left(-\sum_{i=0}^{r} c_i x^{r-i}\right)\left(\sum_{n\geq 0} s(n)x^n\right)\\ &= -\sum_{i=0}^{r}\sum_{n\geq 0} c_i s(n) x^{n+r-i}\\ &= -\sum_{i=0}^{r}\sum_{m\geq -i} c_i s(m+i)x^{m+r}\end{aligned}$$

after the change of variables $n = m+i$. Next we split the inner sum into two pieces so that we can easily change the order of summation:

$$\begin{aligned}\frac{1}{c}a(x) &= -\sum_{i=0}^{r}\sum_{m=-i}^{-1} c_i s(m+i)x^{m+r} - \sum_{i=0}^{r}\sum_{m\geq 0} c_i s(m+i)x^{m+r}\\ &= \text{polynomial} - \sum_{m\geq 0}\left(\sum_{i=0}^{r} c_i s(m+i)\right)x^{m+r}.\end{aligned} \tag{17.3}$$

For all m satisfying $m + r \geq N := \max(1 + \deg a(x), r)$, comparing the coefficients of x^{m+r} on both sides of the equation shows that

$$0 = \sum_{i=0}^{r} c_i s(m+i).$$

Now we are mostly done. Since $c_r = -1$, we move $s(m+r)$ to the other side, obtaining

$$s(m+r) = \sum_{i=0}^{r-1} c_i s(m+i)$$

for all m satisfying $m + r \geq N$. To obtain a recurrence that holds for all $n \geq 0$, we make the change of variables $m = n + N - r$, so that

$$s(n+N) = \sum_{i=0}^{r-1} c_i s(n+N-r+i)$$

for all $n \geq 0$. Therefore $s(n)_{n\geq 0}$ is constant-recursive with rank $\leq N$. □

Notice that if the constant term of $b(x)$ is 0 and the constant term of $a(x)$ is not 0, then the conditions of Theorem 17.14 are not satisfied since $b(x)$ is not invertible.

Theorems 17.12 and 17.14 establish a characterization of constant-recursive sequences in terms of their generating series.

Theorem 17.15. *A sequence is constant-recursive if and only if its generating series is rational.*

A useful observation regarding the previous two examples and previous two proofs is that the recurrence for $s(n)_{n\geq 0}$ is determined by the denominator of the rational expression for $\sum_{n\geq 0} s(n)x^n$ and can be read off directly. The denominator $1 - c_{r-1}x - \cdots - c_1 x^{r-1} - c_0 x^r$ corresponds to the recurrence

$$s(n+r) - c_{r-1}s(n+r-1) - \cdots - c_1 s(n+1) - c_0 s(n) = 0$$

for sufficiently large n. Moreover, the initial conditions are determined by the numerator.

Example 17.16. Let $s(n)_{n\geq 0}$ be the coefficient sequence of the series $\frac{3x-2x^2}{1-2x+4x^2-5x^3+100x^4}$. The recurrence satisfied by $s(n)_{n\geq 0}$ is

$$s(n+4) - 2s(n+3) + 4s(n+2) - 5s(n+1) + 100s(n) = 0.$$

Extracting the initial conditions requires a little more computation. We can use the proof of Theorem 17.14 to get them by comparing the coefficients of $x^0, x^1, \ldots, x^{N-1} = x^3$ on both sides of Equation (17.3) and solving the resulting system. Alternatively, we can compute them one at a time using a variant of polynomial long division. At each step, we divide lowest-order terms rather than highest-order terms. Let $b(x) = 1 - 2x + 4x^2 - 5x^3 + 100x^4$. The (infinitely!) long division is as follows.

$$\begin{array}{r|l}
 & 3x + 4x^2 - 4x^3 - 9x^4 + \cdots \\
\hline
b(x) & 3x - 2x^2 + 0x^3 + 0x^4 + 0x^5 + 0x^6 + 0x^7 + \cdots \\
 & -(3x - 6x^2 + 12x^3 - 15x^4 + 300x^5) \\
\cline{2-2}
 & \quad 4x^2 - 12x^3 + 15x^4 - 300x^5 \\
 & \quad -(4x^2 - 8x^3 + 16x^4 - 20x^5 + 400x^6) \\
\cline{2-2}
 & \qquad -4x^3 - x^4 - 280x^5 - 400x^6 \\
 & \qquad -(-4x^3 + 8x^4 - 16x^5 + 20x^6 - 400x^7) \\
\cline{2-2}
 & \qquad\quad -9x^4 - 264x^5 - 420x^6 + 400x^7 \\
 & \qquad\qquad\qquad \ddots
\end{array}$$

Therefore

$$\frac{3x-2x^2}{1-2x+4x^2-5x^3+100x^4} = 3x + 4x^2 - 4x^3 - 9x^4 + \cdots,$$

so the initial conditions are $s(0) = 0$, $s(1) = 3$, $s(2) = 4$, and $s(3) = -4$.

Example 17.17. Suppose $s(n)_{n\geq 0}$ is a sequence satisfying the recurrence

$$s(n+3) = 5s(n+2) - s(n+1) + 2s(n)$$

for all $n \geq 0$. Immediately this implies that its generating series is of the form

$$\sum_{n\geq 0} s(n)x^n = \frac{\text{polynomial}}{1 - 5x + x^2 - 2x^3}.$$

If we specify that the initial conditions are $s(0) = 1$, $s(1) = -1$, and $s(2) = 4$, then we can compute the numerator. The degree of the numerator is at most 2 by Theorem 17.12, so

$$1 - x + 4x^2 + \cdots = \frac{a_0 + a_1 x + a_2 x^2}{1 - 5x + x^2 - 2x^3}$$

for some coefficients a_0, a_1, a_2. Multiplying by the denominator gives

$$\begin{aligned}\left(1 - 5x + x^2 - 2x^3\right)\left(1 - x + 4x^2 + \cdots\right) &= a_0 + a_1 x + a_2 x^2 \\ 1 - 6x + 10x^2 + \cdots &= a_0 + a_1 x + a_2 x^2,\end{aligned}$$

so $a_0 = 1, a_1, = -6, a_2 = 10$ and the remaining terms on the left side are 0. The generating series of $s(n)_{n\geq 0}$ is therefore $\frac{1-6x+10x^2}{1-5x+x^2-2x^3}$.

Revisiting closure properties

Generating series give us a new way of computing the sum of two constant-recursive sequences, using symbolic algebra rather than linear algebra.

Example 17.18. In Example 15.1 we row-reduced a matrix to compute a recurrence satisfied by $s(n) := F(n) + E(n)$. Using generating series, we compute

$$\begin{aligned}\sum_{n\geq 0} \left(F(n) + E(n)\right) x^n = \sum_{n\geq 0} F(n) x^n + \sum_{n\geq 0} E(n) x^n &= \frac{x}{1 - x - x^2} + \frac{x^2}{1 - x - x^2 - x^3} \\ &= \frac{x - 2x^3 - 2x^4}{1 - 2x - x^2 + x^3 + 2x^4 + x^5}.\end{aligned}$$

We can read off a recurrence directly from the denominator, namely

$$s(n+5) = 2s(n+4) + s(n+3) - s(n+2) - 2s(n+1) - s(n) \tag{17.4}$$

for all $n \geq 0$.

Theorem 17.15 also provides several new closure properties. The first is a converse of Theorem 15.9. It states that "riffling" or "interlacing" multiple constant-recursive sequences produces a constant-recursive sequence. We proved a special case in Theorem 13.20.

Theorem 17.19. *If $s(n)_{n\geq 0}$ is a sequence and $m \geq 1$ is an integer such that $s(mn)_{n\geq 0}$, $s(mn+1)_{n\geq 0}$, $\dots$, and $s(mn+m-1)_{n\geq 0}$ are all constant-recursive sequences with rank $\leq r$, then $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq m^2 r$.*

Proof. For each $i \in \{0, 1, \dots, m-1\}$, let $f_i(x) = \sum_{n\geq 0} s(mn+i) x^n$ be the generating series of $s(mn+i)_{n\geq 0}$. The series $x^i f_i(x^m) = \sum_{n\geq 0} s(mn+i) x^{mn+i}$ is the generating series of the sequence obtained from $s(n)_{n\geq 0}$ by replacing every term that isn't contributed by $s(mn+i)_{n\geq 0}$ with 0. It follows that $x^0 f_0(x^m) + x^1 f_1(x^m) + \cdots + x^{m-1} f_{m-1}(x^m) = \sum_{n\geq 0} s(n) x^n$. By Theorem 17.12, $f_i(x) = \frac{a_i(x)}{b_i(x)}$ for some polynomials with $\deg a_i(x) \leq r - 1$ and $\deg b_i(x) \leq r$. This implies that $x^i f_i(x^m)$ is rational with numerator degree at most $i + m(r-1) \leq mr - 1$ and denominator degree at most mr. Therefore $\sum_{n\geq 0} s(n) x^n = \frac{a(x)}{b(x)}$ for some polynomials with $\deg a(x) \leq (mr-1) + (m-1)mr = m^2 r - 1$ and $\deg b(x) \leq m^2 r$. By Theorem 17.14, $s(n)_{n\geq 0}$ is constant-recursive with rank at most $m^2 r$. $\square$

Since the product of two rational series is rational, if $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ are constant-recursive then $\left(\sum_{i=0}^{n} s(i)t(n-i)\right)_{n\geq 0}$ is also constant-recursive. This sequence is called the *Cauchy*[2] *product* of $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$. In particular, we obtain the following by letting $t(n) = 1$ for all $n \geq 0$.

Corollary 17.20. *If $s(n)_{n\geq 0}$ is a constant-recursive sequence with rank r, then $\left(\sum_{i=0}^{n} s(i)\right)_{n\geq 0}$ is a constant-recursive sequence with rank $\leq r+1$.*

Proof. By Theorem 17.12, $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)}$ for some $a(x), b(x) \in \mathbb{Q}[x]$ such that $\deg a(x) \leq r-1$ and $\deg b(x) \leq r$. Since the generating series of $(1)_{n\geq 0}$ is $\frac{1}{1-x}$, the generating series of $\left(\sum_{i=0}^{n} s(i)\right)_{n\geq 0}$ is $\frac{a(x)}{(1-x)b(x)}$ by Definition 17.2. Theorem 17.14 implies that the rank of $\left(\sum_{i=0}^{n} s(i)\right)_{n\geq 0}$ is at most $\max(1+\deg a(x), 1+\deg b(x)) \leq \max(1+r-1, 1+r) = r+1$. □

The proof of Corollary 17.20 establishes that multiplying a generating series by $\frac{1}{1-x}$ has the effect of taking partial sums of its coefficients.

Example 17.21. Since the generating series of $F(n)_{n\geq 0}$ is $\frac{x}{1-x-x^2}$, the generating series of $\left(\sum_{i=0}^{n} F(i)\right)_{n\geq 0}$ is $\frac{x}{(1-x)(1-x-x^2)} = \frac{x}{1-2x+x^3}$. In particular, $s(n) := \sum_{i=0}^{n} F(i)$ satisfies $s(n+3) = 2s(n+2) - s(n)$ for all $n \geq 0$.

Since multiplying by $\frac{1}{1-x}$ gives partial sums, multiplying by $1-x$ should undo the summation. Indeed, multiplying the generating series of $s(n)_{n\geq 0}$ by $1-x$ produces the generating series of $s(n) - s(n-1)$, where we define $s(-1) = 0$.

This is closely related to the interaction between the difference operator Δ_n and the partial sum $\sum_{i=0}^{n-1} s(i)$ we established in Chapter 10. Removing the constant term from the generating series of $s(n)_{n\geq 0}$ and then multiplying by $\frac{1}{x} - 1 = \frac{1-x}{x}$ gives the generating series of $\Delta_n s(n) = s(n+1) - s(n)$. Since the difference operator serves as an inverse operator of partial summation, the generating series of $\sum_{i=0}^{n-1} s(i)$ is obtained by multiplying by $\frac{x}{1-x}$. To summarize:

sequence operation	nth term	corresponding series operation
shift	$s(n+1)$	subtract $s(0)$ and multiply by $\frac{1}{x}$
difference	$s(n+1) - s(n)$	subtract $s(0)$ and multiply by $\frac{1-x}{x}$
shifted difference	$s(n) - s(n-1)$	multiply by $1-x$
partial sum	$\sum_{i=0}^{n-1} s(i)$	multiply by $\frac{x}{1-x}$
shifted partial sum	$\sum_{i=0}^{n} s(i)$	multiply by $\frac{1}{1-x}$

In particular, generating series give another way (in addition to Propositions 14.2 and 14.3) to see that if $s(n)_{n\geq 0}$ is constant-recursive then $(\Delta_n s(n))_{n\geq 0}$ is also constant-recursive and satisfies the same recurrence; the corresponding operation on the generating series does not affect the denominator.

If the numerator is initially divisible by $1-x$, then the sequence of partial sums can have the same rank as $s(n)_{n\geq 0}$.

Example 17.22. Let $s(n)_{n\geq 0} = 1, 1, 1, 3, 5, 9, 17, 31, \dots$ be the sequence of coefficients in the series $\frac{1-x^2}{1-x-x^2-x^3}$. The generating series of $\left(\sum_{i=0}^{n} s(i)\right)_{n\geq 0}$ is $\frac{1}{1-x} \cdot \frac{1-x^2}{1-x-x^2-x^3} = \frac{1+x}{1-x-x^2-x^3}$.

[2]Augustin-Louis Cauchy (pronounced 'kohshee') was born in 1789 in Paris, France and died in 1857 in Sceaux, France.

The factorizations of the numerator and denominator of a rational series $\frac{a(x)}{b(x)}$ can give us other information as well. If they have a common factor, then $b(x)$ does not correspond to the minimal recurrence of the coefficient sequence. In particular, the factors of a polynomial $b(x)$ determine the possible minimal recurrences satisfied by the coefficient sequence of a series of the form $\frac{a(x)}{b(x)}$.

Corollary 17.23. *Let $s(n)_{n\geq 0}$ be a constant-recursive sequence, and let*

$$s(n+r) = c_{r-1}s(n+r-1) + \cdots + c_1 s(n+1) + c_0 s(n)$$

be the minimal recurrence that holds for all $n \geq 0$. If $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)}$ for some $a(x), b(x) \in \mathbb{Q}[x]$, then $b(x)$ is divisible by $1 - c_{r-1}x - \cdots - c_1 x^{r-1} - c_0 x^r$.

Proof. Let $B(x) = 1 - c_{r-1}x - \cdots - c_1 x^{r-1} - c_0 x^r$. By the proof of Theorem 17.12, $\sum_{n\geq 0} s(n)x^n = \frac{A(x)}{B(x)}$ for some $A(x) \in \mathbb{Q}[x]$ with $\deg A(x) \leq r-1$. Therefore $\frac{a(x)}{b(x)} = \frac{A(x)}{B(x)}$, which implies $a(x)B(x) = A(x)b(x)$. In particular, $B(x)$ divides $A(x)b(x)$. The minimality of the recurrence implies that $A(x)$ and $B(x)$ have no common factors: if they did have a common factor, then canceling that factor in $\frac{A(x)}{B(x)}$ would produce a smaller denominator, corresponding to a smaller recurrence for $s(n)_{n\geq 0}$ by the proof of Theorem 17.14. Since $A(x)$ and $B(x)$ have no common factors, it follows from unique factorization of polynomials that $B(x)$ divides $b(x)$. $\square$

Example 17.24. Suppose $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{1-2x-x^2+x^3+2x^4+x^5}$ for some $a(x) \in \mathbb{Q}[x]$ with $\deg a(x) \leq 4$. Since the divisors of the denominator are itself, $1 - x - x^2$, and $1 - x - x^2 - x^3$, the minimal recurrence of $s(n)_{n\geq 0}$ is either Equation (17.4), the Fibonacci recurrence, the Tribonacci recurrence, or the recurrence $s(n) = 0$.

Generating series of polynomial sequences

Theorem 13.9 states that every polynomial sequence is also constant-recursive, so Theorem 17.15 applies to polynomial sequences in particular. We determined that $\sum_{n\geq 0} nx^n = \frac{x}{(1-x)^2}$ in Example 17.9. What about a quadratic polynomial sequence?

Example 17.25. Let $T(n) = \frac{n(n+1)}{2}$ be the nth triangular number. By Theorem 13.9, we know

$$T(n+3) - 3T(n+2) + 3T(n+1) - T(n) = 0$$

for all $n \geq 0$. Therefore $\sum_{n\geq 0} T(n)x^n$ is of the form $\frac{\text{polynomial}}{1-3x+3x^2-x^3}$. Since the coefficients in the denominator are binomial coefficients, the denominator factors as $(1-x)^3$. The numerator is

$$(1 - 3x + 3x^2 - x^3)(0 + x + 3x^2 + 6x^3 + \cdots) = x + 0x^2 + 0x^3 + \cdots$$

so $\sum_{n\geq 0} T(n)x^n = \frac{x}{(1-x)^3}$.

More generally, we have the following.

Theorem 17.26. *A sequence $s(n)_{n\geq 0}$ of rational numbers is a polynomial sequence with rank $\leq r$ if and only if $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{(1-x)^r}$ for some $a(x) \in \mathbb{Q}[x]$ with $\deg a(x) \leq r-1$.*

PROOF. In one direction, assume that $s(n)_{n\geq 0}$ is a polynomial sequence with rank $\leq r$. By Theorem 13.9,

$$\sum_{i=0}^{r}(-1)^{r-i}\binom{r}{i}s(n+i)=0 \tag{17.5}$$

for all $n\geq 0$. The proof of Theorem 17.12 now implies $\sum_{n\geq 0}s(n)x^n=\frac{a(x)}{b(x)}$ for some $a(x)\in\mathbb{Q}[x]$ with $\deg a(x)\leq r-1$, where

$$b(x)=\sum_{i=0}^{r}(-1)^{r-i}\binom{r}{i}x^{r-i}=(1-x)^r. \qquad \square$$

For the other direction, let $a(x)\in\mathbb{Q}[x]$ with $\deg a(x)\leq r-1$. Let $s(n)_{n\geq 0}$ be the coefficient sequence of the series $\frac{a(x)}{(1-x)^r}$. By Theorem 17.14, Equation (17.5) holds for all $n\geq 0$. By Theorem 14.19, the set of sequences that satisfies Equation (17.5) is $\mathsf{Poly}(r)$, so $s\in\mathsf{Poly}(r)$.

For columns of Pascal's triangle, the numerator of the generating series also takes a particularly simple form.

Example 17.27. What is the generating series of the sequence $\binom{n}{3}_{n\geq 0}$? We use the fact from Corollary 9.3 that $\binom{n+1}{m+1}=\sum_{i=0}^{n}\binom{i}{m}$ to write

$$\binom{n+1}{3}=\sum_{i=0}^{n}\binom{i}{2}=\sum_{i=0}^{n}T(i-1)$$

for all $n\geq -1$, where $T(-1)=0$. The generating series of $T(n)_{n\geq 0}$ is $\sum_{n\geq 0}T(n)x^n=\frac{x}{(1-x)^3}$, so

$$\sum_{n\geq 0}T(n-1)x^n=\sum_{m\geq 0}T(m)x^{m+1}=x\sum_{m\geq 0}T(m)x^m=\frac{x^2}{(1-x)^3}.$$

Therefore, by the proof of Corollary 17.20, $\sum_{n\geq 0}\binom{n+1}{3}x^n=\frac{x^2}{(1-x)^4}$. This implies

$$\sum_{n\geq 0}\binom{n}{3}x^n=\sum_{m\geq 0}\binom{m+1}{3}x^{m+1}=x\sum_{m\geq 0}\binom{m+1}{3}x^m=\frac{x^3}{(1-x)^4}.$$

Theorem 17.28. *For each $r\geq 1$, we have $\sum_{n\geq 0}\binom{n}{r-1}x^n=\frac{x^{r-1}}{(1-x)^r}$.*

PROOF. If $r=1$, then $\sum_{n\geq 0}\binom{n}{r-1}x^n=\sum_{n\geq 0}x^n=\frac{1}{1-x}=\frac{x^{r-1}}{(1-x)^r}$ by Proposition 17.7. Inductively, assume $\sum_{n\geq 0}\binom{n}{r-1}x^n=\frac{x^{r-1}}{(1-x)^r}$. We have

$$\sum_{n\geq 0}\binom{n}{r}x^n=\sum_{m\geq 0}\binom{m+1}{r}x^{m+1}=x\sum_{m\geq 0}\binom{m+1}{r}x^m.$$

By Corollary 9.3, $\binom{m+1}{r}=\sum_{i=0}^{m}\binom{i}{r-1}$. From the proof of Corollary 17.20, it follows that

$$\sum_{n\geq 0}\binom{n}{r}x^n=x\cdot\frac{1}{1-x}\cdot\frac{x^{r-1}}{(1-x)^r}=\frac{x^r}{(1-x)^{r+1}}$$

as desired. $\square$

Equivalently, $\sum_{n\geq 0}\binom{n}{d}x^n=\frac{x^d}{(1-x)^{d+1}}$ for each $d\geq 0$.

Example 17.29. Let $s(n) = \binom{n}{5} - 9\binom{n}{4} - 15\binom{n}{3} - 7\binom{n}{2}$. The generating series of $s(n)_{n\geq 0}$ is

$$\sum_{n\geq 0} s(n)x^n = \frac{x^5}{(1-x)^6} - \frac{9x^4}{(1-x)^5} - \frac{15x^3}{(1-x)^4} - \frac{7x^2}{(1-x)^3}.$$

The 4 rational expressions comprising the generating series in the previous example do not have the same denominator. However, a slightly different basis does produce expressions with the same denominator. Since multiplying a generating series by $\frac{1}{x}$ has the effect of shifting its sequence of coefficients, the series $\frac{x^{r-1}}{(1-x)^r}, \frac{x^{r-2}}{(1-x)^r}, \dots, \frac{1}{(1-x)^r}$ are the generating series of the first several shifts of $\binom{n}{r-1}_{n\geq 0}$:

$$\begin{aligned}
\frac{x^{r-1}}{(1-x)^r} &= \sum_{n\geq 0} \binom{n}{r-1} x^n \\
\frac{x^{r-2}}{(1-x)^r} &= \sum_{n\geq 0} \binom{n+1}{r-1} x^n \\
&\vdots \\
\frac{1}{(1-x)^r} &= \sum_{n\geq 0} \binom{n+r-1}{r-1} x^n.
\end{aligned}$$

Therefore the basis of $\mathsf{Poly}(r)$ consisting of the first r shifts of $\binom{n}{r-1}_{n\geq 0}$ is the same basis of $\mathsf{Poly}(r)$ consisting of the sequences whose generating series are $\frac{x^i}{(1-x)^r}$ for $i \in \{0, 1, \dots, r-1\}$. In other words, for a polynomial sequence, the coefficients in the numerator of its generating series are its coordinates in the basis of shifts of $\binom{n}{r-1}_{n\geq 0}$.

Example 17.30. Let $s(n) = 2\binom{n}{5} + 6\binom{n+2}{5} - 7\binom{n+3}{5}$. We have $\operatorname{rank}(s) = 6$, and we can immediately write down the generating series of $s(n)_{n\geq 0}$:

$$\sum_{n\geq 0} s(n)x^n = \frac{2x^5 + 6x^3 - 7x^2}{(1-x)^6}.$$

Properties of a natural class of integer sequences

We have now seen that the class of constant-recursive sequences has a number of desirable properties.

- Examples occur naturally in mathematics.
- They satisfy several closure properties.
- They have multiple equivalent characterizations (for example, in terms of recurrences, shifts, and generating series).
- They can be guessed from finitely many initial terms by solving a system of linear equations.
- They have simple relationships with other natural classes of sequences.

The bar has been set. As we go forward and study additional classes of integer sequences, we will ideally have each of these properties for each class we encounter.

Questions

Computations.

(1) Let $s \in \mathsf{ShiftSpace}(F)$ with $s(0) = 4$ and $s(1) = 7$. What is the generating series of $s(n)_{n\geq 0}$?

(2) Let $s(n)_{n\geq 0}$ be the sequence of Perrin numbers, defined by $s(0) = 3$, $s(1) = 0$, $s(2) = 2$, and $s(n+3) = s(n+1) + s(n)$ for all $n \geq 0$. What is the generating series of $s(n)_{n\geq 0}$?

(3) What is the generating series of $(n^3 - n^2)_{n\geq 0}$?

(4) For each rational series, what recurrence does the sequence of coefficients satisfy? What are the initial conditions?
   (a) $\frac{2-x}{1-x-x^2}$
   (b) $\frac{1}{1-2x-3x^2-4x^3}$

(5) Let $F(n)$ be the nth Fibonacci number, and let $T(n)$ be the nth triangular number. Determine the series $\sum_{n\geq 0} (F(n) + T(n))x^n$ in two ways — first by rigorously guessing a recurrence for $(F(n) + T(n))_{n\geq 0}$ as in Chapter 16 and second by adding the series $\sum_{n\geq 0} F(n)x^n$ and $\sum_{n\geq 0} T(n)x^n$. Do the two methods produce the same result?

(6) Compute a recurrence satisfied by $F(n) + 2^n$ by adding the two relevant generating series.

(7) What is the generating series of $\sum_{i=0}^{n} s(i)$, where $s(i)$ is the ith Perrin number from Question (2)? What recurrence does this sequence satisfy?

(8) (a) What is the generating series of $F(n-1)_{n\geq 0}$ (where $F(-1) = 1$)?
   (b) What is the generating series of $\left(\sum_{i=0}^{n} F(n-1)\right)_{n\geq 0}$?

(9) Let $s(n)$ be the coefficient of x^n in the series $\frac{1-2x^3+3x^6-6x^9}{(1-x)^{10}}$. What is a simple formula for $s(n)$?

(10) (a) For several non-negative integers m, determine the generating series of $(n^m)_{n\geq 0}$. The numerators are called *Eulerian polynomials.* Plug $x = 1$ into these Eulerian polynomials.
   (b) In a permutation $a_1 a_2 \cdots a_m$, a *descent* is a position $i \in \{1, 2, \dots, m-1\}$ such that $a_i > a_{i+1}$. For example, the set of descents of 4132 is $\{1, 3\}$. For several non-negative integers m, compute the number of length-m permutations on $\{1, 2, \dots, m\}$ with exactly k descents for each k.

Experiments.

(11) Constant-recursive sequences satisfy a number of closure properties. Which of these closure properties also hold for the class of eventual polynomial sequences?

(12) Given polynomials $a(x), b(x) \in \mathbb{Q}[x]$, is it faster to compute the first n terms of the series $\frac{a(x)}{b(x)}$ by solving a system of equations or by using long division?

(13) We have identified 4 bases of $\mathsf{Poly}(r)$: the monomial basis, the Lagrange basis, the binomial coefficient basis, and the basis consisting of the first r shifts of $\binom{n}{r-1}_{n\geq 0}$. What are the 12 matrices that serve as change-of-basis matrices between all pairs of these 4 bases?

(14) We can define generating series of 2-dimensional sequences by considering series in multiple symbols. For example, define $s(n,m)$ by $\frac{1}{1-x-y} = \sum_{n\geq 0}\sum_{m\geq 0} s(n,m)x^n y^m$. What is $s(n,m)$?

Computation proofs.

(15) Define $a(n)_{n\geq 0}$, $b(n)_{n\geq 0}$, and $c(n)_{n\geq 0}$ by their generating series

$$\sum_{n\geq 0} a(n)x^n = \frac{1+53x+9x^2}{1-82x-82x^2+x^3}$$
$$\sum_{n\geq 0} b(n)x^n = \frac{2-26x-12x^2}{1-82x-82x^2+x^3}$$
$$\sum_{n\geq 0} c(n)x^n = \frac{2+8x-10x^2}{1-82x-82x^2+x^3}.$$

Show that $a(n)^3+b(n)^3 = c(n)^3+(-1)^n$ for all $n \geq 0$. This was discovered by Ramanujan[3] and has been discussed by Hirschhorn [**13, 14**].

Programs.

(16) Write a program that computes the generating series (as a rational expression) of a constant-recursive sequence, given a recurrence and initial conditions.

(17) Write a program that computes the recurrence and initial conditions of a constant-recursive sequence, given its generating series as a rational expression.

(18) Given two constant-recursive sequences specified by their recurrences and initial conditions, which method computes their sum (also specified by its recurrence and initial conditions) faster — linear algebra or symbolic algebra?

(19) Write a program that computes the generating series of a polynomial sequence, given the polynomial.

(20) Write a program that can convert the coordinates of a sequence in $\mathsf{Poly}(r)$ given in one of the four major bases to any of the other four major bases.

[3]Srinivasa Ramanujan was born in 1887 in Erode, India and died in 1920 in Kumbakonam, India.

CHAPTER 18

Counting with generating series

Words avoiding a pattern

In Chapter 3 we saw that the number of length-n words on the alphabet $\{0, 1\}$ that do not contain 00 is the Fibonacci number $F(n + 2)$. In this chapter we will use generating series to answer the following generalization of this question. Let Σ be a finite alphabet, and pick a word p on Σ. We say that a word w *avoids* the pattern p if w contains no occurrences of p as a factor. How many words in Σ^n avoid p?

First let's revisit binary words avoiding 00, using generating series.

Example 18.1. Let $s(n)$ be the number of words in $\{0, 1\}^n$ that avoid 00. We will obtain a recurrence for $s(n)_{n\geq 0}$ by computing the generating series $f(x) = \sum_{n\geq 0} s(n)x^n$. The idea is to build words avoiding 00 from smaller words by appending letters. For example, 11011 is built from 1101 by appending 1. The word 1101 itself is built from 110 by appending 1. Given a word avoiding 00, if its last letter is 1 then we can append 0 or 1 and still avoid 00. However, if its last letter is 0 then we cannot append 0; we can only append 1.

Since the last letter of a word determines the ways in which it can be extended, let's consider a generating series for each case. Let $S_0(n)$ be the set of words in $\{0, 1\}^n$ that avoid 00 and end with 0, and let $S_1(n)$ be the set of words in $\{0, 1\}^n$ that avoid 00 and end with 1. Now let $f_0(x)$ be the generating series of the sequence $|S_0(n)|_{n\geq 0}$, so that

$$f_0(x) = \sum_{n\geq 0} |S_0(n)|x^n = 0 + 1x + 1x^2 + 2x^3 + 3x^4 + \cdots$$

$$\begin{array}{cccc} 0 & 10 & 010 & 0110 \\ & & 110 & 1010 \\ & & & 1110 \end{array}$$

where we have listed the words counted by the first few coefficients. Similarly, let $f_0(x)$ be the generating series of the sequence $|S_1(n)|_{n\geq 0}$, so that

$$f_1(x) = \sum_{n\geq 0} |S_1(n)|x^n = 0 + 1x + 2x^2 + 3x^3 + 5x^4 + \cdots .$$

$$\begin{array}{cccc} 1 & 01 & 011 & 0101 \\ & 11 & 101 & 0111 \\ & & 111 & 1011 \\ & & & 1101 \\ & & & 1111 \end{array}$$

How can we obtain $f(x)$ from $f_0(x)$ and $f_1(x)$? Every 00-avoiding word either is empty, ends with 0, or ends with 1. Therefore $f(x) = 1 + f_0(x) + f_1(x)$, thinking coefficientwise.

This isn't an answer, because we don't yet have another way to write $f_0(x)$ and $f_1(x)$. To rewrite $f_0(x)$, we observe that every word in $S_0(n)$ is either the word 0 or built from a unique word $w1 \in S_1(n-1)$ by appending 0. In the language of series, this is

$$f_0(x) = x + f_1(x)x.$$

There is no $f_0(x)x$ on the right side, since we cannot obtain a word avoiding 00 by appending 0 to a word ending with 0. To rewrite $f_1(x)$, use that every word in $S_1(n)$ is either the word 1 or built by appending 1 to a word in $S_0(n-1)$ or $S_1(n-1)$. Therefore

$$f_1(x) = x + f_0(x)x + f_1(x)x.$$

We solve this system by writing it as the matrix equation

$$\begin{bmatrix} 1 & -x \\ -x & 1-x \end{bmatrix} \begin{bmatrix} f_0(x) \\ f_1(x) \end{bmatrix} = \begin{bmatrix} x \\ x \end{bmatrix}.$$

Now form the augmented matrix

$$\left[\begin{array}{cc|c} 1 & -x & x \\ -x & 1-x & x \end{array}\right].$$

The entries are rational expressions rather than rational numbers, but the same row-reduction algorithm works. First clear the lower left $-x$ entry by replacing the second row with $R_2 + xR_1$, where R_i is the ith row; then divide the second row by $1 - x - x^2$:

$$\left[\begin{array}{cc|c} 1 & -x & x \\ 0 & 1-x-x^2 & x^2+x \end{array}\right] \qquad \left[\begin{array}{cc|c} 1 & -x & x \\ 0 & 1 & \frac{x^2+x}{1-x-x^2} \end{array}\right].$$

The second row is now complete. To complete the first row, clear the $-x$ entry by replacing the first row with $R_1 + xR_2$:

$$\left[\begin{array}{cc|c} 1 & 0 & \frac{x}{1-x-x^2} \\ 0 & 1 & \frac{x^2+x}{1-x-x^2} \end{array}\right].$$

This final matrix says $f_0(x) = \frac{x}{1-x-x^2}$ and $f_1(x) = \frac{x^2+x}{1-x-x^2}$. Therefore $f(x) = 1 + f_0(x) + f_1(x) = \frac{x+1}{1-x-x^2}$, and it follows that $s(n+2) = s(n+1) + s(n)$ for all $n \geq 0$. In Example 17.10 we determined that $\sum_{n\geq 0} F(n)x^n = \frac{x}{1-x-x^2}$. Since $s(0) = 1$ and $s(1) = 2$, it follows that $s(n) = F(n+2)$, as in Theorem 3.1. We can also see this from $F(0) + F(1)x + x^2 f(x) = x + x^2 f(x) = \frac{x}{1-x-x^2}$.

In the previous example, we converted a combinatorics question into a linear algebra question. The same procedure works more generally. Suppose we want to count length-n words on an alphabet Σ that avoid p as a factor. The length-$(|p|-1)$ suffix of a word determines which letters can be appended to produce a word that does not end with p. Therefore we'll set up one generating series $f_v(x)$ for each possible suffix $v \in \Sigma^{|p|-1}$.

Example 18.2. How many length-n words on the alphabet $\{0, 1\}$ avoid the word 01? Since the length of 01 is 2, we keep track of length-1 suffixes. The two length-1 suffixes are 0 and 1, so let $f_0(x)$ be the generating series for words avoiding 01 that end with 0, and let $f_1(x)$ be the generating series for words avoiding 01 that end with 1. These series are related by the system

$$\begin{aligned} f_0(x) &= x + f_0(x)x + f_1(x)x \\ f_1(x) &= x + f_1(x)x. \end{aligned}$$

This is a different system than in Example 18.1, so we should probably expect a different recurrence. We convert the system to the augmented matrix

$$\left[\begin{array}{cc|c} 1-x & -x & x \\ 0 & 1-x & x \end{array}\right]$$

and row-reduce to obtain

$$\left[\begin{array}{cc|c} 1 & 0 & \frac{x}{1-2x+x^2} \\ 0 & 1 & \frac{x}{1-x} \end{array}\right].$$

Therefore $f_0(x) = \frac{x}{(1-x)^2}$ and $f_1(x) = \frac{x}{1-x}$, so the generating series for all words avoiding 01 is $f(x) = 1 + f_0(x) + f_1(x) = \frac{1}{(1-x)^2}$. By Theorem 17.28, $f(x) = \sum_{n\geq 0} \binom{n+1}{1} x^n$, so the number of length-n words avoiding 01 is $n + 1$. We can also see this directly, since every word avoiding 01 is of the form $1^k 0^{n-k}$ for some $k \in \{0, 1, \ldots, n\}$.

Example 18.3. Let's determine the number of length-n binary words avoiding 101 by tracking length-2 suffixes. There are four such suffixes, so our generating series are $f_{00}(x), f_{01}(x), f_{10}(x), f_{11}(x)$, where each $f_v(x)$ is the generating series for 101-avoiding words ending with v. These series satisfy the system

$$\begin{aligned} f_{00}(x) &= x^2 + f_{00}(x)x + f_{10}(x)x \\ f_{01}(x) &= x^2 + f_{00}(x)x \\ f_{10}(x) &= x^2 + f_{01}(x)x + f_{11}(x)x \\ f_{11}(x) &= x^2 + f_{01}(x)x + f_{11}(x)x. \end{aligned}$$

For example, every 101-avoiding word ending with 00 is built by appending 0 to a 101-avoiding word ending with 00 or 10. The equation for $f_{01}(x)$ does not involve $f_{10}(x)$ since appending 1 to a word ending with 10 produces the suffix 101. The corresponding augmented matrix is

$$\left[\begin{array}{cccc|c} 1-x & 0 & -x & 0 & x^2 \\ -x & 1 & 0 & 0 & x^2 \\ 0 & -x & 1 & -x & x^2 \\ 0 & -x & 0 & 1-x & x^2 \end{array}\right].$$

Since the generating series for all 101-avoiding words is $f(x) = 1 + 2x + f_{00}(x) + f_{01}(x) + f_{10}(x) + f_{11}(x)$, row-reducing allows us to compute that $f(x) = \frac{x^2+1}{1-2x+x^2-x^3}$. Therefore the number of length-n binary words avoiding 101 satisfies $s(n+3) = 2s(n+2) - s(n+1) + s(n)$. Since all words with length ≤ 2 avoid 101, the initial conditions are $s(0) = 1, s(1) = 2, s(2) = 4$.

In each of the previous examples, we constructed a system of equations in several generating series. We know that solutions to these systems exist, since each of the generating series counts certain combinatorial objects. But why are the solutions unique? In the proof of the following theorem, we show that the entries on the left of the augmented matrix comprise an invertible matrix. By inverting this matrix, we can solve the system and compute rational series.

Theorem 18.4. *Let Σ be a nonempty finite set, and let p be a nonempty word on Σ. Let $s(n)$ be the number of words in Σ^n that avoid p. Then $s(n)_{n\geq 0}$ is constant-recursive.*

PROOF. Let $\ell = |p| - 1$. If $\ell = 0$, then p is a letter, so $s(n) = (|\Sigma| - 1)^n$, and it follows that $s(n)_{n\geq 0}$ is constant-recursive. Therefore let us assume $\ell \geq 1$. For each $v = v_1 v_2 \cdots v_\ell \in \Sigma^\ell$, let $f_v(x)$ be the series whose nth coefficient is the number of words in Σ^n that avoid p and end with v. Each word that avoids p and ends with v is either v or is built in exactly one way by appending the letter v_ℓ to a word that avoids p and ends with $av_1 v_2 \cdots v_{\ell-1}$ for some letter a. Therefore

$$f_v(x) = x^\ell + x \sum_{\substack{a\in\Sigma \\ av\neq p}} f_{av_1v_2\cdots v_{\ell-1}}(x). \tag{18.1}$$

We exclude a from the sum if $av = p$ since in this case appending v_ℓ to a word ending with $av_1 v_2 \cdots v_{\ell-1}$ produces a word containing p. There are $|\Sigma|^\ell$ words $v \in \Sigma^\ell$, so we have an inhomogeneous system of $|\Sigma|^\ell$ linear equations in $|\Sigma|^\ell$ generating series $f_v(x)$. Let $m = |\Sigma|^\ell$, and let M be the $m \times m$ coefficient matrix of this system. We claim that the polynomial $\det M$ has a nonzero constant term and is therefore invertible as a series. Let $M(0)$ be the matrix obtained from M by replacing each x with 0. The nonzero entries in $M(0)$ are precisely the diagonal entries, since the only series in Equation (18.1) with a constant term in its coefficient is $f_v(x)$ itself. Moreover, each of these constant terms is 1. Therefore the constant term of $\det M$ is $\det M(0) = 1^m = 1$. By Proposition 17.4, $\det M$ is invertible. Consequently, M is invertible, and multiplying both sides of the appropriate matrix equation by M^{-1} produces a rational expression for each series $f_v(x)$. It follows that the series

$$\sum_{n\geq 0} s(n)x^n = 1 + |\Sigma|x + |\Sigma|^2x^2 + \cdots + |\Sigma|^{\ell-1}x^{\ell-1} + \sum_{v\in\Sigma^\ell} f_v(x)$$

is rational. By Theorem 17.14, $s(n)_{n\geq 0}$ is constant-recursive. □

Equivalent patterns

The previous section provides an algorithm for computing the generating series for the words on Σ avoiding p, so let's use it to generate some data. As we do this, keep two goals in mind. First, Theorem 18.4 establishes that $s(n)_{n\geq 0}$ is constant-recursive but does not bound its rank, since we did not keep track of the degrees of the numerator and denominator of the generating series. We would like to know the rank so that we know how complex $s(n)_{n\geq 0}$ is. Second, the system of equations in the proof of Theorem 18.4 is huge when $|p|$ is large. Specifically, it consists of $|\Sigma|^{|p|-1}$ equations. We would like to know if there is a faster way to compute the generating series of $s(n)_{n\geq 0}$ so that we can avoid solving this huge system of equations, if possible.

Suppose the alphabet is $\{0, 1\}$. There are four patterns with length 2. For each pattern, we use Theorem 18.4 to compute the corresponding generating series:

pattern p	generating series for words avoiding p
00	$\frac{x+1}{1-x-x^2}$
01	$\frac{1}{1-2x+x^2}$
10	$\frac{1}{1-2x+x^2}$
11	$\frac{x+1}{1-x-x^2}$

There is some duplication! The generating series for 00 and 11 are equal to each other, as are the generating series for 01 and 10. This is explained by the bijection on $\{0, 1\}^*$ that replaces $0 \mapsto 1$ and $1 \mapsto 0$. We say that two patterns p and q are *equivalent* if the number of words in Σ^n avoiding p is equal to the number of words in Σ^n avoiding q for all $n \geq 0$. Here are the equivalence classes for patterns with length 3 (along with a set $B(p)$ that will be defined shortly):

generating series	patterns	$B(p)$
$\frac{x^2+x+1}{1-x-x^2-x^3}$	000, 111	$\{1, 2, 3\}$
$\frac{x^2+1}{1-2x+x^2-x^3}$	010, 101	$\{1, 3\}$
$\frac{1}{1-2x+x^3}$	001, 011, 100, 110	$\{3\}$

The fact that 001 and 100 are equivalent is explained by the bijection that reverses words. However, among patterns with length 4 there are some intriguing equivalences:

generating series	patterns	$B(p)$
$\frac{x^3+x^2+x+1}{1-x-x^2-x^3-x^4}$	0000, 1111	$\{1, 2, 3, 4\}$
$\frac{x^2+1}{1-2x+x^2-2x^3+x^4}$	0101, 1010	$\{2, 4\}$
$\frac{x^3+1}{1-2x+x^3-x^4}$	0010, 0100, 0110, 1001, 1011, 1101	$\{1, 4\}$
$\frac{1}{1-2x+x^4}$	0001, 0011, 0111, 1000, 1100, 1110	$\{4\}$

Is there a combinatorial reason that 0010 and 0110 have the same generating series? In other words, why are the 0010-avoiding words in bijection with the 0110-avoiding words? Here is the data for length-5 patterns, which exhibit several more unexplained equivalences:

generating series	patterns	$B(p)$
$\frac{x^4+x^3+x^2+x+1}{1-x-x^2-x^3-x^4-x^5}$	00000, 11111	$\{1, 2, 3, 4, 5\}$
$\frac{x^4+x^2+1}{1-2x+x^2-2x^3+x^4-x^5}$	01010, 10101	$\{1, 3, 5\}$
$\frac{x^4+x^3+1}{1-2x+x^3-x^4-x^5}$	00100, 11011	$\{1, 2, 5\}$
$\frac{x^3+1}{1-2x+x^3-2x^4+x^5}$	01001, 01101, 10010, 10110	$\{2, 5\}$
$\frac{x^4+1}{1-2x+x^4-x^5}$	00010, 00110, 01000, 01100, 01110, ..., 11101	$\{1, 5\}$
$\frac{1}{1-2x+x^5}$	00001, 00011, 00101, 00111, 01011, ..., 11110	$\{5\}$

Several conjectures are suggested by this data. The numerators seem to have only 0 and 1 as coefficients, with constant term 1. For each length-ℓ pattern p, the monomial $x^{\ell-1}$ seems to appear in the numerator precisely when p begins and ends with the same letter. Further, $x^{\ell-2}$ seems to appear precisely when p begins and ends with the same word of length 2! This suggests the following definitions.

Notation. A *border* of a word p is a nonempty word that is both a prefix and suffix of p. Let $B(p)$ be the set consisting of the lengths of the borders of p.

For example, the borders of 00100 are 0, 00, and 00100. Therefore $B(00100) = \{1, 2, 5\}$. Each border of p that is not p corresponds to a way in which two occurrences of p can overlap. The pattern 00100 can overlap itself in 1 letter as 001$\underline{0}$0100 and in 2 letters as 001$\underline{00}$100.

We conjecture that the generating series counting words avoiding p has numerator $\sum_{i \in B(p)} x^{\ell-i}$. What about the denominator? Its constant term seems to be 1, and the linear term is usually $-2x$. If $B(p) = \{\ell\}$, as is the case for the patterns 0001 and 00001, the denominator appears to be $1 - 2x + x^{\ell}$. If

$B(p) = \{i, \ell\}$ for some $i \neq \ell$, as is the case for 00010 and 01001, the denominator appears to be $(1 - 2x) + \left(x^{\ell-i} - 2x^{\ell-i+1}\right) + x^\ell$. In general the denominator seems to be $\sum_{i \in B(p)} \left(x^{\ell-i} - 2x^{\ell-i+1}\right) + x^\ell$.

Indeed these conjectures are correct, and there is an explicit formula for the generating series in terms of the border lengths! First we establish the denominator. For larger alphabets, coefficient 2 becomes the size of the alphabet.

Theorem 18.5. *Let Σ be a nonempty finite set, and let p be a nonempty word on Σ. Let $k = |\Sigma|$, and let $s(n)$ be the number of words in Σ^n that avoid p. For all $n \geq 0$,*

$$s(n) = \sum_{i \in B(p)} \left(k\, s(n + i - 1) - s(n + i)\right).$$

In particular, $s(n)_{n \geq 0}$ is constant-recursive with rank $\leq |p|$.

This recurrence has a combinatorial interpretation. The product $k\, s(n+i-1)$ is the number of words obtained by taking all length-$(n + i - 1)$ words that avoid p and extending them by a single letter in all possible ways, resulting in words with length $n+i$. Since $s(n+i)$ is the number of length-$(n + i)$ words that avoid p, the difference $k\, s(n+i-1) - s(n+i)$ is the number of length-$(n + i)$ words w that end with p and contain no other occurrences of p. This will lead us to a bijective proof of Theorem 18.5.

Example 18.6. Let $\Sigma = \{0, 1\}$ and $p = 00100$. Let $S_p(n)$ be the set of words $w \in \Sigma^n$ such that w ends with p and contains no other occurrences of p. Fix $n = 5$. For each $i \in B(00100) = \{1, 2, 5\}$, here are the sets $S_p(5 + i)$:

i	$S_p(5+i)$	$\lvert S_p(5+i)\rvert$
1	$\{000100, 100100\}$	2
2	$\{0000100, 0100100, 1000100, 1100100\}$	4
5	$\{00000p, 00011p, 00101p, 00110p, 00111p, 01000p, 01010p,$ $01011p, 01100p, 01101p, 01110p, 01111p, 10000p,$ $10011p, 10100p, 10101p, 10110p, 10111p, 11000p,$ $11010p, 11011p, 11100p, 11101p, 11110p, 11111p\}$	25

The recurrence in Theorem 18.5 suggests that these 31 words correspond bijectively to the 31 words with length 5 that avoid p. Given a word in the previous table with length $5 + i$, we obtain a word with length 5 by deleting the last i letters. This produces a word avoiding p, since $i \geq 1$ and the suffix p is the only occurrence of p. We obtain the following 31 words.

i	set of words obtained from $S_p(5+i)$ by deleting the last i letters
1	$\{00010, 10010\}$
2	$\{00001, 01001, 10001, 11001\}$
5	$\{00000, 00011, 00101, 00110, 00111, 01000, 01010,$ $01011, 01100, 01101, 01110, 01111, 10000,$ $10011, 10100, 10101, 10110, 10111, 11000,$ $11010, 11011, 11100, 11101, 11110, 11111\}$

This example is silly in the sense that there is a simpler way to describe these 31 words — they are precisely the length-5 words that are not p. However, it illustrates the general correspondence. We will prove that the inverse function, defined as follows, is a bijection. Given a word $w \in \Sigma^5$ that avoids p, let $f(w)$ be the shortest prefix of wp that contains exactly one occurrence of p. The following table contains some examples, with the first occurrence of p in wp underlined.

w	wp	$f(w)$	$i = \lvert f(w)\rvert - \lvert w\rvert$
00010	$0\underline{0010 \cdot 0}0100$	000100	1
00001	$00\underline{001 \cdot 00}100$	0000100	2
00000	$00000 \cdot \underline{00100}$	0000000100	5

Since $f(w)$ contains only one occurrence of p, we have $f(w) \in S_p(5+i)$ for the corresponding value of i.

Proof of Theorem 18.5. For each $n \geq 0$, let $A_p(n)$ be the set of words in Σ^n that avoid p. Let $S_p(n)$ be the set of words $w \in \Sigma^n$ such that w ends with p and contains no other occurrences of p. For each $n \geq 0$, we establish a bijection

$$f\colon A_p(n) \to \bigcup_{i \in B(p)} S_p(n+i).$$

For each $w \in A_p(n)$, let $f(w)$ be the shortest prefix of wp that contains exactly one occurrence of p. In particular, $f(w) \in S_p(n+i)$ where $i = |f(w)| - n$. We show that $i \in B(p)$. Since w avoids p, we have $n+1 \leq |f(w)| \leq n + |p|$, so $1 \leq i \leq |p|$. Therefore the first and last occurrences of p in wp overlap in i letters, so $i \in B(p)$ and it follows that $f(w) \in \bigcup_{i \in B(p)} S_p(n+i)$.

To show that f is surjective, let $v \in \bigcup_{i \in B(p)} S_p(n+i)$. Then $v \in S_p(n+i)$ for the unique $i \in B(p)$ such that $|v| = n+i$. Let w be the length-n prefix of v. The only occurrence of p in v is at the end, and $i \geq 1$. This implies that $|w| < |v|$, so $w \in A_p(n)$. We claim $f(w) = v$. Since $i \in B(p)$, the length-$(n+i)$ prefix of wp ends with p and is therefore v. Moreover, this prefix is the shortest prefix of wp that contains exactly one occurrence of p, since v contains exactly one occurrence of p. Therefore $f(w) = v$.

To show that f is injective, let $w, v \in A_p(n)$ such that $f(w) = f(v)$. Then the length-n prefixes of $f(w)$ and $f(v)$ are the same, so $w = v$.

Since f is a bijection, we have

$$s(n) = |A_p(n)| = \sum_{i \in B(p)} |S_p(n+i)| = \sum_{i \in B(p)} \big(k\, s(n+i-1) - s(n+i)\big)$$

for all $n \geq 0$. □

This proof of Theorem 18.5 was directly suggested by the generating series data we computed. It also gives us a new interpretation of the Fibonacci recurrence. For binary words avoiding 00, the set of border lengths is $B(00) = \{1, 2\}$, so Theorem 18.5 implies

$$s(n) = \big(2s(n) - s(n+1)\big) + \big(2s(n+1) - s(n+2)\big).$$

In Chapter 3, and even in Example 18.1, we proved the equivalent recurrence $s(n+2) = s(n+1) + s(n)$ by putting length-$(n+2)$ words avoiding 00 in bijection with shorter words. For example, 0101110 ends with 0, so it is built by appending 10 to 01011. Surprisingly, the proof of Theorem 18.5 works the opposite way — by putting length-n words avoiding 00 in bijection with longer words. For example, $w = 0101110$ corresponds to $f(w) = 01011100$. The generality of Theorem 18.5 suggests that in fact the latter is more natural. Nonetheless, the two bijections are closely related. Namely, let w be a word avoiding 00 with length ≥ 2. The word $f(w)$ ends with 00 and contains only one occurrence of 00, so $f(w)$ ends with 100. Removing the suffix 100 gives the word corresponding to w under the former bijection.

Finally, we prove establish the rational expression for the generating series.

Corollary 18.7. *Let Σ be a nonempty finite set, and let p be a nonempty word on Σ. Let $k = |\Sigma|$, let $\ell = |p|$, let $b(x) = \sum_{i \in B(p)} x^{\ell - i}$, and let $s(n)$ be the number of words in Σ^n that avoid p. The generating series of $s(n)_{n \geq 0}$ is*

$$\sum_{n \geq 0} s(n) x^n = \frac{b(x)}{(1 - kx) b(x) + x^\ell}.$$

PROOF. Let $f(x) = \sum_{n \geq 0} s(n) x^n$ be the generating series of $s(n)_{n \geq 0}$. Theorem 18.5 and the proof of Theorem 17.12 imply that $f(x) = \frac{a(x)}{(1-kx)b(x)+x^\ell}$ for some $a(x) \in \mathbb{Q}[x]$ with $\deg a(x) \leq |p| - 1$. Since $\ell \in B(p)$, the denominator has constant term 1 and is therefore invertible. The numerator is

$$\begin{aligned} a(x) &= \left((1 - kx) b(x) + x^\ell\right) f(x) \\ &= \sum_{i \in B(p)} \left(x^{\ell - i} - k x^{\ell - i + 1}\right) f(x) + \text{higher-order terms} \end{aligned}$$

where all the omitted terms are divisible by x^ℓ. Using the initial conditions $s(n) = k^n$ for $0 \leq n \leq |p| - 1$, we have $f(x) = \sum_{n=0}^{\ell-1} k^n x^n + \sum_{n \geq \ell} s(n) x^n$. Therefore

$$a(x) = \sum_{i \in B(p)} \sum_{n=0}^{\ell - 1} \left(k^n x^{n + \ell - i} - k^{n+1} x^{n + \ell - i + 1}\right) + \text{higher-order terms.}$$

The inner sum telescopes, and we are left with

$$\begin{aligned} a(x) &= \sum_{i \in B(p)} \left(x^{\ell - i} - k^\ell x^{2\ell - i}\right) + \text{higher-order terms} \\ &= \sum_{i \in B(p)} x^{\ell - i} + \text{higher-order terms} \end{aligned}$$

since $k^\ell x^{2\ell - i}$ is divisible by x^ℓ when $i \leq \ell$. Since $\deg a(x) \leq |p| - 1$, we have $a(x) = \sum_{i \in B(p)} x^{\ell - i} = b(x)$. □

Historically, the polynomial $b(x) = \sum_{i \in B(p)} x^{\ell - i}$ seems to have first occurred (in a slightly different form) in a 1966 paper by Solovyov[1]. It appears in an expression for a rational series concerning the expected time at which a given sequence of events first occurs in a sequence of random trials [**28**, Equation (11)]. This answers questions such as the following. If you repeatedly flip a coin, how long must you wait on average to see 5 heads in a row?

Corollary 18.7 implies that if two words p, q in Σ^ℓ satisfy $B(p) = B(q)$ then they are equivalent — they are avoided by the same number of words in Σ^n for all $n \geq 0$. Since they are avoided by the same number of words, we naturally want a bijection between them. Is there a natural bijection? This seems to be a difficult question, and bijections are only known in certain cases.

[1] Александр Соловьёв (Aleksandr Solovyov) was born in 1927 in Moscow, Russia and died in 2001 in Moscow.

Can the rank of $s(n)_{n\geq 0}$ in Theorem 18.5 actually be less than $|p|$? For this to happen, there would need to be cancellation in the rational expression. However, Guibas and Odlyzko [**12**, Theorem 1.8] proved that if $k \geq 3$ and if $g(y)$ is a polynomial whose coefficients belong to $\{0, 1\}$ then $(y - k)g(y) + 1$ is irreducible. By substituting $y = \frac{1}{x}$ and multiplying by $x^{1+\deg g(y)}$, this implies that the rank of $s(n)_{n\geq 0}$ is in fact $|p|$ for alphabets with size ≥ 3. For a 2-letter alphabet, this remains an open question.

Generalizations

The method of computing a rational generating series by setting up and solving a system of equations is quite general. In addition to counting words avoiding a single pattern, it can also be used to count words avoiding multiple patterns simultaneously. Guibas and Odlyzko [**12**, Theorem 1] showed that the corresponding generating series is rational (but note that they write series in $\frac{1}{x}$ instead of x). Even more generally, generating series can be used to count words according to the number of occurrences of each pattern from a finite set. For every set $\{p_1, p_2, \dots, p_t\}$ of words, the cluster method of Goulden and Jackson [**10**] shows that the multivariate series in $x, y_1, y_2, \dots, y_t$, where the coefficient of $x^n y_1^{m_1} y_2^{m_2} \cdots y_t^{m_t}$ is the number of length-n words containing precisely m_i occurrences of p_i for each i, is also rational. For a friendly introduction, see the treatment by Noonan and Zeilberger [**16**].

Questions

Computations.

(1) For each word p, set up and solve the system of equations in the proof of Theorem 18.4 to compute the generating series for words on $\{0, 1, 2\}$ avoiding p, and check that the result agrees with Corollary 18.7.
   (a) 00
   (b) 01
   (c) 212
   (d) 111
   (e) 211
   (f) 012

(2) Let $\Sigma = \{0, 1\}$ and $p = 0101$. Write out the explicit correspondence between words in $\{0, 1\}^4$ that avoid p and $\bigcup_{i\in B(p)} S_p(4 + i)$ under the bijection f in the proof of Theorem 18.5.

(3) How many length-n words on $\{0, 1, 2\}$ avoid
   (a) both 01 and 10?
   (b) both 01 and 00?
   (c) both 01 and 22?

(4) How many words length-n words on $\{0, 1\}$ contain exactly m occurrences of 00?

Experiments.

(5) Let $p = 0001$ and $q = 0011$. Construct an explicit bijection between words in $\{0, 1\}^n$ avoiding p and words in $\{0, 1\}^n$ avoiding q that provides an alternate proof that p and q are equivalent.

Programs.

(6) Write a program that computes $B(p)$, given a word p.

(7) (a) Write a program that, given an alphabet Σ and a word p on Σ, sets up the system of equations in the proof of Theorem 18.4 and then solves the system to compute a rational expression for the generating series $\sum_{n\geq 0} s(n)x^n$, where $s(n)$ is the number of length-n words on Σ that avoid p.

(b) Use your program to check the tables of generating series in this chapter.

(c) How many equivalence classes of length-6 words on $\{0, 1\}$ are there?

(8) Write a program whose inputs are an alphabet Σ and a set $\{p_1, p_2, \ldots, p_t\}$ of words on Σ and whose output is a rational expression for the generating series for words on Σ that avoid all of the words $p_1, p_2, \ldots, p_t$.

CHAPTER 19

Exponential polynomials

In this chapter we will see that the generating series of $s(n)_{n\geq 0}$ provides an explicit formula for the nth term $s(n)$, and this determines the possible growth rates of constant-recursive sequences.

Example 19.1. Define $s(n)_{n\geq 0}$ by $s(0) = 2$, $s(1) = 5$, and $s(n+2) = 5s(n+1) - 6s(n)$ for all $n \geq 0$. Its generating series has denominator $1 - 5x + 6x^2$ and numerator

$$(1 - 5x + 6x^2)(2 + 5x + \cdots) = 2 - 5x.$$

Moreover, the denominator factors as $1-5x+6x^2 = (1-2x)(1-3x)$, so the partial fraction decomposition of the generating series is

$$\sum_{n\geq 0} s(n)x^n = \frac{2-5x}{1-5x+6x^2} = \frac{c_1}{1-2x} + \frac{c_2}{1-3x}$$

for some $c_1, c_2 \in \mathbb{Q}$. To determine c_1 and c_2, we can multiply by $(1-2x)(1-3x)$; this gives $2 - 5x = (1-3x)c_1 + (1-2x)c_2$. Comparing coefficients, we conclude $c_1 = 1$ and $c_2 = 1$. Proposition 17.7 gives the nth terms of the series $\frac{1}{1-2x}$ and $\frac{1}{1-3x}$, so we obtain the explicit formula $s(n) = 2^n + 3^n$. We can even check directly that $2^n + 3^n$ satisfies the recurrence:

$$\begin{aligned} s(n+2) - 5s(n+1) + 6s(n) &= (2^{n+2} + 3^{n+2}) - 5(2^{n+1} + 3^{n+1}) + 6(2^n + 3^n) \\ &= (4 - 10 + 6)2^n + (9 - 15 + 6)3^n \\ &= 0. \end{aligned}$$

In general, the denominator of a rational generating series may not factor into polynomials with rational coefficients, but we shouldn't let that stop us. If we can rewrite the generating series as a sum of series of the form $\frac{c}{1-ax}$ for some numbers a and c, then Proposition 17.7 gives an explicit formula for the nth term.

Example 19.2. From Example 17.10, the generating series of the Fibonacci sequence is

$$\sum_{n\geq 0} F(n)x^n = \frac{x}{1-x-x^2}.$$

We would like a factorization $1 - x - x^2 = (1-\alpha_1 x)(1-\alpha_2 x)$. If we perform the substitution $x = \frac{1}{y}$ and multiply by y^2, we obtain $y^2 - y - 1 = (y-\alpha_1)(y-\alpha_2)$. This shows that the constants α_1 and α_2 are the solutions of $y^2 - y - 1 = 0$. It will be convenient to introduce the notation $\phi := \frac{1+\sqrt{5}}{2}$ and $\bar{\phi} := \frac{1-\sqrt{5}}{2}$ for the two solutions. Therefore

$$y^2 - y - 1 = (y - \phi)(y - \bar{\phi}),$$

so the denominator factors as

$$1 - x - x^2 = (1 - \phi x)(1 - \bar{\phi} x).$$

This allows us to compute the partial fraction decomposition of $\frac{x}{1-x-x^2}$. For some numbers c_1 and c_2, we have

$$\sum_{n \geq 0} F(n) x^n = \frac{x}{1 - x - x^2} = \frac{c_1}{1 - \phi x} + \frac{c_2}{1 - \bar{\phi} x}.$$

Expanding $x = (1 - \bar{\phi} x) c_1 + (1 - \phi x) c_2$ gives $c_1 = \frac{1}{\sqrt{5}}$ and $c_2 = -\frac{1}{\sqrt{5}}$. Therefore

$$\begin{aligned}\sum_{n \geq 0} F(n) x^n &= \frac{1}{\sqrt{5}(1 - \phi x)} - \frac{1}{\sqrt{5}(1 - \bar{\phi} x)} \\ &= \sum_{n \geq 0} \left(\frac{1}{\sqrt{5}} \phi^n - \frac{1}{\sqrt{5}} \bar{\phi}^n \right) x^n.\end{aligned}$$

Comparing coefficients on the two sides of the equation, we obtain the formula

$$F(n) = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n \tag{19.1}$$

for the nth Fibonacci number, for all $n \geq 0$. This formula is known as *Binet's*[1] *formula*.

This approach is quite powerful. However, for the first time in this book, we have been forced outside of the rational numbers. Factoring $1 - x - x^2$ requires the irrational numbers $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$. Factoring other polynomials, such as $1 + x^2$, requires elements from an even larger set — the set $\mathbb{C}$ of complex numbers. Fortunately, the complex numbers are sufficient. The fundamental theorem of algebra states that every polynomial in $\mathbb{C}[x]$ can be factored into linear polynomials in $\mathbb{C}[x]$. In particular, every polynomial in $\mathbb{Q}[x]$ can be factored into linear polynomials in $\mathbb{C}[x]$, so this approach has a chance of working for all constant-recursive sequences.

As in Example 19.2, the numbers of interest aren't the roots of the denominator but of a related polynomial.

Definition 19.3. Let $s(n)_{n \geq 0}$ be a constant-recursive sequence with rank r and minimal recurrence

$$s(n + r) = c_{r-1} s(n + r - 1) + \cdots + c_1 s(n + 1) + c_0 s(n).$$

The *characteristic polynomial* of $s(n)_{n \geq 0}$ is

$$y^r - c_{r-1} y^{r-1} - \cdots - c_1 y - c_0.$$

For the Fibonacci sequence, the characteristic polynomial is $y^2 - y - 1$, whereas the generating series has denominator $1 - x - x^2$. If the generating series of $s(n)_{n \geq 0}$ is $\frac{a(x)}{b(x)}$, then the characteristic polynomial of $s(n)_{n \geq 0}$ has the same coefficients as $b(x)$ but in reversed order relative to the exponents. That is, the characteristic polynomial is $y^{\deg b(x)} b(\frac{1}{y})$.

When the roots $\alpha_1, \alpha_2, \ldots, \alpha_r$ of the characteristic polynomial are all distinct, then the partial fraction decomposition of $\frac{a(x)}{b(x)}$ is $\sum_{i=1}^{r} \frac{c_i}{1 - \alpha_i x}$ for some complex

[1] Jacques Binet (pronounced 'beenay') was born in 1786 in Rennes, France and died in 1856 in Paris, France.

numbers $c_1, c_2, \dots, c_r$, so we obtain the formula $s(n) = \sum_{i=1}^{r} c_i\, \alpha_i^n$ for the nth term of the coefficient sequence.

Example 19.4. Let $s(0) = 3$, $s(1) = -5$, $s(2) = 6$, and $s(n+3) = -2s(n+2) - s(n+1) - 2s(n)$ for all $n \geq 0$. The generating series of $s(n)_{n\geq 0}$ has denominator $1 + 2x + x^2 + 2x^3 = (1+2x)(1+x^2)$ and numerator

$$(1 + 2x + x^2 + 2x^3)(3 - 5x + 6x^2 + \cdots) = 3 + x - x^2.$$

The characteristic polynomial of $s(n)_{n\geq 0}$ is $y^3 + 2y^2 + y + 2 = (y+2)(y^2+1)$. Its roots are $-2, -i, i$. Therefore

$$\sum_{n\geq 0} s(n)x^n = \frac{3 + x - x^2}{1 + 2x + x^2 + 2x^3} = \frac{c_1}{1 - (-2x)} + \frac{c_2}{1 - (-ix)} + \frac{c_3}{1 - ix}$$

for some $c_1, c_2, c_3 \in \mathbb{C}$. These constants are $c_1 = \frac{9}{5}$, $c_2 = \frac{6-7i}{10}$, and $c_3 = \frac{6+7i}{10}$. Therefore

$$s(n) = \tfrac{9}{5} \cdot (-2)^n + \tfrac{6-7i}{10} \cdot (-i)^n + \tfrac{6+7i}{10} \cdot i^n$$

for all $n \geq 0$. Since $i^4 = 1$, the component $\frac{6-7i}{10} \cdot (-i)^n + \frac{6+7i}{10} \cdot i^n$ generates the periodic sequence $\frac{6}{5}, -\frac{7}{5}, -\frac{6}{5}, \frac{7}{5}, \frac{6}{5}, -\frac{7}{5}, -\frac{6}{5}, \frac{7}{5}, \dots$.

Characteristic polynomials with repeated roots

If the roots of the characteristic polynomial are not all distinct, then $s(n)$ is not simply a linear combination of exponential functions.

Example 19.5. Let $s(n)$ be the coefficient of x^n in the series $\frac{1}{(1-3x)^2}$. The characteristic polynomial $(y-3)^2$ has a root 3 with multiplicity 2, so the rational expression $\frac{1}{(1-3x)^2}$ is its own partial fraction decomposition. To determine a formula for $s(n)$, we use the fact that $\frac{1}{(1-x)^2}$ is the generating series of a polynomial sequence with rank 2. We can determine this sequence by computing the first few terms:

$$\left(\frac{1}{1-x}\right)^2 = \left(1 + x + x^2 + x^3 + \cdots\right)^2 = 1 + 2x + 3x^2 + 4x^3 + \cdots.$$

Therefore $\frac{1}{(1-x)^2} = \sum_{n\geq 0} (n+1)x^n$. It follows that $\frac{1}{(1-3x)^2} = \sum_{n\geq 0} 3^n(n+1)x^n$, so $s(n) = 3^n(n+1)$.

The formula for $s(n)$ in the previous example involves both an exponential function and a polynomial.

Definition 19.6. An *exponential polynomial* in the symbol x is an expression of the form $p_1(x)\alpha_1^x + \cdots + p_\ell(x)\alpha_\ell^x$, where $p_i(x) \in \mathbb{C}[x]$ and $\alpha_i \in \mathbb{C}$ for each $i \in \{1, 2, \dots, \ell\}$. A sequence $s(n)_{n\geq 0}$ of complex numbers is an *exponential polynomial sequence* if there exists an exponential polynomial $f(x)$ such that $s(n) = f(n)$ for all $n \geq 0$.

The next result generalizes Example 19.5 to larger exponents. Since the roots of a polynomial with rational coefficients can be complex numbers, we expand our definition of series from Definition 17.1 to allow complex coefficients.

Theorem 19.7. *If r is a positive integer and $\alpha \in \mathbb{C}$, then*

$$\frac{1}{(1-\alpha x)^r} = \sum_{n\geq 0} \alpha^n \binom{n+r-1}{r-1} x^n.$$

In particular, the sequence whose generating series is $\frac{1}{(1-\alpha x)^r}$ *is an exponential polynomial sequence.*

Proof. If $\alpha = 0$, then the equality becomes $1 = 1$, so assume $\alpha \neq 0$. Theorem 17.28 states that $\sum_{n\geq 0} \binom{n}{r-1} x^n = \frac{x^{r-1}}{(1-x)^r}$ if $r \geq 1$. Therefore

$$\frac{(\alpha x)^{r-1}}{(1-\alpha x)^r} = \sum_{n\geq 0} \binom{n}{r-1} \alpha^n x^n.$$

When $n < r-1$, the binomial coefficient $\binom{n}{r-1}$ is 0. This implies

$$\begin{aligned}\frac{1}{(1-\alpha x)^r} &= \sum_{n\geq r-1} \binom{n}{r-1} \alpha^{n-r+1} x^{n-r+1} \\ &= \sum_{m\geq 0} \binom{m+r-1}{r-1} \alpha^m x^m\end{aligned}$$

as desired. $\square$

Example 19.8. Let $\sum_{n\geq 0} s(n)x^n = \frac{3+2x+4x^2}{(1-3x)^3}$. The partial fraction decomposition of this series is of the form $\frac{c_1}{1-3x} + \frac{c_2}{(1-3x)^2} + \frac{c_3}{(1-3x)^3}$. Expanding

$$3 + 2x + 4x^2 = (1-3x)^2 c_1 + (1-3x)c_2 + c_3$$

and comparing coefficients lets us determine

$$\sum_{n\geq 0} s(n)x^n = \frac{3+2x+4x^2}{(1-3x)^3} = \frac{6}{9} \cdot \frac{1}{(1-3x)} - \frac{14}{9} \cdot \frac{1}{(1-3x)^2} + \frac{37}{9} \cdot \frac{1}{(1-3x)^3}.$$

Therefore, by Theorem 19.7,

$$\begin{aligned} s(n) &= 3^n \left(\frac{6}{9}\binom{n}{0} - \frac{14}{9}\binom{n+1}{1} + \frac{37}{9}\binom{n+2}{2} \right) \\ &= 3^n \left(\frac{6}{9} - \frac{14}{9}(n+1) + \frac{37}{18}(n+2)(n+1) \right). \end{aligned}$$

In particular, $s(n)_{n\geq 0}$ is an exponential polynomial sequence.

If we expand our definition of constant-recursive sequences from Definition 12.1 to include sequences whose terms belong to $\mathbb{C}$, then Theorem 19.7 implies that many constant-recursive sequences are exponential polynomial sequences. However, there is one limitation.

Example 19.9. Define $s(n)_{n\geq 0}$ by the recurrence $s(n+3) = s(n+2)$ and initial conditions $s(0) = 0$, $s(1) = 0$, and $s(2) = 1$. Then $s(n)_{n\geq 0}$ is the eventually constant sequence $0, 0, 1, 1, 1, \dots$, but this is not an exponential polynomial sequence. We see from the recurrence that $s(n+3)$ does not depend on $s(n)$. The characteristic polynomial is $y^3 - y^2 = y^2(y-1)$, and $\sum_{n\geq 0} s(n)x^n = \frac{x^2}{1-x} = -1 - x + \frac{1}{1-x}$.

Theorem 19.10. *Let* $s(n)_{n\geq 0}$ *be a sequence of complex numbers. Then* $s(n)_{n\geq 0}$ *is constant-recursive if and only if* $s(n)$ *is given by an exponential polynomial for sufficiently large* n.

PROOF. One direction follows from closure properties. Suppose there exists an exponential polynomial $f(x)$ and an integer N such that $s(n) = f(n)$ for all $n \geq N$. Theorem 15.3 implies that $f(n)_{n\geq 0}$ is constant-recursive. Since $s(n)_{n\geq 0}$ differs from $f(n)_{n\geq 0}$ in only finitely many values, Proposition 13.5 implies that $s(n)_{n\geq 0}$ is constant-recursive as well.

In the other direction, suppose $s(n)_{n\geq 0}$ is constant-recursive. By Theorem 17.12, $\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)}$ for some $a(x), b(x) \in \mathbb{C}[x]$ where the constant term of $b(x)$ is 1. Let $\alpha_1, \dots, \alpha_\ell$ be complex numbers and $m_1, \dots, m_\ell$ be positive integers such that $y^{\deg b(x)} b(\frac{1}{y}) = (y - \alpha_1)^{m_1} \cdots (y - \alpha_\ell)^{m_\ell}$. The partial fraction decomposition of $\frac{a(x)}{b(x)}$ is

$$\sum_{n\geq 0} s(n)x^n = \frac{a(x)}{b(x)} = p(x) + \sum_{i=1}^{\ell} \sum_{j=1}^{m_i} \frac{c_{i,j}}{(1 - \alpha_i x)^j}$$

for some polynomial $p(x) \in \mathbb{C}[x]$ and some $c_{i,j} \in \mathbb{C}$. Therefore, for all $n \geq 1 + \deg p(x)$,

$$s(n) = \sum_{i=1}^{\ell} \left(\alpha_i^n \sum_{j=1}^{m_i} c_{i,j} \binom{n+j-1}{j-1} \right).$$

This expression is an exponential polynomial in n, so $s(n)_{n\geq 0}$ is an eventual exponential polynomial sequence. $\square$

The previous proof shows that, if $a(x)$ and $b(x)$ have no common factor, then the degree of the polynomial $p_i(n)$ that appears with α_i^n is equal to the multiplicity m_i of the root α_i, since $c_{i,m_i} \neq 0$.

Example 19.11. Let $s(n) = (-1)^n \cdot (n + 1) + 2^n \cdot n^2$. The sequence $s(n)_{n\geq 0}$ is constant-recursive by Theorem 15.3. A recurrence for $s(n)_{n\geq 0}$ can be computed using closure properties, and from the recurrence we can construct its generating series. Alternatively, we can compute the generating series by writing n^2 in the basis $\left(\binom{n}{0}, \binom{n+1}{1}, \binom{n+2}{2}\right)$, namely $n^2 = \binom{n}{0} - 3\binom{n+1}{1} + 2\binom{n+2}{2}$. Therefore

$$\begin{aligned}
\sum_{n\geq 0} s(n)x^n &= \sum_{n\geq 0} \left((-1)^n \binom{n+1}{1} + 2^n \binom{n}{0} - 3 \cdot 2^n \binom{n+1}{1} + 2 \cdot 2^n \binom{n+2}{2} \right) x^n \\
&= \frac{1}{(1+x)^2} + \frac{1}{1-2x} - \frac{3}{(1-2x)^2} + \frac{2}{(1-2x)^3} \\
&= \frac{1 - 4x + 20x^2 + 2x^3 + 4x^4}{(1+x)^2(1-2x)^3}.
\end{aligned}$$

Asymptotics

Theorem 19.10 determines the possible asymptotic behavior of a constant-recursive sequence. For example, we can use Binet's formula (19.1) to understand how $F(n)$ grows as n gets large. Since $\frac{1-\sqrt{5}}{2} \approx -.618$, the power $\left(\frac{1-\sqrt{5}}{2}\right)^n$ approaches 0. Therefore $F(n) \approx \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n \approx .447 \cdot 1.618^n$. This shows that the terms of the Fibonacci sequence grow exponentially. We can make this more precise as

$$\frac{F(n)}{\left(\frac{1+\sqrt{5}}{2}\right)^n} = \frac{\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n}{\left(\frac{1+\sqrt{5}}{2}\right)^n} = \frac{1}{\sqrt{5}} - \frac{1}{\sqrt{5}}\frac{\left(\frac{1-\sqrt{5}}{2}\right)^n}{\left(\frac{1+\sqrt{5}}{2}\right)^n} \to \frac{1}{\sqrt{5}}$$

as n gets large. In particular, the ratio $\frac{F(n+1)}{F(n)}$ of consecutive terms approaches $\frac{1+\sqrt{5}}{2}$ as n gets large.

In general, the roots of the characteristic polynomial that have the largest absolute values determine the long-term behavior of a sequence. If $s(n) = 2^n + 3^n$, then

$$\frac{s(n+1)}{s(n)} = \frac{2^{n+1} + 3^{n+1}}{2^n + 3^n} = \frac{2 \cdot \left(\frac{2}{3}\right)^n + 3}{\left(\frac{2}{3}\right)^n + 1} \to 3$$

as $n \to \infty$.

If $s(n)_{n \geq 0}$ is constant-recursive, then $s(n)$ grows at most like $|\alpha|^n \cdot n^d$, for some non-negative integer d, where α is a root of the characteristic polynomial with maximum absolute value. This proves, for example, that $(n!)_{n \geq 0}$ is not constant-recursive since its terms grow super-exponentially.

What isn't known

Given a constant-recursive sequence $s(n)_{n \geq 0}$ and a number $a \in \mathbb{Q}$, can we determine whether there exists n such that $s(n) = a$? At first this sounds like an easy question. For example, if $s(n)_{n \geq 0}$ is the Fibonacci sequence, then $s(n)$ gets large eventually, so there are only finitely many n to check. However, even though the examples we have seen so far suggest that all constant-recursive sequences either grow without bound or are eventually periodic, this is not the case.

Example 19.12. We construct a constant-recursive sequence whose characteristic polynomial has five distinct roots — two pairs of complex roots and one real root. Further, we'll choose complex roots with absolute value 1 and a real root with absolute value < 1. Let $\alpha_1 = \frac{3+4i}{5}$, $\alpha_2 = \frac{3-4i}{5}$, $\alpha_3 = \frac{5+12i}{13}$, $\alpha_4 = \frac{5-12i}{13}$, and $\alpha_5 = \frac{1}{2}$. Then $|\alpha_1| = |\alpha_2| = |\alpha_3| = |\alpha_4| = 1 > |\alpha_5|$. From the polynomial

$$\prod_{i=1}^{5}(y - \alpha_i) = y^5 - \tfrac{321}{130}y^4 + \tfrac{254}{65}y^3 - \tfrac{223}{65}y^2 + \tfrac{129}{65}y - \tfrac{1}{2},$$

we build the recurrence

$$s(n+5) = \tfrac{321}{130}s(n+4) - \tfrac{254}{65}s(n+3) + \tfrac{223}{65}s(n+2) - \tfrac{129}{65}s(n+1) + \tfrac{1}{2}s(n).$$

Let $s(0) = s(1) = s(2) = s(3) = s(4) = 1$. The sequence $s(n)_{n \geq 0}$ is

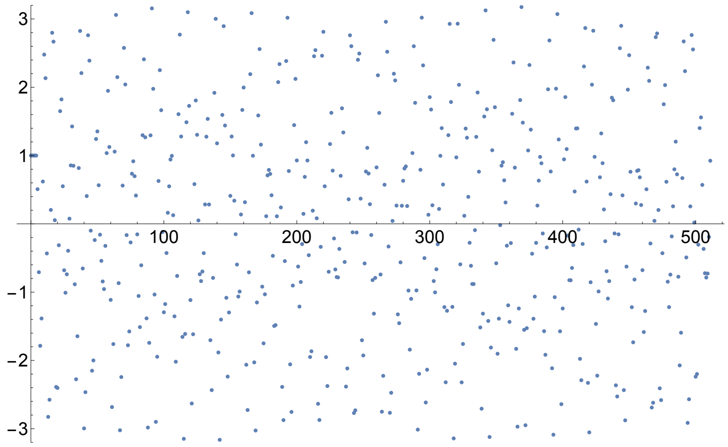
$$1, 1, 1, 1, 1, \tfrac{33}{65}, -\tfrac{2991}{4225}, -\tfrac{490423}{274625}, \ldots.$$

Since the 5 roots are distinct, we have $s(n) = \sum_{i=1}^{5} c_i \alpha_i^n$ for some complex numbers $c_1, \ldots, c_5$. The size of $s(n)$ is bounded, since by the triangle inequality we have

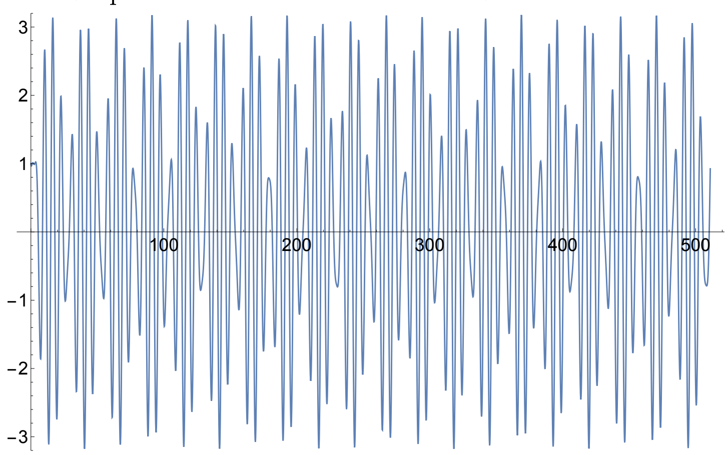
$$\begin{aligned}|s(n)| = \left|\sum_{i=1}^{5} c_i \alpha_i^n\right| \leq \sum_{i=1}^{5}|c_i| \cdot |\alpha_i|^n &= |c_1| + |c_2| + |c_3| + |c_4| + |c_5| \cdot \left(\tfrac{1}{2}\right)^n \\ &\leq |c_1| + |c_2| + |c_3| + |c_4| + |c_5| \approx 4.931.\end{aligned}$$

However, $s(n)_{n \geq 0}$ is not eventually periodic, since the angles $\theta_1 = \arg \alpha_1$ and $\theta_3 = \arg \alpha_3$ (for which $\alpha_1 = e^{i\theta_1}$ and $\alpha_3 = e^{i\theta_3}$) are not rational multiples of π

and since $\left(\frac{1}{2}\right)^n$ approaches 0 as n gets large. Here is a plot of the first 2^9 terms of $s(n)_{n\geq 0}$:



By interpreting the exponential polynomial as a function of a real number, we can see that $s(n)$ consists of values on a curve that is a small perturbation (by $c_5 \cdot \left(\frac{1}{2}\right)^n$) of a sum of two sine waves whose amplitudes are different and whose period lengths are irrational multiples of each other. Here is the curve:



The *Skolem*[2] *problem* is the problem of deciding whether a given constant-recursive sequence contains a term that is 0. It is an open problem to determine whether an algorithm exists for solving this problem. The previous example indicates why the Skolem problem is difficult. Define $t(0) = 0$ and $t(x) = 1$ for all rational numbers $x \neq 0$. In 1933, Skolem proved that if $s(n)_{n\geq 0}$ is a constant-recursive sequence of rational numbers then $t(s(n))_{n\geq 0}$ is eventually periodic. In other words, the zeros of a constant-recursive sequence exhibit eventually periodic behavior. However, the proof does not provide a way to determine whether any zeros occur.

Example 19.13. Let $s(n)_{n\geq 0}$ be the sequence in Example 19.12. The sequence $\left(s(n) - s(2^{100})\right)_{n\geq 0}$ is also constant-recursive. Moreover, this sequence does contain a 0, but how would we detect that, given its recurrence and initial conditions?

[2]Thoralf Skolem was born in 1887 in Sandsvær, Norway and died in 1963 in Oslo, Norway.

Matrix powers

At the end of Chapter 12, we used matrix powers to compute the nth term of a constant-recursive sequence. Specifically, if M is the companion matrix of a size-r recurrence satisfied by $s(n)_{n\geq 0}$, then $s(n)$ is the first entry in the vector

$$M^n \begin{bmatrix} s(0) \\ s(1) \\ \vdots \\ s(r-1) \end{bmatrix}.$$

What if M is not the companion matrix of a recurrence but instead another $r \times r$ matrix?

Example 19.14. Let

$$M = \begin{bmatrix} 3 & 2 & 0 \\ 4 & 8 & 4 \\ 6 & 0 & 8 \end{bmatrix},$$

and let $s(n)$ be the entry in the 1×1 matrix

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} M^n \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix};$$

in other words, $s(n)$ is the top left entry of M^n. The sequence $s(n)_{n\geq 0}$ is

$$1, 3, 17, 187, 2313, 27491, 314689, 3527115, \ldots.$$

To obtain a formula for $s(n)$, we compute a Jordan[3] decomposition of M, since matrices in Jordan form are block diagonal matrices and their nth power has a simple form. The characteristic polynomial of M is $\det(yI - M) = y^3 - 19y^2 - 104y + 176 = (y-4)^2(y-11)$ with roots $4, 4, 11$. Therefore either

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 11 \end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 11 \end{bmatrix}$$

is a Jordan form of M; the first matrix consists of three Jordan blocks (where the two 4s are in different blocks), while the second consists of only two. The eigenvalue 11, which has algebraic multiplicity 1, has eigenvector

$$\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

and therefore has geometric multiplicity 1. The eigenvalue 4, which has algebraic multiplicity 2, has only one eigenvector,

$$\begin{bmatrix} -2 \\ -1 \\ 3 \end{bmatrix},$$

[3]Camille Jordan was born in 1838 in Lyon, France and died in 1922 in Paris, France.

and therefore has geometric multiplicity 1. To obtain a third basis vector, we solve

$$(M-4I)\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=\begin{bmatrix}-2\\-1\\3\end{bmatrix},$$

obtaining

$$\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=\begin{bmatrix}0\\-1\\\frac{3}{4}\end{bmatrix}+v_1\begin{bmatrix}1\\\frac{1}{2}\\-\frac{3}{2}\end{bmatrix}.$$

We may choose any value for v_1; choosing $v_1=0$ gives the change-of-basis matrix

$$S=\begin{bmatrix}-2&0&1\\-1&-1&4\\3&\frac{3}{4}&2\end{bmatrix}\qquad\text{with inverse}\qquad S^{-1}=\begin{bmatrix}-\frac{20}{49}&\frac{3}{49}&\frac{4}{49}\\\frac{8}{7}&-\frac{4}{7}&\frac{4}{7}\\\frac{9}{49}&\frac{6}{49}&\frac{8}{49}\end{bmatrix}.$$

A Jordan form of M is therefore

$$J=S^{-1}MS=\begin{bmatrix}4&1&0\\0&4&0\\0&0&11\end{bmatrix}$$

and not the diagonal matrix with diagonal entries $4,4,11$. Crucially, with the Jordan form, we can now write $M=SJS^{-1}$, so that $M^n=SJ^nS^{-1}$. Powers of J have an explicit formula with exponential polynomial entries:

$$J^n=\begin{bmatrix}4^n&4^{n-1}n&0\\0&4^n&0\\0&0&11^n\end{bmatrix}.$$

Therefore every entry in M^n is an exponential polynomial in n. In particular, $s(n)$ is given by an exponential polynomial, which we compute to be

$$s(n)=\tfrac{4}{49}\cdot 4^n(-7n+10)+\tfrac{9}{49}\cdot 11^n.$$

Theorem 19.15. *A sequence $s(n)_{n\geq 0}$ of rational numbers is constant-recursive if and only if there exists a positive integer r, an $r\times r$ matrix M with rational entries, and an $r\times 1$ vector with rational entries such that $s(n)$ is the first entry in the vector $M^n v$ for all $n\geq 0$.*

Proof. First assume that $s(n)_{n\geq 0}$ is constant-recursive. As in Chapter 12, if $s(n)_{n\geq 0}$ satisfies the recurrence

$$s(n+r)=c_0s(n)+c_1s(n+1)+\cdots+c_{r-1}s(n+r-1),$$

then the companion matrix

$$M=\begin{bmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\c_0&c_1&c_2&\cdots&c_{r-1}\end{bmatrix}.$$

and the vector

$$v = \begin{bmatrix} s(0) \\ s(1) \\ s(2) \\ \vdots \\ s(r-1) \end{bmatrix}$$

have the property that $s(n) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} M^n v$.

In the other direction, let M be an $r \times r$ matrix with rational entries, and let v be an $r \times 1$ vector with rational entries. Let J be a Jordan form of M, and let S be a change-of-basis matrix such that $M = SJS^{-1}$. Since J is a block diagonal matrix, the entries of J^n are exponential polynomials in n. Therefore the entries of $M^n = SJ^nS^{-1}$ are exponential polynomials in n. Let

$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Let $s(n) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} M^n v$. Since the entries of M are rational, $s(n)$ is rational for each $n \geq 0$. By Theorem 19.10, $s(n)_{n \geq 0}$ is constant-recursive. $\square$

Characterizations of constant-recursive sequences

Here we summarize the 6 equivalent characterizations of constant-recursive sequences of rational numbers.

- Recurrence: Definition 12.1
- Difference equation: Theorem 12.12
- Vector space: Theorem 14.15
- Rational generating series: Theorem 17.15
- Eventual exponential polynomial sequence of rational numbers: Theorem 19.10
- Matrix power: Theorem 19.15

Questions

Computations.

(1) Use Binet's formula to compute the first several terms of the Fibonacci sequence.

(2) Verify that the exponential polynomial for $F(n+2)$ given by Binet's formula is equal to the sum of the exponential polynomials for $F(n+1)$ and $F(n)$.

(3) Let $L(n)$ be the nth Lucas number. Use the generating series of $L(n)_{n \geq 0}$ to find an exponential polynomial formula for $L(n)$ analogous to Binet's formula for the Fibonacci sequence.

(4) Let $s(0) = 1$, $s(1) = 3$, and $s(n+2) = 2s(n+1) + s(n)$ for all $n \geq 0$. What is the minimal non-negative integer N such that $s(n)$ is given by an exponential polynomial for all $n \geq N$? What is the exponential polynomial?

(5) Consider the matrix

$$M = \begin{bmatrix} 1 & 3 \\ -2 & 0 \end{bmatrix},$$

and let $s(n)$ be the top left entry of M^n. What is the exponential polynomial for $s(n)$?

Bibliography

[1] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences: Theory, Applications, Generalizations* (2003), Cambridge University Press, Cambridge.

[2] Jean-Paul Allouche and Jeffrey Shallit, The ring of k-regular sequences, *Theoretical Computer Science* **98** (1992) 163–197.

[3] Kenneth Appel and Wolfgang Haken, Every planar map is four colorable I: discharging, *Illinois Journal of Mathematics* **21** (1977) 429–490.

[4] Kenneth Appel, Wolfgang Haken, and John Koch, Every planar map is four colorable II: reducibility, *Illinois Journal of Mathematics* **21** (1977) 491–567.

[5] E. Barcucci, A. Del Lungo, S. Fezzi, and R. Pinzani, Nondecreasing Dyck paths and q-Fibonacci numbers, *Discrete Mathematics* **170** (1997) 211–217.

[6] Émilie Charlier, Narad Rampersad, and Jeffrey Shallit, Enumeration and decidable properties of automatic sequences, *International Journal of Foundations of Computer Science* **23** (2012) 1035–1066.

[7] Gilles Christol, Teturo Kamae, Michel Mendès France, and Gérard Rauzy, Suites algébriques, automates et substitutions, *Bulletin de la Société Mathématique de France* **108** (1980) 401–419.

[8] Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh, Catalan and Motzkin numbers modulo 4 and 8, *European Journal of Combinatorics* **29** (2008) 1449–1466.

[9] Nathan Fine, Binomial coefficients modulo a prime, *The American Mathematical Monthly* **54** (1947) 589–592.

[10] Ian Goulden and David Jackson, An inversion theorem for cluster decompositions of sequences with distinguished subsequences, *Journal of the London Mathematical Society* (2) **20** (1979) 567–576.

[11] Ronald Graham, Donald Knuth, and Oren Patashnik, *Concrete Mathematics* Second Edition (1989), Addison-Wesley Publishing Company.

[12] Leo Guibas and Andrew Odlyzko, String overlaps, pattern matching, and nontransitive games, *Journal of Combinatorial Theory, Series A* **30** (1981) 183–208.

[13] Michael D. Hirschhorn, An amazing identity of Ramanujan, *Mathematics Magazine* **68** (1995) 199–201.

[14] Michael D. Hirschhorn, A proof in the spirit of Zeilberger of an amazing identity of Ramanujan, *Mathematics Magazine* **69** (1996) 267–269.

[15] Manuel Kauers and Peter Paule, *The Concrete Tetrahedron* (2011), Spring-Verlag, Vienna.

[16] John Noonan and Doron Zeilberger, The Goulden–Jackson cluster method: extensions, applications, and implementations, *Journal of Difference Equations and Applications* **5** (1999) 355–377.

[17] Aline Parreau, Michel Rigo, Eric Rowland, and Élise Vandomme, A new approach to the 2-regularity of the ℓ-abelian complexity of 2-automatic sequences, *The Electronic Journal of Combinatorics* **22** (2015) #P1.27.

[18] Bjorn Poonen and Michael Rubinstein, The number of intersection points made by the diagonals of a regular polygon, *SIAM Journal on Discrete Mathematics* **11** (1998) 135–156.

[19] Lara Pudwell and Eric Rowland, Avoiding fractional powers over the natural numbers, *The Electronic Journal of Combinatorics* **25** (2018) #P2.27.

[20] Eric Rowland, IntegerSequences, `https://github.com/ericrowland/IntegerSequences`.

[21] Eric Rowland, IntegerSequences: a package for computing with k-regular sequences, International Congress on Mathematical Software, *Lecture Notes in Computer Science* **10931** (2018) 414–421.

[22] Eric Rowland, A matrix generalization of a theorem of Fine, *Integers* **18A** (2018) A18.

[23] Eric Rowland and Jeffrey Shallit, Avoiding 3/2-powers over the natural numbers, *Discrete Mathematics* **312** (2012) 1282–1288.

[24] Eric Rowland and Reem Yassawi, Automatic congruences for diagonals of rational functions, *Journal de Théorie des Nombres de Bordeaux* **27** (2015) 245–288.

[25] Eric Rowland and Doron Zeilberger, A case study in meta-automation: automatic generation of congruence automata for combinatorial sequences, *Journal of Difference Equations and Applications* **20** (2014) 973–988.

[26] Jeffrey Shallit, Remarks on inferring integer sequences, `https://cs.uwaterloo.ca/~shallit/Talks/infer.ps`.

[27] Neil Sloane et al., The On-Line Encyclopedia of Integer Sequences, `https://oeis.org`.

[28] A. D. Solov'ev, A combinatorial identity and its application to the problem concerning the first occurrence of a rare event (translated from Russian by N. A. Brunswick), *Theory of Probability and Its Applications* **11** (1966) 276–282.

[29] Lukas Spiegelhofer and Michael Wallner, An explicit generating function arising in counting binomial coefficients divisible by powers of primes, *Acta Arithmetica* **181** (2017) 27–55.

[30] Richard Stanley, *Catalan Numbers* (2015), Cambridge University Press, New York.

[31] Hassler Whitney, A logical expansion in mathematics, *Bulletin of the American Mathematical Society* **38** (1932) 572–579.

[32] Herbert Wilf, *Generatingfunctionology* Third Edition (2006), A. K. Peters.

Index