Chapter 3

Counting with Ordinary Generating Functions

This chapter introduces one of the most powerful techniques in the enumerator's toolkit: generating functions. Wilf [**101**] wrote a whole book devoted to their properties. There are several types of generating functions and we will start with the simplest, which are called ordinary generating functions. In Chapters 4, 7, and 8 we will deal with other types. The basic idea in all cases is to take a sequence of numbers in which we are interested and replace it by an algebraic object, namely a polynomial or power series. The advantage of doing this is that one can then bring a host of algebraic techniques to bear in order to study the original sequence. This makes it possible to give proofs of results about the sequence which have the following advantages:

(1) The proofs can be very short.

(2) Many demonstrations can be done by straightforward manipulations which do not require the cleverness of other approaches.

(3) Sometimes no other method is known for obtaining a given result.

3.1. Generating polynomials

Let x be a variable. A sequence

$$a_0, a_1, a_2, \ldots, a_n \tag{3.1}$$

of complex numbers has *ordinary generating polynomial*

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum_{k=0}^{n} a_k x^k.$$

Here, "ordinary" is to distinguish this generating polynomial from other types. Since we will only be dealing with the ordinary case in this chapter, we will usually drop the adjective. Note that $f(x)$ is an element of the algebra $\mathbb{C}[x]$ of polynomials in x

with complex coefficients. We will also often call $f(x)$ the *generating function* for the sequence (3.1) since it is a special case of the generating function for a sequence with a countable, but perhaps not finite, number of terms. This more general setting will be discussed in Section 3.3.

To begin with a simple example, consider the sequence of binomial coefficients found in a row of Pascal's triangle

$$\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n}.$$

The corresponding generating function is

$$f(x) = \sum_{k=0}^{n} \binom{n}{k} x^k.$$

In particular, when $n = 4$ we get

$$f(x) = 1 + 4x + 6x^2 + 4x^3 + x^4 = (1 + x)^4.$$

The power of this generating function is that it can be expressed as a product which is just the well-known Binomial Theorem. We will give two proofs of this result, one combinatorial and one using algebraic manipulations.

Theorem 3.1.1 (Binomial Theorem). *For $n \in \mathbb{N}$ we have*

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

Proof (Combinatorial). Consider expanding the product

$$(1 + x)^n = \overbrace{(1 + x)(1 + x) \cdots (1 + x)}^{n}$$

using the distributive law. One obtains a term x^k in the expansion by picking the x in k of the factors and picking the 1 in the remaining $n - k$. But the number of ways of choosing k objects from n objects is $\binom{n}{k}$. So that is the coefficient of x^k in the product and we are done. □

Proof (Algebraic). We will induct on n. The result is clearly true for $n = 0$ so assume $n \geq 1$. Note that, because of our conventions for binomial coefficients, we can write the generating function as

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

The advantage of doing this is that we will not have to worry about boundary cases when $k = 0$ or $k = n$ and so we will suppress the limits. Now using the binomial recursion in Theorem 1.3.3(a), reindexing, and induction

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

as desired. □

The first proof illustrates the use of the Product Rule for weight-generating functions which will be discussed in Section 3.4. The second proof is an example of the point made in the chapter introduction about how proofs involving generating functions can be based on routine manipulations. And the trick of extending the domain of summation is one which we will often use to simplify demonstrations. We now wish to give an illustration of how a generating function, once derived, can be used to give simple proofs of other results. In particular, setting $x = 1$ in the Binomial Theorem we immediately get

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

which is part (c) of Theorem 1.3.3. Similarly, letting $x = -1$ in Theorem 3.1.1 gives

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = (1-1)^n = 0^n = \delta_{0,n},$$

which is Theorem 1.3.3(d).

We end this section by stating the generating function for the Stirling numbers of the first kind. This result can be proved similarly to the algebraic proof of the Binomial Theorem so its demonstration will be left as an exercise. Finding a generating function for the Stirling numbers of the second kind will have to wait until after we have discussed formal power series in Section 3.3.

Theorem 3.1.2. *For $n \in \mathbb{N}$ we have*

$$\sum_{k=0}^{n} c(n,k) x^k = x(x+1)(x+2)\ldots(x+n-1). \qquad \square$$

Note that by setting $x = 1$ in the previous displayed equation we obtain the special case

$$\#P([n]) = \sum_k c(n,k) = n!\,.$$

So this proposition can be considered a generalization of Theorem 1.2.1. Such extensions are called q-analogues and will be discussed in the next section.

3.2. Statistics and q-analogues

One way of constructing generating functions is through the use of statistics and q-analogues. Because of connections with the theory of hypergeometic series, the variable q is usually used for these generating functions. This is a mnemonic choice since sometimes, as we will see below, q stands for the power of a prime p. There is no formal definition of a q-analogue, so we will start with an example which will illustrate the meta-definition we will eventually give.

A *statistic* on a set S is a function $\operatorname{st}: S \to \mathbb{N}$. Because the range of a statistic is $\mathbb{N}$ we can define, for finite S, a corresponding generating polynomial

$$f(q) = \sum_{s \in S} q^{\operatorname{st} s}.$$

This generating function is sometimes called the *distribution* of st over S because it can also be written

$$f(q) = \sum_{k \geq 0} a_k q^k$$

where a_k is the number of $s \in S$ satisfying $\operatorname{st} s = k$ and this parallels the distribution of a random variable in probability theory. One of the most famous statistics on permutations is the inversion number. A permutation $\pi = \pi_1 \dots \pi_n \in P([n])$ has *inversion set*

$$\operatorname{Inv} \pi = \{(i, j) \mid i < j \text{ and } \pi_i > \pi_j\}.$$

One can think of this as the set of pairs of indices where the corresponding elements of π are out of their natural increasing order. Note that one uses pairs of indices rather than the elements of π because this makes it easier to generalize this concept to words where repetitions are allowed. For example, if $\pi = \pi_1\pi_2\pi_3\pi_4\pi_5 = 41532$, then

$$\operatorname{Inv} \pi = \{(1, 2),\ (1, 4),\ (1, 5),\ (3, 4),\ (3, 5),\ (4, 5)\}.$$

The *inversion number* of π is just

$$\operatorname{inv} \pi = \# \operatorname{Inv} \pi.$$

We will often use the convention of beginning functions having to do with sets with uppercase letters and their corresponding cardinalities with lowercase. Continuing our example, $\operatorname{inv} 41532 = 6$. Clearly $\operatorname{inv}: P([n]) \to \mathbb{N}$ is a statistic and it has a very interesting generating polynomial.

Theorem 3.2.1. *For $n \geq 0$ we have*

$$\sum_{\pi \in P([n])} q^{\operatorname{inv} \pi} = (1)(1 + q)(1 + q + q^2) \cdots (1 + q + q^2 + \cdots + q^{n-1}).$$

Proof. We will induct on n, omitting the trivial base case. Every $\pi \in P([n])$ can be obtained uniquely from a $\sigma \in P([n-1])$ by inserting n into one of the n spaces between the elements of σ (including the space before σ_1 and the space after σ_{n-1}). Let σ^i be the result of placing n in the ith space from the right where the space after σ_{n-1} is considered space zero. Then clearly

$$\operatorname{inv}\sigma^i = i + \operatorname{inv}\sigma.$$

Using this equation and induction we see that

$$\begin{aligned}\sum_{\pi\in P([n])} q^{\operatorname{inv}\pi} &= \sum_{\sigma\in P([n-1])}\sum_{i=0}^{n-1} q^{\operatorname{inv}\sigma^i}\\ &= \sum_{\sigma\in P([n-1])} q^{\operatorname{inv}\sigma}\cdot\sum_{i=0}^{n-1} q^i\\ &= (1+q)(1+q+q^2)\cdots(1+q+q^2+\cdots+q^{n-1})\end{aligned}$$

as we wished to prove. □

Note that by plugging $q = 1$ into this result one obtains

$$\#P([n]) = \sum_{\pi\in P([n])} 1 = n!\,,$$

which is the second statement in Theorem 1.2.1.

Now that we have met some q-analogues (although they have not been named as such), their meta-definition should make more sense. A *q-analogue* of a combinatorial object $\mathcal{O}$ is an object $\mathcal{O}(q)$ such that

$$\lim_{q\to 1}\mathcal{O}(q) = \mathcal{O}.$$

Note that $\mathcal{O}$ could be many things: a number, a definition, or a theorem. For example, one of the standard q-analogues of $n \in \mathbb{N}$ is the polynomial

$$[n]_q = 1 + q + q^2 + \cdots + q^{n-1}. \tag{3.2}$$

Clearly $[n]_1 = n$. Another possible q-analogue of n is the rational function $(1-q^n)/(1-q)$. In this case one cannot just substitute $q = 1$ but must take a limit. Of course, this quotient and $[n]_q$ are equal when $q \neq 1$. Another q-analogue is the *q-factorial*

$$[n]_q! = [1]_q[2]_q\ldots[n]_q.$$

So Theorem 3.2.1 can be restated as

$$\sum_{\pi\in P([n])} q^{\operatorname{inv}\pi} = [n]_q!\,.$$

Note that we will sometimes write $[n]_q$ as just $[n]$. This could cause confusion with the use of $[n]$ as a set, so we will only use this simplification if it is clear which of the two possible meanings is meant. Similarly, we will often drop the q subscript from other q-analogues when convenient.

There is another famous statistic which has $[n]_q!$ as its distribution. The *descent set* of $\pi \in P([n])$ is

$$\operatorname{Des}\pi = \{i \mid \pi_i > \pi_{i+1}\} \tag{3.3}$$

with corresponding *descent number* $\operatorname{des}\pi = \#\operatorname{Des}\pi$. Equivalently $i \in \operatorname{Des}\pi$ if and only if $(i, i+1) \in \operatorname{Inv}\pi$. We also define the *ascent set*, $\operatorname{Asc}\pi$, and *ascent number*, $\operatorname{asc}\pi$, analogously by reversing the inequality in definition (3.3). Using our previous example we have $\operatorname{Des} 41532 = \{1, 3, 4\}$ and $\operatorname{des} 41532 = 3$. The *major index* of π is

$$\operatorname{maj}\pi = \sum_{i \in \operatorname{Des}\pi} i.$$

So $\operatorname{maj} 41532 = 1 + 3 + 4 = 8$. The term "major index" was coined by Dominique Foata [**26**] in honor of Percy MacMahon who first studied this statistic [**61**] and was a major in the British army.

Theorem 3.2.2. *For $n \geq 0$ we have*

$$\sum_{\pi \in P([n])} q^{\operatorname{maj}\pi} = [n]_q!\,.$$

Proof. We start as in the proof of Theorem 3.2.1 but now number the spaces of σ differently. First number the spaces between σ_i and σ_{i+1} where i is a descent, as well as the space after σ_{n-1}, from right to left starting with zero. Now number the remaining spaces, including the one before σ_1, from left to right with the numbers $\operatorname{des}\sigma + 1, \operatorname{des}\sigma + 2, \ldots, n-1$. An example follows this proof.

Let $\sigma^{(j)}$ denote the result of placing n in space j with this maj labeling. We claim that

$$\operatorname{maj}\sigma^{(j)} = j + \operatorname{maj}\sigma. \tag{3.4}$$

Indeed, if space j is in a descent or at the end of σ, then inserting n just moves the j descents to the right of and including the given descent one position to the right. By definition of major index, this adds a total of j to $\operatorname{maj}\sigma$. If space j is in an ascent or at the beginning of σ, then inserting n creates a new descent as well as moving descents to the right of the space one position to the right. It is easy to check for these j that if inserting n in space j caused $\operatorname{maj}\sigma$ to increase by j, then inserting n in place $j+1$ increases $\operatorname{maj}\sigma$ by $j+1$. So, by induction, equation (3.4) continues to hold in this range of j. The completion of the proof is now done exactly as in the demonstration of Theorem 3.2.1. □

Continuing on with $\sigma = 41532$ having $\operatorname{maj}\sigma = 8$, the spaces are labeled using subscripts as follows:

$${}_4 4_3 1_5 5_2 3_1 2_0.$$

Inserting 6 into each space in turn gives

j	0	1	2	3	4	5
$\sigma^{(j)}$	415326	415362	415632	461532	641532	416532
$\operatorname{maj}\sigma^{(j)}$	8	9	10	11	12	13

It turns out that there are many permutation statistics whose distribution is $[n]_q!$ and these statistics were dubbed *Mahonian* by Foata. One can consult the article of Babson and Steingrímsson [**3**] for a list of Mahonian statistics.

Having found q-analogues involving permutations, the reader may suspect that they also exist for combinations. For integers $0 \leq k \leq n$, define the *q-binomial coefficients* or *Gaussian polynomials* to be

$$\left[\begin{array}{c} n \\ k \end{array} \right]_q = \frac{[n]_q!}{[k]_q!\,[n-k]_q!}.$$

As usual, we let this function be zero if $k < 0$ or $k > n$. For example

$$\begin{aligned} \left[\begin{array}{c} 4 \\ 2 \end{array} \right] &= \frac{[4]!}{[2]!\,[2]!} \\ &= \frac{[4][3]}{[2][1]} \\ &= \frac{(1+q+q^2+q^3)(1+q+q^2)}{(1+q)} \\ &= 1 + q + 2q^2 + q^3 + q^4. \end{aligned} \tag{3.5}$$

It is not at all clear from the definition just given that this is actually a polynomial in q rather than just a rational function. But this follows easily using induction and our next result. Note that this theorem gives two q-analogues for the ordinary binomial recursion. This illustrates a general principle that q-analogues are not necessarily unique as we have also seen in the inv and maj interpretations of $[n]_q!$.

Theorem 3.2.3. *We have*

$$\left[\begin{array}{c} 0 \\ k \end{array} \right]_q = \delta_{0,k}$$

and, for $n \geq 1$,

$$\begin{aligned} \left[\begin{array}{c} n \\ k \end{array} \right]_q &= q^k \left[\begin{array}{c} n-1 \\ k \end{array} \right]_q + \left[\begin{array}{c} n-1 \\ k-1 \end{array} \right]_q \\ &= \left[\begin{array}{c} n-1 \\ k \end{array} \right]_q + q^{n-k} \left[\begin{array}{c} n-1 \\ k-1 \end{array} \right]_q. \end{aligned}$$

Proof. The initial condition is trivial. We will prove the first recursion for the q-binomial, leaving the other as an exercise. Using the definition in terms of q-factorials and finding a common denominator gives

$$\begin{aligned} q^k \left[\begin{array}{c} n-1 \\ k \end{array} \right] + \left[\begin{array}{c} n-1 \\ k-1 \end{array} \right] &= \frac{[n-1]!}{[k]!\,[n-k]!} \left(q^k [n-k] + [k] \right) \\ &= \frac{[n-1]!}{[k]!\,[n-k]!} \cdot [n] \\ &= \left[\begin{array}{c} n \\ k \end{array} \right] \end{aligned}$$

as desired. □

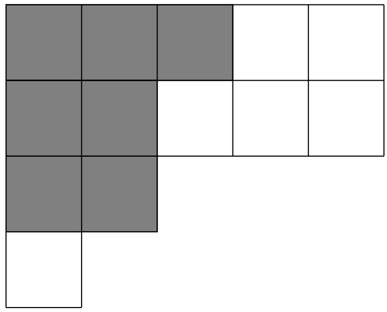

Figure 3.1. The Young diagrams for $(5,5,2,1) \supseteq (3,2,2)$

We will now give a q-analogue of the Binomial Theorem (Theorem 3.1.1). Let q, t be two variables.

Theorem 3.2.4. *For $n \geq 0$ we have*

$$(1+t)(1+qt)(1+q^2t)\cdots(1+q^{n-1}t) = \sum_{k=0}^{n} q^{\binom{k}{2}} \left[\begin{array}{c} n \\ k \end{array}\right]_q t^k. \tag{3.6}$$

Proof. We will induct on n where the case $n = 0$ is easy to check. For $n > 0$ we can use the second recursion in the previous result and the induction hypothesis to write

$$\sum_k q^{\binom{k}{2}} \left[\begin{array}{c} n \\ k \end{array}\right]_q t^k = \sum_k q^{\binom{k}{2}} \left[\begin{array}{c} n-1 \\ k \end{array}\right]_q t^k + \sum_k q^{\binom{k}{2}+n-k} \left[\begin{array}{c} n-1 \\ k-1 \end{array}\right]_q t^k$$

$$= (1+t)(1+qt)\cdots(1+q^{n-2}t) + q^{n-1}t \sum_k q^{\binom{k-1}{2}} \left[\begin{array}{c} n-1 \\ k-1 \end{array}\right]_q t^{k-1}$$

$$= (1+t)(1+qt)\cdots(1+q^{n-2}t) + q^{n-1}t(1+t)(1+qt)\cdots(1+q^{n-2}t)$$

$$= (1+t)(1+qt)\cdots(1+q^{n-1}t),$$

which is what we wished to prove. □

There are many combinatorial interpretations of the q-binomial coefficients. We will content ourselves with presenting two of them here. If $\lambda = (\lambda_1, \ldots, \lambda_k)$ and $\mu = (\mu_1, \ldots, \mu_l)$ are integer partitions, then we say that λ *contains* μ, written $\lambda \supseteq \mu$, if $k \geq l$ and $\lambda_i \geq \mu_i$ for $i \leq l$. Equivalently, the Young diagram of λ contains the Young diagram of μ if they are placed so that their northwest corners align. As an example, $(5,5,2,1) \supseteq (3,2,2)$ and Figure 3.1 shows the diagram of λ with the squares of μ shaded inside. The notation $\mu \subseteq \lambda$ should be self-explanatory. Given $\mu \subseteq \lambda$, one also has the corresponding *skew partition*

$$\lambda/\mu = \{(i,j) \in \lambda \mid (i,j) \notin \mu\}. \tag{3.7}$$

The cells of the skew partition in Figure 3.1 are white.

The $k \times l$ *rectangle* is the integer partition whose multiplicity notation is (k^l). Consider the set of partitions contained in this rectangle

$$\mathcal{R}(k,l) = \{\lambda \mid \lambda \subseteq (k^l)\}.$$

Recalling that $|\lambda|$ is the sum of the parts of λ, we consider the generating function $\sum_{\lambda \in \mathcal{R}(k,l)} q^{|\lambda|}$. For example, if $k = l = 2$, then we have

$\lambda \subseteq (2^2)$	$\emptyset$	(1)	(2)	(1^2)	$(2,1)$	(2^2)
$q^{\|\lambda\|}$	1	q	q^2	q^2	q^3	q^4

which gives

$$\sum_{\lambda \in \mathcal{R}(2,2)} q^{|\lambda|} = 1 + q + 2q^2 + q^3 + q^4.$$

The reader will have noticed the similarity to (3.5), which is not an accident.

Theorem 3.2.5. *For $k, l \geq 0$ we have*

$$\sum_{\lambda \in \mathcal{R}(k,l)} q^{|\lambda|} = \begin{bmatrix} k+l \\ k \end{bmatrix}_q .$$

Proof. We induct on k where the case $k = 0$ is left to the reader. If $k > 0$ and $\lambda \subseteq (k^l)$, then there are two possibilities. Either $\lambda_1 < k$ in which case $\lambda \subseteq ((k-1)^l)$ or $\lambda_1 = k$ so that λ can be written as $\lambda = (k, \lambda')$ where λ' is the partition containing the parts of λ other than λ_1. So $\lambda' \subseteq (k^{l-1})$. Notice that in this case $|\lambda| = |\lambda'| + k$. We now use induction and Theorem 3.2.3 to obtain

$$\begin{aligned} \sum_{\lambda \in \mathcal{R}(k,l)} q^{|\lambda|} &= \sum_{\lambda \in \mathcal{R}(k-1,l)} q^{|\lambda|} + \sum_{\lambda' \in \mathcal{R}(k,l-1)} q^{|\lambda'|+k} \\ &= \begin{bmatrix} k+l-1 \\ k-1 \end{bmatrix} + q^k \begin{bmatrix} k+l-1 \\ k \end{bmatrix} \\ &= \begin{bmatrix} k+l \\ k \end{bmatrix}, \end{aligned}$$

which finishes the proof. □

For our second combinatorial interpretation of the Gaussian polynomials we will need some linear algebra. Let q be a prime power and let $\mathbb{F}_q$ be the Galois field with q elements. Let V be a vector space of dimension $\dim V = n$ over $\mathbb{F}_q$. We will use $W \leq V$ to indicate that W is a subspace of V. Let

$$\begin{bmatrix} V \\ k \end{bmatrix} = \{W \leq V \mid \dim W = k\}.$$

The subspaces of dimension k are in bijective correspondence with $k \times n$ row-reduced echelon matrices of full rank, that is, with no zero rows. For example, if $n = 4$ and

$k = 2$, then the possible matrices are

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix},$$

$$\begin{bmatrix} 1 & * & * & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix}, \begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \end{bmatrix},$$

where the stars represent arbitrary elements of $\mathbb{F}_q$. So the number of subspaces corresponding to one of these star diagrams is q^s where s is the number of stars. Thus

$$\#\begin{bmatrix} \mathbb{F}_q^4 \\ 2 \end{bmatrix} = 1 + q + 2q^2 + q^3 + q^4,$$

which should look very familiar at this point! Note however that, in contrast to previous cases, this actually represents an integer rather than a polynomial since q is a prime power. Of course, this example generalizes. Because of this result people sometimes talk half-jokingly about sets being vector spaces over the (nonexistent) Galois field with one element.

Theorem 3.2.6. *If V is a vector space over $\mathbb{F}_q$ of dimension n, then*

$$\#\begin{bmatrix} V \\ k \end{bmatrix} = \begin{bmatrix} n \\ k \end{bmatrix}_q .$$

Proof. Given $W \le V$ with $\dim W = k$, we first count the number of possible ordered bases $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ for W. Note that since $\dim V = n$ we have $\#V = \#\mathbb{F}_q^n = q^n$. We can pick any nonzero vector for $\mathbf{v}_1$ so the number of choices is $q^n - 1$. For $\mathbf{v}_2$ we can choose any vector in V which is not in the span of $\mathbf{v}_1$, which gives $q^n - q$ possibilities. Continuing in this way, the total count will be

$$(q^n - 1)(q^n - q)(q^n - q^2) \ldots (q^n - q^{k-1}).$$

By a similar argument, the number of different ordered bases which span a given W of dimension k is

$$(q^k - 1)(q^k - q)(q^k - q^2) \ldots (q^k - q^{k-1}).$$

So the number of possible W's is

$$\begin{aligned} \frac{(q^n - 1)(q^n - q) \ldots (q^n - q^{k-1})}{(q^k - 1)(q^k - q) \ldots (q^k - q^{k-1})} &= \frac{q^{\binom{k}{2}}(q^n - 1)(q^{n-1} - 1) \ldots (q^{n-k+1} - 1)}{q^{\binom{k}{2}}(q^k - 1)(q^{k-1} - 1) \ldots (q - 1)} \\ &= \frac{(q-1)^k [n][n-1] \ldots [n-k+1]}{(q-1)^k [k][k-1] \ldots [1]} \\ &= \begin{bmatrix} n \\ k \end{bmatrix}, \end{aligned}$$

as advertised. □

There is a beautiful proof of this result due to Knuth [**50**] using row-reduced echelon matrices as in the previous example. The reader will be asked to supply the details in the exercises.

3.3. The algebra of formal power series

We now wish to generalize the concept of generating function from finite to countably infinite sequences. To do so, we will have to use power series. But we wish to avoid the questions of convergence which come up when using analytic power series. Instead, we will work in the algebra of formal power series. This will mean that we have to be careful since, in an algebra, one is only permitted to apply an operation like addition or multiplication a finite number of times. But there is another concept of convergence which will take care of this issue. We should note that there is a whole branch of combinatorics which uses analytic techniques to extract useful information about a sequence, such as its rate of growth, from the corresponding power series. For information about this approach, see the book of Flajolet and Sedgewick [**25**].

A *formal power series* is an expression of the form

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots = \sum_{n=0}^{\infty} a_nx^n,$$

where the a_n are complex numbers. We also say that $f(x)$ is the *ordinary generating function* or *ogf* for the sequence a_n, $n \geq 0$. Often we will leave out the adjective "ordinary" in this chapter since we will not have met any other type of generating function yet.

Note that these series are considered formal in the sense that the powers of x are just place holders and we are not permitted to substitute a value for x. Because of this rule, analytic convergence is not an issue and we can happily talk about formal power series such as $\sum_{n\geq 0} n!\,x^n$ which converge nowhere except at $x = 0$. We will use the notation

$$\mathbb{C}[[x]] = \left\{ \sum_{n\geq 0} a_nx^n \mid a_n \in \mathbb{C} \text{ for all } n \geq 0 \right\}.$$

The set is an algebra, the *algebra of formal power series*, under the three operations of addition, scalar multiplication, and multiplication defined by

$$\sum_{n\geq 0} a_nx^n + \sum_{n\geq 0} b_nx^n = \sum_{n\geq 0} (a_n + b_n)x^n,$$

$$c\sum_{n\geq 0} a_nx^n = \sum_{n\geq 0} (ca_n)x^n,$$

$$\sum_{n\geq 0} a_nx^n \cdot \sum_{n\geq 0} b_nx^n = \sum_{n\geq 0} c_nx^n,$$

where $c \in \mathbb{C}$ and

$$c_n = \sum_{k=0}^{n} a_kb_{n-k}.$$

The reader may object that, as mentioned earlier, in an algebra one is only permitted a finite number of additions yet the very elements of $\mathbb{C}[[x]]$ seem to involve infinitely many. But this is an illusion. Remember that x is a formal parameter so that the expression $\sum_n a_n x^n$ is only meant to be a mnemonic device which gives intuition to the definitions of the three algebra operations, especially that of multiplication. We could just as easily have defined $\mathbb{C}[[x]]$ to be the set of all complex vectors $(a_0, a_1, a_2, \ldots)$ subject to the operation of vector addition

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots)$$

and similarly for the other two. What *is* true is that one is only permitted to add or multiply a finite number of elements of $\mathbb{C}[[x]]$. So one can only perform operations which will alter the coefficient of a given power of x a finite number of times.

Now given a sequence of complex numbers $a_0, a_1, a_2, \ldots$, we associate with it the *ordinary generating function*

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots \in \mathbb{C}[[x]].$$

We will sometimes say that this series *counts* the objects enumerated by the a_n if appropriate. As with generating polynomials, the reason for doing so is to exploit properties of $\mathbb{C}[[x]]$ to obtain information about the original sequence. We will often write this generating function as $\sum_n a_n x^n$, assuming the that range of indices is $n \geq 0$.

Let us start with a simple example. Consider the sequence $1, 1, 1, \ldots$ with generating function $\sum_n x^n$. We would like to simplify this as a geometric series to

$$1 + x + x^2 + \cdots = \frac{1}{1-x}. \tag{3.8}$$

But what does the right-hand side even mean since $1/(1-x)$ appears to be a rational function and so not an element of $\mathbb{C}[[x]]$? The way out of this conundrum is to remember that given an element a in an algebra A, it is possible for a to have an inverse, namely an element a^{-1} such that $a \cdot a^{-1} = 1$ where 1 is the identity element of A. So to prove (3.8) in this setting we must show that $\sum_n x^n$ and $1 - x$ are inverses. This is easily done by using the distributive law:

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

This example illustrates a general principle that often well-known results about analytic power series carry over to their formal counterparts, although some work may be required to check that this is true. For the most part, we will assume the truth of a standard formula in this setting without further comment. But it would be wise to also give a couple of examples to show that caution may be needed. One illustration is that the expression $1/x$ has no meaning in $\mathbb{C}[[x]]$ because x does not have an inverse. For suppose we have $xf(x) = 1$ for some formal power series $f(x)$. Then on the left-hand side the constant coefficient is 0 while on the right it is 1, a contradiction.

As another example, consider the sequence $1/n!$ for $n \geq 0$. We would like to write

$$e^x = \sum_{n\geq 0} \frac{x^n}{n!}$$

for the corresponding generating function but, again, run into the problem that e^x is not a priori an element of $\mathbb{C}[[x]]$. The solution this time is to *define* e^x to be a formal symbol which stands for this power series. Then, of course, to be complete we would need to verify formally that all the usual rules of exponents hold such as $e^{2x} = (e^x)^2$. We will not take the time to do this. But we will point out a case where the rules do not hold. In particular, in $\mathbb{C}[[x]]$ one can*not* write

$$e^{1+x} = ee^x.$$

This is because the left-hand side is not well-defined. Indeed, when expanding $\sum_n (1+x)^n/n!$ there are infinitely many additions needed to compute the coefficient of any given power of x which, as we have already noted, is not permitted.

Although we will not verify every specific analytic identity needed for formal power series in this text, it would be good to have some general results about which operations are permitted in $\mathbb{C}[[x]]$. First we deal with the issue of when a formal power series is invertible.

Theorem 3.3.1. *If $f(x) = \sum_n a_n x^n$, then $f(x)^{-1}$ exists in $\mathbb{C}[[x]]$ if and only if $a_0 \neq 0$.*

Proof. For the forward direction, suppose $f(x)g(x) = 1$ where $g(x) = \sum_n b_n x^n$. Taking the constant coefficient on both sides gives $a_0 b_0 = 1$. So $a_0 \neq 0$.

Now assume $a_0 \neq 0$. We will construct an inverse $g(x) = \sum_n b_n x^n$. We want $f(x)g(x) = 1$. Comparing coefficients of x^n on both sides we see that we wish to have $a_0 b_0 = 1$ and, for $n \geq 1$,

$$a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 = 0.$$

Since $a_0 \neq 0$ we can take $b_0 = 1/a_0$. By the same token, when $n \geq 1$ we can solve for b_n in the previous displayed equation giving a recursive formula for its value. Thus we can construct such a $g(x)$ and are done. □

Our example with e^x shows that we also need to be careful about substitution. We wish to define the *substitution* of $g(x)$ into $f(x) = \sum_n a_n x^n$ to be

$$f(g(x)) = \sum_{n\geq 0} a_n g(x)^n.$$

But now the right-hand side is an infinite sum of formal power series, not just formal variables. To be able to talk about such sums, we need to introduce a notion of convergence in $\mathbb{C}[[x]]$.

It will be convenient to have the notation that for a formal power series $f(x)$

$$[x^n]f(x) = \text{the coefficient of } x^n \text{ in } f(x),$$

which we have usually been calling a_n. Suppose that we have a sequence $f_0(x), f_1(x), f_2(x), \ldots$ of formal power series. We say that this sequence *converges* to $f(x) \in \mathbb{C}[[x]]$

and write

$$\lim_{k\to\infty} f_k(x) = f(x),$$

if, for any n, the coefficient of x^n in the sequence is eventually constant and equals the coefficient of x^n in $f(x)$. Formally, given n, there exists a corresponding K such that $[x^n]f_k(x) = [x^n]f(x)$ for all $k \geq K$. Otherwise we say that the sequence *diverges* or that the limit *does not exist*.

As an illustration, consider the sequence

$$f_0(x) = 1,\ f_1(x) = 1 + x,\ f_2(x) = 1 + x + x^2,\ \ldots$$

so that $f_k(x) = 1 + x + \cdots + x^k$. Then this sequence has a limit; namely

$$\lim_{k\to\infty} f_k(x) = \sum_{n\geq 0} x^n = \frac{1}{1-x}.$$

To prove this, note that given n we can let $K = n$. So for $k \geq n$ we have $[x^n]f_k = [x^n]f_n = 1$. On the other hand, consider the sequence

$$f_0(x) = 1 + x,\ f_1(x) = 1/2 + x/2,\ f_2(x) = 1/4 + x/4,\ \ldots$$

and in general $f_k(x) = 1/2^k + x/2^k$. This sequence does *not* converge in $\mathbb{C}[[x]]$ since for any n we have that $[x]f_k(x)$ is always different for different k. This is in contrast to the analytic situation where this sequence converges to zero.

As in analysis, we now use convergence of sequences to define convergence of series. Given $f_0(x), f_1(x), f_2(x), \ldots$, we say that their *sum exists and converges to* $f(x)$, written $\sum_{k\geq 0} f_k(x) = f(x)$, if

$$\lim_{k\to\infty} s_k(x) = f(x)$$

where

$$s_k(x) = f_0(x) + f_1(x) + \cdots + f_k(x) \tag{3.9}$$

is the kth partial sum. Divergence is defined as expected. Note that this definition is consistent with our notation for formal power series since given a sequence $a_0, a_1, a_2, \ldots$, we can let $f_k(x) = a_k x^k$ and then prove that $\sum_{k\geq 0} f_k(x) = f(x)$ where $f(x) = \sum_{k\geq 0} a_k x^k$.

To state a criterion for convergence of series, it will be useful to define the *minimum degree* of $f(x) = \sum_n a_n x^n$ to be

$$\operatorname{mdeg} f(x) = \text{smallest } n \text{ such that } a_n \neq 0$$

if $f(x) \neq 0$, and let $\operatorname{mdeg} f(x) = \infty$ if $f(x) = 0$. It turns out that to show a sum of power series converges, it suffices to take a limit of integers.

Theorem 3.3.2. *Given $f_0(x), f_1(x), f_2(x), \ldots \in \mathbb{C}[[x]]$, then $\sum_{k\geq 0} f_k(x)$ exists if and only if*

$$\lim_{k\to\infty} (\operatorname{mdeg} f_k(x)) = \infty.$$

Proof. We will prove the forward direction, leaving the other implication as an exercise. We are given that the sequence $s_k(x)$ as defined by (3.9) converges. So given n, there is a K such that

$$[x^n]s_K(x) = [x^n]s_{K+1}(x) = [x^n]s_{K+2}(x) = \cdots .$$

But for $j \geq 0$ we have

$$s_{K+j}(x) = s_K(x) + f_{K+1}(x) + f_{K+2}(x) + \cdots + f_{K+j}(x).$$

It follows that $[x^n]f_k(x) = 0$ for $k > K$. Now given n, take N to be the maximum of all the K-values associated to integers less than or equal to n. From what we have shown, this forces $\operatorname{mdeg} f_k(x) > n$ for $n > N$. But by definition of a limit of real numbers, this means $\lim_{k\to\infty}(\operatorname{mdeg} f_k(x)) = \infty$. □

We are now on a firm footing with our definition of substitution as we know what it means for a sum of power series to converge. We can use the previous result to give a simple criterion for convergence when substituting one generating function into another.

Theorem 3.3.3. *Given $f(x), g(x) \in \mathbb{C}[[x]]$, then the composition $f(g(x))$ exists if and only if*

(1) *$f(x)$ is a polynomial or*
(2) *$g(x)$ has zero constant term.*

Proof. If $f(x)$ is a polynomial, then $f(g(x))$ is a finite sum and so obviously converges. So assume $f(x) = \sum_n a_n x^n$ is not polynomial.

If $g(x)$ has no constant term, we have $\operatorname{mdeg} a_n g(x)^n \geq n$. So the limit in the previous theorem is infinity and $f(g(x))$ is well-defined.

To finish the proof, consider the remaining case where $[x^0]g(x) \neq 0$. Since $f(x)$ is not a polynomial, there are an infinite number of n such that $a_n \neq 0$. But for these n we have $\operatorname{mdeg} a_n g(x)^n = 0$. So the desired limit cannot be infinity and $f(g(x))$ does not exist in $\mathbb{C}[[x]]$. □

We will also find it useful to consider certain infinite products. We approach their convergence just as we did for infinite sums. Given a sequence $f_0(x), f_1(x), f_2(x), \ldots,$ we say that their *product exists and converges to* $f(x)$, written $\prod_{k\geq 0} f_k(x) = f(x)$, if

$$\lim_{k\to\infty} p_k(x) = f(x)$$

where

$$p_k(x) = f_0(x)f_1(x)\ldots f_k(x).$$

We have the following result whose proof is similar enough to that of Theorem 3.3.2 that we will leave it to the reader.

Theorem 3.3.4. *Let $f_0(x), f_1(x), f_2(x), \ldots$ be power series with zero constant terms. Then $\prod_{k\geq 0}(1 + f_k(x))$ exists if and only if*

$$\lim_{k\to\infty}(\operatorname{mdeg} f_k(x)) = \infty.$$ □

Let us end this section by showing how the previous result can give simple verifications that a product does or does not exist. Consider $\prod_{k\geq 1}(1+x^k)$. In this case $f_k(x) = x^k$ and $\operatorname{mdeg} x^k = k$. So the desired limit is infinity and this product exists. As we will see in Section 3.5, it counts integer partitions with distinct parts. By contrast, the product $\prod_{k\geq 0}(1+x/2^k)$ does not converge since $\operatorname{mdeg} x/2^k = 1$.

3.4. The Sum and Product Rules for ogfs

Just as for sets there is a Sum Rule and a Product Rule for ordinary generating functions. In order to state these results, we need the idea of a weight-generating function. This approach makes it possible to construct generating functions for various sequences in a very combinatorial manner. As a first application, we make a more deep exploration of the Binomial Theorem.

Let S be a set. Then a *weighting* of S is a function $\operatorname{wt}: S \to \mathbb{C}[[x]]$. Most often if $s \in S$, then $\operatorname{wt} s$ will just be a monomial reflecting some property of s. For example, if st is any statistic on S, then we could take $\operatorname{wt} s = x^{\operatorname{st} s}$. For a more concrete illustration which we will continue to use throughout this section, let $S = 2^{[n]}$ and define for $T \in S$

$$\operatorname{wt} T = x^{|T|}. \tag{3.10}$$

Given a weighted set S, we can form the corresponding *weight-generating function*

$$f(x) = f_S(x) = \sum_{s\in S} \operatorname{wt} s.$$

We must be careful that this sum exists in $\mathbb{C}[[x]]$, and if it does, then we say S is a *summable set*. Of course, when S is finite then it is automatically summable. To illustrate for $S = 2^{[3]}$ we have

T	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$\operatorname{wt} T$	1	x	x	x	x^2	x^2	x^2	x^3

so that

$$f_S(x) = 1 + 3x + 3x^2 + x^3.$$

More generally, for $S = 2^{[n]}$ we have

$$\begin{aligned} f_S(x) &= \sum_{T\in 2^{[n]}} x^{|T|} \\ &= \sum_{k=0}^{n} \sum_{T\in\binom{[n]}{k}} x^k \\ &= \sum_{k=0}^{n} \binom{n}{k} x^k \end{aligned}$$

and we have recovered the generating function for a row of Pascal's triangle.

The following theorem will permit us to manipulate weight-generating functions with ease. For the Sum Rule, if S, T are disjoint weighted sets, then we weight $u \in S \uplus T$

using u's weight in S or in T depending on whether $u \in S$ or $u \in T$, respectively. For arbitrary S, T we weight $S \times T$ by letting

$$\operatorname{wt}(s,t) = \operatorname{wt} s \cdot \operatorname{wt} t.$$

Lemma 3.4.1. *Let S, T be summable sets.*

(a) (*Sum Rule*) *The set $S \cup T$ is summable. If $S \cap T = \emptyset$, then*

$$f_{S \uplus T}(x) = f_S(x) + f_T(x).$$

(b) (*Product Rule*) *The set $S \times T$ is summable and*

$$f_{S \times T}(x) = f_S(x) \cdot f_T(x).$$

Proof. (a) Since S is summable, given any $n \in \mathbb{N}$, there are only a finite number of $s \in S$ such that $\operatorname{wt} s$ has a nonzero coefficient of x^n. And the same is true of T. It follows that only finitely many elements of $S \cup T$ have such a coefficient, which mean this set is summable. To prove the desired equality, we compute as follows:

$$f_{S \uplus T}(x) = \sum_{u \in S \uplus T} \operatorname{wt} u = \sum_{u \in S} \operatorname{wt} u + \sum_{u \in T} \operatorname{wt} u = f_S(x) + f_T(x).$$

(b) The statement about summability of $S \times T$ is safely left as an exercise. Computing the weight-generating function gives

$$f_{S \times T}(x) = \sum_{(s,t) \in S \times T} \operatorname{wt}(s,t) = \sum_{s \in S} \operatorname{wt} s \cdot \sum_{t \in T} \operatorname{wt} t = f_S(x) \cdot f_T(x),$$

so we are done. □

We can now use these rules to derive various generating functions in a straightforward manner. We begin by reproving the Binomial Theorem as stated in Theorem 3.1.1. We have already seen that the summation side is the weight-generating function for $S = 2^{[n]}$. For the product side, it will be useful to reformulate S in terms of multiplicity notation. Specifically, consider

$$S' = \{T' = (1^{m_1}, 2^{m_2}, \ldots, n^{m_n}) \mid m_i = 0 \text{ or } 1 \text{ for all } i\}$$

weighted by

$$\operatorname{wt} T' = x^{\sum_i m_i}.$$

Clearly we have a bijection $f : S \to S'$ given by $f(T) = (1^{m_1}, 2^{m_2}, \ldots, n^{m_n})$ where

$$m_i = \begin{cases} 0 & \text{if } i \notin T, \\ 1 & \text{if } i \in T. \end{cases}$$

(In fact, this is the map used in the proof of Theorem 1.3.1.) Furthermore, this bijection is *weight preserving* in that $\operatorname{wt} f(T) = \operatorname{wt} T$. As a concrete example, if $n = 5$ and $T = \{2, 4, 5\}$, then $f(T) = (1^0, 2^1, 3^0, 4^1, 5^1)$ and $\operatorname{wt} f(T) = x^3 = \operatorname{wt} T$. The advantage of using S' is that it is clearly a weighted product of the sets $\{i^0, i^1\}$ for $i \in [n]$ where $\operatorname{wt} i^0 = 1$ and $\operatorname{wt} i^1 = x$. So we can write, where for distinct elements a and b we use the shorthand $a \uplus b$ for $\{a\} \uplus \{b\}$,

$$S' = \{1^0, 1^1\} \times \{2^0, 2^1\} \times \cdots \times \{n^0, n^1\} = (1^0 \uplus 1^1) \times (2^0 \uplus 2^1) \times \cdots \times (n^0 \uplus n^1).$$

Translating this expression using both parts of Lemma 3.4.1, we obtain

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}x^k &= f_S(x)\\ &= f_{S'}(x)\\ &= (\mathrm{wt}\,1^0+\mathrm{wt}\,1^1)(\mathrm{wt}\,2^0+\mathrm{wt}\,2^1)\cdots(\mathrm{wt}\,n^0+\mathrm{wt}\,n^1)\\ &= (1+x)^n.\end{aligned}$$

Since this was the reader's first example of the use of weight-generating functions, we were careful to write out all the details. However, in practice one is usually more concise, for example, making no distinction between S and S' with the understanding that they yield the same weight-generating function and so can be considered the same set in this situation. We also usually omit checking summability, assuming that such details could be filled in if necessary. We are now ready for a more substantial example, namely the Binomial Theorem for negative exponents.

Theorem 3.4.2. *If $n \in \mathbb{N}$, then*

$$\frac{1}{(1-x)^n} = \sum_{k\ge 0}\left(\!\binom{n}{k}\!\right)x^k.$$

Proof. The summation side suggests that we should consider

$$S = \{T \mid T \text{ is a multiset on } [n]\}$$

with weight function given by (3.10). We are rewarded for our choice since

$$f_S(x) = \sum_{T\in S}\mathrm{wt}\,T = \sum_{k\ge 0}\sum_{T\in\left(\!\binom{[n]}{k}\!\right)} x^k = \sum_{k\ge 0}\left(\!\binom{n}{k}\!\right)x^k.$$

We now write

$$\begin{aligned}S &= \{(1^{m_1}, 2^{m_2}, \ldots, n^{m_n}) \mid m_i \ge 0 \text{ for all } i\}\\ &= (1^0 \uplus 1^1 \uplus 1^2 \uplus \ldots) \times (2^0 \uplus 2^1 \uplus 2^2 \uplus \ldots) \times \cdots \times (n^0 \uplus n^1 \uplus n^2 \uplus \ldots)\end{aligned}$$

with weight function $\mathrm{wt}\, i^k = x^k$. Using Lemma 3.4.1 yields

$$\begin{aligned}f_S(x) &= (\mathrm{wt}\,1^0+\mathrm{wt}\,1^1+\mathrm{wt}\,1^2+\cdots)\cdots(\mathrm{wt}\,n^0+\mathrm{wt}\,n^1+\mathrm{wt}\,n^2+\cdots)\\ &= (1+x+x^2+\cdots)^n\\ &= \frac{1}{(1-x)^n}\end{aligned}$$

and the theorem is proved. □

There are several remarks which should be made about this result. First of all, contrast it with our first version of the Binomial Theorem. In Theorem 3.1.1 we are counting subsets of $[n]$ where repeated elements are not allowed and the resulting generating function is $(1+x)^n$. In Theorem 3.4.2 we are counting multisets on $[n]$ so that repetitions are allowed and these are counted by $1/(1-x)^n$. We will see another example of this in the next section.

We can also make Theorem 3.4.2 look almost exactly like Theorem 3.1.1. Indeed, if $n \leq 0$, then by Theorem 3.4.2 and equation (1.6) (with $-n$ substituted for n) we have

$$(1+x)^n = \frac{1}{(1-(-x))^{-n}} = \sum_{k\geq 0}\left(\!\!\binom{-n}{k}\!\!\right)(-x)^k = \sum_{k\geq 0}\binom{n}{k}x^k.$$

This is exactly like Theorem 3.1.1 except that we have an infinite series whereas for positive n we have a polynomial.

Analytically, the Binomial Theorem makes sense for any $n \in \mathbb{C}$ as long as $|x| < 1$ so that the series converges. In $\mathbb{C}[[x]]$ one can make sense of $(1+x)^n$ for any rational $n \in \mathbb{Q}$; see Exercise 12 of this chapter, and prove the following.

Theorem 3.4.3. *For any $n \in \mathbb{Q}$ we have*

$$(1+x)^n = \sum_{k\geq 0}\binom{n}{k}x^k. \qquad \square$$

3.5. Revisiting integer partitions

The theory of integer partitions is one place where ordinary generating functions have played a central role. In this context and others it will be necessary to consider infinite products. But then, as we have seen in the previous section, we must take care that these products converge. There is a corresponding restriction on the sets which we can use to construct weight-generating functions. We begin by discussing this matter.

Let S be a weighted set. We say that S is *rooted* if there is an element $r \in S$ called the *root* satisfying

(1) $\operatorname{wt} r = 1$ and

(2) if $s \in S - \{r\}$, then $\operatorname{wt} s$ has zero constant term.

For example, the sets $(n^0, n^1, n^2, \dots)$ used in the proof of Theorem 3.4.2 were rooted with $r = n^0$ since $\operatorname{wt} n^0 = 1$ and $\operatorname{wt} n^k = x^k$ for $k \geq 1$. Given a sequence $S_1, S_2, S_3 \dots$ of rooted sets with S_i having root r_i, their *direct sum* is defined to be

$$\begin{aligned} &S_1 \oplus S_2 \oplus S_3 \oplus \cdots \\ &\quad = \{(s_1, s_2, s_3, \dots) \mid s_i \in S_i \text{ for all } i \text{ and } s_i \neq r_i \text{ for only finitely many } i\}. \end{aligned}$$

Note that when the number of S_i is finite, then their direct sum is the same as their product. But when their number is infinite the root condition kicks in. Note that, because of this condition, we have a well-defined weighting on $\oplus_{i\geq 1} S_i$ given by

$$\operatorname{wt}(s_1, s_2, s_3, \dots) = \prod_{i\geq 1} \operatorname{wt} s_i$$

since the product has only finitely many factors not equal to 1. In addition, the Product Rule in Lemma 3.4.1 has to be modified appropriately to get convergence. But the proof is similar to the former result and so is left as an exercise.

Theorem 3.5.1. *Let $S_1, S_2, S_3, \dots$ be a sequence of summable, rooted sets such that*

$$\lim_{i\to\infty}[\operatorname{mdeg}(f_{S_i}(x) - 1)] = \infty.$$

Then the direct sum $S_1 \oplus S_2 \oplus S_3 \oplus \cdots$ *is summable and*

$$f_{S_1 \oplus S_2 \oplus S_3 \oplus \cdots}(x) = \prod_{i \geq 1} f_{S_i}(x). \qquad \square$$

We will now prove a theorem of Euler giving the generating function for $p(n)$, the number of integer partitions of n. The reader should contrast the proof with that given for counting multisets in Theorem 3.4.2, which has evident parallels.

Theorem 3.5.2. *We have*

$$\sum_{n \geq 0} p(n) x^n = \prod_{i \geq 1} \frac{1}{1 - x^i}.$$

Proof. Motivated by the sum side we consider the set S of all integer partitions λ of all numbers $n \geq 0$ with weight

$$\operatorname{wt} \lambda = x^{|\lambda|}, \tag{3.11}$$

recalling that $|\lambda|$ is the sum of the parts of λ. It follows that

$$f_S(x) = \sum_{\lambda \in S} \operatorname{wt} \lambda = \sum_{n \geq 0} \sum_{|\lambda| = n} x^n = \sum_{n \geq 0} p(n) x^n.$$

We now express S as a direct sum, using multiplicity notation, as

$$\begin{aligned} S &= \{(1^{m_1}, 2^{m_2}, 3^{m_3}, \ldots) \mid m_i \geq 0 \text{ for all } i \text{ and only finitely many } m_i \neq 0\} \\ &= (1^0 \uplus 1^1 \uplus 1^2 \uplus \cdots) \oplus (2^0 \uplus 2^1 \uplus 2^2 \uplus \cdots) \oplus (3^0 \uplus 3^1 \uplus 3^2 \uplus \cdots) \oplus \cdots. \end{aligned}$$

Note that since we want the exponent on $\operatorname{wt} \lambda$ to be the sum of its parts, and i^k represents a part i repeated k times, we must take

$$\operatorname{wt} i^k = x^{ik}$$

in contrast to the weight used in the proof of Theorem 3.4.2. Translating into generating functions by using the previous theorem gives

$$\begin{aligned} f_S(x) &= \prod_{i \geq 1} (\operatorname{wt} i^0 + \operatorname{wt} i^1 + \operatorname{wt} i^2 + \operatorname{wt} i^3 + \cdots) \\ &= \prod_{i \geq 1} (1 + x^i + x^{2i} + x^{3i} + \cdots) \\ &= \prod_{i \geq 1} \frac{1}{1 - x^i}, \end{aligned}$$

which is the desired result. $\square$

The reader will notice that the previous proof actually shows much more. In particular, the factor $1/(1 - x^i)$ is responsible for keeping track of the parts equal to i in λ. We can make this precise as follows.

Proposition 3.5.3. *Given $n \in \mathbb{N}$ and $P \subseteq \mathbb{P}$, let $p_P(n)$ be the number of partitions of n all of whose parts are in P.*

(a) *We have*

$$\sum_{n\geq 0} p_P(n)x^n = \prod_{i\in P} \frac{1}{1-x^i}.$$

(b) *In particular, for $k \in \mathbb{P}$,*

$$\sum_{n\geq 0} p_{[k]}(n)x^n = \frac{1}{(1-x)(1-x^2)\cdots(1-x^k)}.$$

Proof. For (a), one uses the ideas in the proof of Theorem 3.5.2 except that the elements of S only contain components of the form r^{m_r} for $r \in P$. And (b) follows immediately from (a). □

Instead of restricting the set of parts of a partition, we can restrict the number of parts. Recall that $p(n,k)$ is the number of $\lambda \vdash n$ with length $\ell(\lambda) \leq k$.

Corollary 3.5.4. *For $k \geq 0$ we have*

$$\sum_{n\geq 0} p(n,k)x^n = \frac{1}{(1-x)(1-x^2)\cdots(1-x^k)}.$$

Proof. From the previous result, it suffices to show that there is a size-preserving bijection between the partitions counted by $p_{[k]}(n)$ and those counted by $p(n,k)$. The map $\lambda \to \lambda^t$ is such a map. Indeed, λ only uses parts in $[k]$ if and only if $\lambda_1 \leq k$. In terms of Young diagrams, this means that the first row of λ has length at most k. It follows that the first column of λ^t has length at most k, which is equivalent to λ^t having at most k parts. □

In the previous section we pointed out a relationship between the generating functions for sets and for multisets. The same holds for integer partitions. Let $p_d(n)$ be the number of partitions of n into distinct parts as defined in Section 2.3.

Theorem 3.5.5. *We have*

$$\sum_{n\geq 0} p_d(n)x^n = \prod_{i\geq 1}(1+x^i).$$

Proof. Up to now we have been writing out most of the gory details of our proofs by weight-generating function as the reader gets familiar with the method. But by now it should be sufficient just to write out the highlights. We begin by letting S be all partitions of all $n \in \mathbb{N}$ into distinct parts and using the weighting in (3.11). It is routine to show that $f_S(x)$ results in the sum side of the theorem. To get the product side we write

$$\begin{aligned} S &= \{(1^{m_1}, 2^{m_2}, 3^{m_3}, \dots) \mid m_i = 0 \text{ or } 1 \text{ for all } i, \text{ only finitely many } m_i \neq 0\} \\ &= \bigoplus_{i\geq 1}(i^0 \uplus i^1). \end{aligned}$$

The generating function translation is

$$f_S(x) = \prod_{i\geq 1}(1 + x^i)$$

and we are done. □

As mentioned in the introduction to this chapter, one of the reasons for using generating functions is that they can give quick and easy proofs for various results. Here is an example where we reprove Euler's distinct parts-odd parts result, Theorem 2.3.3, which we restate here for convenience. Let $p_o(n)$ be the number of partitions of n into parts all of which are odd.

Theorem 3.5.6 (Euler). *For all $n \geq 0$,*

$$p_o(n) = p_d(n).$$

Proof. It suffices to show that these two sequence have the same generating function. Using Theorems 3.5.3(a) and 3.5.5 as well as multiplying by a strange name for one, we get

$$\begin{aligned}
\sum_{n\geq 0} p_d(n)x^n &= (1 + x)(1 + x^2)(1 + x^3)\cdots \\
&= (1 + x)(1 + x^2)(1 + x^3)\cdots \frac{(1 - x)(1 - x^2)(1 - x^3)\cdots}{(1 - x)(1 - x^2)(1 - x^3)\cdots} \\
&= \frac{(1 - x^2)(1 - x^4)(1 - x^6)\cdots}{(1 - x)(1 - x^2)(1 - x^3)\cdots} \\
&= \frac{1}{(1 - x)(1 - x^3)(1 - x^5)\cdots} \\
&= \sum_{n\geq 0} p_o(n)x^n,
\end{aligned}$$

which completes this short and slick proof. □

3.6. Recurrence relations and generating functions

The reader may have noticed that many of the combinatorial sequences described in Chapter 1 satisfy recurrence relations. If one has a sequence defined by a recursion, then generating functions can often be used to find an explicit expression for the terms of the sequence. It is also possible to glean information from the generating function derived from a recurrence which is hard to extract from the recurrence itself. This section is devoted to exploring these ideas.

We start with a simple algorithm using generating functions to solve a recurrence relation. Given a sequence $a_0, a_1, a_2, \ldots$ defined by a recursion and boundary conditions, we wish to find a self-contained formula for the nth term.

(1) Multiply the recurrence by x^n; usually a good choice for n is the largest index of all the terms in the recurrence. Sum over all $n \geq d$ where d is the smallest index for which the recurrence is valid.

(2) Let
$$f(x) = \sum_{n\geq 0} a_n x^n$$
and express the equation in step (1) in terms of $f(x)$ using the boundary conditions.

(3) Solve for $f(x)$.

(4) Find a_n as the coefficient of x^n in $f(x)$.

We note that partial fraction expansion can be a useful way to accomplish step (4).

For a simple example, suppose that our sequence is defined by $a_0 = 2$ and $a_n = 3a_{n-1}$ for $n \geq 1$. Calculating the first few values we get $a_1 = 2\cdot 3$, $a_2 = 2\cdot 3^2$, $a_3 = 2\cdot 3^3$. So it is easy to guess and then prove by induction that $a_n = 2 \cdot 3^n$. We would now like to obtain this result using generating functions. Step (1) is easy as we just write $a_n x^n = 3a_{n-1}x^n$ and then sum to get
$$\sum_{n\geq 1} a_n x^n = \sum_{n\geq 1} 3a_{n-1}x^n.$$
Letting $f(x)$ be as in step (2) we see that
$$\sum_{n\geq 1} a_n x^n = f(x) - a_0 = f(x) - 2$$
and
$$\sum_{n\geq 1} 3a_{n-1}x^n = 3x\sum_{n\geq 1} a_{n-1}x^{n-1} = 3xf(x)$$
where the last equality is obtained by substituting n for $n-1$ in the sum. For step (3) we have

$$f(x) - 2 = 3xf(x) \implies f(x) - 3xf(x) = 2 \implies f(x) = \frac{2}{1-3x}. \tag{3.12}$$

As far as step (4), we can now expand $1/(1-3x)$ as a geometric series (that is, use (3.8) and substitute $3x$ for x) to obtain
$$f(x) = 2\sum_{n\geq 0} 3^n x^n = \sum_{n\geq 0} 2\cdot 3^n x^n.$$
Extracting the coefficient of x^n we see that $a_n = 2 \cdot 3^n$ as expected.

In the previous example, it was easier to guess the formula for a_n and then prove it by induction rather than use generating functions. However, there are times when it is impossible to guess the solution this way, but generating functions still give a straightforward method for obtaining the answer. An example of this is given by the Fibonacci sequence. Our result will be slightly nicer if we use the definition of this sequence given by (1.1). Following the algorithm, we write
$$\sum_{n\geq 2} F_n x^n = \sum_{n\geq 2}(F_{n-1} + F_{n-2})x^n.$$

Writing $f(x) = \sum_{n\geq 0} F_n x^n$ we obtain

$$\sum_{n\geq 2} F_n x^n = f(x) - F_0 - F_1 x = f(x) - x$$

and

$$\sum_{n\geq 2} (F_{n-1} + F_{n-2}) x^n = x(f(x) - F_0) + x^2 f(x) = (x + x^2) f(x).$$

Setting the expressions for the left and right sides equal and solving for $f(x)$ yields

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

For the last step we wish to use partial fractions and so must factor $1 - x - x^2$. Using the quadratic formula, we see that the denominator has roots

$$r_1 = \frac{-1 + \sqrt{5}}{2} \quad \text{and} \quad r_2 = \frac{-1 - \sqrt{5}}{2}.$$

It follows that

$$1 - x - x^2 = \left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right)$$

since both sides vanish at $x = r_1, r_2$ and both sides have constant term 1. So we have the partial fraction decomposition

$$f(x) = \frac{x}{\left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right)} = \frac{A}{\left(1 - \frac{x}{r_1}\right)} + \frac{B}{\left(1 - \frac{x}{r_2}\right)} \tag{3.13}$$

for constants A, B. Clearing denominators gives

$$x = A\left(1 - \frac{x}{r_2}\right) + B\left(1 - \frac{x}{r_1}\right).$$

Setting $x = r_1$ reduces this equation to $r_1 = A(1 - r_1/r_2)$ and solving for A shows that $A = 1/\sqrt{5}$. Similarly letting $x = r_2$ yields $B = -1/\sqrt{5}$. Plugging these values back into (3.13) and expanding the series,

$$f(x) = \frac{1}{\sqrt{5}} \cdot \sum_{n\geq 0} \frac{x^n}{r_1^n} - \frac{1}{\sqrt{5}} \cdot \sum_{n\geq 0} \frac{x^n}{r_2^n}.$$

By rationalizing denominators one can check that $1/r_1 = (1 + \sqrt{5})/2$ and $1/r_2 = (1 - \sqrt{5})/2$. So taking the coefficient of x^n on both sides of the previous displayed equation gives

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1 - \sqrt{5}}{2}\right)^n. \tag{3.14}$$

This example shows the true power of the generating function method. It would be impossible to guess the formula in (3.14) from just computing values of F_n. In fact, it is not even obvious that the right-hand side is an integer!

Our algorithm can be used to derive generating functions in the case where one has a triangle of numbers rather than just a sequence. Here we illustrate this using the

Stirling numbers. Recall that the signless Stirling numbers of the first kind satisfy the recurrence relation and boundary conditions in Theorem 1.5.2. Translating these to the signed version gives $s(0, k) = \delta_{0,k}$ and

$$s(n, k) = s(n-1, k-1) - (n-1)s(n-1, k)$$

for $n \geq 1$. We wish to find the generating function $f_n(x) = \sum_k s(n, k)x^k$ where we are using the fact that $s(n, k) = 0$ for $k < 0$ or $k > n$ to sum over all integers n. Applying our algorithm we have

$$\begin{aligned} f_n(x) &= \sum_k s(n,k)x^k \\ &= \sum_k [s(n-1,k-1) - (n-1)s(n-1,k)]x^k \\ &= xf_{n-1}(x) - (n-1)f_{n-1}(x) \\ &= (x-n+1)f_{n-1}(x), \end{aligned}$$

giving us a recursion for the sequence of generating functions $f_n(x)$. From the boundary condition for $s(0, k)$ we have $f_0(x) = 1$. It is now easy to guess a formula for $f_n(x)$ by writing out the first few values and proving that pattern holds by induction to obtain the theorem below, which also follows easily from Theorem 3.1.2.

Theorem 3.6.1. *For $n \geq 0$ we have*

$$\sum_k s(n,k)x^k = x(x-1)\cdots(x-n+1). \qquad \square$$

In an entirely analogous manner, one can obtain a generating function for the Stirling numbers of the second kind. Because of the similarity, the proof is left to the reader.

Theorem 3.6.2. *For $k \geq 0$ we have*

$$\sum_n S(n,k)x^n = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}. \qquad \square$$

Comparing the previous two results, the reader will note a similar relationship as between generating functions for objects without repetitions (sets, distinct partitions) and those where repetitions are allowed (multisets, ordinary partitions). As already mentioned, this will be explained in Section 3.9.

So far, all the generating functions we have derived from recurrences have been rational functions. This is because the recursions are linear and we will prove a general result to this effect in the next section. We will end this section by illustrating that more complicated generating functions, for example algebraic ones, do arise in practice. Let us consider the Catalan numbers $C(n)$ and the generating function $c(x) = \sum_{n\geq 0} C(n)x^n$. Using the recursion and boundary condition in Theorem 1.11.2 and computing in the way we have become accustomed to, we obtain

$$c(x) = 1 + \sum_{n\geq 1} C(n)x^n = 1 + \sum_{n\geq 1}\left(\sum_{i+j=n-1} C(i)C(j)\right)x^n = 1 + xc(x)^2.$$

Writing $xc(x)^2 - c(x) + 1 = 0$ and solving for $c(x)$ using the quadratic formula yields

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

Two things seem to be wrong with this formula for $c(x)$. First of all, we don't know whether the plus or minus solution is the correct one. And second, we seem to have left the ring of formal power series because we are dividing by x which has no inverse. Both of these can be solved simultaneously by choosing the sign so that the numerator has no constant term. Then one can divide by x simply by reducing the power of each term in the top by one. By Theorem 3.4.3 we see that the generating function for $\sqrt{1-4x} = (1-4x)^{1/2}$ has constant term $\binom{1/2}{0} = 1$. So the correct sign is negative and we have proved the following.

Theorem 3.6.3. *We have*

$$\sum_{n\geq 0} C(n)x^n = \frac{1-\sqrt{1-4x}}{2x}. \qquad \square$$

One can use this generating function to rederive the explicit expression for $C(n)$ in Theorem 1.11.3, and the reader will be asked to carry out the details in the exercises.

3.7. Rational generating functions and linear recursions

The reader may have noticed in the previous section that, both in the initial example and for the Fibonacci sequence, the solution of the recursion for a_n was a linear combination of functions of the form r^n where r varied over the reciprocals of the roots of the denominator of the corresponding generating function. This happens for a wide variety of recursions which we will study in this section. Before giving a theorem which characterizes this situation, we will study one more example to illustrate what can happen.

Consider the sequence defined by $a_0 = 1, a_1 = -4$, and

$$a_n = 4a_{n-1} - 4a_{n-2} \quad \text{for } n \geq 2. \tag{3.15}$$

Following the usual four-step program, we have, for $f(x) = \sum_{n\geq 0} a_n x^n$,

$$\begin{aligned} f(x) - 1 + 4x &= \sum_{n\geq 2} a_n x^n \\ &= 4x \sum_{n\geq 2} a_{n-1}x^{n-1} - 4x^2 \sum_{n\geq 2} a_{n-2}x^{n-2} \\ &= 4x(f(x)-1) - 4x^2 f(x). \end{aligned}$$

Solving for $f(x)$ and evaluating the constants in the partial fraction expansion yields

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

Taking the coefficient of x^n on both sides using the Theorem 3.4.2 (interchanging the roles of n and k) together with the fact that

$$\left(\!\!\binom{k}{n}\!\!\right) = \binom{n+k-1}{n} = \binom{n+k-1}{k-1} \tag{3.16}$$

gives a final answer of

$$a_n = 4 \cdot 2^n - 3\binom{n+1}{1}2^n = (1-3n)2^n. \tag{3.17}$$

So now, instead of a constant times r^n we have a polynomial in n as the coefficient. And that polynomial has degree less than the multiplicity of $1/r$ as a root of the denominator. These observations generalize.

Consider a sequence of complex numbers a_n for $n \geq 0$. We say that the sequence satisfies a (*homogeneous*) *linear recursion of degree d with constant coefficients* if there is a $d \in \mathbb{P}$ and constants $c_1, \ldots, c_d \in \mathbb{C}$ with $c_d \neq 0$ such that

$$a_{n+d} + c_1 a_{n+d-1} + c_2 a_{n+d-2} + \cdots + c_d a_n = 0. \tag{3.18}$$

To simplify things later, we have put all the terms of the recursion on the left-hand side of the equation and made a_{n+d} the term of highest index rather than a_n. One can also consider the nonhomogeneous case where one has a summand c_{d+1} which does not multiply any term of the sequence, but we will have no cause to do so here. It turns out that the sequences satisfying a recursion (3.18) are exactly the ones having rational generating functions.

Theorem 3.7.1. *Given a sequence a_n for $n \geq 0$ and $d \in \mathbb{P}$, the following are equivalent.*

(a) *The sequence satisfies* (3.18).

(b) *The generating function $f(x) = \sum_{n\geq 0} a_n x^n$ has the form*

$$f(x) = \frac{p(x)}{q(x)} \tag{3.19}$$

where

$$q(x) = 1 + c_1 x + c_2 x^2 + \cdots + c_d x^d \tag{3.20}$$

and $\deg p(x) < d$.

(c) *We can write*

$$a_n = \sum_{i=1}^{k} p_i(n) r_i^n$$

where the r_i are distinct, nonzero complex numbers satisfying

$$1 + c_1 x + c_2 x^2 + \cdots + c_d x^d = \prod_{i=1}^{k} (1 - r_i x)^{d_i} \tag{3.21}$$

and $p_i(n)$ is a polynomial with $\deg p_i(n) < d_i$ *for all i.*

Proof. We first prove the equivalence of (a) and (b). Showing that (a) implies (b) is essentially an application of our algorithm. Multiplying (3.18) by x^{n+d} and summing over $n \geq 0$ gives

$$\begin{aligned} 0 &= \sum_{n\geq 0} a_{n+d}x^{n+d} + c_1 x \sum_{n\geq 0} a_{n+d-1}x^{n+d-1} + \cdots + c_d x^d \sum_{n\geq 0} a_n x^n \\ &= \left[f(x) - \sum_{n=0}^{d-1} a_n x^n \right] + c_1 x \left[f(x) - \sum_{n=0}^{d-2} a_n x^n \right] + \cdots + c_d x^d f(x) \\ &= q(x)f(x) - p(x) \end{aligned}$$

where $q(x)$ is given by (3.20) and $p(x)$ is the sum of the remaining terms, which implies $\deg p(x) < d$. Solving for $f(x)$ completes this direction.

To prove (b) implies (a), cross multiply (3.19) and use (3.20) to write

$$p(x) = q(x)f(x) = (1 + c_1 x + c_2 x^2 + \cdots + c_d x^d)f(x).$$

Since $\deg p(x) < d$ we have that $[x^{n+d}]p(x) = 0$ for all $n \geq 0$. So taking the coefficient of x^{n+d} on both sides of the previous displayed equation gives the recursion (3.18).

We now show that (b) and (c) are equivalent. The fact that (b) implies (c) again follows from the algorithm. Specifically, using equations (3.19), (3.20), and (3.21), as well as partial fraction expansion, we have

$$f(x) = \frac{p(x)}{\prod_{i=1}^{k}(1 - r_i x)^{d_i}} = \sum_{i=1}^{k}\sum_{j=1}^{d_i} \frac{A_{i,j}}{(1 - r_i x)^j} \tag{3.22}$$

for certain constants $A_{i,j}$. But by Theorem 3.4.2 and equation (3.16) we have that

$$[x^n]\frac{1}{(1 - r_i x)^j} = \left(\!\!\binom{j}{n}\!\!\right) r_i^n = \binom{n+j-1}{j-1} r_i^n$$

where

$$\binom{n+j-1}{j-1} = \frac{(n+j-1)(n+j-2)\cdots(n+1)}{(j-1)!}$$

is a polynomial in n of degree $j-1$ for any given j. Now taking the coefficient of x^n on both sides of (3.22) gives

$$a_n = \sum_{i=1}^{k} \left[\sum_{j=1}^{d_i} A_{i,j} \binom{n+j-1}{j-1} \right] r_i^n.$$

Calling the polynomial inside the brackets $p_i(n)$, we have derived the desired expansion.

The proof that (c) implies (b) essentially reverses the steps of the forward direction. So it is left as an exercise. □

We note that the preceding theorem is not just of theoretical significance but is also very useful computationally. In particular, because of the equivalence of (a) and (c), one can solve a linear, constant coefficient recursion in a more direct manner without having to deal with generating functions. To illustrate, suppose we have a sequence satisfying (3.18). Then we know that the solution in (c) is in terms of the r_i which are the reciprocals of the roots of $q(x)$ as given by (3.20). To simplify things, we consider the polynomial

$$r(x) = x^d q(1/x) = x^d + c_1 x^{d-1} + c_2 x^{d-2} + \cdots + c_d.$$

Comparison with (3.21) shows that the r_i are the roots of $r(x)$. We now find the $p_i(n)$ by solving for the coefficients of these polynomials using the initial conditions. To be quite concrete, consider again the example (3.15) with which we began this section. Since $a_n - 4a_{n-1} + 4a_{n-2} = 0$ for $n \geq 2$ we factor $r(x) = x^2 - 4x + 4 = (x-2)^2$. So $a_n = p(n)2^n$ where $\deg p(n) < 2$. It follows that $p(n) = A + Bn$ for constants A, B. Plugging in $n = 0$ we get $1 = a_0 = A2^0 = A$. Now letting $n = 1$ gives $-4 = a_1 = (1+B)2^1$ or $B = -3$. Thus a_n is again given as in (3.17). But this solution is clearly simpler than the first one given. This is called the *method of undetermined coefficients*. Of course, the advantage of using generating functions is that they can be used to solve recursions even when they are not linear and constant coefficient.

There is a striking resemblance between the theory we have developed in this section and the method of undetermined coefficients for solving linear differential equations with constant coefficients. This is not an accident and the material in this section may be considered as part of the theory of finite differences which is a discrete analogue of the theory of differential equations. We will have more to say about finite differences when we study Möbius inversion in Section 5.5.

3.8. Chromatic polynomials

Sometimes generating functions or polynomials appear in unexpected ways. We now illustrate this phenomenon using the chromatic polynomial of a graph.

Let $G = (V, E)$ be a graph. A (*vertex*) *coloring* of G from a set S is a function $c : V \to S$. We refer to S as the *color set*. Figure 3.2 contains a graph which we will be using as our running example together with two colorings using the set $S = \{\text{white, gray, black}\}$. We say that c is *proper* if, for all edges $uv \in E$ we have $c(u) \neq c(v)$. The first coloring in Figure 3.2 is proper while the second is not since the edge vx has both endpoints colored gray. The *chromatic number* of G, denoted $\chi(G)$, is the minimum cardinality of a set S such that there is a proper coloring $c : V \to S$. In our example, $\chi(G) = 3$ because we have displayed a proper coloring with three colors in Figure 3.2 (black, white, and gray), and one cannot use fewer colors because of the triangle uvx.

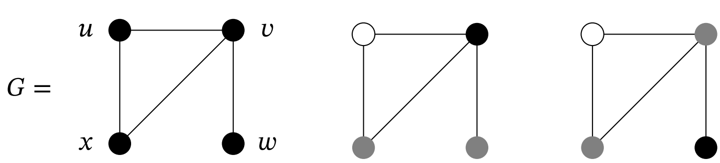

Figure 3.2. A graph and two colorings

The chromatic number is an important invariant in graph theory. But by its definition, it belongs more to extremal combinatorics (which studies structures which minimize or maximize a constraint) than the enumerative side of the subject. Although we will not have much more to say about $\chi(G)$ here, we would be remiss if we did not state one of the most famous mathematical theorems in which it plays a part. Call a graph *planar* if it can be drawn in the plane without any pair of edges crossing.

Theorem 3.8.1 (The Four Color Theorem). *If G is planar, then*

$$\chi(G) \le 4. \qquad \square$$

Note that this result is in stark contrast to ordinary graphs which can have arbitrarily large chromatic number. Complete graphs, for example, have $\chi(K_n) = n$. The Four Color Theorem caused quite a stir when it was proved in 1977 by Appel and Haken (with the help of Koch) [**1**,**2**]. For one thing, it had been the Four Color Conjecture for over 100 years. Also their proof was the first to make heavy use of computers to do the calculations for all the various cases and the demonstration could not be completely checked by a human.

We now turn to the enumerating graph colorings. Let $t \in \mathbb{N}$. The *chromatic polynomial* of G is defined to be

$$P(G;t) = \text{the number of proper colorings } c\colon\ V \to [t].$$

This concept was introduced by George Birkhoff [**13**]. It is not clear at this point why $P(G;t)$ should be called a polynomial, but let us compute it for the graph in Figure 3.2. Consider coloring the vertices of G in the order u, v, w, x. There are t choices for the color of u. This leaves $t-1$ possibilities for v since it cannot be the same color as u. By the same token, the number of choices for w is $t-1$. Finally, x can be colored in $t-2$ ways since it cannot have the colors of u or v and these are different. So the final count is

$$P(G) = P(G;t) = t(t-1)(t-1)(t-2) = t^4 - 4t^3 + 5t^2 - 2t. \tag{3.23}$$

This is a polynomial in t, the number of colors! Before proving that this is always the case, we have a couple of remarks. First of all, there is a close relationship between $P(G;t)$ and $\chi(G)$; namely $P(G;t) = 0$ if $0 \le t < \chi(G)$ but $P(G;\chi(G)) > 0$. This follows from the definitions of P and χ since the latter is the smallest nonnegative integer for which proper colorings of G exist and the former counts such colorings. Secondly, it is not always possible to compute $P(G;t)$ in the manner above and express it as a product of factors $t-k$ for integers k. For example, consider the cycle C_4 with vertices labeled clockwise as u, v, w, x. If we try to use this method to compute $P(C_4;t)$, then everything is fine until we get to coloring x, for x is adjacent to both u and w. But we cannot be sure whether u and w have the same color or not since they themselves are not adjacent.

It turns out that the same ideas can be used both for proving that $P(G;t)$ is always a polynomial in t and to rectify the difficulty in computing $P(C_4;t)$. Consider a graph $G = (V,E)$ and an edge $e \in E$. The graph obtained by *deleting* e from G is denoted $G \setminus e$ and has vertices V and edges $E - \{e\}$. The middle graph in Figure 3.3 is obtained from our running example by deleting $e = vx$. The graph obtained by *contracting* e in G is

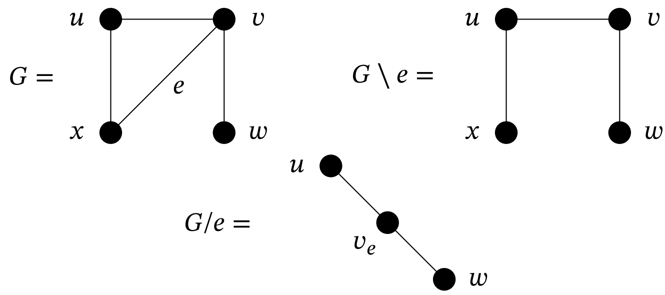

Figure 3.3. Deletion and contraction

denoted G/e and is obtained by shrinking e to a new vertex v_e, making v_e adjacent to the vertices which were adjacent to either endpoint of e and leaving all other vertices and edges of G the same. Contracting vx in our example graph results in the graph on the right in Figure 3.3. The next lemma is crucial in the study of $P(G;t)$. It is ideally set up for induction on $\#E$ since both $G \setminus e$ and G/e have fewer edges than G.

Lemma 3.8.2 (Deletion-Contraction Lemma)**.** *If G is a graph, then for any $e \in E$ we have*

$$P(G;t) = P(G \setminus e;t) - P(G/e;t).$$

Proof. We will prove this in the form $P(G \setminus e) = P(G) + P(G/e)$. Suppose $e = uv$. Since e is no longer present in $G \setminus e$, its proper colorings are of two types: those where $c(u) \neq c(v)$ and those where $c(u) = c(v)$. If $c(u) \neq c(v)$, then properly coloring $G \setminus e$ is the same as properly coloring G. So there are $P(G)$ colorings of the first type. There is also a bijection between the proper colorings of $G \setminus e$ where $c(u) = c(v)$ and those of G/e; namely color v_e with the common color of u and v and leave all the other colors the same. It follows that there are $P(G/e)$ colorings of the second type and the lemma is proved. □

We can now easily show that $P(G;t)$ lives up to its name.

Theorem 3.8.3. *We have that $P(G;t)$ is a polynomial in t for any graph G.*

Proof. We induct on $\#E$. If G has no edges, then clearly $P(G;t) = t^{\#V}$ which is a polynomial in t. If $\#E \geq 1$, then pick $e \in E$. By deletion-contraction $P(G;t) = P(G \setminus e;t) - P(G/e;t)$. And by induction we have that both $P(G \setminus e;t)$ and $P(G/e;t)$ are polynomials in t. Thus the same is true of their difference. □

We can also use Lemma 3.8.2 to compute the chromatic polynomial of C_4. Recall our notation of P_n and K_n for paths and complete graphs on n vertices, respectively. Now picking any edge $e \in E(C_4)$ we can use deletion-contraction and then determine the polynomials of the resulting graphs via coloring vertex by vertex to obtain

$$P(C_4) = P(P_4) - P(K_3) = t(t-1)^3 - t(t-1)(t-2) = t(t-1)(t^2 - 3t + 3).$$

Note that the quadratic factor has complex roots, thus substantiating our claim that $P(G)$ does not always have roots which are integers.

One can use induction and Lemma 3.8.2 to prove a host of results about $P(G;t)$. Since these demonstrations are all similar, we will leave them to the reader. We will use a nonstandard way of writing down the coefficients of this polynomial, which will turn out to be convenient later.

Theorem 3.8.4. *Let* $G = (V, E)$ *and write*

$$P(G;t) = a_0 t^n - a_1 t^{n-1} + a_2 t^{n-2} - \cdots + (-1)^n a_n.$$

(a) $n = \#V$.

(b) $\operatorname{mdeg} P(G;t) =$ *the number of components of* G.

(c) $a_i \geq 0$ *for all* i *and* $a_i > 0$ *for* $0 \leq i \leq n - \operatorname{mdeg} P(G;t)$.

(d) $a_0 = 1$ *and* $a_1 = \#E$. □

Now that we know $P(G;t)$ is a polynomial we can ask if there is any combinatorial interpretation for its coefficients, the reverse of our approach up to now, which has been to start with a sequence and then find its generating function. Put a total order on the edge set E, writing $e < f$ if e is less than f in this order and similarly for other notation. If C is a cycle in G, then the corresponding *broken circuit* B is the set of edges obtained from $E(C)$ by removing the smallest edge in the total order. Returning to the graph in Figure 3.2, let $b = uv$, $c = ux$, $d = vw$, $e = vx$ and impose the order $b < c < d < e$. The only cycle has edges b, c, e and the corresponding broken circuit is $B = \{c, e\}$ which are the edges of a path. Say that a set of edges $A \subseteq E$ *contains no broken circuit* or is an *NBC set* if $A \not\supseteq B$ for any broken circuit B. Let

$$\mathrm{NBC}_k = \mathrm{NBC}_k(G) = \{A \subseteq E \mid \#A = k \text{ and } A \text{ is an NBC set}\}$$

and $\mathrm{nbc}_k = \mathrm{nbc}_k(G) = \#\,\mathrm{NBC}_k(G)$. In our example graph

k	$\mathrm{NBC}_k(G)$	$\mathrm{nbc}_k(G)$
0	$\{\emptyset\}$	1
1	$\{\{b\}, \{c\}, \{d\}, \{e\}\}$	4
2	$\{\{b,c\}, \{b,d\}, \{b,e\}, \{c,d\}, \{d,e\}\}$	5
3	$\{\{b,c,d\}, \{b,d,e\}\}$	2
4	$\emptyset$	0

Comparison of the last column of this table with the coefficients of $P(G;t)$ presages our next result, which is due to Whitney [**99**]. It is surprising that the conclusion does not depend on the total order given to the edges. The proof we give is based on the demonstration of Blass and Sagan [**16**].

Theorem 3.8.5. *If* $\#V = n$, *then, given any ordering of* E,

$$P(G;t) = \sum_{k=0}^{n} (-1)^k \,\mathrm{nbc}_k(G)\, t^{n-k}.$$

Proof. Identify each $A \in \mathrm{NBC}_k(G)$ with the associated spanning subgraph. Then A is acyclic since any cycle contains a broken circuit. It follows from Theorem 1.10.2 that

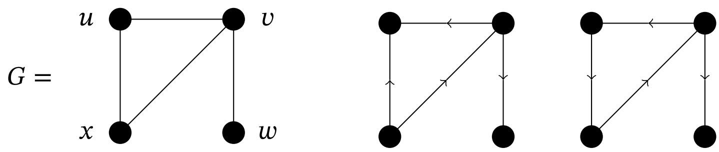

Figure 3.4. Two orientations of a graph

A is a forest with $n - k$ component trees. Hence $\mathrm{nbc}_k(G)\, t^{n-k}$ is the number of pairs (A, c) where $A \in \mathrm{NBC}_k(G)$ and $c\colon V \to [t]$ is a coloring constant on each component of A. We call such a coloring *A-improper*. Make the set of such pairs into a signed set by letting $\mathrm{sgn}(A, c) = (-1)^{\#A}$. So the theorem will be proved if we exhibit a sign-reversion involution ι on these pairs whose fixed points have positive sign and are in bijection with the proper coloring of G.

Define the fixed points of ι to be the (A, c) such that $A = \emptyset$ and c is proper. These pairs clearly have the desired characteristics. For any other pair, c is not a proper coloring so there must be an edge $e = uv$ with $c(u) = c(v)$. Let e be the smallest such edge in the total order. Now define $\iota(A, c) = (A \Delta \{e\}, c) := (A', c)$. It is clear that ι reverses signs. And it is an involution because c does not change, and so the smallest monochromatic edge is the same in a pair and its image. We just need to check that ι is well-defined. If $A' = A - \{e\}$, then obviously A is still an NBC set and c is A'-improper. If $A' = A \cup \{e\}$, then, since e joined two vertices of the same color, c is still A'-improper. But assume, towards a contradiction, that A' is no longer NBC. Then $A' \supseteq B$ where B is a broken circuit, and $e \in B$ since A is NBC. Since c is A'-improper, all edges in B have vertices colored $c(u)$. But e is the smallest edge having that property, and so the smaller edge removed from a cycle to get B cannot exist. Thus A' is NBC, ι is well-defined, and we are done with the proof. □

One of the amazing things about the chromatic polynomial is that it often appears in places where a priori it has no business being because no graph coloring is involved. We now give two illustrations of this. Recall from Section 2.6 that an *orientation O* of a graph G is a digraph with the same vertex set obtained by replacing each edge uv of G by one of the possible arcs $\overrightarrow{uv}$ or $\overrightarrow{vu}$. See Figure 3.4 for two orientations of our usual graph. Call O *acyclic* if it does not contain any directed cycles and let $\mathcal{A}(G)$ and $a(G)$ denote the set and number of acyclic orientations of G, respectively. The first of the orientations just given is acyclic while the second is not. The total number of orientations of the cycle u, v, x, u is 2^3 and the number of those which produce a cycle is 2 (clockwise and counterclockwise). Since neither orientation of vw can produce a cycle, we see that $a(G) = 2(2^3 - 2) = 12$. We now do something very strange and plug $t = -1$ into the chromatic polynomial (3.23) and get $P(G; -1) = (-1)(-2)(-2)(-3) = 12$. Although it is *not* at all clear what it means to color a graph with -1 colors, we have just seen an example of the following celebrated theorem of Stanley [**85**].

Theorem 3.8.6. *For any graph G with $\#V = n$ we have*

$$P(G; -1) = (-1)^n\, a(G).$$

Proof. We induct on $\#E$ where the base case is an easy check. It suffices to show that $(-1)^n P(G;-1)$ and $a(G)$ satisfy the same recursion. Using the Deletion-Contraction Lemma for the former, we see that we need to show $a(G) = a(G \setminus e) + a(G/e)$ for a fixed $e = uv \in E$. Consider the map

$$\phi : A(G) \to A(G \setminus e)$$

which sends $O \mapsto O'$ where O' is obtained from O by removing the arc corresponding to e. Clearly O' is still acyclic so the function is well-defined.

We claim ϕ is onto. Suppose to the contrary that there is some $O' \in A(G \setminus e)$ such that adding back $\vec{uv}$ creates a directed cycle C, and similarly with $\vec{vu}$ creating a cycle C'. Then $(C - \vec{uv}) \cup (C' - \vec{vu})$ is a closed, directed walk which must contain a directed cycle by Exercise 14(b) of Chapter 1. But this third directed cycle is contained in O', which is the desired contradiction.

If $O' \in A(G \setminus e)$, then by definition of the map $\#\phi^{-1}(O') \leq 2$. And from the previous paragraph $\#\phi^{-1}(O') \geq 1$. So $a(G) = x + 2y$ where $x = \#\{O' \mid \phi^{-1}(O') = 1\}$ and $y = \#\{O' \mid \phi^{-1}(O') = 2\}$. Since $a(G \setminus e) = x + y$ it suffices to show that $a(G/e) = y$. We will do this by constructing a bijection

$$\psi : \{O' \in A(G \setminus e) \mid \phi^{-1}(O') = 2\} \to A(G/e).$$

Let Y be the domain of ψ. If there are a pair of edges $wu, wv \in E(G)$, then any $O' \in Y$ contains either both $\vec{wu}$ and $\vec{wv}$ or both $\vec{uw}$ and $\vec{vw}$. This is because in all other cases adding back one of the orientations of e would create an orientation of G with a cycle, contradicting the fact that $\phi^{-1}(O') = 2$. So define $O'' = \psi(O')$ to be the orientation of G/e which agrees with O' on all arcs not containing the new vertex v_e, and on any edge of the form wv_e uses the same orientation as either wu or wv. (As just shown, these two orientations are either both towards or both away from e). Proving that ψ is a well-defined bijection is left as an exercise. □

We should mention that Stanley actually proved a stronger result giving a combinatorial interpretation to $P(G;-t)$ for all negative integers $-t$. See Exercise 28(b) for details. So, as we saw with the binomial coefficients in (1.6), we have another instance of combinatorial reciprocity. We will study this phenomenon more generally in the next section.

For our second example of the protean nature of the chromatic polynomial, we will need to assume that our graphs G have vertex set $[n]$ so that one has a total order on the vertices. Let F be a spanning forest of G and root each component tree T of F at its smallest vertex r. Say that F is *increasing* if the integers on any path starting at r form an increasing sequence for all roots r. In Figure 3.5 the reader will find the usual

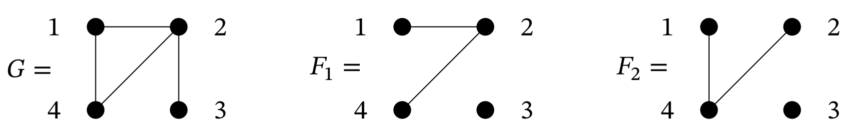

Figure 3.5. A graph and two spanning forests

graph, now labeled with [4], and two spanning forests. We have that F_1 is increasing as any singleton node is increasing, and in the nontrivial tree the only maximal path from the root is $1, 2, 4$, which is an increasing sequence. On the other hand F_2 is not increasing because of the path $1, 4, 2$.

For a graph $G = (V, E)$ we define

$$\mathrm{ISF}_m(G) = \{F \mid F \text{ is an increasing spanning forest of } G \text{ with } m \text{ edges}\}$$

and $\mathrm{isf}_m(G) = \#\,\mathrm{ISF}_m(G)$. If $\#V = n$, then consider the corresponding generating polynomial

$$\mathrm{isf}(G) = \mathrm{isf}(G;t) = \sum_{m=0}^{n} (-1)^m \,\mathrm{isf}_m(G) t^{n-m}.$$

Let us compute this for our example graph. Any tree with zero or one edge is increasing so that $\mathrm{isf}_0(G) = 1$ and $\mathrm{isf}_1(G) = \#E = 4$. Any of the pairs of edges of G form an increasing forest except for the pair giving F_2 in Figure 3.5. So $\mathrm{isf}_2(G) = \binom{4}{2} - 1 = 5$. Similarly one checks that $\mathrm{isf}_3(G) = 2$. And $\mathrm{isf}_4(G) = 0$ since G itself is not a forest. So

$$\mathrm{isf}(G;t) = t^4 - 4t^3 + 5t^2 - 2t = t(t-1)^2(t-2) = P(G;t).$$

We cannot always have $\mathrm{isf}(G) = P(G)$ because the former depends on the labeling of the vertices (even though our notation conceals that fact) while the latter does not. So we will put aside deciding when they are equal for now and concentrate on the factorization over $\mathbb{Z}$ which we have just seen and which, as we will see, is not a coincidence. In fact, the roots will be the cardinalities of the edge sets defined by

$$E_k = \{ik \in E \mid i < k\} \tag{3.24}$$

for $1 \le k \le n$. In our example $E_1 = \emptyset$ (since there is no vertex smaller than 1), $E_2 = \{12\}$, $E_3 = \{23\}$, and $E_4 = \{14, 24\}$.

Lemma 3.8.7. *If G has $V = [n]$, then a spanning subgraph F is an increasing forest if and only if it is obtained by picking at most one edge from each E_k for $k \in [n]$.*

Proof. For the forward direction assume, towards a contradiction, that F contains both ik and jk with $i, j < k$. So if r is the root of the tree containing i, j, k, then, by the increasing condition, i must be the vertex preceding k on the unique r–k path. But the same must be true of j, which is a contradiction.

For the reverse implication, we must first verify that F is acyclic. But if F contains a cycle C, then let k be its maximum vertex. It follows that there are $ik, jk \in E(C)$ and, by the maximum requirement, $i, j < k$. This contradicts the assumption in this direction. Similarly one can show that if F is not increasing, then one can produce two edges from the same E_k and so we are done. □

It is now a simple matter to prove the following result of Hallam and Sagan [**41**]. The proof given here was obtained by Hallam, Martin, and Sagan [**40**].

Theorem 3.8.8. *If G has $V = [n]$, then*

$$\mathrm{isf}(G;t) = \prod_{k=1}^{n} (t - |E_k|).$$

Proof. The coefficient of t^{n-m} in the product is, up to sign, the sum of all terms of the form $|E_{i_1}||E_{i_2}|\cdots|E_{i_m}|$ where the i_j are distinct indices. But this product is the number of ways to pick one edge out of each of the sets $E_{i_1}, E_{i_2}, \dots, E_{i_m}$. So, by the previous lemma, the sum is the number of increasing forests with m edges, finishing the proof. □

Returning to the question of when the chromatic and increasing spanning forest polynomials are equal, we need the following definition. Graph G has a *perfect elimination order* (peo) if there is a total ordering of V as $v_1, v_2, \dots, v_n$ such that, for all k, the set of vertices coming before v_k in this order and adjacent to v_k form the vertices of a clique (complete subgraph of G). This definition may seem strange at first glance, but it has been useful in various graph-theoretic contexts. Returning to our running example graph we see that the order $1, 2, 3, 4$ is a peo since 1 is adjacent to no earlier vertex, 2 and 3 are both adjacent to a single previous vertex which is a K_1, and 4 is adjacent to 1 and 2 which form an edge also known as a K_2. We can now prove another result from [**40**].

Lemma 3.8.9. *Let G have $V = [n]$. Write the edges of G as ij with $i < j$ and order them lexicographically. For all $m \geq 0$ we have*

$$\mathrm{ISF}_m(G) \subseteq \mathrm{NBC}_m(G).$$

Furthermore, we have equality for all m if and only if the natural order on $[n]$ is a peo.

Proof. To prove the inclusion we suppose that F is an increasing spanning forest which contains a broken circuit B and derive a contradiction. By the lexicographic ordering of the edges, B must be a path of the form $v_1, v_2, \dots, v_l$ where $v_1 = \min\{v_1, \dots, v_l\}$ and $v_2 > v_l$. So there must be a smallest index $i \geq 2$ such that $v_i > v_{i+1}$. It follows that $v_{i-1}, v_{i+1} < v_i$ so that the two corresponding edges of B contradict Lemma 3.8.7.

For the forward direction of the second statement we must show that if $i, j < k$ and $ik, jk \in E(G)$, then $ij \in E(G)$. By Lemma 3.8.7 again, $\{ik, jk\}$ is not the edge set of an increasing spanning forest. So, by the assumed equality, this set must contain a broken circuit. Since there are only two edges, this set must actually be a broken circuit, and $ij \in E(G)$ must be the edge used to complete the cycle. The reverse implication is left as an exercise. □

From this result we immediately conclude the following.

Theorem 3.8.10. *Let G have $V = [n]$. Then* $\mathrm{isf}(G;t) = P(G;t)$ *if and only if the natural order on $[n]$ is a peo.* □

3.9. Combinatorial reciprocity

When plugging a negative parameter into a counting function results in a sign times another enumerative function, then this is called *combinatorial reciprocity*. This concept was introduced and studied by Stanley [**86**]. We have already seen two examples of this in equation (1.6) and Theorem 3.8.6 (and, more generally, Exercise 28 of this chapter). Here we will make a connection with recurrences and rational generating functions. See the text of Beck and Sanyal [**5**] for a whole book devoted to this subject.

Before stating a general theorem, we return to the example with which we began Section 3.6. This was the sequence defined by $a_0 = 2$ and $a_n = 3a_{n-1}$ for $n \geq 1$. One can extend the domain of this recursion to all integral n, in which case one gets $2 = a_0 = 3a_{-1}$ so that $a_{-1} = 2/3$. Then $2/3 = a_{-1} = 3a_{-2}$ yielding $a_{-2} = 2/3^2$, and so forth. An easy induction shows that for $n \leq 0$ we have $a_n = 2 \cdot 3^n$ just as for $n \geq 0$. We can also compute the generating function for the negatively indexed part of the sequence, where it is convenient to start with a_{-1}, which is the geometric series

$$\sum_{n\geq 1} a_{-n}x^n = \sum_{n\geq 1} \frac{2x^n}{3^n} = \frac{2x/3}{1-x/3} = \frac{2x}{3-x}.$$

Comparing this to $f(x) = \sum_{n\geq 0} a_n x^n$ as found in (3.12) we see that

$$-f(1/x) = \frac{-2}{1-3/x} = \frac{2x}{3-x} = \sum_{n\geq 1} a_{-n}x^n.$$

The reader should have some qualms about writing $f(1/x)$ since we have gone to great pains to point out that x has no inverse in $\mathbb{C}[[x]]$. Indeed, if we use the definition that $f(x) = \sum_{n\geq 0} a_n x^n$, then $f(1/x) = \sum_{n\geq 0} a_n/x^n$, which is not a formal power series! But if $f(x)$ can be expressed as a rational function $f(x) = p(x)/q(x)$ where $\deg p(x) \leq \deg q(x) \coloneqq d$, then we can make sense of this substitution as follows. Since $q(x)$ has degree d we have that $x^d q(1/x)$ is also a polynomial and is invertible since its constant coefficient is nonzero (Theorem 3.3.1). Furthermore, $x^d p(1/x)$ is also a polynomial since $d \geq \deg p(x)$. So we can *define*

$$f(1/x) = \frac{x^d p(1/x)}{x^d q(1/x)} \tag{3.25}$$

and stay inside the formal power series ring. With this convention, the following result makes sense.

Theorem 3.9.1. *Suppose that a_n is a sequence defined for all $n \in \mathbb{Z}$ and satisfying the linear recurrence relation with constant coefficients* (3.18) *for all such n. Letting $f(x) = \sum_{n\geq 0} a_n x^n$, we have*

$$\sum_{n\geq 1} a_{-n}x^n = -f(1/x).$$

Proof. By (3.25), to prove the theorem it suffices to show that

$$x^d q(1/x) \sum_{n\geq 1} a_{-n}x^n = -x^d p(1/x).$$

Note that by (3.20) we have

$$x^d q(1/x) = x^d + c_1 x^{d-1} + c_2 x^{d-2} + \cdots + c_d.$$

So if $m \geq 1$, then, using (3.18) and the fact that $x^d p(1/x)$ has degree at most d,

$$\begin{aligned}[x^{m+d}]x^d q(1/x) \sum_{n\geq 1} a_{-n}x^n &= a_{-m} + c_1 a_{-m-1} + c_2 a_{-m-2} + \cdots + c_d a_{-m-d} \\ &= 0 \\ &= [x^{m+d}](-x^d p(1/x)).\end{aligned}$$

Similarly we can prove that this equality of coefficients continues to hold for $-d \leq m \leq 0$. This completes the proof. □

To illustrate this theorem, we consider the negative binomial expansion. So if one fixes $n \geq 1$, then using Theorem 3.4.2 and equation (3.16)

$$f(x) := \frac{1}{(1-x)^n} = \sum_{k\geq 0} \left(\!\!\binom{n}{k}\!\!\right) x^k = \sum_{k\geq 0} \binom{n+k-1}{n-1} x^k.$$

Note that since n is fixed we are thinking of $\binom{n+k-1}{n-1}$ as a function of k. Substituting $-k$ for k in the binomial coefficient, we wish to consider the corresponding generating function

$$g(x) = \sum_{k\geq 1} \binom{n-k-1}{n-1} x^k.$$

We note that $\binom{n-k-1}{n-1} = 0$ for $1 \leq k < n$ since then $0 \leq n-k-1 < n-1$. So x^n can be factored out from $g(x)$ and, using (1.6) and the above expression for the negative binomial expansion,

$$\begin{aligned}g(x) &= x^n \sum_{k\geq n} \binom{n-k-1}{n-1} x^{k-n} \\ &= x^n \sum_{j\geq 0} \binom{-j-1}{n-1} x^j \\ &= (-1)^{n-1} x^n \sum_{j\geq 0} \left(\!\!\binom{j+1}{n-1}\!\!\right) x^j \\ &= (-1)^{n-1} x^n \sum_{j\geq 0} \binom{n+j-1}{n-1} x^j \\ &= \frac{(-1)^{n-1}x^n}{(1-x)^n}.\end{aligned}$$

On the other hand, we could apply Theorem 3.9.1 and write

$$g(x) = \frac{-1}{(1-1/x)^n} = \frac{-x^n}{(x-1)^n} = \frac{(-1)^{n-1}x^n}{(1-x)^n}$$

giving the same result but with less computation.

Exercises

(1) Prove that for $n \in \mathbb{N}$

$$\sum_{k=0}^{n} 2^k \binom{n}{k} = 3^n$$

in two ways.

(a) Use the Binomial Theorem.

(b) Use a combinatorial argument.

(c) Generalize this identity by replacing 2^k by c^k for any $c \in \mathbb{N}$, giving both a proof using the Binomial Theorem and one which is combinatorial.

(2) For $m, n, k \in \mathbb{N}$ show that

$$\binom{m+n}{k} = \sum_{l \geq 0} \binom{m}{l} \binom{n}{k-l}$$

in three ways: by induction, using the Binomial Theorem, and using a combinatorial argument.

(3) Let $x_1, \ldots, x_m$ be variables. Prove the multinomial coefficient identity

$$\sum_{n_1 + \cdots + n_m = n} \binom{n}{n_1, \ldots, n_m} x_1^{n_1} \cdots x_m^{n_m} = (x_1 + \cdots + x_m)^n,$$

in three ways:

(a) by inducting on n,

(b) by using the Binomial Theorem and inducting on m,

(c) by a combinatorial argument.

(4) (a) Prove Theorem 3.1.2.

(b) Use this generating function to rederive Corollary 1.5.3.

(5) (a) Recall that an inversion of $\pi \in P([n])$ is a pair (i, j) with $i < j$ and $\pi_i > \pi_j$ and we call π_i the *maximum* of the inversion. The *inversion table* of π is $I(\pi) = (a_1, a_2, \ldots, a_n)$ where a_k is the number of inversions with maximum k. Show that $0 \leq a_k < k$ for all k and that

$$\operatorname{inv} \pi = \sum_{k=1}^{n} a_k.$$

(b) Let

$$\mathcal{I}_n = \{(a_1, a_2, \ldots, a_n) \mid 0 \leq a_k < k \text{ for all } k\}.$$

Show that the map $\pi \mapsto I(\pi)$ is a bijection $P([n]) \to \mathcal{I}_n$.

(c) Use part (b) and weight-generating functions to rederive Theorem 3.2.1.

(6) Let $\operatorname{st} : \mathfrak{S}_n \to \mathbb{N}$ be a statistic on permutations. Recall that for a permutation π we let $\operatorname{Av}_n(\pi)$ denote the set of permutations in $\mathfrak{S}_n$ avoiding π. Say that permutations π, σ are st-*Wilf equivalent*, written $\pi \overset{\text{st}}{\equiv} \sigma$, if for all $n \geq 0$ we have equality of the

generating functions

$$\sum_{\tau \in \mathrm{Av}_n(\pi)} q^{\mathrm{st}\,\tau} = \sum_{\tau \in \mathrm{Av}_n(\sigma)} q^{\mathrm{st}\,\tau}.$$

(a) Show that if π and σ are st-Wilf equivalent, then they are Wilf equivalent.

(b) Show that

$$132 \overset{\mathrm{inv}}{\equiv} 213$$

and

$$231 \overset{\mathrm{inv}}{\equiv} 312$$

and that there are no other inv-Wilf equivalences between two permutations in $\mathfrak{S}_3$.

(c) Show that

$$132 \overset{\mathrm{maj}}{\equiv} 231$$

and

$$213 \overset{\mathrm{maj}}{\equiv} 312$$

and that there are no other maj-Wilf equivalences between two permutations in $\mathfrak{S}_3$.

(7) (a) Prove that

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n \\ n-k \end{bmatrix}$$

in three ways: using the q-factorial definition, using the interpretation in terms of integer partitions, and using the interpretation in terms of subspaces.

(b) Prove the second recursion in Theorem 3.2.3 in two ways: by mimicking the proof of the first recursion and by using the first recursion in combination with part (a).

(c) If $S \subseteq [n]$, then let ΣS be the sum of the elements of S. Give two proofs of the following q-analogue of the fact that $\#\binom{[n]}{k} = \binom{n}{k}$:

$$\sum_{S \in \binom{[n]}{k}} q^{\Sigma S} = q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q,$$

one proof by induction and the other using Theorem 3.2.5.

(d) Give two other proofs of Theorem 3.2.6: one by inducting on n and one by using Theorem 3.2.5.

(8) (a) Reprove Theorem 3.2.4 in two way: using integer partitions and using subspaces.

(b) The Negative q-Binomial Theorem states that

$$\frac{1}{(1-t)(1-qt)(1-q^2t)\dots(1-q^{n-1}t)} = \sum_{k \geq 0} \begin{bmatrix} n+k-1 \\ k \end{bmatrix} t^k.$$

Give three proofs of this result: inductive, using integer partitions, and using subspaces.

(9) (a) Given $n_1+n_2+\cdots+n_m = n$, define the corresponding *q-multinomial coefficient* to be

$$\begin{bmatrix} n \\ n_1, n_2, \ldots, n_m \end{bmatrix}_q = \frac{[n]_q!}{[n_1]_q!\,[n_2]_q!\,\ldots\,[n_m]_q!}$$

if all $n_i \geq 0$ or zero otherwise. Prove that

$$\begin{bmatrix} n \\ n_1, n_2, \ldots, n_m \end{bmatrix}_q = \sum_{i=1}^{m} q^{n_1+n_2+\cdots+n_{i-1}} \begin{bmatrix} n-1 \\ n_1, \ldots, n_{i-1}, n_i - 1, n_{i+1}, \ldots, n_m \end{bmatrix}_q.$$

(b) Define inversions, descents, and the major index for permutations (linear orderings) of the multiset $M = \{\{1^{n_1}, 2^{n_2}, \ldots, m^{n_m}\}\}$ exactly the way as was done for permutations without repetition. Let $P(M)$ be the set of permutations of M. Prove that

$$\sum_{\pi \in P(M)} q^{\operatorname{inv} \pi} = \sum_{\pi \in P(M)} q^{\operatorname{maj} \pi} = \begin{bmatrix} n \\ n_1, n_2, \ldots, n_m \end{bmatrix}_q.$$

(c) Let V be a vector space over $\mathbb{F}_q$ of dimension n. Assume $S = \{s_1 < \cdots < s_m\} \subseteq \{0, 1, \ldots, n\}$. Then a *flag of type S* is a chain of subspaces

$$F : W_1 < W_2 < \cdots < W_m \leq V$$

such that $\dim W_i = s_i$ for all i. The reason for this terminology is that when $n = 2$ and $S = \{0, 1, 2\}$, then F consists of a point (the origin) contained in a line contained in a plane which could be viewed as a drawing of a physical flag with the point being the hole in the ground, the line being the flag pole, and the plane being the cloth flag itself. Give two proofs that

$$\#\{F \text{ of type } S\} = \begin{bmatrix} n \\ s_1,\ s_2 - s_1,\ s_3 - s_2,\ \ldots,\ s_m - s_{m-1},\ n - s_m \end{bmatrix}_q,$$

one by mimicking the proof of Theorem 3.2.6 and one by induction on m.

(10) (a) Prove that in $\mathbb{C}[[x]]$ we have $e^{kx} = (e^x)^k$ for any $k \in \mathbb{N}$.

(b) Define formal power series for the trigonometric functions using their usual Taylor expansions. Prove that in $\mathbb{C}[[x]]$ we have $\sin^2 x + \cos^2 x = 1$ and $\sin 2x = 2 \sin x \cos x$.

(c) If one is given a sequence $a_0, a_1, a_2, \ldots$ and defines

$$f_k(x) = a_k x^k,$$

then show that $\sum_{k\geq 0} f_k(x) = f(x)$ where $f(x) = \sum_{k\geq 0} a_k x^k$.

(d) Prove the backwards direction of Theorem 3.3.2.

(e) Use Theorems 3.3.1 and 3.3.3 to reprove that $1/x$ and e^{1+x} are not well-defined in $\mathbb{C}[[x]]$.

(f) Prove Theorem 3.3.4

(11) Prove that if S, T are summable sets, then so is $S \times T$.

(12) Say that $f(x) \in \mathbb{C}[[x]]$ has a *square root* if there is $g(x) \in \mathbb{C}[[x]]$ such that $f(x) = g(x)^2$.

(a) Prove that $f(x)$ has a square root if and only if $\operatorname{mdeg} f(x)$ is even.

(b) Show that as formal power series

$$(1+x)^{1/2} = \sum_{k\geq 0} \binom{1/2}{k} x^k.$$

(c) Show that as formal power series

$$(e^x)^{1/2} = \sum_{k\geq 0} \frac{x^k}{2^k\, k!}.$$

(d) Generalize the previous parts of this exercise to mth roots for $m \in \mathbb{P}$.

(13) Give a second proof of Theorem 3.4.2 by using induction.

(14) Prove Theorem 3.4.3.

(15) Prove Theorem 3.5.1.

(16) (a) Finish the proof of Theorem 3.5.3(a).

(b) Give a second proof of Theorem 3.5.3(b) by using Theorem 3.2.5.

(c) Show that the generating function for the number of partitions of n with largest part k equals the generating function for the number of partitions of n with exactly k parts and that both are equal to the product

$$\frac{x^k}{(1-x)(1-x^2)\cdots(1-x^k)}.$$

(d) The *Durfee square* of a Young diagram λ is the largest square partition (d^d) such that $(d^d) \subseteq \lambda$. Use this concept to prove that

$$\sum_{n\geq 0} p(n)x^n = \sum_{d\geq 0} \frac{x^{d^2}}{(1-x)^2(1-x^2)^2\cdots(1-x^d)^2}.$$

(17) Let a_n be the number of integer partitions of n such that any part i is repeated fewer than i times and let b_n be the number of integer partitions of n such that no part is a square. Use generating functions to show that $a_n = b_n$ for all n.

(18) Given $m \geq 2$, use generating functions to show that the number of partitions of n where each part is repeated fewer than m times equals the number of partitions of n into parts not divisible by m. Note that bijective proofs of this result were given in Exercise 15 of Chapter 2.

(19) (a) Show that F_n is the closest integer to

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

for $n \geq 1$.

(b) Prove that

$$F_n^2 = \begin{cases} F_{n-1}F_{n+1} + 1 & \text{if } n \text{ is odd,} \\ F_{n-1}F_{n+1} - 1 & \text{if } n \text{ is even} \end{cases}$$

in two ways: using equation (3.14) and by a combinatorial argument.

(20) (a) Use the algorithm in Section 3.6 to rederive Theorem 3.1.1.
(b) Complete the proof of Theorem 3.6.1.
(c) Give a second proof of Theorem 3.6.1 using Theorem 3.1.2
(d) Prove Theorem 3.6.2.
(e) Let s be the infinite matrix with rows and columns indexed by $\mathbb{N}$ and with $s(n,k)$ being the entry in row n and column k. Similarly define S with entries $S(n,k)$. Show that $Ss = sS = I$ where I is the $\mathbb{N} \times \mathbb{N}$ identity matrix. Hint: Use Theorems 3.6.1 and 3.6.2.

(21) Given $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, define the corresponding *q-Stirling number of the second kind* by $S[0,k] = \delta_{0,k}$ and, for $n \geq 1$,

$$S[n,k] = S[n-1,k-1] + [k]_q S[n-1,k].$$

(a) Show that

$$\sum_{n\geq 0} S[n,k]x^n = \frac{x^k}{(1-[1]_q x)(1-[2]_q x)\cdots(1-[k]_q x)}.$$

(b) All set partitions $\rho = B_1/B_2/\ldots/B_k$ in the rest of this problem will be written in *standard form*, which means that

$$1 = \min B_1 < \min B_2 < \cdots < \min B_k.$$

An *inversion* of ρ is a pair (b, B_j) where $b \in B_i$ for some $i < j$ and $b > \min B_j$. We let $\operatorname{inv}\rho$ be the number of inversions of ρ. For example, $\rho = B_1/B_2/B_3 = 136/25/4$ has inversions $(3,B_2)$, $(6,B_2)$, $(6,B_3)$, and $(5,B_3)$ so that $\operatorname{inv}\rho = 4$. Show that

$$S[n,k] = \sum_{\rho\in S([n],k)} q^{\operatorname{inv}\rho}.$$

(c) A *descent* of a set partition ρ is a pair (b, B_{i+1}) where $b \in B_i$ and $b > \min B_j$. We let $\operatorname{des}\rho$ be the number of descents of ρ. In the previous example, ρ has descents $(3,B_2)$, $(6,B_2)$, and $(5,B_3)$ so that $\operatorname{des}\rho = 3$. The *descent multiset* of ρ is denoted $\operatorname{Des}\rho$ and is the multiset

$$\{\{1^{d_1}, 2^{d_2}, \ldots, k^{d_k} \mid \text{for all } i,\ d_i = \text{number of descents } (b, B_{i+1})\}\}.$$

The *major index* of ρ is

$$\operatorname{maj}\rho = \sum_{i\in\operatorname{Des}\rho} i = d_1 + 2d_2 + \cdots + kd_k.$$

In our running example $\operatorname{Des}\rho = \{\{1^2, 2^1\}\}$ so that $\operatorname{maj}\rho = 1+1+2 = 4$. Show that

$$S[n,k] = \sum_{\rho\in S([n],k)} q^{\operatorname{maj}\rho}.$$

(22) Given $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, define the corresponding *signless q-Stirling number of the first kind* by $c[0,k] = \delta_{0,k}$ and, for $n \geq 1$,

$$c[n,k] = c[n-1,k-1] + [n-1]_q c[n-1,k]$$

(a) Show that

$$\sum_{k\geq 0} c[n,k]x^k = x(x+[1]_q)(x+[2]_q)\cdots(x+[n-1]_q).$$

(b) The *standard form* of $\pi \in \mathfrak{S}_n$ is $\pi = \kappa_1\kappa_2\cdots\kappa_k$ where the κ_i are the cycles of π,

$$\min\kappa_1 < \min\kappa_2 < \cdots < \min\kappa_k,$$

and each κ_i is written beginning with $\min\kappa_i$. Define the *cycle major index* of π to be $\operatorname{maj}_c\pi = \operatorname{maj}\pi'$ where π' is the permutation in one-line form obtained by removing the cycle parentheses in the standard form of π. If, for example, $\pi = (1,7,2)(3,6,8)(4,5)$, then $\pi' = 17236845$ so that $\operatorname{maj}_c\pi = 2+6 = 8$. Show that

$$c[n,k] = \sum_{\pi\in c([n],k)} q^{\operatorname{maj}_c\pi}.$$

(23) Reprove the formula

$$C(n) = \frac{1}{n+1}\binom{2n}{n}$$

by using Theorem 3.6.3.

(24) (a) Show that if $k, l \in \mathbb{N}$ are constants, then $\binom{n+l}{k}$ is a polynomial in n of degree k.
(b) Show that the polynomials

$$\binom{n}{0}, \binom{n+1}{1}, \binom{n+2}{2}, \ldots$$

form a basis for the algebra of polynomials in n.
(c) Use part (a) to complete the proof of Theorem 3.7.1.

(25) Redo the solution for the first recursion in Section 3.6 as well as the one for F_n using the method of undetermined coefficients.

(26) Prove for the n-cycle that

$$P(C_n;t) = (t-1)^n + (-1)^n(t-1)$$

in two ways: using deletion-contraction and using NBC sets.

(27) Prove Theorem 3.8.4 using induction and give a second proof of parts (b)–(d) using NBC sets.

(28) (a) Complete the proof of Theorem 3.8.6.
(b) Let $G = (V,E)$ be a graph and $t \in \mathbb{P}$. Call an acyclic orientation O and a (not necessarily proper) coloring $c : V \to [t]$ *compatible* if, for all arcs $\overrightarrow{uv}$ of O, we have $c(u) \le c(v)$. Show that if $\#V = n$, then

$$P(G;-t) = (-1)^n(\text{number of compatible pairs } (O,c)).$$

(c) Show that Theorem 3.8.6 is a special case of part (b).

(29) Finish the proofs of Lemma 3.8.7 and Lemma 3.8.9.

(30) (a) Call a permutation σ which avoids $\Pi = \{231, 312, 321\}$ *tight*. Show that σ is tight if and only if σ is an involution having only 2-cycles of the form $(i, i+1)$ for some i.
(b) Let G be a graph with $V = [n]$. Call a spanning forest F of G *tight* if the sequence of labels on any path starting at a root of F avoids Π as in part (a). Let

$$\operatorname{TSF}_m(G) = \{F \mid F \text{ is a tight spanning forest of } G \text{ with } m \text{ edges}\}.$$

Show that if G has no 3-cycles, then for all $m \geq 0$

$$\mathrm{TSF}_m(G) \subseteq \mathrm{NBC}_m(G).$$

(c) A *candidate path* in a graph G with no 3-cycles is a path of the form

$$a, c, b, v_1, v_2, \ldots, v_m = d$$

such that $a < b < c$, $m \geq 1$, and v_m is the only v_i smaller than c. A total order on $V(G)$ is called a *quasiperfect ordering* (qpo) if every candidate path satisfies the following condition: either $ad \in E(G)$, or $d < b$ and $cd \in E(G)$. Consider the generating function

$$\mathrm{tsf}(G;t) = \sum_{m \geq 0} (-1)^m \, \mathrm{tsf}_m(G) t^{n-m}.$$

Show that $\mathrm{tsf}(G;t) = P(G;t)$ if and only if the natural order on $[n]$ is a qpo.

(31) Fill in the details of the case $-d \leq m \leq 0$ in the proof of Theorem 3.9.1.

(32) (a) Extend the Fibonacci numbers F_n to all $n \in \mathbb{Z}$ by insisting that their recursion continue to hold for $n < 0$. Show that if $n \geq 0$, then

$$F_{-n} = (-1)^{n-1} F_n.$$

(b) Find $\sum_{n \geq 1} F_{-n} x^n$ in two ways: by using part (a) and by using Theorem 3.9.1.