12 Generating Functions

Generating Functions are one of the most surprising and useful inventions in Discrete Math. Roughly speaking, generating functions transform problems about *sequences* into problems about *functions*. This is great because we've got piles of mathematical machinery for manipulating functions. Thanks to generating functions, we can then apply all that machinery to problems about sequences. In this way, we can use generating functions to solve all sorts of counting problems. They can also be used to find closed-form expressions for sums and to solve recurrences. In fact, many of the problems we addressed in Chapters 9–11 can be formulated and solved using generating functions.

12.1 Definitions and Examples

The *ordinary generating function* for the sequence[1] $\langle g_0, g_1, g_2, g_3 \dots \rangle$ is the power series:

$$G(x) = g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots .$$

There are a few other kinds of generating functions in common use, but ordinary generating functions are enough to illustrate the power of the idea, so we'll stick to them and from now on, *generating function* will mean the ordinary kind.

A generating function is a "formal" power series in the sense that we usually regard x as a placeholder rather than a number. Only in rare cases will we actually evaluate a generating function by letting x take a real number value, so we generally ignore the issue of convergence.

Throughout this chapter, we'll indicate the correspondence between a sequence and its generating function with a double-sided arrow as follows:

$$\langle g_0, g_1, g_2, g_3, \dots \rangle \quad \longleftrightarrow \quad g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots .$$

For example, here are some sequences and their generating functions:

$$\begin{aligned}
\langle 0, 0, 0, 0, \dots \rangle \quad &\longleftrightarrow \quad 0 + 0x + 0x^2 + 0x^3 + \cdots = 0 \\
\langle 1, 0, 0, 0, \dots \rangle \quad &\longleftrightarrow \quad 1 + 0x + 0x^2 + 0x^3 + \cdots = 1 \\
\langle 3, 2, 1, 0, \dots \rangle \quad &\longleftrightarrow \quad 3 + 2x + 1x^2 + 0x^3 + \cdots = 3 + 2x + x^2
\end{aligned}$$

[1] In this chapter, we'll put sequences in angle brackets to more clearly distinguish them from the many other mathematical expressions floating around.

The pattern here is simple: the ith term in the sequence (indexing from 0) is the coefficient of x^i in the generating function.

Recall that the sum of an infinite geometric series is:

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

This equation does not hold when $|z| \geq 1$, but as remarked, we won't worry about convergence issues for now. This formula gives closed form generating functions for a whole range of sequences. For example:

$$\begin{aligned}
\langle 1, 1, 1, 1, \ldots \rangle \quad &\longleftrightarrow 1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x} \\
\langle 1, -1, 1, -1, \ldots \rangle \quad &\longleftrightarrow 1 - x + x^2 - x^3 + x^4 - \cdots = \frac{1}{1+x} \\
\left\langle 1, a, a^2, a^3, \ldots \right\rangle \quad &\longleftrightarrow 1 + ax + a^2x^2 + a^3x^3 + \cdots = \frac{1}{1-ax} \\
\langle 1, 0, 1, 0, 1, 0, \ldots \rangle \quad &\longleftrightarrow 1 + x^2 + x^4 + x^6 + x^8 + \cdots = \frac{1}{1-x^2}
\end{aligned}$$

12.2 Operations on Generating Functions

The magic of generating functions is that we can carry out all sorts of manipulations on sequences by performing mathematical operations on their associated generating functions. Let's experiment with various operations and characterize their effects in terms of sequences.

12.2.1 Scaling

Multiplying a generating function by a constant scales every term in the associated sequence by the same constant. For example, we noted above that:

$$\langle 1, 0, 1, 0, 1, 0, \ldots \rangle \quad \longleftrightarrow 1 + x^2 + x^4 + x^6 + \cdots = \frac{1}{1-x^2}.$$

Multiplying the generating function by 2 gives

$$\frac{2}{1-x^2} = 2 + 2x^2 + 2x^4 + 2x^6 + \cdots$$

which generates the sequence:

$$\langle 2, 0, 2, 0, 2, 0, \ldots \rangle .$$

Rule 12.2.1 (Scaling Rule)**.** *If*

$$\langle f_0, f_1, f_2, \ldots \rangle \quad \longleftrightarrow F(x),$$

then

$$\langle cf_0,\ cf_1,\ cf_2,\ \ldots \rangle \quad \longleftrightarrow c \cdot F(x).$$

The idea behind this rule is that:

$$\begin{aligned} \langle cf_0, cf_1, cf_2, \ldots \rangle \quad &\longleftrightarrow cf_0 + cf_1 x + cf_2 x^2 + \cdots \\ &= c \cdot (f_0 + f_1 x + f_2 x^2 + \cdots) \\ &= cF(x). \end{aligned}$$

12.2.2 Addition

Adding generating functions corresponds to adding the two sequences term by term. For example, adding two of our earlier examples gives:

$$\begin{array}{rl} \langle 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ \ldots \rangle & \longleftrightarrow \dfrac{1}{1-x} \\ +\ \langle 1, -1, 1, -1, 1, -1, \ldots \rangle & \longleftrightarrow \dfrac{1}{1+x} \\ \hline \langle 2,\ 0,\ 2,\ 0,\ 2,\ 0,\ \ldots \rangle & \longleftrightarrow \dfrac{1}{1-x} + \dfrac{1}{1+x} \end{array}$$

We've now derived two different expressions that both generate the sequence $\langle 2, 0, 2, 0, \ldots \rangle$. They are, of course, equal:

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

Rule 12.2.2 (Addition Rule)**.** *If*

$$\begin{aligned} \langle f_0, f_1, f_2, \ldots \rangle &\longleftrightarrow F(x) \qquad \textit{and} \\ \langle g_0, g_1, g_2, \ldots \rangle &\longleftrightarrow G(x), \end{aligned}$$

then

$$\langle f_0 + g_0,\ f_1 + g_1,\ f_2 + g_2,\ \ldots \rangle \longleftrightarrow F(x) + G(x).$$

The idea behind this rule is that:

$$\begin{aligned}\langle f_0 + g_0,\ f_1 + g_1,\ f_2 + g_2,\ \ldots\rangle \quad &\longleftrightarrow \sum_{n=0}^{\infty} (f_n + g_n)x^n \\ &= \left(\sum_{n=0}^{\infty} f_n x^n\right) + \left(\sum_{n=0}^{\infty} g_n x^n\right) \\ &= F(x) + G(x).\end{aligned}$$

12.2.3 Right Shifting

Let's start over again with a simple sequence and its generating function:

$$\langle 1, 1, 1, 1, \ldots\rangle \quad \longleftrightarrow \frac{1}{1-x}.$$

Now let's *right-shift* the sequence by adding k leading zeros:

$$\begin{aligned}\langle \overbrace{0, 0, \ldots, 0}^{k \text{ zeroes}}, 1, 1, 1, \ldots\rangle \quad &\longleftrightarrow x^k + x^{k+1} + x^{k+2} + x^{k+3} + \cdots \\ &= x^k \cdot (1 + x + x^2 + x^3 + \cdots) \\ &= \frac{x^k}{1-x}.\end{aligned}$$

Evidently, adding k leading zeros to the sequence corresponds to multiplying the generating function by x^k. This holds true in general.

Rule 12.2.3 (Right-Shift Rule)**.** *If* $\langle f_0, f_1, f_2, \ldots\rangle \longleftrightarrow F(x)$*, then:*

$$\langle \overbrace{0, 0, \ldots, 0}^{k \text{ zeroes}}, f_0, f_1, f_2, \ldots\rangle \quad \longleftrightarrow x^k \cdot F(x).$$

The idea behind this rule is that:

$$\begin{aligned}\langle \overbrace{0, 0, \ldots, 0}^{k \text{ zeroes}}, f_0, f_1, f_2, \ldots\rangle \quad &\longleftrightarrow f_0 x^k + f_1 x^{k+1} + f_2 x^{k+2} + \cdots \\ &= x^k \cdot (f_0 + f_1 x + f_2 x^2 + f_3 x^3 + \cdots) \\ &= x^k \cdot F(x).\end{aligned}$$

12.2.4 Differentiation

What happens if we take the *derivative* of a generating function? As an example, let's differentiate the now-familiar generating function for an infinite sequence of 1's:

$$
\begin{aligned}
& & 1 + x + x^2 + x^3 + x^4 + \cdots &= \frac{1}{1-x} \\
&\text{IMPLIES} & \frac{d}{dx}\,(1 + x + x^2 + x^3 + x^4 + \cdots) &= \frac{d}{dx}\left(\frac{1}{1-x}\right) \\
&\text{IMPLIES} & 1 + 2x + 3x^2 + 4x^3 + \cdots &= \frac{1}{(1-x)^2} \\
&\text{IMPLIES} & \langle 1, 2, 3, 4, \ldots \rangle &\longleftrightarrow \frac{1}{(1-x)^2}. \qquad (12.1)
\end{aligned}
$$

We found a generating function for the sequence $\langle 1, 2, 3, 4, \ldots \rangle$ of positive integers!

In general, differentiating a generating function has two effects on the corresponding sequence: each term is multiplied by its index and the entire sequence is shifted left one place.

Rule 12.2.4 (Derivative Rule)**.** *If*

$$\langle f_0, f_1, f_2, f_3, \ldots \rangle \quad \longleftrightarrow F(x),$$

then

$$\langle f_1, 2f_2, 3f_3, \ldots \rangle \quad \longleftrightarrow F'(x).$$

The idea behind this rule is that:

$$
\begin{aligned}
\langle f_1, 2f_2, 3f_3, \ldots \rangle \quad &\longleftrightarrow f_1 + 2f_2x + 3f_3x^2 + \cdots \\
&= \frac{d}{dx}\,(f_0 + f_1x + f_2x^2 + f_3x^3 + \cdots) \\
&= \frac{d}{dx}\,F(x).
\end{aligned}
$$

The Derivative Rule is very useful. In fact, there is frequent, independent need for each of differentiation's two effects, multiplying terms by their index and left-shifting one place. Typically, we want just one effect and must somehow cancel out the other. For example, let's try to find the generating function for the sequence of squares, $\langle 0, 1, 4, 9, 16, \ldots \rangle$. If we could start with the sequence $\langle 1, 1, 1, 1, \ldots \rangle$ and multiply each term by its index two times, then we'd have the desired result:

$$\langle 0 \cdot 0,\ 1 \cdot 1,\ 2 \cdot 2,\ 3 \cdot 3,\ \ldots \rangle = \langle 0, 1, 4, 9, \ldots \rangle\,.$$

A challenge is that differentiation not only multiplies each term by its index, but also shifts the whole sequence left one place. However, the Right-Shift Rule 12.2.3 tells how to cancel out this unwanted left-shift: multiply the generating function by x.

Our procedure, therefore, is to begin with the generating function for $\langle 1, 1, 1, 1, \ldots \rangle$, differentiate, multiply by x, and then differentiate and multiply by x once more. Then

$$
\begin{array}{rlll}
 & \langle 1, 1, 1, 1, \ldots \rangle & \longleftrightarrow \dfrac{1}{1-x} & \\
\text{Derivative Rule:} & \langle 1, 2, 3, 4, \ldots \rangle & \longleftrightarrow \dfrac{d}{dx}\dfrac{1}{1-x} & = \dfrac{1}{(1-x)^2} \\
\text{Right-shift Rule:} & \langle 0, 1, 2, 3, \ldots \rangle & \longleftrightarrow x \cdot \dfrac{1}{(1-x)^2} & = \dfrac{x}{(1-x)^2} \\
\text{Derivative Rule:} & \langle 1, 4, 9, 16, \ldots \rangle & \longleftrightarrow \dfrac{d}{dx}\dfrac{x}{(1-x)^2} & = \dfrac{1+x}{(1-x)^3} \\
\text{Right-shift Rule:} & \langle 0, 1, 4, 9, \ldots \rangle & \longleftrightarrow x \cdot \dfrac{1+x}{(1-x)^3} & = \dfrac{x(1+x)}{(1-x)^3}
\end{array}
$$

Thus, the generating function for squares is:

$$\frac{x(1+x)}{(1-x)^3}. \tag{12.2}$$

12.2.5 Products

Rule 12.2.5 (Product Rule)**.** *If*

$$\langle a_0, a_1, a_2, \ldots \rangle \quad \longleftrightarrow A(x), \quad \textit{and} \quad \langle b_0, b_1, b_2, \ldots \rangle \quad \longleftrightarrow B(x),$$

then

$$\langle c_0,\ c_1,\ c_2,\ \ldots \rangle \quad \longleftrightarrow A(x) \cdot B(x),$$

where

$$c_n ::= a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0.$$

To understand this rule, let

$$C(x) ::= A(x) \cdot B(x) = \sum_{n=0}^{\infty} c_n x^n.$$

We can evaluate the product $A(x) \cdot B(x)$ by using a table to identify all the cross-terms from the product of the sums:

	b_0x^0	b_1x^1	b_2x^2	b_3x^3	$\ldots$
a_0x^0	$a_0b_0x^0$	$a_0b_1x^1$	$a_0b_2x^2$	$a_0b_3x^3$	$\ldots$
a_1x^1	$a_1b_0x^1$	$a_1b_1x^2$	$a_1b_2x^3$	$\ldots$	
a_2x^2	$a_2b_0x^2$	$a_2b_1x^3$	$\ldots$		
a_3x^3	$a_3b_0x^3$	$\ldots$			
$\vdots$	$\ldots$				

Notice that all terms involving the same power of x lie on a diagonal. Collecting these terms together, we find that the coefficient of x^n in the product is the sum of all the terms on the $(n+1)$st diagonal, namely,

$$a_0b_n + a_1b_{n-1} + a_2b_{n-2} + \cdots + a_nb_0. \tag{12.3}$$

This expression (12.3) may be familiar from a signal processing course; the sequence $\langle c_0, c_1, c_2, \ldots \rangle$ is called the *convolution* of sequences $\langle a_0, a_1, a_2, \ldots \rangle$ and $\langle b_0, b_1, b_2, \ldots \rangle$.

12.3 Evaluating Sums

The product rule looks complicated. But it is surprisingly useful. For example, suppose that we set

$$B(x) = \frac{1}{1-x}.$$

Then $b_i = 1$ for $i \geq 0$ and the nth coefficient of $A(x)B(x)$ is

$$a_0 \cdot 1 + a_1 \cdot 1 + a_2 \cdot 1 + \cdots + a_n \cdot 1 = \sum_{i=0}^{n} a_i.$$

In other words, given any sequence $\langle a_0, a_1, a_2, \ldots \rangle$, we can compute

$$s_n = \sum_{i=0}^{n} a_i$$

for all n by simply multiplying the sequence's generating function by $1/(1-x)$. This is the Summation Rule.

Rule 12.3.1 (Summation Rule)**.** *If*

$$\langle a_0, a_1, a_2, \dots \rangle \quad \longleftrightarrow A(x),$$

then

$$\langle s_0, s_1, s_2, \dots \rangle \quad \longleftrightarrow \frac{A(x)}{1-x}$$

where

$$s_n = \sum_{i=0}^{n} a_i \quad \textit{for } n \geq 0.$$

The Summation Rule sounds powerful, and it is! We know from Chapter 9 that computing sums is often not easy. But multiplying by $1/(1-x)$ is about as easy as it gets.

For example, suppose that we want to compute the sum of the first n squares

$$s_n = \sum_{i=0}^{n} i^2$$

and we forgot the method in Chapter 9. All we need to do is compute the generating function for $\langle 0, 1, 4, 9, \dots \rangle$ and multiply by $1/(1-x)$. We already computed the generating function for $\langle 0, 1, 4, 9, \dots \rangle$ in Equation 12.2—it is

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

Hence, the generating function for $\langle s_0, s_1, s_2, \dots \rangle$ is

$$\frac{x(1+x)}{(1-x)^4}.$$

This means that $\sum_{i=0}^{n} i^2$ is the coefficient of x^n in $x(1+x)/(1-x)^4$.

That was pretty easy, but there is one problem—we have no idea how to determine the coefficient of x^n in $x(1+x)/(1-x)^4$! And without that, this whole endeavor (while magical) would be useless. Fortunately, there is a straightforward way to produce the sequence of coefficients from a generating function.

12.4 Extracting Coefficients

12.4.1 Taylor Series

Given a sequence of coefficients $\langle f_0, f_1, f_2, \ldots \rangle$, computing the generating function $F(x)$ is easy since

$$F(x) = f_0 + f_1 x + f_2 x^2 + \cdots .$$

To compute the sequence of coefficients from the generating function, we need to compute the *Taylor Series* for the generating function.

Rule 12.4.1 (Taylor Series)**.** *Let $F(x)$ be the generating function for the sequence*

$$\langle f_0, f_1, f_2, \ldots \rangle.$$

Then

$$f_0 = F(0)$$

and

$$f_n = \frac{F^{(n)}(0)}{n!}$$

for $n \geq 1$, where $F^{(n)}(0)$ is the nth derivative of $F(x)$ evaluated at $x = 0$.

This is because if

$$F(x) = f_0 + f_1 x + f_2 x^2 + \cdots ,$$

then

$$\begin{aligned} F(0) &= f_0 + f_1 \cdot 0 + f_2 \cdot 0^2 + \cdots \\ &= f_0. \end{aligned}$$

Also,

$$\begin{aligned} F'(x) &= \frac{d}{dx}(F(x)) \\ &= f_1 + 2f_2 x + 3f_3 x^2 + 4f_4 x^3 + \cdots \end{aligned}$$

and so

$$F'(0) = f_1,$$

as desired. Taking second derivatives, we find that

$$\begin{aligned} F''(x) &= \frac{d}{dx}(F'(x)) \\ &= 2f_2 + 3 \cdot 2f_3 x + 4 \cdot 3 f_4 x^2 + \cdots \end{aligned}$$

and so

$$F''(0) = 2f_2,$$

which means that

$$f_2 = \frac{F''(0)}{2}.$$

In general,

$$\begin{aligned} F^{(n)} = n! f_n + (n+1)! f_{n+1} x + \frac{(n+2)!}{2} f_{n+2} x^2 + \cdots \\ + \frac{(n+k)!}{k!} f_{n+k} x^k + \cdots \end{aligned}$$

and so

$$F^{(n)}(0) = n! f_n$$

and

$$f_n = \frac{F^{(n)}(0)}{n!},$$

as claimed.

This means that

$$\left\langle F(0), F'(0), \frac{F''(0)}{2!}, \frac{F'''(0)}{3!}, \ldots, \frac{F^{(n)}(0)}{n!}, \ldots \right\rangle \longleftrightarrow F(x). \tag{12.4}$$

The sequence on the left-hand side of Equation 12.4 gives the well-known Taylor Series expansion for a function

$$F(x) = F(0) + F'(0)x + \frac{F''(0)}{2!}x^2 + \frac{F'''(0)}{3!}x^3 + \cdots + \frac{F^{(n)}(0)}{n!}x^n + \cdots.$$

12.4.2 Examples

Let’s try this out on a familiar example:

$$F(x) = \frac{1}{1-x}.$$

Computing derivatives, we find that

$$
\begin{aligned}
F'(x) &= \frac{1}{(1-x)^2}, \\
F''(x) &= \frac{2}{(1-x)^3}, \\
F'''(x) &= \frac{2 \cdot 3}{(1-x)^4}, \\
&\vdots \\
F^{(n)} &= \frac{n!}{(1-x)^{n+1}}.
\end{aligned}
$$

This means that the coefficient of x^n in $1/(1-x)$ is

$$\frac{F^{(n)}(0)}{n!} = \frac{n!}{n!\,(1-0)^{n+1}} = 1.$$

In other words, we have reconfirmed what we already knew; namely, that

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

Using a similar approach, we can establish some other well-known series:

$$
\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots , \\
e^{ax} &= 1 + ax + \frac{a^2}{2!}x^2 + \frac{a^3}{3!}x^3 + \cdots + \frac{a^n}{n!}x^n + \cdots , \\
\ln(1-x) &= -ax - \frac{a^2}{2}x^2 - \frac{a^3}{3}x^3 - \cdots - \frac{a^n}{n}x^n - \cdots .
\end{aligned}
$$

But what about the series for

$$F(x) = \frac{x(1+x)}{(1-x)^4}? \tag{12.5}$$

In particular, we need to know the coefficient of x^n in $F(x)$ to determine

$$s_n = \sum_{i=0}^{n} i^2.$$

While it is theoretically possible to compute the nth derivative of $F(x)$, the result is a bloody mess. Maybe these generating functions weren't such a great idea after all. . . .

12.4.3 Massage Helps

In times of stress, a little massage can often help relieve the tension. The same is true for polynomials with painful derivatives. For example, let's take a closer look at Equation 12.5. If we massage it a little bit, we find that

$$F(x) = \frac{x + x^2}{(1-x)^4} = \frac{x}{(1-x)^4} + \frac{x^2}{(1-x)^4}. \tag{12.6}$$

The goal is to find the coefficient of x^n in $F(x)$. If you stare at Equation 12.6 long enough (or if you combine the Right-Shift Rule with the Addition Rule), you will notice that the coefficient of x^n in $F(x)$ is just the sum of

$$\text{the coefficient of } x^{n-1} \text{ in } \frac{1}{(1-x)^4} \text{ and}$$
$$\text{the coefficient of } x^{n-2} \text{ in } \frac{1}{(1-x)^4}.$$

Maybe there is some hope after all. Let's see if we can produce the coefficients for $1/(1-x)^4$. We'll start by looking at the derivatives:

$$\begin{aligned}
F'(x) &= \frac{4}{(1-x)^5}, \\
F''(x) &= \frac{4 \cdot 5}{(1-x)^6}, \\
F'''(x) &= \frac{4 \cdot 5 \cdot 6}{(1-x)^7}, \\
&\vdots \\
F^{(n)}(x) &= \frac{(n+3)!}{6(1-x)^{n+4}}.
\end{aligned}$$

This means that the nth coefficient of $1/(1-x)^4$ is

$$\frac{F^{(n)}(0)}{n!} = \frac{(n+3)!}{6n!} = \frac{(n+3)(n+2)(n+1)}{6}. \tag{12.7}$$

We are now almost done. Equation 12.7 means that the coefficient of x^{n-1} in $1/(1-x)^4$ is

$$\frac{(n+2)(n+1)n}{6} \tag{12.8}$$

and the coefficient[2] of x^{n-2} is

$$\frac{(n+1)n(n-1)}{6}. \tag{12.9}$$

Adding these values produces the desired sum

$$\begin{aligned}\sum_{i=0}^{n} i^2 &= \frac{(n+2)(n+1)n}{6} + \frac{(n+1)n(n-1)}{6}\\ &= \frac{(2n+1)(n+1)n}{6}.\end{aligned}$$

This matches Equation 9.14 from Chapter 9. Using generating functions to get the result may have seemed to be more complicated, but at least there was no need for guessing or solving a linear system of equations over 4 variables.

You might argue that the massage step was a little tricky. After all, how were you supposed to know that by converting $F(x)$ into the form shown in Equation 12.6, it would be sufficient to compute derivatives of $1/(1-x)^4$, which is easy, instead of derivatives of $x(1+x)/(1-x)^4$, which could be harder than solving a 64-disk Tower of Hanoi problem step-by-step?

The good news is that this sort of massage works for any generating function that is a ratio of polynomials. Even better, you probably already know how to do it from calculus—it's the method of *partial fractions*!

12.4.4 Partial Fractions

The idea behind partial fractions is to express a ratio of polynomials as a sum of a polynomial and terms of the form

$$\frac{cx^a}{(1-\alpha x)^b} \tag{12.10}$$

where a and b are integers and $b > a \geq 0$. That's because it is easy to compute derivatives of $1/(1-\alpha x)^b$ and thus it is easy to compute the coefficients of Equation 12.10. Let's see why.

Lemma 12.4.2. *If $b \in \mathbb{N}^+$, then the nth derivative of $1/(1-\alpha x)^b$ is*

$$\frac{(n+b-1)!\,\alpha^n}{(b-1)!\,(1-\alpha x)^{b+n}}.$$

[2] To be precise, Equation 12.8 holds for $n \geq 1$ and Equation 12.9 holds for $n \geq 2$. But since Equation 12.8 is 0 for $n = 1$ and Equation 12.9 is 0 for $n = 1, 2$, both equations hold for all $n \geq 0$.

Proof. The proof is by induction on n. The induction hypothesis $P(n)$ is the statement of the lemma.

Base case ($n = 1$): The first derivative is

$$\frac{b\alpha}{(1-\alpha x)^{b+1}}.$$

This matches

$$\frac{(1+b-1)!\,\alpha^1}{(b-1)!\,(1-\alpha x)^{b+1}} = \frac{b\alpha}{(1-\alpha x)^{b+1}},$$

and so $P(1)$ is true.

Induction step: We next assume $P(n)$ to prove $P(n+1)$ for $n \geq 1$. $P(n)$ implies that the nth derivative of $1/(1-\alpha x)^b$ is

$$\frac{(n+b-1)!\,\alpha^n}{(b-1)!\,(1-\alpha x)^{b+n}}.$$

Taking one more derivative reveals that the $(n+1)$st derivative is

$$\frac{(n+b-1)!\,(b+n)\alpha^{n+1}}{(b-1)!\,(1-\alpha x)^{b+n+1}} = \frac{(n+b)!\,\alpha^{n+1}}{(b-1)!\,(1-\alpha x)^{b+n+1}},$$

which means that $P(n+1)$ is true. Hence, the induction is complete. ■

Corollary 12.4.3. *If $a, b \in \mathbb{N}$ and $b > a \geq 0$, then for any $n \geq a$, the coefficient of x^n in*

$$\frac{cx^a}{(1-\alpha x)^b}$$

is

$$\frac{c\,(n-a+b-1)!\,\alpha^{n-a}}{(n-a)!\,(b-1)!}.$$

Proof. By the Taylor Series Rule, the nth coefficient of

$$\frac{1}{(1-\alpha x)^b}$$

is the nth derivative of this expression evaluated at $x = 0$ and then divided by $n!$. By Lemma 12.4.2, this is

$$\frac{(n+b-1)!\,\alpha^n}{n!\,(b-1)!\,(1-0)^{b+n}} = \frac{(n+b-1)!\,\alpha^n}{n!\,(b-1)!}.$$

By the Scaling Rule and the Right-Shift Rule, the coefficient of x^n in

$$\frac{cx^\alpha}{(1-\alpha x)^b}$$

is thus

$$\frac{c\,(n-a+b-1)!\,\alpha^{n-a}}{(n-a)!\,(b-1)!}.$$

as claimed. ■

Massaging a ratio of polynomials into a sum of a polynomial and terms of the form in Equation 12.10 takes a bit of work but is generally straightforward. We will show you the process by means of an example.

Suppose our generating function is the ratio

$$F(x) = \frac{4x^3+2x^2+3x+6}{2x^3-3x^2+1}. \tag{12.11}$$

The first step in massaging $F(x)$ is to get the degree of the numerator to be less than the degree of the denominator. This can be accomplished by dividing the numerator by the denominator and taking the remainder, just as in the Fundamental Theorem of Arithmetic—only now we have polynomials instead of numbers. In this case we have

$$\frac{4x^3+2x^2+3x+6}{2x^3-3x^2+1} = 2 + \frac{8x^2+3x+4}{2x^3-3x^2+1}.$$

The next step is to factor the denominator. This will produce the values of α for Equation 12.10. In this case,

$$\begin{aligned}
2x^3-3x^2+1 &= (2x+1)(x^2-2x+1)\\
&= (2x+1)(x-1)^2\\
&= (1-x)^2(1+2x).
\end{aligned}$$

We next find values c_1, c_2, c_3 so that

$$\frac{8x^2+3x+4}{2x^3-3x^2+1} = \frac{c_1}{1+2x} + \frac{c_2}{(1-x)^2} + \frac{c_3 x}{(1-x)^2}. \tag{12.12}$$

This is done by cranking through the algebra:

$$\begin{aligned}
\frac{c_1}{1+2x} + \frac{c_2}{(1-x)^2} + \frac{c_3 x}{(1-x)^2} &= \frac{c_1(1-x)^2 + c_2(1+2x) + c_3x(1+2x)}{(1+2x)(1-x)^2}\\
&= \frac{c_1 - 2c_1x + c_1x^2 + c_2 + 2c_2x + c_3x + 2c_3x^2}{2x^3-3x^2+1}\\
&= \frac{c_1 + c_2 + (-2c_1 + 2c_2 + c_3)x + (c_1 + 2c_3)x^2}{2x^3-3x^2+1}.
\end{aligned}$$

For Equation 12.12 to hold, we need

$$\begin{aligned}
8 &= c_1 + 2c_3, \\
3 &= -2c_1 + 2c_2 + c_3, \\
4 &= c_1 + c_2.
\end{aligned}$$

Solving these equations, we find that $c_1 = 2,\ c_2 = 2$, and $c_3 = 3$. Hence,

$$\begin{aligned}
F(x) &= \frac{4x^3 + 2x^2 + 3x + 6}{2x^3 - 3x^2 + 1} \\
&= 2 + \frac{2}{1 + 2x} + \frac{2}{(1 - x)^2} + \frac{3x}{(1 - x)^2}.
\end{aligned}$$

Our massage is done! We can now compute the coefficients of $F(x)$ using Corollary 12.4.3 and the Sum Rule. The result is

$$f_0 = 2 + 2 + 2 = 6$$

and

$$\begin{aligned}
f_n &= \frac{2(n - 0 + 1 - 1)!\,(-2)^{n-0}}{(n - 0)!\,(1 - 1)!} \\
&\quad + \frac{2(n - 0 + 2 - 1)!\,(1)^{n-0}}{(n - 0)!\,(2 - 1)!} \\
&\quad + \frac{3(n - 1 + 2 - 1)!\,(1)^{n-1}}{(n - 1)!\,(2 - 1)!} \\
&= (-1)^n 2^{n+1} + 2(n + 1) + 3n \\
&= (-1)^n 2^{n+1} + 5n + 2
\end{aligned}$$

for $n \geq 1$.

Aren't you glad that you know that? Actually, this method turns out to be useful in solving linear recurrences, as we'll see in the next section.

12.5 Solving Linear Recurrences

Generating functions can be used to find a solution to any linear recurrence. We'll show you how this is done by means of a familiar example, the Fibonacci recurrence, so that you can more easily understand the similarities and differences of this approach and the method we showed you in Chapter 10.

12.5.1 Finding the Generating Function

Let's begin by recalling the definition of the Fibonacci numbers:

$$\begin{aligned} f_0 &= 0 \\ f_1 &= 1 \\ f_n &= f_{n-1} + f_{n-2} \qquad \text{for } n \geq 2. \end{aligned}$$

We can expand the final clause into an infinite sequence of equations. Thus, the Fibonacci numbers are defined by:

$$\begin{aligned} f_0 &= 0 \\ f_1 &= 1 \\ f_2 &= f_1 + f_0 \\ f_3 &= f_2 + f_1 \\ f_4 &= f_3 + f_2 \\ &\vdots \end{aligned}$$

The overall plan is to *define* a function $F(x)$ that generates the sequence on the left side of the equality symbols, which are the Fibonacci numbers. Then we *derive* a function that generates the sequence on the right side. Finally, we equate the two and solve for $F(x)$. Let's try this. First, we define:

$$F(x) = f_0 + f_1 x + f_2 x^2 + f_3 x^3 + f_4 x^4 + \cdots .$$

Now we need to derive a generating function for the sequence:

$$\langle 0,\ 1,\ f_1 + f_0,\ f_2 + f_1,\ f_3 + f_2,\ \ldots \rangle .$$

One approach is to break this into a sum of three sequences for which we know generating functions and then apply the Addition Rule:

$$\begin{array}{rllllll} \langle 0, & 1, & 0, & 0, & 0, & \ldots \rangle & \longleftrightarrow x \\ \langle 0, & f_0, & f_1, & f_2, & f_3, & \ldots \rangle & \longleftrightarrow xF(x) \\ + \ \langle 0, & 0, & f_0, & f_1, & f_2, & \ldots \rangle & \longleftrightarrow x^2 F(x) \\ \hline \langle 0, & 1 + f_0, & f_1 + f_0, & f_2 + f_1, & f_3 + f_2, & \ldots \rangle & \longleftrightarrow x + xF(x) + x^2 F(x) \end{array}$$

This sequence is almost identical to the right sides of the Fibonacci equations. The one blemish is that the second term is $1 + f_0$ instead of simply 1. However, this amounts to nothing, since $f_0 = 0$ anyway.

If we equate $F(x)$ with the new function $x + xF(x) + x^2F(x)$, then we're implicitly writing down *all* of the equations that define the Fibonacci numbers in one fell swoop:

$$\begin{array}{ccccccccc} F(x) & = & f_0 & + & f_1 \; x & + & f_2 \; x^2 & + & f_3 \; x^3 + \cdots \\ \parallel & & \parallel & & \parallel & & \parallel & & \parallel \\ x + xF(x) + x^2F(x) & = & 0 & + & (1 + f_0)x & + & (f_1 + f_0)x^2 & + & (f_2 + f_1)x^3 + \cdots \end{array}$$

Solving for $F(x)$ gives the generating function for the Fibonacci sequence:

$$F(x) = x + xF(x) + x^2F(x)$$

so

$$F(x) = \frac{x}{1 - x - x^2}. \tag{12.13}$$

This is pretty cool. After all, who would have thought that the Fibonacci numbers are precisely the coefficients of such a simple function? Even better, this function is a ratio of polynomials and so we can use the method of partial fractions from Section 12.4.4 to find a closed-form expression for the nth Fibonacci number.

12.5.2 Extracting the Coefficients

Repeated differentiation of Equation 12.13 would be very painful. But it is easy to use the method of partial fractions to compute the coefficients. Since the degree of the numerator in Equation 12.13 is less than the degree of the denominator, the first step is to factor the denominator:

$$1 - x - x^2 = (1 - \alpha_1 x)(1 - \alpha_2 x)$$

where $\alpha_1 = (1 + \sqrt{5})/2$ and $\alpha_2 = (1 - \sqrt{5})/2$. These are the same as the roots of the characteristic equation for the Fibonacci recurrence that we found in Chapter 10. That is not a coincidence.

The next step is to find c_1 and c_2 that satisfy

$$\begin{aligned} \frac{x}{1 - x - x^2} &= \frac{c_1}{1 - \alpha_1 x} + \frac{c_2}{1 - \alpha_2 x} \\ &= \frac{c_1(1 - \alpha_2 x) + c_2(1 - \alpha_1 x)}{(1 - \alpha_1 x)(1 - \alpha_2 x)} \\ &= \frac{c_1 + c_2 - (c_1\alpha_2 + c_2\alpha_1)x}{1 - x - x^2}. \end{aligned}$$

Hence,

$$c_1 + c_2 = 0 \qquad \text{and} \qquad -(c_1\alpha_2 + c_2\alpha_1) = 1.$$

Solving these equations, we find that

$$c_1 = \frac{1}{\alpha_1 - \alpha_2} = \frac{1}{\sqrt{5}}$$
$$c_2 = \frac{-1}{\alpha_1 - \alpha_2} = \frac{-1}{\sqrt{5}}.$$

We can now use Corollary 12.4.3 and the Sum Rule to conclude that

$$\begin{aligned} f_n &= \frac{\alpha_1^n}{\sqrt{5}} - \frac{\alpha_2^n}{\sqrt{5}} \\ &= \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \end{aligned}$$

This is exactly the same formula we derived for the nth Fibonacci number in Chapter 10.

12.5.3 General Linear Recurrences

The method that we just used to solve the Fibonacci recurrence can also be used to solve general linear recurrences of the form

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_d f_{n-d} + g_n$$

for $n \geq d$. The generating function for $\langle f_0, f_1, f_2, \ldots \rangle$ is

$$F(x) = \frac{h(x) + G(x)}{1 - a_1 x - a_2 x^2 - \cdots - a_d x^d}$$

where $G(x)$ is the generating function for the sequence

$$\langle \overbrace{0, 0, \ldots, 0}^{d}, g_d, g_{d+1}, g_{d+2}, \ldots \rangle$$

and $h(x)$ is a polynomial of degree at most $d - 1$ that is based on the values of f_0, $f_1, \ldots, f_{d-1}$. In particular,

$$h(x) = \sum_{i=0}^{d-1} h_i x^i$$

where

$$h_i = f_0 - a_1 f_{i-1} - a_2 f_{i-2} - \cdots - a_i f_0$$

for $0 \leq i < d$.

To solve the recurrence, we use the method of partial fractions described in Section 12.4.4 to find a closed-form expression for $F(x)$. This can be easy or hard to do depending on $G(x)$.

12.6 Counting with Generating Functions

Generating functions are particularly useful for solving counting problems. In particular, problems involving choosing items from a set often lead to nice generating functions by letting the coefficient of x^n be the number of ways to choose n items.

12.6.1 Choosing Distinct Items from a Set

The generating function for binomial coefficients follows directly from the Binomial Theorem:

$$\begin{aligned}\left\langle \binom{k}{0}, \binom{k}{1}, \binom{k}{2}, \ldots, \binom{k}{k}, 0, 0, 0, \ldots \right\rangle &\longleftrightarrow \binom{k}{0} + \binom{k}{1}x + \binom{k}{2}x^2 + \cdots + \binom{k}{k}x^k \\ &= (1+x)^k\end{aligned}$$

Thus, the coefficient of x^n in $(1+x)^k$ is $\binom{k}{n}$, the number of ways to choose n distinct items[3] from a set of size k. For example, the coefficient of x^2 is $\binom{k}{2}$, the number of ways to choose 2 items from a set with k elements. Similarly, the coefficient of x^{k+1} is the number of ways to choose $k+1$ items from a size k set, which is zero.

12.6.2 Building Generating Functions that Count

Often we can translate the description of a counting problem directly into a generating function for the solution. For example, we could figure out that $(1+x)^k$ generates the number of ways to select n distinct items from a k-element set without resorting to the Binomial Theorem or even fussing with binomial coefficients! Let's see how.

First, consider a single-element set $\{a_1\}$. The generating function for the number of ways to select n elements from this set is simply $1+x$: we have 1 way to select zero elements, 1 way to select one element, and 0 ways to select more than one element. Similarly, the number of ways to select n elements from the set $\{a_2\}$ is also given by the generating function $1+x$. The fact that the elements differ in the two cases is irrelevant.

Now here is the main trick: *the generating function for choosing elements from a union of disjoint sets is the product of the generating functions for choosing from each set.* We'll justify this in a moment, but let's first look at an example. According to this principle, the generating function for the number of ways to select

[3]Watch out for the reversal of the roles that k and n played in earlier examples; we're led to this reversal because we've been using n to refer to the power of x in a power series.

n elements from the $\{a_1, a_2\}$ is:

$$\underbrace{(1+x)}_{\text{select from } \{a_1\}} \cdot \underbrace{(1+x)}_{\text{select from } \{a_2\}} = \underbrace{(1+x)^2}_{\text{select from } \{a_1, a_2\}} = 1 + 2x + x^2.$$

Sure enough, for the set $\{a_1, a_2\}$, we have 1 way to select zero elements, 2 ways to select one element, 1 way to select two elements, and 0 ways to select more than two elements.

Repeated application of this rule gives the generating function for selecting n items from a k-element set $\{a_1, a_2, \ldots, a_k\}$:

$$\underbrace{(1+x)}_{\text{select from } \{a_1\}} \cdot \underbrace{(1+x)}_{\text{select from } \{a_2\}} \cdots \underbrace{(1+x)}_{\text{select from } \{a_k\}} = \underbrace{(1+x)^k}_{\substack{\text{select from} \\ \{a_1, a_2, \ldots, a_k\}}}$$

This is the same generating function that we obtained by using the Binomial Theorem. But this time around, we translated directly from the counting problem to the generating function.

We can extend these ideas to a general principle:

Rule 12.6.1 (Convolution Rule)**.** *Let $A(x)$ be the generating function for selecting items from set $\mathcal{A}$, and let $B(x)$ be the generating function for selecting items from set $\mathcal{B}$. If $\mathcal{A}$ and $\mathcal{B}$ are disjoint, then the generating function for selecting items from the union $\mathcal{A} \cup \mathcal{B}$ is the product $A(x) \cdot B(x)$.*

This rule is rather ambiguous: what exactly are the rules governing the selection of items from a set? Remarkably, the Convolution Rule remains valid under *many* interpretations of selection. For example, we could insist that distinct items be selected or we might allow the same item to be picked a limited number of times or any number of times. Informally, the only restrictions are that (1) the order in which items are selected is disregarded and (2) restrictions on the selection of items from sets $\mathcal{A}$ and $\mathcal{B}$ also apply in selecting items from $\mathcal{A} \cup \mathcal{B}$. (Formally, there must be a bijection between n-element selections from $\mathcal{A} \cup \mathcal{B}$ and ordered pairs of selections from $\mathcal{A}$ and $\mathcal{B}$ containing a total of n elements.)

To count the number of ways to select n items from $\mathcal{A} \cup \mathcal{B}$, we observe that we can select n items by choosing j items from $\mathcal{A}$ and $n - j$ items from $\mathcal{B}$, where j is any number from 0 to n. This can be done in $a_j b_{n-j}$ ways. Summing over all the possible values of j gives a total of

$$a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0$$

ways to select n items from $\mathcal{A} \cup \mathcal{B}$. By the Product Rule, this is precisely the coefficient of x^n in the series for $A(x)B(x)$.

12.6.3 Choosing Items with Repetition

The first counting problem we considered was the number of ways to select a dozen doughnuts when five flavors were available. We can generalize this question as follows: in how many ways can we select n items from a k-element set if we're allowed to pick the same item multiple times? In these terms, the doughnut problem asks how many ways we can select $n = 12$ doughnuts from the set of $k = 5$ flavors

$$\{\text{chocolate, lemon-filled, sugar, glazed, plain}\}$$

where, of course, we're allowed to pick several doughnuts of the same flavor. Let's approach this question from a generating functions perspective.

Suppose we make n choices (with repetition allowed) of items from a set containing a single item. Then there is one way to choose zero items, one way to choose one item, one way to choose two items, etc. Thus, the generating function for choosing n elements with repetition from a 1-element set is:

$$\langle 1, 1, 1, 1, \dots \rangle \quad \longleftrightarrow \quad 1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}.$$

The Convolution Rule says that the generating function for selecting items from a union of disjoint sets is the product of the generating functions for selecting items from each set:

$$\underbrace{\frac{1}{1-x}}_{\text{choose } a_1\text{'s}} \cdot \underbrace{\frac{1}{1-x}}_{\text{choose } a_2\text{'s}} \cdots \underbrace{\frac{1}{1-x}}_{\text{choose } a_k\text{'s}} = \underbrace{\frac{1}{(1-x)^k}}_{\substack{\text{repeatedly choose from} \\ \{a_1, a_2, \dots, a_k\}}}$$

Therefore, the generating function for choosing items from a k-element set with repetition allowed is $1/(1-x)^k$. Computing derivatives and applying the Taylor Series Rule, we can find that the coefficient of x^n in $1/(1-x)^k$ is

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

This is the Bookkeeper Rule from Chapter 11—namely there are $\binom{n+k-1}{n}$ ways to select n items with replication from a set of k items.

12.6.4 Fruit Salad

In this chapter, we have covered a lot of methods and rules for using generating functions. We'll now do an example that demonstrates how the rules and methods can be combined to solve a more challenging problem—making fruit salad.

In how many ways can we make a salad with n fruits subject to the following constraints?

- The number of apples must be even.
- The number of bananas must be a multiple of 5.
- There can be at most four oranges.
- There can be at most one pear.

For example, there are 7 ways to make a salad with 6 fruits:

Apples	6	4	4	2	2	0	0
Bananas	0	0	0	0	0	5	5
Oranges	0	2	1	4	3	1	0
Pears	0	0	1	0	1	0	1

These constraints are so complicated that the problem seems hopeless! But generating functions can solve the problem in a straightforward way.

Let's first construct a generating function for choosing apples. We can choose a set of 0 apples in one way, a set of 1 apple in zero ways (since the number of apples must be even), a set of 2 apples in one way, a set of 3 apples in zero ways, and so forth. So we have:

$$A(x) = 1 + x^2 + x^4 + x^6 + \cdots = \frac{1}{1 - x^2}.$$

Similarly, the generating function for choosing bananas is:

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

We can choose a set of 0 oranges in one way, a set of 1 orange in one way, and so on. However, we can not choose more than four oranges, so we have the generating function:

$$O(x) = 1 + x + x^2 + x^3 + x^4 = \frac{1 - x^5}{1 - x}.$$

Here we're using the geometric sum formula. Finally, we can choose only zero or one pear, so we have:

$$P(x) = 1 + x.$$

The Convolution Rule says that the generating function for choosing from among all four kinds of fruit is:

$$\begin{aligned} A(x)B(x)O(x)P(x) &= \frac{1}{1-x^2}\frac{1}{1-x^5}\frac{1-x^5}{1-x}(1+x) \\ &= \frac{1}{(1-x)^2} \\ &= 1 + 2x + 3x^2 + 4x^3 + \cdots . \end{aligned}$$

Almost everything cancels! We're left with $1/(1-x)^2$, which we found a power series for earlier: the coefficient of x^n is simply $n+1$. Thus, the number of ways to make a salad with n fruits is just $n+1$. This is consistent with the example we worked out at the start, since there were 7 different salads containing 6 fruits. *Amazing!*