5

Binomial Coefficients

Lucky us!

LET'S TAKE A BREATHER. The previous chapters have seen some heavy going, with sums involving floor, ceiling, mod, phi, and mu functions. Now we're going to study binomial coefficients, which turn out to be (a) more important in applications, and (b) easier to manipulate, than all those other quantities.

5.1 BASIC IDENTITIES

Otherwise known as combinations of n things, k at a time.

The symbol $\binom{n}{k}$ is a binomial coefficient, so called because of an important property we look at later this section, the binomial theorem. But we read the symbol "n choose k." This incantation arises from its combinatorial interpretation — it is the number of ways to choose a k-element subset from an n-element set. For example, from the set $\{1,2,3,4\}$ we can choose two elements in six ways,

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

so $\binom{4}{2} = 6$.

To express the number $\binom{n}{k}$ in more familiar terms it's easiest to first determine the number of k-element *sequences*, rather than subsets, chosen from an n-element set; for sequences, the order of the elements counts. We use the same argument we used in Chapter 4 to show that $n!$ is the number of permutations of n objects. There are n choices for the first element of the sequence; for each, there are $n-1$ choices for the second; and so on, until there are $n-k+1$ choices for the kth. This gives $n(n-1)\ldots(n-k+1) = n^{\underline{k}}$ choices in all. And since each k-element subset has exactly $k!$ different orderings, this number of *sequences* counts each *subset* exactly $k!$ times. To get our answer, we simply divide by $k!$:

$$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots(1)}.$$

For example,

$$\binom{4}{2} = \frac{4\cdot 3}{2\cdot 1} = 6\,;$$

this agrees with our previous enumeration.

We call n the *upper index* and k the *lower index*. The indices are restricted to be nonnegative integers by the combinatorial interpretation, because sets don't have negative or fractional numbers of elements. But the binomial coefficient has many uses besides its combinatorial interpretation, so we will remove some of the restrictions. It's most useful, it turns out, to allow an arbitrary real (or even complex) number to appear in the upper index, and to allow an arbitrary integer in the lower. Our formal definition therefore takes the following form:

$$\binom{r}{k} = \begin{cases} \dfrac{r(r-1)\ldots(r-k+1)}{k(k-1)\ldots(1)} = \dfrac{r^{\underline{k}}}{k!}\,, & \text{integer } k \geqslant 0; \\ 0\,, & \text{integer } k < 0. \end{cases} \tag{5.1}$$

This definition has several noteworthy features. First, the upper index is called r, not n; the letter r emphasizes the fact that binomial coefficients make sense when any real number appears in this position. For instance, we have $\binom{-1}{3} = (-1)(-2)(-3)/(3\cdot 2\cdot 1) = -1$. There's no combinatorial interpretation here, but $r = -1$ turns out to be an important special case. A noninteger index like $r = -1/2$ also turns out to be useful.

Second, we can view $\binom{r}{k}$ as a kth-degree polynomial in r. We'll see that this viewpoint is often helpful.

Third, we haven't defined binomial coefficients for noninteger lower indices. A reasonable definition can be given, but actual applications are rare, so we will defer this generalization to later in the chapter.

Final note: We've listed the restrictions 'integer $k \geqslant 0$' and 'integer $k < 0$' at the right of the definition. Such restrictions will be listed in all the identities we will study, so that the range of applicability will be clear. In general the fewer restrictions the better, because an unrestricted identity is most useful; still, any restrictions that apply are an important part of the identity. When we manipulate binomial coefficients, it's easier to ignore difficult-to-remember restrictions temporarily and to check later that nothing has been violated. But the check needs to be made.

For example, almost every time we encounter $\binom{n}{n}$ it equals 1, so we can get lulled into thinking that it's always 1. But a careful look at definition (5.1) tells us that $\binom{n}{n}$ is 1 only when $n \geqslant 0$ (assuming that n is an integer); when $n < 0$ we have $\binom{n}{n} = 0$. Traps like this can (and will) make life adventuresome.

Before getting to the identities that we will use to tame binomial coefficients, let's take a peek at some small values. The numbers in Table 155 form the beginning of *Pascal's triangle*, named after Blaise Pascal (1623–1662)

Table 155 Pascal's triangle.

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

Binomial coefficients were well known in Asia, many centuries before Pascal was born [90], but he had no way to know that.

because he wrote an influential treatise about them [285]. The empty entries in this table are actually 0's, because of a zero in the numerator of (5.1); for example, $\binom{1}{2} = (1\cdot 0)/(2\cdot 1) = 0$. These entries have been left blank simply to help emphasize the rest of the table.

It's worthwhile to memorize formulas for the first three columns,

$$\binom{r}{0} = 1\,, \qquad \binom{r}{1} = r\,, \qquad \binom{r}{2} = \frac{r(r-1)}{2}\,; \tag{5.2}$$

these hold for arbitrary reals. (Recall that $\binom{n+1}{2} = \frac{1}{2}n(n+1)$ is the formula we derived for triangular numbers in Chapter 1; triangular numbers are conspicuously present in the $\binom{n}{2}$ column of Table 155.) It's also a good idea to memorize the first five rows or so of Pascal's triangle, so that when the pattern 1, 4, 6, 4, 1 appears in some problem we will have a clue that binomial coefficients probably lurk nearby.

In Italy it's called Tartaglia's triangle.

The numbers in Pascal's triangle satisfy, practically speaking, infinitely many identities, so it's not too surprising that we can find some surprising relationships by looking closely. For example, there's a curious "hexagon property," illustrated by the six numbers 56, 28, 36, 120, 210, 126 that surround 84 in the lower right portion of Table 155. Both ways of multiplying alternate numbers from this hexagon give the same product: $56\cdot 36\cdot 210 = 28\cdot 120\cdot 126 = 423360$. The same thing holds if we extract such a hexagon from any other part of Pascal's triangle.

And now the identities. Our goal in this section will be to learn a few simple rules by which we can solve the vast majority of practical problems involving binomial coefficients.

"C'est une chose estrange combien il est fertile en proprietez."
—B. Pascal [285]

Definition (5.1) can be recast in terms of factorials in the common case that the upper index r is an integer, n, that's greater than or equal to the lower index k:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}\,, \qquad \text{integers } n \geqslant k \geqslant 0. \tag{5.3}$$

To get this formula, we just multiply the numerator and denominator of (5.1) by $(n-k)!$. It's occasionally useful to expand a binomial coefficient into this factorial form (for example, when proving the hexagon property). And we often want to go the other way, changing factorials into binomials.

The factorial representation hints at a symmetry in Pascal's triangle: Each row reads the same left-to-right as right-to-left. The identity reflecting this—called the *symmetry* identity—is obtained by changing k to $n-k$:

$$\binom{n}{k} = \binom{n}{n-k}\,, \qquad \begin{array}{c}\text{integer } n \geqslant 0,\\ \text{integer } k.\end{array} \tag{5.4}$$

This formula makes combinatorial sense, because by specifying the k chosen things out of n we're in effect specifying the $n-k$ unchosen things.

The restriction that n and k be integers in identity (5.4) is obvious, since each lower index must be an integer. But why can't n be negative? Suppose, for example, that $n = -1$. Is

$$\binom{-1}{k} \overset{?}{=} \binom{-1}{-1-k}$$

a valid equation? No. For instance, when $k = 0$ we get 1 on the left and 0 on the right. In fact, for any integer $k \geqslant 0$ the left side is

$$\binom{-1}{k} = \frac{(-1)(-2)\ldots(-k)}{k!} = (-1)^k\,,$$

which is either 1 or -1; but the right side is 0, because the lower index is negative. And for negative k the left side is 0 but the right side is

$$\binom{-1}{-1-k} = (-1)^{-1-k}\,,$$

which is either 1 or -1. So the equation '$\binom{-1}{k} = \binom{-1}{-1-k}$' is always false!

The symmetry identity fails for all other negative integers n, too. But unfortunately it's all too easy to forget this restriction, since the expression in the upper index is sometimes negative only for obscure (but legal) values

I just hope I don't fall into this trap during the midterm.

of its variables. Everyone who's manipulated binomial coefficients much has fallen into this trap at least three times.

But the symmetry identity does have a big redeeming feature: It works for all values of k, even when $k < 0$ or $k > n$. (Because both sides are zero in such cases.) Otherwise $0 \leqslant k \leqslant n$, and symmetry follows immediately from (5.3):

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n!}{\bigl(n-(n-k)\bigr)!\;(n-k)!} = \binom{n}{n-k}.$$

Our next important identity lets us move things in and out of binomial coefficients:

$$\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}, \qquad \text{integer } k \neq 0. \tag{5.5}$$

The restriction on k prevents us from dividing by 0 here. We call (5.5) an *absorption* identity, because we often use it to absorb a variable into a binomial coefficient when that variable is a nuisance outside. The equation follows from definition (5.1), because $r^{\underline{k}} = r(r-1)^{\underline{k-1}}$ and $k! = k(k-1)!$ when $k > 0$; both sides are zero when $k < 0$.

If we multiply both sides of (5.5) by k, we get an absorption identity that works even when $k = 0$:

$$k\binom{r}{k} = r\binom{r-1}{k-1}, \qquad \text{integer } k. \tag{5.6}$$

This one also has a companion that keeps the lower index intact:

$$(r-k)\binom{r}{k} = r\binom{r-1}{k}, \qquad \text{integer } k. \tag{5.7}$$

We can derive (5.7) by sandwiching an application of (5.6) between two applications of symmetry:

$$\begin{aligned}
(r-k)\binom{r}{k} &= (r-k)\binom{r}{r-k} && \text{(by symmetry)}\\
&= r\binom{r-1}{r-k-1} && \text{(by (5.6))}\\
&= r\binom{r-1}{k}. && \text{(by symmetry)}
\end{aligned}$$

But wait a minute. We've claimed that the identity holds for *all* real r, yet the derivation we just gave holds only when r is a positive integer. (The upper index $r-1$ must be a nonnegative integer if we're to use the symmetry property (5.4) with impunity.) Have we been cheating? No. It's true that the derivation is valid only for positive integers r; but we can claim that the identity holds for all values of r, because both sides of (5.7) are polynomials in r of degree $k+1$. A nonzero polynomial of degree d or less can have at most d distinct zeros; therefore the difference of two such polynomials, which also has degree d or less, cannot be zero at more than d points unless it is identically zero. In other words, if two polynomials of degree d or less agree at more than d points, they must agree everywhere. We have shown that $(r-k)\binom{r}{k} = r\binom{r-1}{k}$ whenever r is a positive integer; so these two polynomials agree at infinitely many points, and they must be identically equal.

(Well, not here anyway.)

The proof technique in the previous paragraph, which we will call the *polynomial argument*, is useful for extending many identities from integers to reals; we'll see it again and again. Some equations, like the symmetry identity (5.4), are not identities between polynomials, so we can't always use this method. But many identities do have the necessary form.

For example, here's another polynomial identity, perhaps the most important binomial identity of all, known as the *addition formula*:

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}, \qquad \text{integer } k. \tag{5.8}$$

When r is a positive integer, the addition formula tells us that every number in Pascal's triangle is the sum of two numbers in the previous row, one directly above it and the other just to the left. And the formula applies also when r is negative, real, or complex; the only restriction is that k be an integer, so that the binomial coefficients are defined.

One way to prove the addition formula is to assume that r is a positive integer and to use the combinatorial interpretation. Recall that $\binom{r}{k}$ is the number of possible k-element subsets chosen from an r-element set. If we have a set of r eggs that includes exactly one bad egg, there are $\binom{r}{k}$ ways to select k of the eggs. Exactly $\binom{r-1}{k}$ of these selections involve nothing but good eggs; and $\binom{r-1}{k-1}$ of them contain the bad egg, because such selections have $k-1$ of the $r-1$ good eggs. Adding these two numbers together gives (5.8). This derivation assumes that r is a positive integer, and that $k \geqslant 0$. But both sides of the identity are zero when $k < 0$, and the polynomial argument establishes (5.8) in all remaining cases.

We can also derive (5.8) by adding together the two absorption identities (5.7) and (5.6):

$$(r-k)\binom{r}{k} + k\binom{r}{k} = r\binom{r-1}{k} + r\binom{r-1}{k-1};$$

the left side is $r\binom{r}{k}$, and we can divide through by r. This derivation is valid for everything but $r = 0$, and it's easy to check that remaining case.

Those of us who tend not to discover such slick proofs, or who are otherwise into tedium, might prefer to derive (5.8) by a straightforward manipulation of the definition. If $k > 0$,

$$\begin{aligned}\binom{r-1}{k}+\binom{r-1}{k-1} &= \frac{(r-1)^{\underline{k}}}{k!}+\frac{(r-1)^{\underline{k-1}}}{(k-1)!}\\ &= \frac{(r-1)^{\underline{k-1}}\,(r-k)}{k!}+\frac{(r-1)^{\underline{k-1}}\,k}{k!}\\ &= \frac{(r-1)^{\underline{k-1}}\,r}{k!} = \frac{r^{\underline{k}}}{k!} = \binom{r}{k}.\end{aligned}$$

Again, the cases for $k \leqslant 0$ are easy to handle.

We've just seen three rather different proofs of the addition formula. This is not surprising; binomial coefficients have many useful properties, several of which are bound to lead to proofs of an identity at hand.

The addition formula is essentially a recurrence for the numbers of Pascal's triangle, so we'll see that it is especially useful for proving other identities by induction. We can also get a new identity immediately by unfolding the recurrence. For example,

$$\begin{aligned}\binom{5}{3} &= \binom{4}{3}+\binom{4}{2}\\ &= \binom{4}{3}+\binom{3}{2}+\binom{3}{1}\\ &= \binom{4}{3}+\binom{3}{2}+\binom{2}{1}+\binom{2}{0}\\ &= \binom{4}{3}+\binom{3}{2}+\binom{2}{1}+\binom{1}{0}+\binom{1}{-1}.\end{aligned}$$

Since $\binom{1}{-1} = 0$, that term disappears and we can stop. This method yields the general formula

$$\begin{aligned}\sum_{k\leqslant n}\binom{r+k}{k} &= \binom{r}{0}+\binom{r+1}{1}+\cdots+\binom{r+n}{n}\\ &= \binom{r+n+1}{n}, \qquad \text{integer } n. \qquad (5.9)\end{aligned}$$

Notice that we don't need the lower limit $k \geqslant 0$ on the index of summation, because the terms with $k < 0$ are zero.

This formula expresses one binomial coefficient as the sum of others whose upper and lower indices stay the same distance apart. We found it by repeatedly expanding the binomial coefficient with the smallest lower index: first

$\binom{5}{3}$, then $\binom{4}{2}$, then $\binom{3}{1}$, then $\binom{2}{0}$. What happens if we unfold the other way, repeatedly expanding the one with largest lower index? We get

$$\begin{aligned}\binom{5}{3} &= \binom{4}{3} + \binom{4}{2}\\ &= \binom{3}{3} + \binom{3}{2} + \binom{4}{2}\\ &= \binom{2}{3} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2}\\ &= \binom{1}{3} + \binom{1}{2} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2}\\ &= \binom{0}{3} + \binom{0}{2} + \binom{1}{2} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2}.\end{aligned}$$

Now $\binom{0}{3}$ is zero (so are $\binom{0}{2}$ and $\binom{1}{2}$, but these make the identity nicer), and we can spot the general pattern:

$$\begin{aligned}\sum_{0\leqslant k\leqslant n}\binom{k}{m} &= \binom{0}{m} + \binom{1}{m} + \cdots + \binom{n}{m}\\ &= \binom{n+1}{m+1}, \qquad \text{integers } m,n \geqslant 0. \end{aligned} \tag{5.10}$$

This identity, which we call *summation on the upper index*, expresses a binomial coefficient as the sum of others whose lower indices are constant. In this case the sum needs the lower limit $k \geqslant 0$, because the terms with $k < 0$ *aren't* zero. Also, m and n can't in general be negative.

Identity (5.10) has an interesting combinatorial interpretation. If we want to choose $m+1$ tickets from a set of $n+1$ tickets numbered 0 through n, there are $\binom{k}{m}$ ways to do this when the largest ticket selected is number k.

We can prove both (5.9) and (5.10) by induction using the addition formula, but we can also prove them from each other. For example, let's prove (5.9) from (5.10); our proof will illustrate some common binomial coefficient manipulations. Our general plan will be to massage the left side $\sum\binom{r+k}{k}$ of (5.9) so that it looks like the left side $\sum\binom{k}{m}$ of (5.10); then we'll invoke that identity, replacing the sum by a single binomial coefficient; finally we'll transform that coefficient into the right side of (5.9).

We can assume for convenience that r and n are nonnegative integers; the general case of (5.9) follows from this special case, by the polynomial argument. Let's write m instead of r, so that this variable looks more like a nonnegative integer. The plan can now be carried out systematically as follows:

$$\begin{aligned}\sum_{k\leqslant n}\binom{m+k}{k} &= \sum_{-m\leqslant k\leqslant n}\binom{m+k}{k}\\ &= \sum_{-m\leqslant k\leqslant n}\binom{m+k}{m}\\ &= \sum_{0\leqslant k\leqslant m+n}\binom{k}{m}\\ &= \binom{m+n+1}{m+1} \;=\; \binom{m+n+1}{n}.\end{aligned}$$

Let's look at this derivation blow by blow. The key step is in the second line, where we apply the symmetry law (5.4) to replace $\binom{m+k}{k}$ by $\binom{m+k}{m}$. We're allowed to do this only when $m+k \geqslant 0$, so our first step restricts the range of k by discarding the terms with $k < -m$. (This is legal because those terms are zero.) Now we're almost ready to apply (5.10); the third line sets this up, replacing k by $k-m$ and tidying up the range of summation. This step, like the first, merely plays around with $\sum$-notation. Now k appears by itself in the upper index and the limits of summation are in the proper form, so the fourth line applies (5.10). One more use of symmetry finishes the job.

Certain sums that we did in Chapters 1 and 2 were actually special cases of (5.10), or disguised versions of this identity. For example, the case $m = 1$ gives the sum of the nonnegative integers up through n:

$$\binom{0}{1}+\binom{1}{1}+\cdots+\binom{n}{1} \;=\; 0+1+\cdots+n \;=\; \frac{(n+1)n}{2} \;=\; \binom{n+1}{2}.$$

And the general case is equivalent to Chapter 2's rule

$$\sum_{0\leqslant k\leqslant n} k^{\underline{m}} \;=\; \frac{(n+1)^{\underline{m+1}}}{m+1}, \qquad \text{integers } m,n \geqslant 0,$$

if we divide both sides of this formula by $m!$. In fact, the addition formula (5.8) tells us that

$$\Delta\left(\binom{x}{m}\right) \;=\; \binom{x+1}{m}-\binom{x}{m} \;=\; \binom{x}{m-1},$$

if we replace r and k respectively by $x+1$ and m. Hence the methods of Chapter 2 give us the handy indefinite summation formula

$$\sum \binom{x}{m}\,\delta x \;=\; \binom{x}{m+1} + C. \tag{5.11}$$

Binomial coefficients get their name from the *binomial theorem*, which deals with powers of the binomial expression $x+y$. Let's look at the smallest cases of this theorem:

"At the age of twenty-one he [Moriarty] wrote a treatise upon the Binomial Theorem, which has had a European vogue. On the strength of it, he won the Mathematical Chair at one of our smaller Universities."
—S. Holmes [84]

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

It's not hard to see why these coefficients are the same as the numbers in Pascal's triangle: When we expand the product

$$(x+y)^n = \overbrace{(x+y)(x+y)\ldots(x+y)}^{n\text{ factors}},$$

every term is itself the product of n factors, each either an x or y. The number of such terms with k factors of x and $n-k$ factors of y is the coefficient of x^ky^{n-k} after we combine like terms. And this is exactly the number of ways to choose k of the n binomials from which an x will be contributed; that is, it's $\binom{n}{k}$.

Some textbooks leave the quantity 0^0 undefined, because the functions x^0 and 0^x have different limiting values when x decreases to 0. But this is a mistake. We must define

$$x^0 = 1, \qquad \text{for all } x,$$

if the binomial theorem is to be valid when $x=0$, $y=0$, and/or $x=-y$. The theorem is too important to be arbitrarily restricted! By contrast, the function 0^x is quite unimportant. (See [220] for further discussion.)

But what exactly is the binomial theorem? In its full glory it is the following identity:

$$(x+y)^r = \sum_k \binom{r}{k} x^ky^{r-k}, \qquad \begin{matrix}\text{integer } r\geqslant 0\\ \text{or } |x/y|<1.\end{matrix} \tag{5.12}$$

The sum is over all integers k; but it is really a finite sum when r is a nonnegative integer, because all terms are zero except those with $0\leqslant k\leqslant r$. On the other hand, the theorem is also valid when r is negative, or even when r is an arbitrary real or complex number. In such cases the sum really is infinite, and we must have $|x/y|<1$ to guarantee the sum's absolute convergence.

Two special cases of the binomial theorem are worth special attention, even though they are extremely simple. If $x = y = 1$ and $r = n$ is nonnegative, we get

$$2^n = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n}, \qquad \text{integer } n \geqslant 0.$$

This equation tells us that row n of Pascal's triangle sums to 2^n. And when x is -1 instead of $+1$, we get

$$0^n = \binom{n}{0} - \binom{n}{1} + \cdots + (-1)^n\binom{n}{n}, \qquad \text{integer } n \geqslant 0.$$

For example, $1 - 4 + 6 - 4 + 1 = 0$; the elements of row n sum to zero if we give them alternating signs, except in the top row (when $n = 0$ and $0^0 = 1$).

When r is not a nonnegative integer, we most often use the binomial theorem in the special case $y = 1$. Let's state this special case explicitly, writing z instead of x to emphasize the fact that an arbitrary complex number can be involved here:

$$(1+z)^r = \sum_k \binom{r}{k} z^k, \qquad |z| < 1. \tag{5.13}$$

The general formula in (5.12) follows from this one if we set $z = x/y$ and multiply both sides by y^r.

We have proved the binomial theorem only when r is a nonnegative integer, by using a combinatorial interpretation. We can't deduce the general case from the nonnegative-integer case by using the polynomial argument, because the sum is infinite in the general case. But when r is arbitrary, we can use Taylor series and the theory of complex variables:

$$\begin{aligned} f(z) &= \frac{f(0)}{0!}z^0 + \frac{f'(0)}{1!}z^1 + \frac{f''(0)}{2!}z^2 + \cdots \\ &= \sum_{k\geqslant 0} \frac{f^{(k)}(0)}{k!} z^k . \end{aligned}$$

The derivatives of the function $f(z) = (1+z)^r$ are easily evaluated; in fact, $f^{(k)}(z) = r^{\underline{k}}(1+z)^{r-k}$. Setting $z = 0$ gives (5.13).

(Chapter 9 tells the meaning of O.)

We also need to prove that the infinite sum converges, when $|z| < 1$. It does, because $\binom{r}{k} = O(k^{-1-r})$ by equation (5.83) below.

Now let's look more closely at the values of $\binom{n}{k}$ when n is a negative integer. One way to approach these values is to use the addition law (5.8) to fill in the entries that lie above the numbers in Table 155, thereby obtaining Table 164. For example, we must have $\binom{-1}{0} = 1$, since $\binom{0}{0} = \binom{-1}{0} + \binom{-1}{-1}$ and $\binom{-1}{-1} = 0$; then we must have $\binom{-1}{1} = -1$, since $\binom{0}{1} = \binom{-1}{1} + \binom{-1}{0}$; and so on.

Table 164 Pascal's triangle, extended upward.

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
−4	1	−4	10	−20	35	−56	84	−120	165	−220	286
−3	1	−3	6	−10	15	−21	28	−36	45	−55	66
−2	1	−2	3	−4	5	−6	7	−8	9	−10	11
−1	1	−1	1	−1	1	−1	1	−1	1	−1	1
0	1	0	0	0	0	0	0	0	0	0	0

All these numbers are familiar. Indeed, the rows and columns of Table 164 appear as columns in Table 155 (but minus the minus signs). So there must be a connection between the values of $\binom{n}{k}$ for negative n and the values for positive n. The general rule is

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, \qquad \text{integer } k; \tag{5.14}$$

it is easily proved, since

$$\begin{aligned} r^{\underline{k}} &= r(r-1)\ldots(r-k+1) \\ &= (-1)^k(-r)(1-r)\ldots(k-1-r) \;=\; (-1)^k(k-r-1)^{\underline{k}} \end{aligned}$$

when $k \geqslant 0$, and both sides are zero when $k < 0$.

Identity (5.14) is particularly valuable because it holds without any restriction. (Of course, the lower index must be an integer so that the binomial coefficients are defined.) The transformation in (5.14) is called *negating the upper index*, or "upper negation."

But how can we remember this important formula? The other identities we've seen — symmetry, absorption, addition, etc. — are pretty simple, but this one looks rather messy. Still, there's a mnemonic that's not too bad: To negate the upper index, we begin by writing down $(-1)^k$, where k is the lower index. (The lower index doesn't change.) Then we immediately write k again, twice, in both lower and upper index positions. Then we negate the original upper index by *subtracting* it from the new upper index. And we complete the job by *subtracting* 1 more (always subtracting, not adding, because this is a negation process).

You call this a mnemonic? I'd call it pneumatic — full of air. It does help me remember, though.

Let's negate the upper index twice in succession, for practice. We get

(Now is a good time to do warmup exercise 4.)

$$\begin{aligned} \binom{r}{k} &= (-1)^k \binom{k-r-1}{k} \\ &= (-1)^{2k} \binom{k-(k-r-1)-1}{k} \;=\; \binom{r}{k}, \end{aligned}$$

so we're right back where we started. This is probably not what the framers of the identity intended; but it's reassuring to know that we haven't gone astray.

It's also frustrating, if we're trying to get somewhere else.

Some applications of (5.14) are, of course, more useful than this. We can use upper negation, for example, to move quantities between upper and lower index positions. The identity has a symmetric formulation,

$$(-1)^m\binom{-n-1}{m} = (-1)^n\binom{-m-1}{n}, \qquad \text{integers } m,n \geqslant 0, \tag{5.15}$$

which holds because both sides are equal to $\binom{m+n}{n}$.

Upper negation can also be used to derive the following interesting sum:

$$\begin{aligned}\sum_{k\leqslant m}\binom{r}{k}(-1)^k &= \binom{r}{0}-\binom{r}{1}+\cdots+(-1)^m\binom{r}{m}\\ &= (-1)^m\binom{r-1}{m}, \qquad \text{integer } m.\end{aligned} \tag{5.16}$$

The idea is to negate the upper index, then apply (5.9), and negate again:

(Here double negation helps, because we've sandwiched another operation in between.)

$$\begin{aligned}\sum_{k\leqslant m}\binom{r}{k}(-1)^k &= \sum_{k\leqslant m}\binom{k-r-1}{k}\\ &= \binom{-r+m}{m}\\ &= (-1)^m\binom{r-1}{m}.\end{aligned}$$

This formula gives us a partial sum of the rth row of Pascal's triangle, provided that the entries of the row have been given alternating signs. For instance, if $r=5$ and $m=2$ the formula gives $1-5+10=6=(-1)^2\binom{4}{2}$.

Notice that if $m\geqslant r$, (5.16) gives the alternating sum of the entire row, and this sum is zero when r is a positive integer. We proved this before, when we expanded $(1-1)^r$ by the binomial theorem; it's interesting to know that the partial sums of this expression can also be evaluated in closed form.

How about the simpler partial sum,

$$\sum_{k\leqslant m}\binom{n}{k} = \binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{m}; \tag{5.17}$$

surely if we can evaluate the corresponding sum with alternating signs, we ought to be able to do this one? But no; there is no closed form for the partial sum of a row of Pascal's triangle. We can do columns — that's (5.10) — but not rows. Curiously, however, there is a way to partially sum the row elements if they have been multiplied by their distance from the center:

$$\sum_{k\leqslant m}\binom{r}{k}\Bigl(\frac{r}{2}-k\Bigr) \;=\; \frac{m+1}{2}\binom{r}{m+1}, \qquad \text{integer } m. \tag{5.18}$$

(This formula is easily verified by induction on m.) The relation between these partial sums with and without the factor of $(r/2-k)$ in the summand is analogous to the relation between the integrals

$$\int_{-\infty}^{\alpha} xe^{-x^2}\,dx \;=\; -\tfrac{1}{2}e^{-\alpha^2} \qquad \text{and} \qquad \int_{-\infty}^{\alpha} e^{-x^2}\,dx\,.$$

The apparently more complicated integral on the left, with the factor of x, has a closed form, while the simpler-looking integral on the right, without the factor, has none. Appearances can be deceiving.

(Well, the right-hand integral is $\frac{1}{2}\sqrt{\pi}(1+\operatorname{erf}\alpha)$, a constant plus a multiple of the "error function" of α, if we're willing to accept that as a closed form.)

Near the end of this chapter, we'll study a method by which it's possible to determine whether or not there is a closed form for the partial sums of a given series involving binomial coefficients, in a fairly general setting. This method is capable of discovering identities (5.16) and (5.18), and it also will tell us that (5.17) is a dead end.

Partial sums of the binomial series lead to a curious relationship of another kind:

$$\sum_{k\leqslant m}\binom{m+r}{k}x^k y^{m-k} \;=\; \sum_{k\leqslant m}\binom{-r}{k}(-x)^k(x+y)^{m-k}\,, \quad \text{integer } m. \tag{5.19}$$

This identity isn't hard to prove by induction: Both sides are zero when $m<0$ and 1 when $m=0$. If we let S_m stand for the sum on the left, we can apply the addition formula (5.8) and show easily that

$$S_m \;=\; \sum_{k\leqslant m}\binom{m-1+r}{k}x^k y^{m-k} + \sum_{k\leqslant m}\binom{m-1+r}{k-1}x^k y^{m-k}\,;$$

and

$$\sum_{k\leqslant m}\binom{m-1+r}{k}x^k y^{m-k} \;=\; yS_{m-1} + \binom{m-1+r}{m}x^m\,,$$

$$\sum_{k\leqslant m}\binom{m-1+r}{k-1}x^k y^{m-k} \;=\; xS_{m-1}\,,$$

when $m>0$. Hence

$$S_m \;=\; (x+y)S_{m-1} + \binom{-r}{m}(-x)^m\,,$$

and this recurrence is satisfied also by the right-hand side of (5.19). By induction, both sides must be equal; QED.

But there's a neater proof. When r is an integer in the range $0 \geqslant r \geqslant -m$, the binomial theorem tells us that both sides of (5.19) are $(x+y)^{m+r}y^{-r}$. And since both sides are polynomials in r of degree m or less, agreement at $m+1$ different values is enough (but just barely!) to prove equality in general.

It may seem foolish to have an identity where one sum equals another. Neither side is in closed form. But sometimes one side turns out to be easier to evaluate than the other. For example, if we set $x = -1$ and $y = 1$, we get

$$\sum_{k \leqslant m} \binom{m+r}{k}(-1)^k = \binom{-r}{m}, \qquad \text{integer } m \geqslant 0,$$

an alternative form of identity (5.16). And if we set $x = y = 1$ and $r = m+1$, we get

$$\sum_{k \leqslant m} \binom{2m+1}{k} = \sum_{k \leqslant m} \binom{m+k}{k} 2^{m-k}.$$

The left-hand side sums just half of the binomial coefficients with upper index $2m+1$, and these are equal to their counterparts in the other half because Pascal's triangle has left-right symmetry. Hence the left-hand side is just $\frac{1}{2}2^{2m+1} = 2^{2m}$. This yields a formula that is quite unexpected,

(There's a nice combinatorial proof of this formula [247].)

$$\sum_{k \leqslant m} \binom{m+k}{k} 2^{-k} = 2^m, \qquad \text{integer } m \geqslant 0. \tag{5.20}$$

Let's check it when $m = 2$: $\binom{2}{0} + \frac{1}{2}\binom{3}{1} + \frac{1}{4}\binom{4}{2} = 1 + \frac{3}{2} + \frac{6}{4} = 4$. Astounding.

So far we've been looking either at binomial coefficients by themselves or at sums of terms in which there's only one binomial coefficient per term. But many of the challenging problems we face involve products of two or more binomial coefficients, so we'll spend the rest of this section considering how to deal with such cases.

Here's a handy rule that often helps to simplify the product of two binomial coefficients:

$$\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}, \qquad \text{integers } m, k. \tag{5.21}$$

We've already seen the special case $k = 1$; it's the absorption identity (5.6). Although both sides of (5.21) are products of binomial coefficients, one side often is easier to sum because of interactions with the rest of a formula. For example, the left side uses m twice, the right side uses it only once. Therefore we usually want to replace $\binom{r}{m}\binom{m}{k}$ by $\binom{r}{k}\binom{r-k}{m-k}$ when summing on m.

Equation (5.21) holds primarily because of cancellation between $m!$'s in the factorial representations of $\binom{r}{m}$ and $\binom{m}{k}$. If all variables are integers and $r \geqslant m \geqslant k \geqslant 0$, we have

$$\begin{aligned}\binom{r}{m}\binom{m}{k} &= \frac{r!}{m!\,(r-m)!}\,\frac{m!}{k!\,(m-k)!}\\ &= \frac{r!}{k!\,(m-k)!\,(r-m)!}\\ &= \frac{r!}{k!\,(r-k)!}\,\frac{(r-k)!}{(m-k)!\,(r-m)!} \;=\; \binom{r}{k}\binom{r-k}{m-k}.\end{aligned}$$

That was easy. Furthermore, if $m < k$ or $k < 0$, both sides of (5.21) are zero; so the identity holds for all integers m and k. Finally, the polynomial argument extends its validity to all real r.

Yeah, right.

A binomial coefficient $\binom{r}{k} = r!/(r-k)!\,k!$ can be written in the form $(a+b)!/a!\,b!$ after a suitable renaming of variables. Similarly, the quantity in the middle of the derivation above, $r!/k!\,(m-k)!\,(r-m)!$, can be written in the form $(a+b+c)!/a!\,b!\,c!$. This is a "trinomial coefficient," which arises in the "trinomial theorem":

$$\begin{aligned}(x+y+z)^n &= \sum_{\substack{0\leqslant a,b,c\leqslant n\\ a+b+c=n}} \frac{(a+b+c)!}{a!\,b!\,c!}x^a y^b z^c\\ &= \sum_{\substack{0\leqslant a,b,c\leqslant n\\ a+b+c=n}} \binom{a+b+c}{b+c}\binom{b+c}{c}x^a y^b z^c .\end{aligned}$$

"Excogitavi autem olim mirabilem regulam pro numeris coefficientibus potestatum, non tantum a binomio $x+y$, sed et a trinomio $x+y+z$, imo a polynomio quocunque, ut data potentia gradus cujuscunque v. gr. decimi, et potentia in ejus valore comprehensa, ut $x^5y^3z^2$, possim statim assignare numerum coefficientem, quem habere debet, sine ulla Tabula jam calculata."
— G. W. Leibniz [245]

So $\binom{r}{m}\binom{m}{k}$ is really a trinomial coefficient in disguise. Trinomial coefficients pop up occasionally in applications, and we can conveniently write them as

$$\binom{a+b+c}{a,\,b,\,c} = \frac{(a+b+c)!}{a!\,b!\,c!}$$

in order to emphasize the symmetry present.

Binomial and trinomial coefficients generalize to *multinomial coefficients*, which are always expressible as products of binomial coefficients:

$$\begin{aligned}\binom{a_1+a_2+\cdots+a_m}{a_1,a_2,\ldots,a_m} &= \frac{(a_1+a_2+\cdots+a_m)!}{a_1!\,a_2!\,\ldots\,a_m!}\\ &= \binom{a_1+a_2+\cdots+a_m}{a_2+\cdots+a_m}\ldots\binom{a_{m-1}+a_m}{a_m}.\end{aligned}$$

Therefore, when we run across such a beastie, our standard techniques apply.

Table 169 Sums of products of binomial coefficients.

$$\sum_k \binom{r}{m+k}\binom{s}{n-k} = \binom{r+s}{m+n}, \qquad \text{integers } m, n. \quad (5.22)$$

$$\sum_k \binom{l}{m+k}\binom{s}{n+k} = \binom{l+s}{l-m+n}, \qquad \begin{array}{c}\text{integer } l \geqslant 0,\\ \text{integers } m, n.\end{array} \quad (5.23)$$

$$\sum_k \binom{l}{m+k}\binom{s+k}{n}(-1)^k = (-1)^{l+m}\binom{s-m}{n-l}, \qquad \begin{array}{c}\text{integer } l \geqslant 0,\\ \text{integers } m, n.\end{array} \quad (5.24)$$

$$\sum_{k\leqslant l} \binom{l-k}{m}\binom{s}{k-n}(-1)^k = (-1)^{l+m}\binom{s-m-1}{l-m-n}, \qquad \begin{array}{c}\text{integers}\\ l, m, n \geqslant 0.\end{array} \quad (5.25)$$

$$\sum_{0\leqslant k\leqslant l} \binom{l-k}{m}\binom{q+k}{n} = \binom{l+q+1}{m+n+1}, \qquad \begin{array}{c}\text{integers } l, m \geqslant 0,\\ \text{integers } n \geqslant q \geqslant 0.\end{array} \quad (5.26)$$

Now we come to Table 169, which lists identities that are among the most important of our standard techniques. These are the ones we rely on when struggling with a sum involving a product of two binomial coefficients. Each of these identities is a sum over k, with one appearance of k in each binomial coefficient; there also are four nearly independent parameters, called m, n, r, etc., one in each index position. Different cases arise depending on whether k appears in the upper or lower index, and on whether it appears with a plus or minus sign. Sometimes there's an additional factor of $(-1)^k$, which is needed to make the terms summable in closed form.

Fold down the corner on this page, so you can find the table quickly later. You'll need it!

Table 169 is far too complicated to memorize in full; it is intended only for reference. But the first identity in this table is by far the most memorable, and it should be remembered. It states that the sum (over all integers k) of the product of two binomial coefficients, in which the upper indices are constant and the lower indices have a constant sum for all k, is the binomial coefficient obtained by summing both lower and upper indices. This identity is known as *Vandermonde's convolution*, because Alexandre Vandermonde wrote a significant paper about it in the late 1700s [357]; it was, however, known to Chu Shih-Chieh in China as early as 1303. All of the other identities in Table 169 can be obtained from Vandermonde's convolution by doing things like negating upper indices or applying the symmetry law, etc., with care; therefore Vandermonde's convolution is the most basic of all.

We can prove Vandermonde's convolution by giving it a nice combinatorial interpretation. If we replace k by $k-m$ and n by $n-m$, we can assume that $m = 0$; hence the identity to be proved is

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \qquad \text{integer } n. \tag{5.27}$$

Let r and s be nonnegative integers; the general case then follows by the polynomial argument. On the right side, $\binom{r+s}{n}$ is the number of ways to choose n people from among r men and s women. On the left, each term of the sum is the number of ways to choose k of the men and $n-k$ of the women. Summing over all k counts each possibility exactly once.

Sexist! You mentioned men first.

Much more often than not we use these identities left to right, since that's the direction of simplification. But every once in a while it pays to go the other direction, temporarily making an expression more complicated. When this works, we've usually created a double sum for which we can interchange the order of summation and then simplify.

Before moving on let's look at proofs for two more of the identities in Table 169. It's easy to prove (5.23); all we need to do is replace the first binomial coefficient by $\binom{l}{l-m-k}$, then Vandermonde's (5.22) applies.

The next one, (5.24), is a bit more difficult. We can reduce it to Vandermonde's convolution by a sequence of transformations, but we can just as easily prove it by resorting to the old reliable technique of mathematical induction. Induction is often the first thing to try when nothing else obvious jumps out at us, and induction on l works just fine here.

For the basis $l = 0$, all terms are zero except when $k = -m$; so both sides of the equation are $(-1)^m\binom{s-m}{n}$. Now suppose that the identity holds for all values less than some fixed l, where $l > 0$. We can use the addition formula to replace $\binom{l}{m+k}$ by $\binom{l-1}{m+k} + \binom{l-1}{m+k-1}$; the original sum now breaks into two sums, each of which can be evaluated by the induction hypothesis:

$$\sum_k \binom{l-1}{m+k}\binom{s+k}{n}(-1)^k + \sum_k \binom{l-1}{m+k-1}\binom{s+k}{n}(-1)^k$$
$$= (-1)^{l-1+m}\binom{s-m}{n-l+1} + (-1)^{l+m}\binom{s-m+1}{n-l+1}.$$

And this simplifies to the right-hand side of (5.24), if we apply the addition formula once again.

Two things about this derivation are worthy of note. First, we see again the great convenience of summing over all integers k, not just over a certain range, because there's no need to fuss over boundary conditions. Second, the addition formula works nicely with mathematical induction, because it's a recurrence for binomial coefficients. A binomial coefficient whose upper index is l is expressed in terms of two whose upper indices are $l-1$, and that's exactly what we need to apply the induction hypothesis.

So much for Table 169. What about sums with three or more binomial coefficients? If the index of summation is spread over all the coefficients, our chances of finding a closed form aren't great: Only a few closed forms are known for sums of this kind, hence the sum we need might not match the given specs. One of these rarities, proved in exercise 43, is

$$\begin{aligned}\sum_k \binom{m-r+s}{k}\binom{n+r-s}{n-k}\binom{r+k}{m+n}\\ = \binom{r}{m}\binom{s}{n}, \qquad \text{integers } m, n \geqslant 0. \qquad (5.28)\end{aligned}$$

Here's another, more symmetric example:

$$\begin{aligned}\sum_k \binom{a+b}{a+k}\binom{b+c}{b+k}\binom{c+a}{c+k}(-1)^k\\ = \frac{(a+b+c)!}{a!\,b!\,c!}, \qquad \text{integers } a, b, c \geqslant 0. \qquad (5.29)\end{aligned}$$

This one has a two-coefficient counterpart,

$$\sum_k \binom{a+b}{a+k}\binom{b+a}{b+k}(-1)^k = \frac{(a+b)!}{a!\,b!}, \qquad \text{integers } a, b \geqslant 0, \qquad (5.30)$$

which incidentally doesn't appear in Table 169. The analogous four-coefficient sum doesn't have a closed form, but a similar sum does:

$$\begin{aligned}\sum_k (-1)^k \binom{a+b}{a+k}\binom{b+c}{b+k}\binom{c+d}{c+k}\binom{d+a}{d+k}\Big/\binom{2a+2b+2c+2d}{a+b+c+d+k}\\ = \frac{(a+b+c+d)!\,(a+b+c)!\,(a+b+d)!\,(a+c+d)!\,(b+c+d)!}{(2a+2b+2c+2d)!\,(a+c)!\,(b+d)!\,a!\,b!\,c!\,d!},\\ \text{integers } a, b, c, d \geqslant 0.\end{aligned}$$

This was discovered by John Dougall [82] early in the twentieth century.

Is Dougall's identity the hairiest sum of binomial coefficients known? No! The champion so far is

$$\begin{aligned}\sum_{k_{ij}} (-1)^{\Sigma_{i<j}k_{ij}} \Biggl(\prod_{1\leqslant i<j<n}\binom{a_i+a_j}{a_j+k_{ij}}\Biggr)\Biggl(\prod_{1\leqslant j<n}\binom{a_j+a_n}{a_n+\Sigma_{i<j}k_{ij}-\Sigma_{i>j}k_{ji}}\Biggr)\\ = \binom{a_1+\cdots+a_n}{a_1, a_2, \ldots, a_n}, \qquad \text{integers } a_1, a_2, \ldots, a_n \geqslant 0. \qquad (5.31)\end{aligned}$$

Here the sum is over $\binom{n-1}{2}$ index variables k_{ij} for $1 \leqslant i < j < n$. Equation (5.29) is the special case $n = 3$; the case $n = 4$ can be written out as follows, if we use (a,b,c,d) for (a_1,a_2,a_3,a_4) and (i,j,k) for (k_{12},k_{13},k_{23}):

$$\sum_{i,j,k}(-1)^{i+j+k}\binom{a+b}{b+i}\binom{a+c}{c+j}\binom{b+c}{c+k}\binom{a+d}{d-i-j}\binom{b+d}{d+i-k}\binom{c+d}{d+j+k}$$
$$= \frac{(a+b+c+d)!}{a!\,b!\,c!\,d!}\,, \qquad \text{integers } a,b,c,d \geqslant 0.$$

The left side of (5.31) is the coefficient of $z_1^0 z_2^0 \ldots z_n^0$ after the product of $n(n-1)$ fractions

$$\prod_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}} \left(1-\frac{z_i}{z_j}\right)^{a_i}$$

has been fully expanded into positive and negative powers of the z's. The right side of (5.31) was conjectured by Freeman Dyson in 1962 and proved by several people shortly thereafter. Exercise 89 gives a "simple" proof of (5.31).

Another noteworthy identity involving lots of binomial coefficients is

$$\sum_{j,k}(-1)^{j+k}\binom{j+k}{k+l}\binom{r}{j}\binom{n}{k}\binom{s+n-j-k}{m-j}$$
$$= (-1)^l\binom{n+r}{n+l}\binom{s-r}{m-n-l}, \quad \text{integers } l,m,n; \quad n\geqslant 0. \quad (5.32)$$

This one, proved in exercise 83, even has a chance of arising in practical applications. But we're getting far afield from our theme of "basic identities," so we had better stop and take stock of what we've learned.

We've seen that binomial coefficients satisfy an almost bewildering variety of identities. Some of these, fortunately, are easily remembered, and we can use the memorable ones to derive most of the others in a few steps. Table 174 collects ten of the most useful formulas, all in one place; these are the best identities to know.

5.2 BASIC PRACTICE

In the previous section we derived a bunch of identities by manipulating sums and plugging in other identities. It wasn't too tough to find those derivations — we knew what we were trying to prove, so we could formulate a general plan and fill in the details without much trouble. Usually, however, out in the real world, we're not faced with an identity to prove; we're faced with a sum to simplify. And we don't know what a simplified form might look like (or even if one exists). By tackling many such sums in this section and the next, we will hone our binomial coefficient tools.

To start, let's try our hand at a few sums involving a single binomial coefficient.

Algorithm *self-teach:*
1 read problem
2 attempt solution
3 skim book solution
4 if attempt failed goto 1
else goto next problem

Problem 1: A sum of ratios.

We'd like to have a closed form for

$$\sum_{k=0}^{m} \binom{m}{k} \Big/ \binom{n}{k}, \qquad \text{integers } n \geqslant m \geqslant 0.$$

At first glance this sum evokes panic, because we haven't seen any identities that deal with a quotient of binomial coefficients. (Furthermore the sum involves two binomial coefficients, which seems to contradict the sentence preceding this problem.) However, just as we can use the factorial representations to reexpress a product of binomial coefficients as another product — that's how we got identity (5.21) — we can do likewise with a quotient. In fact we can avoid the grubby factorial representations by letting $r = n$ and dividing both sides of equation (5.21) by $\binom{n}{k}\binom{n}{m}$; this yields

Unfortunately, that algorithm can put you in an infinite loop.
Suggested patches:
0 set $c \leftarrow 0$
3a set $c \leftarrow c+1$
3b if $c = N$ *goto your TA*

$$\binom{m}{k} \Big/ \binom{n}{k} = \binom{n-k}{m-k} \Big/ \binom{n}{m}.$$

So we replace the quotient on the left, which appears in our sum, by the one on the right; the sum becomes

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

We still have a quotient, but the binomial coefficient in the denominator doesn't involve the index of summation k, so we can remove it from the sum. We'll restore it later.

~~goto~~

—E. W. Dijkstra

We can also simplify the boundary conditions by summing over all $k \geqslant 0$; the terms for $k > m$ are zero. The sum that's left isn't so intimidating:

$$\sum_{k \geqslant 0} \binom{n-k}{m-k}.$$

It's similar to the one in identity (5.9), because the index k appears twice with the same sign. But here it's $-k$ and in (5.9) it's not. The next step should therefore be obvious; there's only one reasonable thing to do:

... But this subchapter is called BASIC practice.

$$\sum_{k \geqslant 0} \binom{n-k}{m-k} = \sum_{m-k \geqslant 0} \binom{n-(m-k)}{m-(m-k)} = \sum_{k \leqslant m} \binom{n-m+k}{k}.$$

Table 174 The top ten binomial coefficient identities.

$\displaystyle\binom{n}{k} = \frac{n!}{k!\,(n-k)!}\,,$	integers $n \geqslant k \geqslant 0$.	*factorial expansion*
$\displaystyle\binom{n}{k} = \binom{n}{n-k}\,,$	integer $n \geqslant 0$, integer k.	*symmetry*
$\displaystyle\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}\,,$	integer $k \neq 0$.	*absorption/extraction*
$\displaystyle\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}\,,$	integer k.	*addition/induction*
$\displaystyle\binom{r}{k} = (-1)^k\binom{k-r-1}{k}\,,$	integer k.	*upper negation*
$\displaystyle\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}\,,$	integers m, k.	*trinomial revision*
$\displaystyle\sum_k \binom{r}{k} x^k y^{r-k} = (x+y)^r\,,$	integer $r \geqslant 0$, or $\lvert x/y \rvert < 1$.	*binomial theorem*
$\displaystyle\sum_{k \leqslant n} \binom{r+k}{k} = \binom{r+n+1}{n}\,,$	integer n.	*parallel summation*
$\displaystyle\sum_{0 \leqslant k \leqslant n} \binom{k}{m} = \binom{n+1}{m+1}\,,$	integers $m, n \geqslant 0$.	*upper summation*
$\displaystyle\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}\,,$	integer n.	*Vandermonde convolution*

And now we can apply the parallel summation identity, (5.9):

$$\sum_{k \leqslant m} \binom{n-m+k}{k} = \binom{(n-m)+m+1}{m} = \binom{n+1}{m}\,.$$

Finally we reinstate the $\binom{n}{m}$ in the denominator that we removed from the sum earlier, and then apply (5.7) to get the desired closed form:

$$\binom{n+1}{m} \Big/ \binom{n}{m} = \frac{n+1}{n+1-m}\,.$$

This derivation actually works for any real value of n, as long as no division by zero occurs; that is, as long as n isn't one of the integers $0, 1, \ldots, m-1$.

The more complicated the derivation, the more important it is to check the answer. This one wasn't too complicated but we'll check anyway. In the small case $m = 2$ and $n = 4$ we have

$$\binom{2}{0}\bigg/\binom{4}{0} \;+\; \binom{2}{1}\bigg/\binom{4}{1} \;+\; \binom{2}{2}\bigg/\binom{4}{2} \;=\; 1+\frac{1}{2}+\frac{1}{6} \;=\; \frac{5}{3}\,;$$

yes, this agrees perfectly with our closed form $(4+1)/(4+1-2)$.

Problem 2: From the literature of sorting.

Our next sum appeared way back in ancient times (the early 1970s) before people were fluent with binomial coefficients. A paper that introduced an improved merging technique [196] concludes with the following remarks: "It can be shown that the expected number of saved transfers ... is given by the expression

$$T \;=\; \sum_{r=0}^{n} r\,\frac{{}_{m-r-1}C_{m-n-1}}{{}_{m}C_{n}}$$

Here m and n are as defined above, and ${}_mC_n$ is the symbol for the number of combinations of m objects taken n at a time. ... The author is grateful to the referee for reducing a more complex equation for expected transfers saved to the form given here."

We'll see that this is definitely not a final answer to the author's problem. It's not even a midterm answer.

Please, don't remind me of the midterm.

First we should translate the sum into something we can work with; the ghastly notation ${}_{m-r-1}C_{m-n-1}$ is enough to stop anybody, save the enthusiastic referee (please). In our language we'd write

$$T \;=\; \sum_{k=0}^{n} k\binom{m-k-1}{m-n-1}\bigg/\binom{m}{n}\,, \qquad \text{integers } m > n \geqslant 0.$$

The binomial coefficient in the denominator doesn't involve the index of summation, so we can remove it and work with the new sum

$$S \;=\; \sum_{k=0}^{n} k\binom{m-k-1}{m-n-1}\,.$$

What next? The index of summation appears in the upper index of the binomial coefficient but not in the lower index. So if the other k weren't there, we could massage the sum and apply summation on the upper index (5.10). With the extra k, though, we can't. If we could somehow absorb that k into the binomial coefficient, using one of our absorption identities, we could then sum on the upper index. Unfortunately those identities don't work here. But if the k were instead $m-k$, we could use absorption identity (5.6):

$$(m-k)\binom{m-k-1}{m-n-1} = (m-n)\binom{m-k}{m-n}.$$

So here's the key: We'll rewrite k as $m-(m-k)$ and split the sum S into two sums:

$$\begin{aligned}
\sum_{k=0}^{n} k\binom{m-k-1}{m-n-1} &= \sum_{k=0}^{n}\bigl(m-(m-k)\bigr)\binom{m-k-1}{m-n-1}\\
&= \sum_{k=0}^{n} m\binom{m-k-1}{m-n-1} - \sum_{k=0}^{n}(m-k)\binom{m-k-1}{m-n-1}\\
&= m\sum_{k=0}^{n}\binom{m-k-1}{m-n-1} - \sum_{k=0}^{n}(m-n)\binom{m-k}{m-n}\\
&= mA-(m-n)B\,,
\end{aligned}$$

where

$$A = \sum_{k=0}^{n}\binom{m-k-1}{m-n-1}, \qquad B = \sum_{k=0}^{n}\binom{m-k}{m-n}.$$

The sums A and B that remain are none other than our old friends in which the upper index varies while the lower index stays fixed. Let's do B first, because it looks simpler. A little bit of massaging is enough to make the summand match the left side of (5.10):

$$\begin{aligned}
\sum_{0\leqslant k\leqslant n}\binom{m-k}{m-n} &= \sum_{0\leqslant m-k\leqslant n}\binom{m-(m-k)}{m-n}\\
&= \sum_{m-n\leqslant k\leqslant m}\binom{k}{m-n}\\
&= \sum_{0\leqslant k\leqslant m}\binom{k}{m-n}.
\end{aligned}$$

In the last step we've included the terms with $0\leqslant k< m-n$ in the sum; they're all zero, because the upper index is less than the lower. Now we sum on the upper index, using (5.10), and get

$$B = \sum_{0\leqslant k\leqslant m}\binom{k}{m-n} = \binom{m+1}{m-n+1}.$$

The other sum A is the same, but with m replaced by $m-1$. Hence we have a closed form for the given sum S, which can be further simplified:

$$\begin{aligned} S \;=\; mA-(m-n)B \;&=\; m\binom{m}{m-n} - (m-n)\binom{m+1}{m-n+1} \\ &=\; \left(m-(m-n)\frac{m+1}{m-n+1}\right)\binom{m}{m-n} \\ &=\; \left(\frac{n}{m-n+1}\right)\binom{m}{m-n}. \end{aligned}$$

And this gives us a closed form for the original sum:

$$\begin{aligned} T \;&=\; S\bigg/\binom{m}{n} \\ &=\; \frac{n}{m-n+1}\binom{m}{m-n}\bigg/\binom{m}{n} \\ &=\; \frac{n}{m-n+1}. \end{aligned}$$

Even the referee can't simplify this.

Again we use a small case to check the answer. When $m=4$ and $n=2$, we have

$$T \;=\; 0\cdot\tbinom{3}{1}/\tbinom{4}{2} \;+\; 1\cdot\tbinom{2}{1}/\tbinom{4}{2} \;+\; 2\cdot\tbinom{1}{1}/\tbinom{4}{2} \;=\; 0+\tfrac{2}{6}+\tfrac{2}{6} \;=\; \tfrac{2}{3},$$

which agrees with our formula $2/(4-2+1)$.

Problem 3: From an old exam.

Let's do one more sum that involves a single binomial coefficient. This one, unlike the last, originated in the halls of academia; it was a problem on a take-home test. We want the value of $Q_{1000000}$, when

Do old exams ever die?

$$Q_n \;=\; \sum_{k\leqslant 2^n}\binom{2^n-k}{k}(-1)^k, \qquad \text{integer } n \geqslant 0.$$

This one's harder than the others; we can't apply *any* of the identities we've seen so far. And we're faced with a sum of $2^{1000000}$ terms, so we can't just add them up. The index of summation k appears in both indices, upper and lower, but with opposite signs. Negating the upper index doesn't help, either; it removes the factor of $(-1)^k$, but it introduces a $2k$ in the upper index.

When nothing obvious works, we know that it's best to look at small cases. If we can't spot a pattern and prove it by induction, at least we'll have some data for checking our results. Here are the nonzero terms and their sums for the first four values of n.

n			Q_n
0	$\binom{1}{0}$	$= 1$	$= 1$
1	$\binom{2}{0} - \binom{1}{1}$	$= 1 - 1$	$= 0$
2	$\binom{4}{0} - \binom{3}{1} + \binom{2}{2}$	$= 1 - 3 + 1$	$= -1$
3	$\binom{8}{0} - \binom{7}{1} + \binom{6}{2} - \binom{5}{3} + \binom{4}{4}$	$= 1 - 7 + 15 - 10 + 1$	$= 0$

We'd better not try the next case, $n = 4$; the chances of making an arithmetic error are too high. (Computing terms like $\binom{12}{4}$ and $\binom{11}{5}$ by hand, let alone combining them with the others, is worthwhile only if we're desperate.)

So the pattern starts out $1, 0, -1, 0$. Even if we knew the next term or two, the closed form wouldn't be obvious. But if we could find and prove a recurrence for Q_n we'd probably be able to guess and prove its closed form. To find a recurrence, we need to relate Q_n to Q_{n-1} (or to $Q_{\text{smaller values}}$); but to do this we need to relate a term like $\binom{128-13}{13}$, which arises when $n = 7$ and $k = 13$, to terms like $\binom{64-13}{13}$. This doesn't look promising; we don't know any neat relations between entries in Pascal's triangle that are 64 rows apart. The addition formula, our main tool for induction proofs, only relates entries that are one row apart.

But this leads us to a key observation: There's no need to deal with entries that are 2^{n-1} rows apart. The variable n never appears by itself, it's always in the context 2^n. So the 2^n is a red herring! If we replace 2^n by m, all we need to do is find a closed form for the more general (but easier) sum

Oh, the sneakiness of the instructor who set that exam.

$$R_m = \sum_{k \leqslant m} \binom{m-k}{k} (-1)^k , \qquad \text{integer } m \geqslant 0;$$

then we'll also have a closed form for $Q_n = R_{2^n}$. And there's a good chance that the addition formula will give us a recurrence for the sequence R_m.

Values of R_m for small m can be read from Table 155, if we alternately add and subtract values that appear in a southwest-to-northeast diagonal. The results are:

m	0	1	2	3	4	5	6	7	8	9	10
R_m	1	1	0	-1	-1	0	1	1	0	-1	-1

There seems to be a lot of cancellation going on.

Let's look now at the formula for R_m and see if it defines a recurrence. Our strategy is to apply the addition formula (5.8) and to find sums that have the form R_k in the resulting expression, somewhat as we did in the perturbation method of Chapter 2:

$$\begin{aligned}
R_m &= \sum_{k\leqslant m}\binom{m-k}{k}(-1)^k\\
&= \sum_{k\leqslant m}\binom{m-1-k}{k}(-1)^k + \sum_{k\leqslant m}\binom{m-1-k}{k-1}(-1)^k\\
&= \sum_{k\leqslant m}\binom{m-1-k}{k}(-1)^k + \sum_{k+1\leqslant m}\binom{m-2-k}{k}(-1)^{k+1}\\
&= \sum_{k\leqslant m-1}\binom{m-1-k}{k}(-1)^k + \binom{-1}{m}(-1)^m\\
&\qquad - \sum_{k\leqslant m-2}\binom{m-2-k}{k}(-1)^k - \binom{-1}{m-1}(-1)^{m-1}\\
&= R_{m-1} + (-1)^{2m} - R_{m-2} - (-1)^{2(m-1)} = R_{m-1} - R_{m-2}\,.
\end{aligned}$$

(In the next-to-last step we've used the formula $\binom{-1}{m} = (-1)^m$, which we know is true when $m \geqslant 0$.) This derivation is valid for $m \geqslant 2$.

Anyway those of us who've done warmup exercise 4 know it.

From this recurrence we can generate values of R_m quickly, and we soon perceive that the sequence is periodic. Indeed,

$$R_m = \begin{cases} 1 \\ 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{cases} \quad \text{if } m \bmod 6 = \begin{cases} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{cases}.$$

The proof by induction is by inspection. Or, if we must give a more academic proof, we can unfold the recurrence one step to obtain

$$R_m = (R_{m-2} - R_{m-3}) - R_{m-2} = -R_{m-3}\,,$$

whenever $m \geqslant 3$. Hence $R_m = R_{m-6}$ whenever $m \geqslant 6$.

Finally, since $Q_n = R_{2^n}$, we can determine Q_n by determining $2^n \bmod 6$ and using the closed form for R_m. When $n = 0$ we have $2^0 \bmod 6 = 1$; after that we keep multiplying by 2 (mod 6), so the pattern 2, 4 repeats. Thus

$$Q_n = R_{2^n} = \begin{cases} R_1 = 1, & \text{if } n = 0;\\ R_2 = 0, & \text{if } n \text{ is odd};\\ R_4 = -1, & \text{if } n > 0 \text{ is even}.\end{cases}$$

This closed form for Q_n agrees with the first four values we calculated when we started on the problem. We conclude that $Q_{1000000} = R_4 = -1$.

Problem 4: A sum involving two binomial coefficients.

Our next task is to find a closed form for

$$\sum_{k=0}^{n} k\binom{m-k-1}{m-n-1}, \qquad \text{integers } m > n \geqslant 0.$$

Wait a minute. Where's the second binomial coefficient promised in the title of this problem? And why should we try to simplify a sum we've already simplified? (This is the sum S from Problem 2.)

Well, this is a sum that's easier to simplify if we view the summand as a product of two binomial coefficients, and then use one of the general identities found in Table 169. The second binomial coefficient materializes when we rewrite k as $\binom{k}{1}$:

$$\sum_{k=0}^{n} k\binom{m-k-1}{m-n-1} = \sum_{0\leqslant k\leqslant n} \binom{k}{1}\binom{m-k-1}{m-n-1}.$$

And identity (5.26) is the one to apply, since its index of summation appears in both upper indices and with opposite signs.

But our sum isn't quite in the correct form yet. The upper limit of summation should be $m-1$, if we're to have a perfect match with (5.26). No problem; the terms for $n < k \leqslant m-1$ are zero. So we can plug in, with $(l, m, n, q) \leftarrow (m-1, m-n-1, 1, 0)$; the answer is

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

This is cleaner than the formula we got before. We can convert it to the previous formula by using (5.7):

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

Similarly, we can get interesting results by plugging special values into the other general identities we've seen. Suppose, for example, that we set $m = n = 1$ and $q = 0$ in (5.26). Then the identity reads

$$\sum_{0\leqslant k\leqslant l} (l-k)k = \binom{l+1}{3}.$$

The left side is $l\big((l+1)l/2\big) - (1^2+2^2+\cdots+l^2)$, so this gives us a brand new way to solve the sum-of-squares problem that we beat to death in Chapter 2.

The moral of this story is: Special cases of very general sums are sometimes best handled in the general form. When learning general forms, it's wise to learn their simple specializations.

Problem 5: A sum with three factors.

Here's another sum that isn't too bad. We wish to simplify

$$\sum_k \binom{n}{k}\binom{s}{k}k\,, \qquad \text{integer } n \geqslant 0.$$

The index of summation k appears in both lower indices and with the same sign; therefore identity (5.23) in Table 169 looks close to what we need. With a bit of manipulation, we should be able to use it.

The biggest difference between (5.23) and what we have is the extra k in our sum. But we can absorb k into one of the binomial coefficients by using one of the absorption identities:

$$\begin{aligned}\sum_k \binom{n}{k}\binom{s}{k}k &= \sum_k \binom{n}{k}\binom{s-1}{k-1}s \\ &= s\sum_k \binom{n}{k}\binom{s-1}{k-1}.\end{aligned}$$

We don't care that the s appears when the k disappears, because it's constant. And now we're ready to apply the identity and get the closed form,

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

If we had chosen in the first step to absorb k into $\binom{n}{k}$, not $\binom{s}{k}$, we wouldn't have been allowed to apply (5.23) directly, because $n-1$ might be negative; the identity requires a nonnegative value in at least one of the upper indices.

Problem 6: A sum with menacing characteristics.

The next sum is more challenging. We seek a closed form for

$$\sum_{k\geqslant 0} \binom{n+k}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}\,, \qquad \text{integer } n \geqslant 0.$$

So we should deep six this sum, right?

One useful measure of a sum's difficulty is the number of times the index of summation appears. By this measure we're in deep trouble — k appears six times. Furthermore, the key step that worked in the previous problem — to absorb something outside the binomial coefficients into one of them — won't work here. If we absorb the $k+1$ we just get another occurrence of k in its place. And not only that: Our index k is twice shackled with the coefficient 2 inside a binomial coefficient. Multiplicative constants are usually harder to remove than additive constants.

We're lucky this time, though. The $2k$'s are right where we need them for identity (5.21) to apply, so we get

$$\sum_{k\geqslant 0}\binom{n+k}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1} \;=\; \sum_{k\geqslant 0}\binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1}\,.$$

The two 2's disappear, and so does one occurrence of k. So that's one down and five to go.

The $k+1$ in the denominator is the most troublesome characteristic left, and now we can absorb it into $\binom{n}{k}$ using identity (5.6):

$$\begin{aligned}\sum_{k\geqslant 0}\binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1} &= \sum_{k}\binom{n+k}{k}\binom{n+1}{k+1}\frac{(-1)^k}{n+1}\\ &= \frac{1}{n+1}\sum_{k}\binom{n+k}{k}\binom{n+1}{k+1}(-1)^k\,.\end{aligned}$$

(Recall that $n\geqslant 0$.) Two down, four to go.

To eliminate another k we have two promising options. We could use symmetry on $\binom{n+k}{k}$; or we could negate the upper index $n+k$, thereby eliminating that k as well as the factor $(-1)^k$. Let's explore both possibilities, starting with the symmetry option:

$$\frac{1}{n+1}\sum_{k}\binom{n+k}{k}\binom{n+1}{k+1}(-1)^k \;=\; \frac{1}{n+1}\sum_{k}\binom{n+k}{n}\binom{n+1}{k+1}(-1)^k\,.$$

Third down, three to go, and we're in position to make a big gain by plugging into (5.24): Replacing (l,m,n,s) by $(n+1,1,n,n)$, we get

For a minute I thought we'd have to punt.

$$\frac{1}{n+1}\sum_{k}\binom{n+k}{n}\binom{n+1}{k+1}(-1)^k \;=\; \frac{1}{n+1}(-1)^n\binom{n-1}{-1} \;=\; 0\,.$$

Zero, eh? After all that work? Let's check it when $n=2$: $\binom{2}{0}\binom{0}{0}\frac{1}{1}-\binom{3}{2}\binom{2}{1}\frac{1}{2}+\binom{4}{4}\binom{4}{2}\frac{1}{3}=1-\frac{6}{2}+\frac{6}{3}=0$. It checks.

Just for the heck of it, let's explore our other option, negating the upper index of $\binom{n+k}{k}$:

$$\frac{1}{n+1}\sum_{k}\binom{n+k}{k}\binom{n+1}{k+1}(-1)^k \;=\; \frac{1}{n+1}\sum_{k}\binom{-n-1}{k}\binom{n+1}{k+1}\,.$$

Now (5.23) applies, with $(l,m,n,s)\leftarrow(n+1,1,0,-n-1)$, and

$$\frac{1}{n+1}\sum_{k}\binom{-n-1}{k}\binom{n+1}{k+1} \;=\; \frac{1}{n+1}\binom{0}{n}\,.$$

Hey wait. This is zero when $n > 0$, but it's 1 when $n = 0$. Our other path to the solution told us that the sum was zero in all cases! What gives? The sum actually does turn out to be 1 when $n = 0$, so the correct answer is '$[n=0]$'. We must have made a mistake in the previous derivation.

Try binary search: Replay the middle formula first, to see if the mistake was early or late.

Let's do an instant replay on that derivation when $n = 0$, in order to see where the discrepancy first arises. Ah yes; we fell into the old trap mentioned earlier: We tried to apply symmetry when the upper index could be negative! We were not justified in replacing $\binom{n+k}{k}$ by $\binom{n+k}{n}$ when k ranges over all integers, because this converts zero into a nonzero value when $k < -n$. (Sorry about that.)

The other factor in the sum, $\binom{n+1}{k+1}$, turns out to be zero when $k < -n$, except when $n = 0$ and $k = -1$. Hence our error didn't show up when we checked the case $n = 2$. Exercise 6 explains what we should have done.

Problem 7: A new obstacle.

This one's even tougher; we want a closed form for

$$\sum_{k\geqslant 0}\binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}\,, \qquad \text{integers } m,n>0.$$

If m were 0 we'd have the sum from the problem we just finished. But it's not, and we're left with a real mess—nothing we used in Problem 6 works here. (Especially not the crucial first step.)

However, if we could somehow get rid of the m, we could use the result just derived. So our strategy is: Replace $\binom{n+k}{m+2k}$ by a sum of terms like $\binom{l+k}{2k}$ for some nonnegative integer l; the summand will then look like the summand in Problem 6, and we can interchange the order of summation.

What should we substitute for $\binom{n+k}{m+2k}$? A painstaking examination of the identities derived earlier in this chapter turns up only one suitable candidate, namely equation (5.26) in Table 169. And one way to use it is to replace the parameters (l, m, n, q, k) by $(n+k-1, 2k, m-1, 0, j)$, respectively:

$$\begin{aligned}
&\sum_{k\geqslant 0}\binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}\\
&\qquad=\sum_{k\geqslant 0}\;\sum_{0\leqslant j\leqslant n+k-1}\binom{n+k-1-j}{2k}\binom{j}{m-1}\binom{2k}{k}\frac{(-1)^k}{k+1}\\
&\qquad=\sum_{j\geqslant 0}\binom{j}{m-1}\sum_{\substack{k\geqslant j-n+1\\ k\geqslant 0}}\binom{n+k-1-j}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}\,.
\end{aligned}$$

In the last step we've changed the order of summation, manipulating the conditions below the $\sum$'s according to the rules of Chapter 2.

We can't quite replace the inner sum using the result of Problem 6, because it has the extra condition $k \geqslant j-n+1$. But this extra condition is superfluous unless $j-n+1>0$; that is, unless $j \geqslant n$. And when $j \geqslant n$, the first binomial coefficient of the inner sum is zero, because its upper index is between 0 and $k-1$, thus strictly less than the lower index $2k$. We may therefore place the additional restriction $j<n$ on the outer sum, without affecting which nonzero terms are included. This makes the restriction $k \geqslant j-n+1$ superfluous, and we can use the result of Problem 6. The double sum now comes tumbling down:

$$\begin{aligned}
\sum_{j\geqslant 0}\binom{j}{m-1} & \sum_{\substack{k\geqslant j-n+1\\ k\geqslant 0}}\binom{n+k-1-j}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}\\
&= \sum_{0\leqslant j<n}\binom{j}{m-1}\sum_{k\geqslant 0}\binom{n+k-1-j}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}\\
&= \sum_{0\leqslant j<n}\binom{j}{m-1}[n-1-j=0] \;=\; \binom{n-1}{m-1}.
\end{aligned}$$

The inner sums vanish except when $j=n-1$, so we get a simple closed form as our answer.

Problem 8: A different obstacle.

Let's branch out from Problem 6 in another way by considering the sum

$$S_m \;=\; \sum_{k\geqslant 0}\binom{n+k}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1+m}\,, \qquad \text{integers } m,n\geqslant 0.$$

Again, when $m=0$ we have the sum we did before; but now the m occurs in a different place. This problem is a bit harder yet than Problem 7, but (fortunately) we're getting better at finding solutions. We can begin as in Problem 6,

$$S_m \;=\; \sum_{k\geqslant 0}\binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1+m}\,.$$

Now (as in Problem 7) we try to expand the part that depends on m into terms that we know how to deal with. When m was zero, we absorbed $k+1$ into $\binom{n}{k}$; if $m>0$, we can do the same thing if we expand $1/(k+1+m)$ into absorbable terms. And our luck still holds: We proved a suitable identity

$$\sum_{j=0}^{m}\binom{m}{j}\binom{r}{j}^{-1} \;=\; \frac{r+1}{r+1-m}\,, \qquad \begin{array}{c}\text{integer } m\geqslant 0,\\ r\notin\{0,1,\ldots,m-1\}.\end{array} \tag{5.33}$$

in Problem 1. Replacing r by $-k-2$ gives the desired expansion,

$$S_m = \sum_{k\geqslant 0}\binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1}\sum_{j\geqslant 0}\binom{m}{j}\binom{-k-2}{j}^{-1}.$$

Now the $(k+1)^{-1}$ can be absorbed into $\binom{n}{k}$, as planned. In fact, it could also be absorbed into $\binom{-k-2}{j}^{-1}$. Double absorption suggests that even more cancellation might be possible behind the scenes. Yes — expanding everything in our new summand into factorials and going back to binomial coefficients gives a formula that we can sum on k:

They expect us to check this on a sheet of scratch paper.

$$\begin{aligned} S_m &= \frac{m!\,n!}{(m+n+1)!}\sum_{j\geqslant 0}(-1)^j\binom{m+n+1}{n+1+j}\sum_k\binom{n+1+j}{k+j+1}\binom{-n-1}{k} \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{j\geqslant 0}(-1)^j\binom{m+n+1}{n+1+j}\binom{j}{n}. \end{aligned}$$

The sum over all integers j is zero, by (5.24). Hence $-S_m$ is the sum for $j<0$.

To evaluate $-S_m$ for $j<0$, let's replace j by $-k-1$ and sum for $k\geqslant 0$:

$$\begin{aligned} S_m &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\geqslant 0}(-1)^k\binom{m+n+1}{n-k}\binom{-k-1}{n} \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\leqslant n}(-1)^{n-k}\binom{m+n+1}{k}\binom{k-n-1}{n} \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\leqslant n}(-1)^k\binom{m+n+1}{k}\binom{2n-k}{n} \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\leqslant 2n}(-1)^k\binom{m+n+1}{k}\binom{2n-k}{n}. \end{aligned}$$

Finally (5.25) applies, and we have our answer:

$$S_m = (-1)^n\frac{m!\,n!}{(m+n+1)!}\binom{m}{n} = (-1)^n m^{\underline{n}} m^{\underline{-n-1}}.$$

Whew; we'd better check it. When $n=2$ we find

$$S_m = \frac{1}{m+1}-\frac{6}{m+2}+\frac{6}{m+3} = \frac{m(m-1)}{(m+1)(m+2)(m+3)}.$$

Our derivation requires m to be an integer, but the result holds for all real m, because the quantity $(m+1)^{\overline{n+1}}S_m$ is a polynomial in m of degree $\leqslant n$.

5.3 TRICKS OF THE TRADE

Let's look next at three techniques that significantly amplify the methods we have already learned.

Trick 1: Going halves.

This should really be called Trick 1/2.

Many of our identities involve an arbitrary real number r. When r has the special form "integer minus one half," the binomial coefficient $\binom{r}{k}$ can be written as a quite different-looking product of binomial coefficients. This leads to a new family of identities that can be manipulated with surprising ease.

One way to see how this works is to begin with the *duplication formula*

$$r^{\underline{k}}\,(r-\tfrac{1}{2})^{\underline{k}} \;=\; (2r)^{\underline{2k}}/2^{2k}\,, \qquad \text{integer } k \geqslant 0. \tag{5.34}$$

This identity is obvious if we expand the falling powers and interleave the factors on the left side:

$$\begin{aligned} r(r-\tfrac{1}{2})(r-1)(r-\tfrac{3}{2})\ldots(r-k+1)(r-k+\tfrac{1}{2}) \\ = \frac{(2r)(2r-1)\ldots(2r-2k+1)}{2\cdot 2\cdot \ldots \cdot 2}\,. \end{aligned}$$

Now we can divide both sides by $k!^2$, and we get

$$\binom{r}{k}\binom{r-1/2}{k} \;=\; \binom{2r}{2k}\binom{2k}{k}\Big/2^{2k}\,, \qquad \text{integer } k. \tag{5.35}$$

If we set $k = r = n$, where n is an integer, this yields

$$\binom{n-1/2}{n} \;=\; \binom{2n}{n}\Big/2^{2n}\,, \qquad \text{integer } n. \tag{5.36}$$

And negating the upper index gives yet another useful formula,

$$\binom{-1/2}{n} \;=\; \left(\frac{-1}{4}\right)^{n}\binom{2n}{n}\,, \qquad \text{integer } n. \tag{5.37}$$

For example, when $n = 4$ we have

. . . we halve . . .

$$\begin{aligned} \binom{-1/2}{4} &= \frac{(-1/2)(-3/2)(-5/2)(-7/2)}{4!} \\ &= \left(\frac{-1}{2}\right)^{4}\frac{1\cdot 3\cdot 5\cdot 7}{1\cdot 2\cdot 3\cdot 4} \\ &= \left(\frac{-1}{4}\right)^{4}\frac{1\cdot 3\cdot 5\cdot 7\cdot 2\cdot 4\cdot 6\cdot 8}{1\cdot 2\cdot 3\cdot 4\cdot 1\cdot 2\cdot 3\cdot 4} \;=\; \left(\frac{-1}{4}\right)^{4}\binom{8}{4}\,. \end{aligned}$$

Notice how we've changed a product of odd numbers into a factorial.

Identity (5.35) has an amusing corollary. Let $r = \frac{1}{2}n$, and take the sum over all integers k. The result is

$$\sum_k \binom{n}{2k}\binom{2k}{k}2^{-2k} = \sum_k \binom{n/2}{k}\binom{(n-1)/2}{k}$$
$$= \binom{n-1/2}{\lfloor n/2 \rfloor}, \qquad \text{integer } n \geqslant 0 \tag{5.38}$$

by (5.23), because either $n/2$ or $(n-1)/2$ is $\lfloor n/2 \rfloor$, a nonnegative integer!

We can also use Vandermonde's convolution (5.27) to deduce that

$$\sum_k \binom{-1/2}{k}\binom{-1/2}{n-k} = \binom{-1}{n} = (-1)^n, \qquad \text{integer } n \geqslant 0.$$

Plugging in the values from (5.37) gives

$$\binom{-1/2}{k}\binom{-1/2}{n-k} = \left(\frac{-1}{4}\right)^k \binom{2k}{k} \left(\frac{-1}{4}\right)^{n-k} \binom{2(n-k)}{n-k}$$
$$= \frac{(-1)^n}{4^n}\binom{2k}{k}\binom{2n-2k}{n-k};$$

this is what sums to $(-1)^n$. Hence we have a remarkable property of the "middle" elements of Pascal's triangle:

$$\sum_k \binom{2k}{k}\binom{2n-2k}{n-k} = 4^n, \qquad \text{integer } n \geqslant 0. \tag{5.39}$$

For example, $\binom{0}{0}\binom{6}{3}+\binom{2}{1}\binom{4}{2}+\binom{4}{2}\binom{2}{1}+\binom{6}{3}\binom{0}{0} = 1\cdot20+2\cdot6+6\cdot2+20\cdot1 = 64 = 4^3$.

These illustrations of our first trick indicate that it's wise to try changing binomial coefficients of the form $\binom{2k}{k}$ into binomial coefficients of the form $\binom{n-1/2}{k}$, where n is some appropriate integer (usually 0, 1, or k); the resulting formula might be much simpler.

Trick 2: High-order differences.

We saw earlier that it's possible to evaluate partial sums of the series $\binom{n}{k}(-1)^k$, but not of the series $\binom{n}{k}$. It turns out that there are many important applications of binomial coefficients with alternating signs, $\binom{n}{k}(-1)^k$. One of the reasons for this is that such coefficients are intimately associated with the difference operator Δ defined in Section 2.6.

The difference Δf of a function f at the point x is

$$\Delta f(x) = f(x+1) - f(x);$$

if we apply Δ again, we get the second difference

$$\begin{aligned}\Delta^2 f(x) &= \Delta f(x+1) - \Delta f(x) = \big(f(x{+}2) - f(x{+}1)\big) - \big(f(x{+}1) - f(x)\big)\\ &= f(x+2) - 2f(x+1) + f(x)\,,\end{aligned}$$

which is analogous to the second derivative. Similarly, we have

$$\begin{aligned}\Delta^3 f(x) &= f(x+3) - 3f(x+2) + 3f(x+1) - f(x)\,;\\ \Delta^4 f(x) &= f(x+4) - 4f(x+3) + 6f(x+2) - 4f(x+1) + f(x)\,;\end{aligned}$$

and so on. Binomial coefficients enter these formulas with alternating signs.

In general, the nth difference is

$$\Delta^n f(x) = \sum_k \binom{n}{k} (-1)^{n-k} f(x+k)\,, \qquad \text{integer } n \geqslant 0. \tag{5.40}$$

This formula is easily proved by induction, but there's also a nice way to prove it directly using the elementary theory of operators. Recall that Section 2.6 defines the shift operator E by the rule

$$Ef(x) = f(x+1)\,;$$

hence the operator Δ is $E - 1$, where 1 is the identity operator defined by the rule $1f(x) = f(x)$. By the binomial theorem,

$$\Delta^n = (E-1)^n = \sum_k \binom{n}{k} E^k (-1)^{n-k}\,.$$

This is an equation whose elements are operators; it is equivalent to (5.40), since E^k is the operator that takes $f(x)$ into $f(x+k)$.

An interesting and important case arises when we consider negative falling powers. Let $f(x) = (x-1)^{\underline{-1}} = 1/x$. Then, by rule (2.45), we have $\Delta f(x) = (-1)(x-1)^{\underline{-2}}$, $\Delta^2 f(x) = (-1)(-2)(x-1)^{\underline{-3}}$, and in general

$$\Delta^n\big((x-1)^{\underline{-1}}\big) = (-1)^{\underline{n}}(x-1)^{\underline{-n-1}} = (-1)^n \frac{n!}{x(x+1)\ldots(x+n)}\,.$$

Equation (5.40) now tells us that

$$\begin{aligned}\sum_k \binom{n}{k} \frac{(-1)^k}{x+k} &= \frac{n!}{x(x+1)\ldots(x+n)}\\ &= x^{-1}\binom{x+n}{n}^{-1}, \qquad x \notin \{0,-1,\ldots,-n\}.\end{aligned} \tag{5.41}$$

For example,

$$\frac{1}{x}-\frac{4}{x+1}+\frac{6}{x+2}-\frac{4}{x+3}+\frac{1}{x+4}$$
$$=\frac{4!}{x(x+1)(x+2)(x+3)(x+4)} = 1\Big/x\binom{x+4}{4}.$$

The sum in (5.41) is the partial fraction expansion of $n!/\bigl(x(x+1)\ldots(x+n)\bigr)$.

Significant results can be obtained from positive falling powers too. If $f(x)$ is a polynomial of degree d, the difference $\Delta f(x)$ is a polynomial of degree $d-1$; therefore $\Delta^d f(x)$ is a constant, and $\Delta^n f(x)=0$ if $n>d$. This extremely important fact simplifies many formulas.

A closer look gives further information: Let

$$f(x) = a_d x^d + a_{d-1}x^{d-1} + \cdots + a_1 x^1 + a_0 x^0$$

be any polynomial of degree d. We will see in Chapter 6 that we can express ordinary powers as sums of falling powers (for example, $x^2 = x^{\underline{2}} + x^{\underline{1}}$); hence there are coefficients $b_d, b_{d-1}, \ldots, b_1, b_0$ such that

$$f(x) = b_d x^{\underline{d}} + b_{d-1}x^{\underline{d-1}} + \cdots + b_1 x^{\underline{1}} + b_0 x^{\underline{0}}.$$

(It turns out that $b_d = a_d$ and $b_0 = a_0$, but the intervening coefficients are related in a more complicated way.) Let $c_k = k!\,b_k$ for $0 \leqslant k \leqslant d$. Then

$$f(x) = c_d\binom{x}{d} + c_{d-1}\binom{x}{d-1} + \cdots + c_1\binom{x}{1} + c_0\binom{x}{0};$$

thus, any polynomial can be represented as a sum of multiples of binomial coefficients. Such an expansion is called the *Newton series* of $f(x)$, because Isaac Newton used it extensively.

We observed earlier in this chapter that the addition formula implies

$$\Delta\left(\binom{x}{k}\right) = \binom{x}{k-1}.$$

Therefore, by induction, the nth difference of a Newton series is very simple:

$$\Delta^n f(x) = c_d\binom{x}{d-n} + c_{d-1}\binom{x}{d-1-n} + \cdots + c_1\binom{x}{1-n} + c_0\binom{x}{-n}.$$

If we now set $x=0$, all terms $c_k\binom{x}{k-n}$ on the right side are zero, except the term with $k-n=0$; hence

$$\Delta^n f(0) = \begin{cases} c_n, & \text{if } n \leqslant d; \\ 0, & \text{if } n > d. \end{cases}$$

The Newton series for $f(x)$ is therefore

$$f(x) \;=\; \Delta^d f(0)\binom{x}{d} + \Delta^{d-1} f(0)\binom{x}{d-1} + \cdots + \Delta f(0)\binom{x}{1} + f(0)\binom{x}{0}.$$

For example, suppose $f(x) = x^3$. It's easy to calculate

$$\begin{array}{ccccccc} f(0)=0, & & f(1)=1, & & f(2)=8, & & f(3)=27; \\ & \Delta f(0)=1, & & \Delta f(1)=7, & & \Delta f(2)=19; & \\ & & \Delta^2 f(0)=6, & & \Delta^2 f(1)=12; & & \\ & & & \Delta^3 f(0)=6. & & & \end{array}$$

So the Newton series is $x^3 = 6\binom{x}{3} + 6\binom{x}{2} + 1\binom{x}{1} + 0\binom{x}{0}$.

Our formula $\Delta^n f(0) = c_n$ can also be stated in the following way, using (5.40) with $x = 0$:

$$\sum_k \binom{n}{k}(-1)^k\left(c_0\binom{k}{0} + c_1\binom{k}{1} + c_2\binom{k}{2} + \cdots\right) \;=\; (-1)^n c_n\,, \qquad \text{integer } n \geqslant 0.$$

Here $\langle c_0, c_1, c_2, \ldots\rangle$ is an arbitrary sequence of coefficients; the infinite sum $c_0\binom{k}{0} + c_1\binom{k}{1} + c_2\binom{k}{2} + \cdots$ is actually finite for all $k \geqslant 0$, so convergence is not an issue. In particular, we can prove the important identity

$$\sum_k \binom{n}{k}(-1)^k(a_0 + a_1 k + \cdots + a_n k^n) \;=\; (-1)^n n!\, a_n\,, \qquad \text{integer } n \geqslant 0, \tag{5.42}$$

because the polynomial $a_0 + a_1 k + \cdots + a_n k^n$ can always be written as a Newton series $c_0\binom{k}{0} + c_1\binom{k}{1} + \cdots + c_n\binom{k}{n}$ with $c_n = n!\, a_n$.

Many sums that appear to be hopeless at first glance can actually be summed almost trivially by using the idea of nth differences. For example, let's consider the identity

$$\sum_k \binom{n}{k}\binom{r-sk}{n}(-1)^k \;=\; s^n\,, \qquad \text{integer } n \geqslant 0. \tag{5.43}$$

This looks very impressive, because it's quite different from anything we've seen so far. But it really is easy to understand, once we notice the telltale factor $\binom{n}{k}(-1)^k$ in the summand, because the function

$$f(k) \;=\; \binom{r-sk}{n} \;=\; \frac{1}{n!}(-1)^n s^n k^n + \cdots \;=\; (-1)^n s^n\binom{k}{n} + \cdots$$

is a polynomial in k of degree n, with leading coefficient $(-1)^n s^n/n!$. Therefore (5.43) is nothing more than an application of (5.42).

We have discussed Newton series under the assumption that $f(x)$ is a polynomial. But we've also seen that infinite Newton series

$$f(x) \;=\; c_0\binom{x}{0} + c_1\binom{x}{1} + c_2\binom{x}{2} + \cdots$$

make sense too, because such sums are always finite when x is a nonnegative integer. Our derivation of the formula $\Delta^n f(0) = c_n$ works in the infinite case, just as in the polynomial case; so we have the general identity

$$f(x) \;=\; f(0)\binom{x}{0} + \Delta f(0)\binom{x}{1} + \Delta^2 f(0)\binom{x}{2} + \Delta^3 f(0)\binom{x}{3} + \cdots\,, \qquad \text{integer } x \geqslant 0. \tag{5.44}$$

This formula is valid for any function $f(x)$ that is defined for nonnegative integers x. Moreover, if the right-hand side converges for other values of x, it defines a function that "interpolates" $f(x)$ in a natural way. (There are infinitely many ways to interpolate function values, so we cannot assert that (5.44) is true for all x that make the infinite series converge. For example, if we let $f(x) = \sin(\pi x)$, we have $f(x) = 0$ at all integer points, so the right-hand side of (5.44) is identically zero; but the left-hand side is nonzero at all noninteger x.)

A Newton series is finite calculus's answer to infinite calculus's Taylor series. Just as a Taylor series can be written

$$g(a+x) \;=\; \frac{g(a)}{0!}x^0 + \frac{g'(a)}{1!}x^1 + \frac{g''(a)}{2!}x^2 + \frac{g'''(a)}{3!}x^3 + \cdots\,,$$

(Since $E = 1 + \Delta$, $E^x = \sum_k \binom{x}{k}\Delta^k$; *and* $E^x g(a) = g(a+x)$.*)*

the Newton series for $f(x) = g(a+x)$ can be written

$$g(a+x) \;=\; \frac{g(a)}{0!}x^{\underline{0}} + \frac{\Delta g(a)}{1!}x^{\underline{1}} + \frac{\Delta^2 g(a)}{2!}x^{\underline{2}} + \frac{\Delta^3 g(a)}{3!}x^{\underline{3}} + \cdots\,. \tag{5.45}$$

(This is the same as (5.44), because $\Delta^n f(0) = \Delta^n g(a)$ for all $n \geqslant 0$ when $f(x) = g(a+x)$.) Both the Taylor and Newton series are finite when g is a polynomial, or when $x = 0$; in addition, the Newton series is finite when x is a positive integer. Otherwise the sums may or may not converge for particular values of x. If the Newton series converges when x is not a nonnegative integer, it might actually converge to a value that's *different* from $g(a+x)$, because the Newton series (5.45) depends only on the spaced-out function values $g(a)$, $g(a+1)$, $g(a+2)$, ...

One example of a convergent Newton series is provided by the binomial theorem. Let $g(x) = (1+z)^x$, where z is a fixed complex number such that $|z| < 1$. Then $\Delta g(x) = (1+z)^{x+1} - (1+z)^x = z(1+z)^x$, hence $\Delta^n g(x) = z^n(1+z)^x$. In this case the infinite Newton series

$$g(a+x) \;=\; \sum_n \Delta^n g(a)\binom{x}{n} \;=\; (1+z)^a \sum_n \binom{x}{n} z^n$$

converges to the "correct" value $(1+z)^{a+x}$, for all x.

James Stirling tried to use Newton series to generalize the factorial function to noninteger values. First he found coefficients S_n such that

$$x! \;=\; \sum_n S_n \binom{x}{n} \;=\; S_0\binom{x}{0} + S_1\binom{x}{1} + S_2\binom{x}{2} + \cdots \tag{5.46}$$

is an identity for $x = 0$, $x = 1$, $x = 2$, etc. But he discovered that the resulting series doesn't converge except when x is a nonnegative integer. So he tried again, this time writing

"Forasmuch as these terms increase very fast, their differences will make a diverging progression, which hinders the ordinate of the parabola from approaching to the truth; therefore in this and the like cases, I interpolate the logarithms of the terms, whose differences constitute a series swiftly converging."
—J. Stirling [343]

$$\ln x! \;=\; \sum_n s_n \binom{x}{n} \;=\; s_0\binom{x}{0} + s_1\binom{x}{1} + s_2\binom{x}{2} + \cdots\,. \tag{5.47}$$

Now $\Delta(\ln x!) = \ln(x+1)! - \ln x! = \ln(x+1)$, hence

$$\begin{aligned} s_n \;&=\; \Delta^n(\ln x!)\big|_{x=0} \\ &=\; \Delta^{n-1}\bigl(\ln(x+1)\bigr)\big|_{x=0} \\ &=\; \sum_k \binom{n-1}{k}(-1)^{n-1-k}\ln(k+1) \end{aligned}$$

by (5.40). The coefficients are therefore $s_0 = s_1 = 0$; $s_2 = \ln 2$; $s_3 = \ln 3 - 2\ln 2 = \ln\frac{3}{4}$; $s_4 = \ln 4 - 3\ln 3 + 3\ln 2 = \ln\frac{32}{27}$; etc. In this way Stirling obtained a series that does converge (although he didn't prove it); in fact, his series converges for all $x > -1$. He was thereby able to evaluate $\frac{1}{2}!$ satisfactorily. Exercise 88 tells the rest of the story.

(Proofs of convergence were not invented until the nineteenth century.)

Trick 3: Inversion.

A special case of the rule (5.45) we've just derived for Newton's series can be rewritten in the following way:

$$g(n) \;=\; \sum_k \binom{n}{k}(-1)^k f(k) \quad\Longleftrightarrow\quad f(n) \;=\; \sum_k \binom{n}{k}(-1)^k g(k)\,. \tag{5.48}$$

This dual relationship between f and g is called an *inversion formula*; it's rather like the Möbius inversion formulas (4.56) and (4.61) that we encountered in Chapter 4. Inversion formulas tell us how to solve "implicit recurrences," where an unknown sequence is embedded in a sum.

Invert this: 'zınb ppo'.

For example, $g(n)$ might be a known function, and $f(n)$ might be unknown; and we might have found a way to show that $g(n) = \sum_k \binom{n}{k}(-1)^k f(k)$. Then (5.48) lets us express $f(n)$ as a sum of known values.

We can prove (5.48) directly by using the basic methods at the beginning of this chapter. If $g(n) = \sum_k \binom{n}{k}(-1)^k f(k)$ for all $n \geqslant 0$, then

$$\begin{aligned}
\sum_k \binom{n}{k}(-1)^k g(k) &= \sum_k \binom{n}{k}(-1)^k \sum_j \binom{k}{j}(-1)^j f(j) \\
&= \sum_j f(j) \sum_k \binom{n}{k}(-1)^{k+j}\binom{k}{j} \\
&= \sum_j f(j) \sum_k \binom{n}{j}(-1)^{k+j}\binom{n-j}{k-j} \\
&= \sum_j f(j)\binom{n}{j} \sum_k (-1)^k \binom{n-j}{k} \\
&= \sum_j f(j)\binom{n}{j}[n-j=0] \quad = \quad f(n)\,.
\end{aligned}$$

The proof in the other direction is, of course, the same, because the relation between f and g is symmetric.

Let's illustrate (5.48) by applying it to the "football victory problem": A group of n fans of the winning football team throw their hats high into the air. The hats come back randomly, one hat to each of the n fans. How many ways $h(n,k)$ are there for exactly k fans to get their own hats back?

For example, if $n = 4$ and if the hats and fans are named A, B, C, D, the $4! = 24$ possible ways for hats to land generate the following numbers of rightful owners:

ABCD	4	BACD	2	CABD	1	DABC	0
ABDC	2	BADC	0	CADB	0	DACB	1
ACBD	2	BCAD	1	CBAD	2	DBAC	1
ACDB	1	BCDA	0	CBDA	1	DBCA	2
ADBC	1	BDAC	0	CDAB	0	DCAB	0
ADCB	2	BDCA	1	CDBA	0	DCBA	0

Therefore $h(4,4) = 1$; $h(4,3) = 0$; $h(4,2) = 6$; $h(4,1) = 8$; $h(4,0) = 9$.

We can determine $h(n,k)$ by noticing that it is the number of ways to choose k lucky hat owners, namely $\binom{n}{k}$, times the number of ways to arrange the remaining $n-k$ hats so that none of them goes to the right owner, namely $h(n-k,0)$. A permutation is called a *derangement* if it moves every item, and the number of derangements of n objects is sometimes denoted by the symbol '$n¡$', read "n subfactorial." Therefore $h(n-k,0) = (n-k)¡$, and we have the general formula

$$h(n,k) \;=\; \binom{n}{k} h(n-k,0) \;=\; \binom{n}{k}(n-k)¡\,.$$

(Subfactorial notation isn't standard, and it's not clearly a great idea; but let's try it awhile to see if we grow to like it. We can always resort to 'D_n' or something, if '$n¡$' doesn't work out.)

Our problem would be solved if we had a closed form for $n¡$, so let's see what we can find. There's an easy way to get a recurrence, because the sum of $h(n,k)$ for all k is the total number of permutations of n hats:

$$\begin{aligned} n! \;=\; \sum_k h(n,k) \;&=\; \sum_k \binom{n}{k}(n-k)¡ \\ &=\; \sum_k \binom{n}{k} k¡\,, \qquad \text{integer } n \geqslant 0. \end{aligned} \tag{5.49}$$

(We've changed k to $n-k$ and $\binom{n}{n-k}$ to $\binom{n}{k}$ in the last step.) With this implicit recurrence we can compute all the $h(n,k)$'s we like:

n	$h(n,0)$	$h(n,1)$	$h(n,2)$	$h(n,3)$	$h(n,4)$	$h(n,5)$	$h(n,6)$
0	1						
1	0	1					
2	1	0	1				
3	2	3	0	1			
4	9	8	6	0	1		
5	44	45	20	10	0	1	
6	265	264	135	40	15	0	1

For example, here's how the row for $n=4$ can be computed: The two rightmost entries are obvious—there's just one way for all hats to land correctly, and there's no way for just three fans to get their own. (Whose hat would the fourth fan get?) When $k=2$ and $k=1$, we can use our equation for $h(n,k)$, giving $h(4,2) = \binom{4}{2}h(2,0) = 6\cdot 1 = 6$, and $h(4,1) = \binom{4}{1}h(3,0) = 4\cdot 2 = 8$. We can't use this equation for $h(4,0)$; rather, we can, but it gives us $h(4,0) = \binom{4}{0}h(4,0)$, which is true but useless. Taking another tack, we can use the relation $h(4,0)+8+6+0+1 = 4!$ to deduce that $h(4,0) = 9$; this is the value of $4¡$. Similarly $n¡$ depends on the values of $k¡$ for $k<n$.

The art of mathematics, as of life, is knowing which truths are useless.

How can we solve a recurrence like (5.49)? Easy; it has the form of (5.48), with $g(n) = n!$ and $f(k) = (-1)^k k¡$. Hence its solution is

$$n¡ \;=\; (-1)^n \sum_k \binom{n}{k} (-1)^k k!\,.$$

Well, this isn't really a solution; it's a sum that should be put into closed form if possible. But it's better than a recurrence. The sum can be simplified, since $k!$ cancels with a hidden $k!$ in $\binom{n}{k}$, so let's try that: We get

$$n¡ \;=\; \sum_{0\leqslant k\leqslant n} \frac{n!}{(n-k)!}(-1)^{n+k} \;=\; n! \sum_{0\leqslant k\leqslant n} \frac{(-1)^k}{k!}\,. \qquad (5.50)$$

The remaining sum converges rapidly to the number $\sum_{k\geqslant 0}(-1)^k/k! = e^{-1}$. In fact, the terms that are excluded from the sum are

$$\begin{aligned} n! \sum_{k>n} \frac{(-1)^k}{k!} \;&=\; \frac{(-1)^{n+1}}{n+1} \sum_{k\geqslant 0} (-1)^k \frac{(n+1)!}{(k+n+1)!} \\ &=\; \frac{(-1)^{n+1}}{n+1}\left(1 - \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} - \cdots\right), \end{aligned}$$

and the parenthesized quantity lies between 1 and $1 - \frac{1}{n+2} = \frac{n+1}{n+2}$. Therefore the difference between $n¡$ and $n!/e$ is roughly $1/n$ in absolute value; more precisely, it lies between $1/(n+1)$ and $1/(n+2)$. But $n¡$ is an integer. Therefore it must be what we get when we round $n!/e$ to the nearest integer, if $n > 0$. So we have the closed form we seek:

$$n¡ \;=\; \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor \;+\; [n=0]\,. \qquad (5.51)$$

This is the number of ways that no fan gets the right hat back. When n is large, it's more meaningful to know the *probability* that this happens. If we assume that each of the $n!$ arrangements is equally likely — because the hats were thrown extremely high — this probability is

Baseball fans: .367 is also Ty Cobb's lifetime batting average, the all-time record. Can this be a coincidence?

$$\frac{n¡}{n!} \;=\; \frac{n!/e + O(1)}{n!} \;\sim\; \frac{1}{e} \;=\; .367\ldots\,.$$

(Hey wait, you're fudging. Cobb's average was $4191/11429 \approx .366699$, while $1/e \approx .367879$. But maybe if Wade Boggs has a few really good seasons. . .)

So when n gets large the probability that all hats are misplaced is almost 37%.

Incidentally, recurrence (5.49) for subfactorials is exactly the same as (5.46), the first recurrence considered by Stirling when he was trying to generalize the factorial function. Hence $S_k = k¡$. These coefficients are so large, it's no wonder the infinite series (5.46) diverges for noninteger x.

Before leaving this problem, let's look briefly at two interesting patterns that leap out at us in the table of small $h(n,k)$. First, it seems that the numbers 1, 3, 6, 10, 15, . . . below the all-0 diagonal are the triangular numbers.

This observation is easy to prove, since those table entries are the $h(n, n-2)$'s, and we have

$$h(n, n-2) \;=\; \binom{n}{n-2} 2¡ \;=\; \binom{n}{2}.$$

It also seems that the numbers in the first two columns differ by ± 1. Is this always true? Yes,

$$\begin{aligned} h(n,0) - h(n,1) \;&=\; n¡ - n(n-1)¡ \\ &=\; \Bigl(n! \sum_{0 \leqslant k \leqslant n} \frac{(-1)^k}{k!}\Bigr) - \Bigl(n(n-1)! \sum_{0 \leqslant k \leqslant n-1} \frac{(-1)^k}{k!}\Bigr) \\ &=\; n! \frac{(-1)^n}{n!} \;=\; (-1)^n. \end{aligned}$$

In other words, $n¡ = n(n-1)¡ + (-1)^n$. This is a much simpler recurrence for the derangement numbers than we had before.

Now let's invert something else. If we apply inversion to the formula

But inversion is the source of smog.

$$\sum_k \binom{n}{k} \frac{(-1)^k}{x+k} \;=\; \frac{1}{x} \binom{x+n}{n}^{-1}$$

that we derived in (5.41), we find

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

This is interesting, but not really new. If we negate the upper index in $\binom{x+k}{k}$, we have merely discovered identity (5.33) again.

5.4 GENERATING FUNCTIONS

We come now to the most important idea in this whole book, the notion of a *generating function*. An infinite sequence $\langle a_0, a_1, a_2, \ldots \rangle$ that we wish to deal with in some way can conveniently be represented as a *power series* in an auxiliary variable z,

$$A(z) \;=\; a_0 + a_1 z + a_2 z^2 + \cdots \;=\; \sum_{k \geqslant 0} a_k z^k. \tag{5.52}$$

It's appropriate to use the letter z as the name of the auxiliary variable, because we'll often be thinking of z as a complex number. The theory of complex variables conventionally uses 'z' in its formulas; power series (a.k.a. analytic functions or holomorphic functions) are central to that theory.

We will be seeing lots of generating functions in subsequent chapters. Indeed, Chapter 7 is entirely devoted to them. Our present goal is simply to introduce the basic concepts, and to demonstrate the relevance of generating functions to the study of binomial coefficients.

A generating function is useful because it's a single quantity that represents an entire infinite sequence. We can often solve problems by first setting up one or more generating functions, then by fooling around with those functions until we know a lot about them, and finally by looking again at the coefficients. With a little bit of luck, we'll know enough about the function to understand what we need to know about its coefficients.

(See [223] for a discussion of the history and usefulness of this notation.)

If $A(z)$ is any power series $\sum_{k\geqslant 0} a_k z^k$, we will find it convenient to write

$$[z^n]\,A(z) \;=\; a_n\,; \tag{5.53}$$

in other words, $[z^n]\,A(z)$ denotes the coefficient of z^n in $A(z)$.

Let $A(z)$ be the generating function for $\langle a_0, a_1, a_2, \ldots\rangle$ as in (5.52), and let $B(z)$ be the generating function for another sequence $\langle b_0, b_1, b_2, \ldots\rangle$. Then the product $A(z)B(z)$ is the power series

$$\begin{aligned}(a_0 + a_1 z + a_2 z^2 + \cdots)&(b_0 + b_1 z + b_2 z^2 + \cdots)\\ &= a_0 b_0 \;+\; (a_0 b_1 + a_1 b_0)z \;+\; (a_0 b_2 + a_1 b_1 + a_2 b_0)z^2 \;+\; \cdots\,;\end{aligned}$$

the coefficient of z^n in this product is

$$a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 \;=\; \sum_{k=0}^{n} a_k b_{n-k}\,.$$

Therefore if we wish to evaluate any sum that has the general form

$$c_n \;=\; \sum_{k=0}^{n} a_k b_{n-k}\,, \tag{5.54}$$

and if we know the generating functions $A(z)$ and $B(z)$, we have

$$c_n \;=\; [z^n]\,A(z)B(z)\,.$$

The sequence $\langle c_n\rangle$ defined by (5.54) is called the *convolution* of the sequences $\langle a_n\rangle$ and $\langle b_n\rangle$; two sequences are "convolved" by forming the sums of all products whose subscripts add up to a given amount. The gist of the previous paragraph is that convolution of sequences corresponds to multiplication of their generating functions.

Generating functions give us powerful ways to discover and/or prove identities. For example, the binomial theorem tells us that $(1+z)^r$ is the generating function for the sequence $\langle\binom{r}{0},\binom{r}{1},\binom{r}{2},\ldots\rangle$:

$$(1+z)^r \;=\; \sum_{k\geqslant 0}\binom{r}{k}z^k\,.$$

Similarly,

$$(1+z)^s \;=\; \sum_{k\geqslant 0}\binom{s}{k}z^k\,.$$

If we multiply these together, we get another generating function:

$$(1+z)^r(1+z)^s \;=\; (1+z)^{r+s}\,.$$

And now comes the punch line: Equating coefficients of z^n on both sides of this equation gives us

$$\sum_{k=0}^{n}\binom{r}{k}\binom{s}{n-k} \;=\; \binom{r+s}{n}\,.$$

We've discovered Vandermonde's convolution, (5.27)!

(5.27)! = (5.27)(4.27) (3.27)(2.27) (1.27)(0.27)!.

That was nice and easy; let's try another. This time we use $(1-z)^r$, which is the generating function for the sequence $\langle(-1)^n\binom{r}{n}\rangle = \langle\binom{r}{0},-\binom{r}{1},\binom{r}{2},\ldots\rangle$. Multiplying by $(1+z)^r$ gives another generating function whose coefficients we know:

$$(1-z)^r(1+z)^r \;=\; (1-z^2)^r\,.$$

Equating coefficients of z^n now gives the equation

$$\sum_{k=0}^{n}\binom{r}{k}\binom{r}{n-k}(-1)^k \;=\; (-1)^{n/2}\binom{r}{n/2}[n\text{ even}]\,. \tag{5.55}$$

We should check this on a small case or two. When $n=3$, for example, the result is

$$\binom{r}{0}\binom{r}{3}-\binom{r}{1}\binom{r}{2}+\binom{r}{2}\binom{r}{1}-\binom{r}{3}\binom{r}{0} \;=\; 0\,.$$

Each positive term is cancelled by a corresponding negative term. And the same thing happens whenever n is odd, in which case the sum isn't very interesting. But when n is even, say $n = 2$, we get a nontrivial sum that's different from Vandermonde's convolution:

$$\binom{r}{0}\binom{r}{2} - \binom{r}{1}\binom{r}{1} + \binom{r}{2}\binom{r}{0} = 2\binom{r}{2} - r^2 = -r\,.$$

So (5.55) checks out fine when $n = 2$. It turns out that (5.30) is a special case of our new identity (5.55).

Binomial coefficients also show up in some other generating functions, most notably the following important identities in which the lower index stays fixed and the upper index varies:

If you have a highlighter pen, these two equations have got to be marked.

$$\frac{1}{(1-z)^{n+1}} = \sum_{k\geqslant 0}\binom{n+k}{n}z^k\,, \qquad \text{integer } n \geqslant 0 \tag{5.56}$$

$$\frac{z^n}{(1-z)^{n+1}} = \sum_{k\geqslant 0}\binom{k}{n}z^k\,, \qquad \text{integer } n \geqslant 0. \tag{5.57}$$

The second identity here is just the first one multiplied by z^n, that is, "shifted right" by n places. The first identity is just a special case of the binomial theorem in slight disguise: If we expand $(1-z)^{-n-1}$ by (5.13), the coefficient of z^k is $\binom{-n-1}{k}(-1)^k$, which can be rewritten as $\binom{k+n}{k}$ or $\binom{n+k}{n}$ by negating the upper index. These special cases are worth noting explicitly, because they arise so frequently in applications.

When $n = 0$ we get a special case of a special case, the geometric series:

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

This is the generating function for the sequence $\langle 1, 1, 1, \ldots\rangle$, and it is especially useful because the convolution of any other sequence with this one is the sequence of sums: When $b_k = 1$ for all k, (5.54) reduces to

$$c_n = \sum_{k=0}^{n} a_k\,.$$

Therefore if $A(z)$ is the generating function for the summands $\langle a_0, a_1, a_2, \ldots\rangle$, then $A(z)/(1-z)$ is the generating function for the sums $\langle c_0, c_1, c_2, \ldots\rangle$.

The problem of derangements, which we solved by inversion in connection with hats and football fans, can be resolved with generating functions in an interesting way. The basic recurrence

$$n! = \sum_k \binom{n}{k}(n-k)¡$$

can be put into the form of a convolution if we expand $\binom{n}{k}$ in factorials and divide both sides by $n!$:

$$1 \;=\; \sum_{k=0}^{n} \frac{1}{k!}\,\frac{(n-k)¡}{(n-k)!}\,.$$

The generating function for the sequence $\langle \frac{1}{0!}, \frac{1}{1!}, \frac{1}{2!}, \ldots \rangle$ is e^z; hence if we let

$$D(z) \;=\; \sum_{k\geqslant 0} \frac{k¡}{k!} z^k\,,$$

the convolution/recurrence tells us that

$$\frac{1}{1-z} \;=\; e^z\, D(z)\,.$$

Solving for $D(z)$ gives

$$D(z) \;=\; \frac{1}{1-z}\, e^{-z} \;=\; \frac{1}{1-z}\left(\frac{1}{0!}z^0 - \frac{1}{1!}z^1 + \frac{1}{2!}z^2 + \cdots\right).$$

Equating coefficients of z^n now tells us that

$$\frac{n¡}{n!} \;=\; \sum_{k=0}^{n} \frac{(-1)^k}{k!}\,;$$

this is the formula we derived earlier by inversion.

So far our explorations with generating functions have given us slick proofs of things that we already knew how to derive by more cumbersome methods. But we haven't used generating functions to obtain any new results, except for (5.55). Now we're ready for something new and more surprising. There are two families of power series that generate an especially rich class of binomial coefficient identities: Let us define the *generalized binomial series* $\mathcal{B}_t(z)$ and the *generalized exponential series* $\mathcal{E}_t(z)$ as follows:

$$\mathcal{B}_t(z) \;=\; \sum_{k\geqslant 0} (tk)^{\underline{k-1}}\,\frac{z^k}{k!}\,; \qquad \mathcal{E}_t(z) \;=\; \sum_{k\geqslant 0} (tk+1)^{k-1}\,\frac{z^k}{k!}\,. \tag{5.58}$$

It can be shown that these functions satisfy the identities

$$\mathcal{B}_t(z)^{1-t} - \mathcal{B}_t(z)^{-t} \;=\; z\,; \qquad \mathcal{E}_t(z)^{-t}\ln \mathcal{E}_t(z) \;=\; z\,. \tag{5.59}$$

In the special case $t=0$, we have

$$\mathcal{B}_0(z) \;=\; 1+z\,; \qquad \mathcal{E}_0(z) \;=\; e^z\,;$$

this explains why the series with parameter t are called "generalized" binomials and exponentials.

The generalized binomial series $\mathcal{B}_t(z)$ was discovered in the 1750s by J. H. Lambert [236, §38], who noticed a few years later [237] that its powers satisfy the first identity in (5.60).

The following pairs of identities are valid for all real r:

$$\mathcal{B}_t(z)^r = \sum_{k\geqslant 0}\binom{tk+r}{k}\frac{r}{tk+r}z^k\,;$$

$$\mathcal{E}_t(z)^r = \sum_{k\geqslant 0} r\frac{(tk+r)^{k-1}}{k!}z^k\,; \qquad (5.60)$$

$$\frac{\mathcal{B}_t(z)^r}{1-t+t\mathcal{B}_t(z)^{-1}} = \sum_{k\geqslant 0}\binom{tk+r}{k}z^k\,;$$

$$\frac{\mathcal{E}_t(z)^r}{1-zt\mathcal{E}_t(z)^t} = \sum_{k\geqslant 0}\frac{(tk+r)^k}{k!}z^k\,. \qquad (5.61)$$

(When $tk+r=0$, we have to be a little careful about how the coefficient of z^k is interpreted; each coefficient is a polynomial in r. For example, the constant term of $\mathcal{E}_t(z)^r$ is $r(0+r)^{-1}$, and this is equal to 1 even when $r=0$.)

Since equations (5.60) and (5.61) hold for all r, we get very general identities when we multiply together the series that correspond to different powers r and s. For example,

$$\begin{aligned}\mathcal{B}_t(z)^r\frac{\mathcal{B}_t(z)^s}{1-t+t\mathcal{B}_t(z)^{-1}} &= \sum_{k\geqslant 0}\binom{tk+r}{k}\frac{r}{tk+r}z^k\sum_{j\geqslant 0}\binom{tj+s}{j}z^j\\ &= \sum_{n\geqslant 0}z^n\sum_{k\geqslant 0}\binom{tk+r}{k}\frac{r}{tk+r}\binom{t(n-k)+s}{n-k}\,.\end{aligned}$$

This power series must equal

$$\frac{\mathcal{B}_t(z)^{r+s}}{1-t+t\mathcal{B}_t(z)^{-1}} = \sum_{n\geqslant 0}\binom{tn+r+s}{n}z^n\,;$$

hence we can equate coefficients of z^n and get the identity

$$\sum_k\binom{tk+r}{k}\binom{t(n-k)+s}{n-k}\frac{r}{tk+r} = \binom{tn+r+s}{n}\,, \qquad \text{integer } n,$$

valid for all real r, s, and t. When $t=0$ this identity reduces to Vandermonde's convolution. (If by chance $tk+r$ happens to equal zero in this formula, the denominator factor $tk+r$ should be considered to cancel with the $tk+r$ in the numerator of the binomial coefficient. Both sides of the identity are polynomials in r, s, and t.) Similar identities hold when we multiply $\mathcal{B}_t(z)^r$ by $\mathcal{B}_t(z)^s$, etc.; Table 202 presents the results.

Table 202 General convolution identities, valid for integer $n \geqslant 0$.

$$\sum_k \binom{tk+r}{k}\binom{tn-tk+s}{n-k}\frac{r}{tk+r} = \binom{tn+r+s}{n}. \tag{5.62}$$

$$\sum_k \binom{tk+r}{k}\binom{tn-tk+s}{n-k}\frac{r}{tk+r}\cdot\frac{s}{tn-tk+s}$$
$$= \binom{tn+r+s}{n}\frac{r+s}{tn+r+s}. \tag{5.63}$$

$$\sum_k \binom{n}{k}(tk+r)^k(tn-tk+s)^{n-k}\frac{r}{tk+r} = (tn+r+s)^n. \tag{5.64}$$

$$\sum_k \binom{n}{k}(tk+r)^k(tn-tk+s)^{n-k}\frac{r}{tk+r}\cdot\frac{s}{tn-tk+s}$$
$$= (tn+r+s)^n\frac{r+s}{tn+r+s}. \tag{5.65}$$

We have learned that it's generally a good idea to look at special cases of general results. What happens, for example, if we set $t = 1$? The generalized binomial $\mathcal{B}_1(z)$ is very simple—it's just

$$\mathcal{B}_1(z) = \sum_{k\geqslant 0} z^k = \frac{1}{1-z};$$

therefore $\mathcal{B}_1(z)$ doesn't give us anything we didn't already know from Vandermonde's convolution. But $\mathcal{E}_1(z)$ is an important function,

$$\mathcal{E}(z) = \sum_{k\geqslant 0}(k+1)^{k-1}\frac{z^k}{k!} = 1 + z + \frac{3}{2}z^2 + \frac{8}{3}z^3 + \frac{125}{24}z^4 + \cdots \tag{5.66}$$

which we haven't seen before; it satisfies the basic identity

$$\mathcal{E}(z) = e^{z\mathcal{E}(z)}. \tag{5.67}$$

Aha! This is the iterated power function $\mathcal{E}(\ln z) = z^{z^{z^{\cdot^{\cdot^{\cdot}}}}}$ *that I've often wondered about.*

Zzzzzz. . .

This function, first studied by Euler [117] and Eisenstein [91], arises in a great many applications [203, 193].

The special cases $t = 2$ and $t = -1$ of the generalized binomial are of particular interest, because their coefficients occur again and again in problems that have a recursive structure. Therefore it's useful to display these

The power series for $\mathcal{B}_{1/2}(z)^r = (\sqrt{z^2+4}+z)^{2r}/4^r$ is noteworthy too.

series explicitly for future reference:

$$\begin{aligned}\mathcal{B}_2(z) &= \sum_k \binom{2k}{k}\frac{z^k}{1+k}\\ &= \sum_k \binom{2k+1}{k}\frac{z^k}{1+2k} = \frac{1-\sqrt{1-4z}}{2z}\,. \end{aligned}\qquad(5.68)$$

$$\begin{aligned}\mathcal{B}_{-1}(z) &= \sum_k \binom{1-k}{k}\frac{z^k}{1-k}\\ &= \sum_k \binom{2k-1}{k}\frac{(-z)^k}{1-2k} = \frac{1+\sqrt{1+4z}}{2}\,. \end{aligned}\qquad(5.69)$$

$$\mathcal{B}_2(z)^r = \sum_k \binom{2k+r}{k}\frac{r}{2k+r}\,z^k\,. \qquad(5.70)$$

$$\mathcal{B}_{-1}(z)^r = \sum_k \binom{r-k}{k}\frac{r}{r-k}\,z^k\,. \qquad(5.71)$$

$$\frac{\mathcal{B}_2(z)^r}{\sqrt{1-4z}} = \sum_k \binom{2k+r}{k} z^k\,. \qquad(5.72)$$

$$\frac{\mathcal{B}_{-1}(z)^{r+1}}{\sqrt{1+4z}} = \sum_k \binom{r-k}{k} z^k\,. \qquad(5.73)$$

The coefficients $\binom{2n}{n}\frac{1}{n+1}$ of $\mathcal{B}_2(z)$ are called the *Catalan numbers* C_n, because Eugène Catalan wrote an influential paper about them in the 1830s [52]. The sequence begins as follows:

n	0	1	2	3	4	5	6	7	8	9	10
C_n	1	1	2	5	14	42	132	429	1430	4862	16796

The coefficients of $\mathcal{B}_{-1}(z)$ are essentially the same, but there's an extra 1 at the beginning and the other numbers alternate in sign: $\langle 1, 1, -1, 2, -5, 14, \ldots\rangle$. Thus $\mathcal{B}_{-1}(z) = 1 + z\mathcal{B}_2(-z)$. We also have $\mathcal{B}_{-1}(z) = \mathcal{B}_2(-z)^{-1}$.

Let's close this section by deriving an important consequence of (5.72) and (5.73), a relation that shows further connections between the functions $\mathcal{B}_{-1}(z)$ and $\mathcal{B}_2(-z)$:

$$\frac{\mathcal{B}_{-1}(z)^{n+1} - (-z)^{n+1}\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} = \sum_{k\leqslant n}\binom{n-k}{k}z^k\,.$$

This holds because the coefficient of z^k in $(-z)^{n+1}\mathcal{B}_2(-z)^{n+1}/\sqrt{1+4z}$ is

$$\begin{aligned}
[z^k]\,\frac{(-z)^{n+1}\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} &= (-1)^{n+1}[z^{k-n-1}]\,\frac{\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} \\
&= (-1)^{n+1}(-1)^{k-n-1}[z^{k-n-1}]\,\frac{\mathcal{B}_2(z)^{n+1}}{\sqrt{1-4z}} \\
&= (-1)^k\binom{2(k-n-1)+n+1}{k-n-1} \\
&= (-1)^k\binom{2k-n-1}{k-n-1} \;=\; (-1)^k\binom{2k-n-1}{k} \\
&= \binom{n-k}{k} \;=\; [z^k]\,\frac{\mathcal{B}_{-1}(z)^{n+1}}{\sqrt{1+4z}}
\end{aligned}$$

when $k > n$. The terms nicely cancel each other out. We can now use (5.68) and (5.69) to obtain the closed form

$$\sum_{k\leqslant n}\binom{n-k}{k}z^k \;=\; \frac{1}{\sqrt{1+4z}}\left(\left(\frac{1+\sqrt{1+4z}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{1+4z}}{2}\right)^{n+1}\right), \qquad \text{integer } n \geqslant 0. \tag{5.74}$$

(The special case $z = -1$ came up in Problem 3 of Section 5.2. Since the numbers $\frac{1}{2}(1 \pm \sqrt{-3}\,)$ are sixth roots of unity, the sums $\sum_{k\leqslant n}\binom{n-k}{k}(-1)^k$ have the periodic behavior we observed in that problem.) Similarly we can combine (5.70) with (5.71) to cancel the large coefficients and get

$$\sum_{k<n}\binom{n-k}{k}\frac{n}{n-k}z^k \;=\; \left(\frac{1+\sqrt{1+4z}}{2}\right)^{n} + \left(\frac{1-\sqrt{1+4z}}{2}\right)^{n}, \qquad \text{integer } n > 0. \tag{5.75}$$

5.5 HYPERGEOMETRIC FUNCTIONS

The methods we've been applying to binomial coefficients are very effective, when they work, but we must admit that they often appear to be ad hoc — more like tricks than techniques. When we're working on a problem, we often have many directions to pursue, and we might find ourselves going around in circles. Binomial coefficients are like chameleons, changing their appearance easily. Therefore it's natural to ask if there isn't some unifying principle that will systematically handle a great variety of binomial coefficient summations all at once. Fortunately, the answer is yes. The unifying principle is based on the theory of certain infinite sums called *hypergeometric series*.

They're even more versatile than chameleons; we can dissect them and put them back together in different ways.

The study of hypergeometric series was launched many years ago by Euler, Gauss, and Riemann; such series, in fact, are still the subject of considerable research. But hypergeometrics have a somewhat formidable notation, which takes a little time to get used to.

Anything that has survived for centuries with such awesome notation must be really useful.

The general hypergeometric series is a power series in z with $m+n$ parameters, and it is defined as follows in terms of rising factorial powers:

$$F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n} \middle| z\right) = \sum_{k \geqslant 0} \frac{a_1^{\overline{k}} \ldots a_m^{\overline{k}}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}}} \frac{z^k}{k!}. \tag{5.76}$$

To avoid division by zero, none of the b's may be zero or a negative integer. Other than that, the a's and b's may be anything we like. The notation '$F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$' is also used as an alternative to the two-line form (5.76), since a one-line form sometimes works better typographically. The a's are said to be *upper parameters*; they occur in the numerator of the terms of F. The b's are *lower parameters*, and they occur in the denominator. The final quantity z is called the *argument*.

Standard reference books often use '${}_mF_n$' instead of 'F' as the name of a hypergeometric with m upper parameters and n lower parameters. But the extra subscripts tend to clutter up the formulas and waste our time, if we're compelled to write them over and over. We can count how many parameters there are, so we usually don't need extra additional unnecessary redundancy.

Many important functions occur as special cases of the general hypergeometric; indeed, that's why hypergeometrics are so powerful. For example, the simplest case occurs when $m = n = 0$: There are no parameters at all, and we get the familiar series

$$F\left({} \middle| z\right) = \sum_{k \geqslant 0} \frac{z^k}{k!} = e^z.$$

Actually the notation looks a bit unsettling when m or n is zero. We can add an extra '1' above and below in order to avoid this:

$$F\left({1 \atop 1} \middle| z\right) = e^z.$$

In general we don't change the function if we cancel a parameter that occurs in both numerator and denominator, or if we insert two identical parameters.

The next simplest case has $m = 1$, $a_1 = 1$, and $n = 0$; we change the parameters to $m = 2$, $a_1 = a_2 = 1$, $n = 1$, and $b_1 = 1$, so that $n > 0$. This series also turns out to be familiar, because $1^{\overline{k}} = k!$:

$$F\left({1,1 \atop 1} \middle| z\right) = \sum_{k \geqslant 0} z^k = \frac{1}{1-z}.$$

It's our old friend, the geometric series; $F(a_1,\ldots,a_m;b_1,\ldots,b_n;z)$ is called hypergeometric because it includes the geometric series $F(1,1;1;z)$ as a very special case.

The general case $m=1$ and $n=0$ is, in fact, easy to sum in closed form,

$$F\left({a,1 \atop 1}\middle|\,z\right) = \sum_{k\geqslant 0} a^{\overline{k}}\,\frac{z^k}{k!} = \sum_k \binom{a+k-1}{k} z^k = \frac{1}{(1-z)^a}\,, \qquad (5.77)$$

using (5.56). If we replace a by $-a$ and z by $-z$, we get the binomial theorem,

$$F\left({-a,1 \atop 1}\middle|\,-z\right) = (1+z)^a\,.$$

A negative integer as upper parameter causes the infinite series to become finite, since $(-a)^{\overline{k}}=0$ whenever $k>a\geqslant 0$ and a is an integer.

The general case $m=0$, $n=1$ is another famous series, but it's not as well known in the literature of discrete mathematics:

$$F\left({1 \atop b,1}\middle|\,z\right) = \sum_{k\geqslant 0} \frac{(b-1)!}{(b-1+k)!}\,\frac{z^k}{k!} = I_{b-1}\big(2\sqrt{z}\,\big)\,\frac{(b-1)!}{z^{(b-1)/2}}\,. \qquad (5.78)$$

This function I_{b-1} is called a "modified Bessel function" of order $b-1$. The special case $b=1$ gives us $F\left({1 \atop 1,1}\middle|\,z\right)=I_0(2\sqrt{z})$, which is the interesting series $\sum_{k\geqslant 0} z^k/k!^2$.

The special case $m=n=1$ is called a "confluent hypergeometric series" and often denoted by the letter M:

$$F\left({a \atop b}\middle|\,z\right) = \sum_{k\geqslant 0} \frac{a^{\overline{k}}}{b^{\overline{k}}}\,\frac{z^k}{k!} = M(a,b,z)\,. \qquad (5.79)$$

This function, which has important applications to engineering, was introduced by Ernst Kummer.

By now a few of us are wondering why we haven't discussed convergence of the infinite series (5.76). The answer is that we can ignore convergence if we are using z simply as a formal symbol. It is not difficult to verify that formal infinite sums of the form $\sum_{k\geqslant n}\alpha_k z^k$ form a field, if the coefficients α_k lie in a field. We can add, subtract, multiply, divide, differentiate, and do functional composition on such formal sums without worrying about convergence; any identities we derive will still be formally true. For example, the hypergeometric $F\left({1,1,1 \atop 1}\middle|\,z\right)=\sum_{k\geqslant 0}k!\,z^k$ doesn't converge for any nonzero z; yet we'll see in Chapter 7 that we can still use it to solve problems. On the other hand, whenever we replace z by a particular numerical value, we do have to be sure that the infinite sum is well defined.

The next step up in complication is actually the most famous hypergeometric of all. In fact, it was *the* hypergeometric series until about 1870, when everything was generalized to arbitrary m and n. This one has two upper parameters and one lower parameter:

$$F\left({a, b \atop c} \middle| z\right) = \sum_{k \geqslant 0} \frac{a^{\overline{k}} b^{\overline{k}} z^k}{c^{\overline{k}}\, k!}. \tag{5.80}$$

It is often called the Gaussian hypergeometric, because many of its subtle properties were first proved by Gauss in his doctoral dissertation of 1812 [143], although Euler [118] and Pfaff [292] had already discovered some remarkable things about it. One of its important special cases is

"There must be many universities to-day where 95 per cent, if not 100 per cent, of the functions studied by physics, engineering, and even mathematics students, are covered by this single symbol $F(a, b; c; x)$."
— W. W. Sawyer [318]

$$\begin{aligned} \ln(1+z) &= z\,F\left({1, 1 \atop 2} \middle| -z\right) = z \sum_{k \geqslant 0} \frac{k!\,k!}{(k+1)!} \frac{(-z)^k}{k!} \\ &= z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots . \end{aligned}$$

Notice that $z^{-1}\ln(1+z)$ is a hypergeometric function, but $\ln(1+z)$ itself cannot be hypergeometric, since a hypergeometric series always has the value 1 when $z = 0$.

So far hypergeometrics haven't actually done anything for us except provide an excuse for name-dropping. But we've seen that several very different functions can all be regarded as hypergeometric; this will be the main point of interest in what follows. We'll see that a large class of sums can be written as hypergeometric series in a "canonical" way, hence we will have a good filing system for facts about binomial coefficients.

What series are hypergeometric? It's easy to answer this question if we look at the ratio between consecutive terms:

$$F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n} \middle| z\right) = \sum_{k \geqslant 0} t_k, \qquad t_k = \frac{a_1^{\overline{k}} \ldots a_m^{\overline{k}}\, z^k}{b_1^{\overline{k}} \ldots b_n^{\overline{k}}\, k!}.$$

The first term is $t_0 = 1$, and the other terms have ratios given by

$$\begin{aligned} \frac{t_{k+1}}{t_k} &= \frac{a_1^{\overline{k+1}} \ldots a_m^{\overline{k+1}}}{a_1^{\overline{k}} \ldots a_m^{\overline{k}}} \frac{b_1^{\overline{k}} \ldots b_n^{\overline{k}}}{b_1^{\overline{k+1}} \ldots b_n^{\overline{k+1}}} \frac{k!}{(k+1)!} \frac{z^{k+1}}{z^k} \\ &= \frac{(k+a_1) \ldots (k+a_m)\, z}{(k+b_1) \ldots (k+b_n)(k+1)}. \end{aligned} \tag{5.81}$$

This is a *rational function* of k, that is, a quotient of polynomials in k. According to the Fundamental Theorem of Algebra, any rational function of k can be factored over the complex numbers and put into this form. The a's are the negatives of the roots of the polynomial in the numerator, and the b's are the negatives of the roots of the polynomial in the denominator. If the denominator doesn't already contain the special factor $(k+1)$, we can include $(k+1)$ in both numerator and denominator. A constant factor remains, and we can call it z. Therefore hypergeometric series are precisely those series whose first term is 1 and whose term ratio t_{k+1}/t_k is a rational function of k.

Suppose, for example, that we're given an infinite series with term ratio

$$\frac{t_{k+1}}{t_k} = \frac{k^2+7k+10}{4k^2+1},$$

a rational function of k. The numerator polynomial splits nicely into two factors, $(k+2)(k+5)$, and the denominator is $4(k+i/2)(k-i/2)$. Since the denominator is missing the required factor $(k+1)$, we write the term ratio as

$$\frac{t_{k+1}}{t_k} = \frac{(k+2)(k+5)(k+1)(1/4)}{(k+i/2)(k-i/2)(k+1)},$$

and we can read off the results: The given series is

$$\sum_{k\geqslant 0} t_k = t_0\, F\left({2,\,5,\,1 \atop i/2,\,-i/2}\,\middle|\,1/4\right).$$

Thus, we have a general method for finding the hypergeometric representation of a given quantity S, when such a representation is possible: First we write S as an infinite series whose first term is nonzero. We choose a notation so that the series is $\sum_{k\geqslant 0} t_k$ with $t_0 \neq 0$. Then we calculate t_{k+1}/t_k. If the term ratio is not a rational function of k, we're out of luck. Otherwise we express it in the form (5.81); this gives parameters $a_1, \ldots, a_m, b_1, \ldots, b_n$, and an argument z, such that $S = t_0\, F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$.

(Now is a good time to do warmup exercise 11.)

Gauss's hypergeometric series can be written in the recursively factored form

$$F\left({a,b \atop c}\,\middle|\,z\right) = 1 + \frac{a}{1}\frac{b}{c}z\left(1 + \frac{a+1}{2}\frac{b+1}{c+1}z\left(1 + \frac{a+2}{3}\frac{b+2}{c+2}z\,(1+\cdots)\right)\right)$$

if we wish to emphasize the importance of term ratios.

Let's try now to reformulate the binomial coefficient identities derived earlier in this chapter, expressing them as hypergeometrics. For example, let's figure out what the parallel summation law,

$$\sum_{k\leqslant n}\binom{r+k}{k} = \binom{r+n+1}{n}, \qquad \text{integer } n,$$

looks like in hypergeometric notation. We need to write the sum as an infinite series that starts at $k = 0$, so we replace k by $n - k$:

$$\sum_{k\geqslant 0}\binom{r+n-k}{n-k} = \sum_{k\geqslant 0}\frac{(r+n-k)!}{r!\,(n-k)!} = \sum_{k\geqslant 0} t_k\,.$$

This series is formally infinite but actually finite, because the $(n-k)!$ in the denominator will make $t_k = 0$ when $k > n$. (We'll see later that $1/x!$ is defined for all x, and that $1/x! = 0$ when x is a negative integer. But for now, let's blithely disregard such technicalities until we gain more hypergeometric experience.) The term ratio is

$$\begin{aligned}\frac{t_{k+1}}{t_k} = \frac{(r+n-k-1)!\,r!\,(n-k)!}{r!\,(n-k-1)!\,(r+n-k)!} &= \frac{n-k}{r+n-k}\\ &= \frac{(k+1)(k-n)(1)}{(k-n-r)(k+1)}\,.\end{aligned}$$

Furthermore $t_0 = \binom{r+n}{n}$. Hence the parallel summation law is equivalent to the hypergeometric identity

$$\binom{r+n}{n} F\left({1,\,-n \atop -n-r}\,\Big|\,1\right) = \binom{r+n+1}{n}.$$

Dividing through by $\binom{r+n}{n}$ gives a slightly simpler version,

$$F\left({1,\,-n \atop -n-r}\,\Big|\,1\right) = \frac{r+n+1}{r+1}\,,\qquad \text{if } \binom{r+n}{n} \neq 0. \tag{5.82}$$

Let's do another one. The term ratio of identity (5.16),

$$\sum_{k\leqslant m}\binom{r}{k}(-1)^k = (-1)^m\binom{r-1}{m}\,,\qquad \text{integer } m,$$

is $(k-m)/(r-m+k+1) = (k+1)(k-m)(1)/(k-m+r+1)(k+1)$, after we replace k by $m-k$; hence (5.16) gives a closed form for

$$F\left({1,\,-m \atop -m+r+1}\,\Big|\,1\right).$$

This is essentially the same as the hypergeometric function on the left of (5.82), but with m in place of n and $r+1$ in place of $-r$. Therefore identity (5.16) could have been derived from (5.82), the hypergeometric version of (5.9). (No wonder we found it easy to prove (5.16) by using (5.9).)

First derangements, now degenerates.

Before we go further, we should think about degenerate cases, because hypergeometrics are not defined when a lower parameter is zero or a negative integer. We usually apply the parallel summation identity when r and n are positive integers; but then $-n-r$ is a negative integer and the hypergeometric (5.76) is undefined. How then can we consider (5.82) to be legitimate? The answer is that we can take the limit of $F\left({1,-n \atop -n-r+\epsilon}\middle|1\right)$ as $\epsilon \to 0$.

We will look at such things more closely later in this chapter, but for now let's just be aware that some denominators can be dynamite. It is interesting, however, that the very first sum we've tried to express hypergeometrically has turned out to be degenerate.

(We proved the identities originally for integer r, and used the polynomial argument to show that they hold in general. Now we're proving them first for irrational r, and using a limiting argument to show that they hold for integers!)

Another possibly sore point in our derivation of (5.82) is that we expanded $\binom{r+n-k}{n-k}$ as $(r+n-k)!/r!\,(n-k)!$. This expansion fails when r is a negative integer, because $(-m)!$ has to be ∞ if the law

$$0! \;=\; 0\cdot(-1)\cdot(-2)\cdot\ldots\cdot(-m+1)\cdot(-m)!$$

is going to hold. Again, we need to approach integer results by considering a limit of $r+\epsilon$ as $\epsilon \to 0$.

But we defined the factorial representation $\binom{r}{k} = r!/k!\,(r-k)!$ only when r is an integer! If we want to work effectively with hypergeometrics, we need a factorial function that is defined for all complex numbers. Fortunately there is such a function, and it can be defined in many ways. Here's one of the most useful definitions of $z!$, actually a definition of $1/z!$:

$$\frac{1}{z!} \;=\; \lim_{n\to\infty}\binom{n+z}{n}n^{-z}. \tag{5.83}$$

(See exercise 21. Euler [99, 100, 72] discovered this when he was 22 years old.) The limit can be shown to exist for all complex z, and it is zero only when z is a negative integer. Another significant definition is

$$z! \;=\; \int_0^\infty t^z e^{-t}\,dt\,, \qquad \text{if } \Re z > -1. \tag{5.84}$$

This integral exists only when the real part of z exceeds -1, but we can use the formula

$$z! \;=\; z\,(z-1)! \tag{5.85}$$

to extend the definition to all complex z (except negative integers). Still another definition comes from Stirling's interpolation of $\ln z!$ in (5.47). All of these approaches lead to the same generalized factorial function.

There's a very similar function called the *Gamma function*, which relates to ordinary factorials somewhat as rising powers relate to falling powers. Standard reference books often use factorials and Gamma functions simultaneously, and it's convenient to convert between them if necessary using the following formulas:

$$\Gamma(z+1) \;=\; z!\,; \tag{5.86}$$

$$(-z)!\,\Gamma(z) \;=\; \frac{\pi}{\sin \pi z}\,. \tag{5.87}$$

How do you write z to the $\overline{w}$ power, when $\overline{w}$ is the complex conjugate of w?

$z^{(\overline{w})}$.

We can use these generalized factorials to define generalized factorial powers, when z and w are arbitrary complex numbers:

$$z^{\underline{w}} \;=\; \frac{z!}{(z-w)!}\,; \tag{5.88}$$

$$z^{\overline{w}} \;=\; \frac{\Gamma(z+w)}{\Gamma(z)}\,. \tag{5.89}$$

The only proviso is that we must use appropriate limiting values when these formulas give ∞/∞. (The formulas never give $0/0$, because factorials and Gamma-function values are never zero.) A binomial coefficient can be written

$$\binom{z}{w} \;=\; \lim_{\zeta\to z}\,\lim_{\omega\to w}\,\frac{\zeta!}{\omega!\,(\zeta-\omega)!} \tag{5.90}$$

when z and w are any complex numbers whatever.

I see, the lower index arrives at its limit first. That's why $\binom{z}{w}$ is zero when w is a negative integer.

Armed with generalized factorial tools, we can return to our goal of reducing the identities derived earlier to their hypergeometric essences. The binomial theorem (5.13) turns out to be neither more nor less than (5.77), as we might expect. So the next most interesting identity to try is Vandermonde's convolution (5.27):

$$\sum_k \binom{r}{k}\binom{s}{n-k} \;=\; \binom{r+s}{n}, \qquad \text{integer } n.$$

The kth term here is

$$t_k \;=\; \frac{r!}{(r-k)!\,k!}\,\frac{s!}{(s-n+k)!\,(n-k)!}\,,$$

and we are no longer too shy to use generalized factorials in these expressions. Whenever t_k contains a factor like $(\alpha+k)!$, with a plus sign before the k, we get $(\alpha+k+1)!/(\alpha+k)! = k+\alpha+1$ in the term ratio t_{k+1}/t_k, by (5.85); this contributes the parameter '$\alpha+1$' to the corresponding hypergeometric — as an upper parameter if $(\alpha+k)!$ was in the numerator of t_k, but as a lower parameter otherwise. Similarly, a factor like $(\alpha-k)!$ leads to $(\alpha-k-1)!/(\alpha-k)! = (-1)/(k-\alpha)$; this contributes '$-\alpha$' to the opposite set of parameters (reversing the roles of upper and lower), and negates the hypergeometric argument. Factors like $r!$, which are independent of k, go into t_0 but disappear from the term ratio. Using such tricks we can predict without further calculation that the term ratio of (5.27) is

$$\frac{t_{k+1}}{t_k} = \frac{k-r}{k+1}\,\frac{k-n}{k+s-n+1}$$

times $(-1)^2 = 1$, and Vandermonde's convolution becomes

$$\binom{s}{n} F\left({-r,\,-n \atop s-n+1}\,\middle|\,1\right) = \binom{r+s}{n}. \tag{5.91}$$

We can use this equation to determine $F(a,b;c;z)$ in general, when $z=1$ and when b is a negative integer.

Let's rewrite (5.91) in a form so that table lookup is easy when a new sum needs to be evaluated. The result turns out to be

$$F\left({a,b \atop c}\,\middle|\,1\right) = \frac{\Gamma(c-a-b)\,\Gamma(c)}{\Gamma(c-a)\,\Gamma(c-b)}\,; \qquad \begin{matrix}\text{integer } b \leqslant 0\\ \text{or } \Re c > \Re a + \Re b.\end{matrix} \tag{5.92}$$

Vandermonde's convolution (5.27) covers only the case that one of the upper parameters, say b, is a nonpositive integer; but Gauss proved that (5.92) is valid also when a, b, c are complex numbers whose real parts satisfy $\Re c > \Re a + \Re b$. In other cases, the infinite series $F\left({a,b \atop c}\,\middle|\,1\right)$ doesn't converge. When $b = -n$, the identity can be written more conveniently with factorial powers instead of Gamma functions:

A few weeks ago, we were studying what Gauss had done in kindergarten.
Now we're studying stuff beyond his Ph.D. thesis.
Is this intimidating or what?

$$F\left({a,\,-n \atop c}\,\middle|\,1\right) = \frac{(c-a)^{\overline{n}}}{c^{\overline{n}}} = \frac{(a-c)^{\underline{n}}}{(-c)^{\underline{n}}}\,, \qquad \text{integer } n \geqslant 0. \tag{5.93}$$

It turns out that all five of the identities in Table 169 are special cases of Vandermonde's convolution; formula (5.93) covers them all, when proper attention is paid to degenerate situations.

Notice that (5.82) is just the special case $a=1$ of (5.93). Therefore we don't really need to remember (5.82); and we don't really need the identity (5.9) that led us to (5.82), even though Table 174 said that it was memorable. A computer program for formula manipulation, faced with the problem of evaluating $\sum_{k\leqslant n}\binom{r+k}{k}$, could convert the sum to a hypergeometric and plug into the general identity for Vandermonde's convolution.

Problem 1 in Section 5.2 asked for the value of

$$\sum_{k\geqslant 0} \binom{m}{k} \Big/ \binom{n}{k}.$$

This problem is a natural for hypergeometrics, and after a bit of practice any hypergeometer can read off the parameters immediately as $F(1,-m;-n;1)$. Hmmm; that problem was yet another special takeoff on Vandermonde!

The sum in Problem 2 and Problem 4 likewise yields $F(2, 1-n; 2-m; 1)$. (We need to replace k by $k+1$ first.) And the "menacing" sum in Problem 6 turns out to be just $F(n+1, -n; 2; 1)$. Is there nothing more to sum, besides disguised versions of Vandermonde's powerful convolution?

Well, yes, Problem 3 is a bit different. It deals with a special case of the general sum $\sum_k \binom{n-k}{k} z^k$ considered in (5.74), and this leads to a closed-form expression for

$$F\left(\begin{matrix}1+2\lceil n/2\rceil,\ -n\\ 1/2\end{matrix}\,\middle|\,-z/4\right).$$

We also proved something new in (5.55), when we looked at the coefficients of $(1-z)^r(1+z)^r$:

$$F\left(\begin{matrix}1-c-2n,\ -2n\\ c\end{matrix}\,\middle|\,-1\right) = (-1)^n \frac{(2n)!}{n!}\frac{(c-1)!}{(c+n-1)!}, \qquad \text{integer } n \geqslant 0.$$

Kummer was a summer.

This is called *Kummer's formula* when it's generalized to complex numbers:

$$F\left(\begin{matrix}a,\ b\\ 1+b-a\end{matrix}\,\middle|\,-1\right) = \frac{(b/2)!}{b!}(b-a)^{\underline{b/2}}. \tag{5.94}$$

The summer of '36.

(Ernst Kummer [229] proved this in 1836.)

It's interesting to compare these two formulas. Replacing c by $1-2n-a$, we find that the results are consistent if and only if

$$(-1)^n \frac{(2n)!}{n!} = \lim_{b\to -2n} \frac{(b/2)!}{b!} = \lim_{x\to -n} \frac{x!}{(2x)!} \tag{5.95}$$

when n is a positive integer. Suppose, for example, that $n = 3$; then we should have $-6!/3! = \lim_{x\to-3} x!/(2x)!$. We know that $(-3)!$ and $(-6)!$ are both infinite; but we might choose to ignore that difficulty and to imagine that $(-3)! = (-3)(-4)(-5)(-6)!$, so that the two occurrences of $(-6)!$ will cancel. Such temptations must, however, be resisted, because they lead to the wrong answer! The limit of $x!/(2x)!$ as $x \to -3$ is not $(-3)(-4)(-5)$ but rather $-6!/3! = (-4)(-5)(-6)$, according to (5.95).

The right way to evaluate the limit in (5.95) is to use equation (5.87), which relates negative-argument factorials to positive-argument Gamma functions. If we replace x by $-n-\epsilon$ and let $\epsilon \to 0$, two applications of (5.87) give

$$\frac{(-n-\epsilon)!}{(-2n-2\epsilon)!}\,\frac{\Gamma(n+\epsilon)}{\Gamma(2n+2\epsilon)} = \frac{\sin(2n+2\epsilon)\pi}{\sin(n+\epsilon)\pi}.$$

Now $\sin(x+y) = \sin x \cos y + \cos x \sin y$; so this ratio of sines is

$$\frac{\cos 2n\pi \,\sin 2\epsilon\pi}{\cos n\pi \,\sin \epsilon\pi} \;=\; (-1)^n\bigl(2 + O(\epsilon)\bigr)\,,$$

by the methods of Chapter 9. Therefore, by (5.86), we have

$$\lim_{\epsilon\to 0}\frac{(-n-\epsilon)!}{(-2n-2\epsilon)!} \;=\; 2(-1)^n\frac{\Gamma(2n)}{\Gamma(n)} \;=\; 2(-1)^n\frac{(2n-1)!}{(n-1)!} \;=\; (-1)^n\frac{(2n)!}{n!}\,,$$

as desired.

Let's complete our survey by restating the other identities we've seen so far in this chapter, clothing them in hypergeometric garb. The triple-binomial sum in (5.29) can be written

$$F\left({1-a-2n,\ 1-b-2n,\ -2n \atop a,\ b}\,\Big|\,1\right) \;=\; (-1)^n\frac{(2n)!}{n!}\frac{(a+b+2n-1)^{\overline{n}}}{a^{\overline{n}}\,b^{\overline{n}}}\,,\qquad \text{integer } n \geqslant 0.$$

When this one is generalized to complex numbers, it is called *Dixon's formula*:

$$F\left({a,\ b,\ c \atop 1+c-a,\ 1+c-b}\,\Big|\,1\right) \;=\; \frac{(c/2)!}{c!}\,\frac{(c-a)^{\underline{c/2}}\,(c-b)^{\underline{c/2}}}{(c-a-b)^{\underline{c/2}}}\,, \qquad \Re a + \Re b < 1 + \Re c/2. \tag{5.96}$$

One of the most general formulas we've encountered is the triple-binomial sum (5.28), which yields *Saalschütz's identity*:

(Historical note: Saalschütz [315] independently discovered this formula almost 100 years after Pfaff [292] had first published it. Taking the limit as $n \to \infty$ yields equation (5.92).)

$$\begin{aligned} F\left({a,\ b,\ -n \atop c,\ a+b-c-n+1}\,\Big|\,1\right) \;&=\; \frac{(c-a)^{\overline{n}}\,(c-b)^{\overline{n}}}{c^{\overline{n}}\,(c-a-b)^{\overline{n}}} \qquad\qquad (5.97)\\ &=\; \frac{(a-c)^{\underline{n}}\,(b-c)^{\underline{n}}}{(-c)^{\underline{n}}\,(a+b-c)^{\underline{n}}}\,, \qquad \text{integer } n \geqslant 0. \end{aligned}$$

This formula gives the value at $z = 1$ of the general hypergeometric series with three upper parameters and two lower parameters, provided that one of the upper parameters is a nonpositive integer and that $b_1 + b_2 = a_1 + a_2 + a_3 + 1$. (If the sum of the lower parameters exceeds the sum of the upper parameters by 2 instead of by 1, the formula of exercise 25 can be used to express $F(a_1, a_2, a_3; b_1, b_2; 1)$ in terms of two hypergeometrics that satisfy Saalschütz's identity.)

Our hard-won identity in Problem 8 of Section 5.2 reduces to

$$\frac{1}{1+x}\,F\left({x+1,\ n+1,\ -n \atop 1,\ x+2}\,\Big|\,1\right) \;=\; (-1)^n x^{\underline{n}}\, x^{\underline{-n-1}}\,.$$

Sigh. This is just the special case $c = 1$ of Saalschütz's identity (5.97), so we could have saved a lot of work by going to hypergeometrics directly!

What about Problem 7? That extra-menacing sum gives us the formula

$$F\left({n+1,\ m-n,\ 1,\ \frac{1}{2} \atop \frac{1}{2}m+1,\ \frac{1}{2}m+\frac{1}{2},\ 2}\,\middle|\,1\right) \;=\; \frac{m}{n}\,,$$

which is the first case we've seen with three lower parameters. So it looks new. But it really isn't; the left-hand side can be replaced by

$$F\left({n,\ m-n-1,\ -\frac{1}{2} \atop \frac{1}{2}m,\ \frac{1}{2}m-\frac{1}{2}}\,\middle|\,1\right) - 1\,,$$

using exercise 26, and Saalschütz's identity wins again.

(Historical note: The great relevance of hypergeometric series to binomial coefficient identities was first pointed out by George Andrews in 1974 [9, section 5].)

Well, that's another deflating experience, but it's also another reason to appreciate the power of hypergeometric methods.

The convolution identities in Table 202 do not have hypergeometric equivalents, because their term ratios are rational functions of k only when t is an integer. Equations (5.64) and (5.65) aren't hypergeometric even when $t = 1$. But we can take note of what (5.62) tells us when t has small integer values:

$$F\left({\frac{1}{2}r,\ \frac{1}{2}r+\frac{1}{2},\ -n,\ -n-s \atop r+1,\ -n-\frac{1}{2}s,\ -n-\frac{1}{2}s+\frac{1}{2}}\,\middle|\,1\right) \;=\; \binom{r+s+2n}{n}\Big/\binom{s+2n}{n}\,;$$

$$F\left({\frac{1}{3}r,\ \frac{1}{3}r+\frac{1}{3},\ \frac{1}{3}r+\frac{2}{3},\ -n,\ -n-\frac{1}{2}s,\ -n-\frac{1}{2}s-\frac{1}{2} \atop \frac{1}{2}r+\frac{1}{2},\ \frac{1}{2}r+1,\ -n-\frac{1}{3}s,\ -n-\frac{1}{3}s+\frac{1}{3},\ -n-\frac{1}{3}s+\frac{2}{3}}\,\middle|\,1\right)$$
$$\;=\; \binom{r+s+3n}{n}\Big/\binom{s+3n}{n}\,.$$

The first of these formulas gives the result of Problem 7 again, when the quantities (r, s, n) are replaced respectively by $(1, 2n+1-m, -1-n)$.

Finally, the "unexpected" sum (5.20) gives us an unexpected hypergeometric identity that turns out to be quite instructive. Let's look at it in slow motion. First we convert to an infinite sum,

$$\sum_{k\leqslant m}\binom{m+k}{k}2^{-k} \;=\; 2^m \qquad\Longleftrightarrow\qquad \sum_{k\geqslant 0}\binom{2m-k}{m-k}2^k \;=\; 2^{2m}\,.$$

The term ratio from $(2m-k)!\,2^k/m!\,(m-k)!$ is $2(k-m)/(k-2m)$, so we have a hypergeometric identity with $z = 2$:

$$\binom{2m}{m} F\left({1,\ -m \atop -2m}\,\middle|\,2\right) \;=\; 2^{2m}\,, \qquad \text{integer } m \geqslant 0. \tag{5.98}$$

But look at the lower parameter '$-2m$'. Negative integers are verboten, so this identity is undefined!

It's high time to look at such limiting cases carefully, as promised earlier, because degenerate hypergeometrics can often be evaluated by approaching them from nearby nondegenerate points. We must be careful when we do this, because different results can be obtained if we take limits in different ways. For example, here are two limits that turn out to be quite different when one of the upper parameters is increased by ϵ:

$$\begin{aligned}\lim_{\epsilon\to 0} F\left({-1+\epsilon,\,-3 \atop -2+\epsilon}\middle|\,1\right) &= \lim_{\epsilon\to 0}\Big(1+\tfrac{(-1+\epsilon)(-3)}{(-2+\epsilon)\,1!}+\tfrac{(-1+\epsilon)(\epsilon)(-3)(-2)}{(-2+\epsilon)(-1+\epsilon)\,2!}\\ &\qquad\qquad+\tfrac{(-1+\epsilon)(\epsilon)(1+\epsilon)(-3)(-2)(-1)}{(-2+\epsilon)(-1+\epsilon)(\epsilon)\,3!}\Big)\\ &= 1-\tfrac{3}{2}+0+\tfrac{1}{2} \;=\; 0\,;\\ \lim_{\epsilon\to 0} F\left({-1,-3 \atop -2+\epsilon}\middle|\,1\right) &= \lim_{\epsilon\to 0}\Big(1+\tfrac{(-1)(-3)}{(-2+\epsilon)\,1!}+0+0\Big)\\ &= 1-\tfrac{3}{2}+0+0 \;=\; -\tfrac{1}{2}\,.\end{aligned}$$

Similarly, we have defined $\binom{-1}{-1}=0=\lim_{\epsilon\to 0}\binom{-1+\epsilon}{-1}$; this is not the same as $\lim_{\epsilon\to 0}\binom{-1+\epsilon}{-1+\epsilon}=1$. The proper way to treat (5.98) as a limit is to realize that the upper parameter $-m$ is being used to make all terms of the series $\sum_{k\geqslant 0}\binom{2m-k}{m-k}2^k$ zero for $k>m$; this means that we want to make the following more precise statement:

$$\binom{2m}{m}\lim_{\epsilon\to 0} F\left({1,\,-m \atop -2m+\epsilon}\middle|\,2\right) \;=\; 2^{2m}\,, \qquad \text{integer } m\geqslant 0. \tag{5.99}$$

Each term of this limit is well defined, because the denominator factor $(-2m)^{\overline{k}}$ does not become zero until $k>2m$. Therefore this limit gives us exactly the sum (5.20) we began with.

5.6 HYPERGEOMETRIC TRANSFORMATIONS

It should be clear by now that a database of known hypergeometric closed forms is a useful tool for doing sums of binomial coefficients. We simply convert any given sum into its canonical hypergeometric form, then look it up in the table. If it's there, fine, we've got the answer. If not, we can add it to the database if the sum turns out to be expressible in closed form. We might also include entries in the table that say, "This sum does not have a simple closed form in general." For example, the sum $\sum_{k\leqslant m}\binom{n}{k}$

corresponds to the hypergeometric

$$\binom{n}{m} F\left({1,\,-m \atop n-m+1}\,\Big|\,-1\right), \qquad \text{integers } n \geqslant m \geqslant 0; \tag{5.100}$$

this has a simple closed form only if m is near 0, $\frac{1}{2}n$, or n.

But there's more to the story, since hypergeometric functions also obey identities of their own. This means that every closed form for hypergeometrics leads to additional closed forms and to additional entries in the database. For example, the identities in exercises 25 and 26 tell us how to transform one hypergeometric into two others with similar but different parameters. These can in turn be transformed again.

The hypergeometric database should really be a "knowledge base."

In 1797, J. F. Pfaff [292] discovered a surprising *reflection law*,

$$\frac{1}{(1-z)^a} F\left({a,\,b \atop c}\,\Big|\,\frac{-z}{1-z}\right) = F\left({a,\,c-b \atop c}\,\Big|\,z\right), \tag{5.101}$$

which is a transformation of another type. This is a formal identity in power series, if the quantity $(-z)^k/(1-z)^{k+a}$ is replaced by the infinite series $(-z)^k\bigl(1+\binom{k+a}{1}z+\binom{k+a+1}{2}z^2+\cdots\bigr)$ when the left-hand side is expanded (see exercise 50). We can use this law to derive new formulas from the identities we already know, when $z \neq 1$.

For example, Kummer's formula (5.94) can be combined with the reflection law (5.101) if we choose the parameters so that both identities apply:

$$\begin{aligned} 2^{-a} F\left({a,\,1-a \atop 1+b-a}\,\Big|\,\frac{1}{2}\right) &= F\left({a,\,b \atop 1+b-a}\,\Big|\,-1\right) \\ &= \frac{(b/2)!}{b!}(b-a)^{\underline{b/2}}. \end{aligned} \tag{5.102}$$

We can now set $a = -n$ and go back from this equation to a new identity in binomial coefficients that we might need some day:

$$\begin{aligned} \sum_{k\geqslant 0} \frac{(-n)^{\overline{k}}\,(1+n)^{\overline{k}}}{(1+b+n)^{\overline{k}}}\,\frac{2^{-k}}{k!} &= \sum_k \binom{n}{k}\left(\frac{-1}{2}\right)^k \binom{n+k}{k}\Big/\binom{n+b+k}{k} \\ &= 2^{-n}\frac{(b/2)!\,(b+n)!}{b!\,(b/2+n)!}, \quad \text{integer } n \geqslant 0. \end{aligned} \tag{5.103}$$

For example, when $n = 3$ this identity says that

$$\begin{aligned} 1 - 3\,\frac{4}{2(4+b)} + 3\,\frac{4\cdot 5}{4(4+b)(5+b)} - \frac{4\cdot 5\cdot 6}{8(4+b)(5+b)(6+b)} \\ = \frac{(b+3)(b+2)(b+1)}{(b+6)(b+4)(b+2)}. \end{aligned}$$

It's almost unbelievable, but true, for all b. (Except when a factor in the denominator vanishes.)

This is fun; let's try again. Maybe we'll find a formula that will really astonish our friends. What does Pfaff's reflection law tell us if we apply it to the strange form (5.99), where $z = 2$? In this case we set $a = -m$, $b = 1$, and $c = -2m + \epsilon$, obtaining

$$\begin{aligned}(-1)^m \lim_{\epsilon\to 0} F\left({-m, 1 \atop -2m+\epsilon}\middle|\, 2\right) &= \lim_{\epsilon\to 0} F\left({-m, -2m-1+\epsilon \atop -2m+\epsilon}\middle|\, 2\right)\\ &= \lim_{\epsilon\to 0} \sum_{k\geqslant 0} \frac{(-m)^{\overline{k}}\,(-2m-1+\epsilon)^{\overline{k}}}{(-2m+\epsilon)^{\overline{k}}}\,\frac{2^k}{k!}\\ &= \sum_{k\leqslant m} \binom{m}{k} \frac{(2m+1)^{\underline{k}}}{(2m)^{\underline{k}}}(-2)^k\,,\end{aligned}$$

because none of the limiting terms is close to zero. This leads to another miraculous formula,

(Hysterical note: See exercise 51 if you get a different result.)

$$\begin{aligned}\sum_{k\leqslant m} \binom{m}{k} \frac{2m+1}{2m+1-k}(-2)^k &= (-1)^m 2^{2m} \bigg/ \binom{2m}{m}\\ &= 1 \bigg/ \binom{-1/2}{m}\,, \qquad \text{integer } m \geqslant 0. \qquad (5.104)\end{aligned}$$

When $m = 3$, for example, the sum is

$$1 - 7 + \frac{84}{5} - 14 = -\frac{16}{5}\,,$$

and $\binom{-1/2}{3}$ is indeed equal to $-\frac{5}{16}$.

When we looked at our binomial coefficient identities and converted them to hypergeometric form, we overlooked (5.19) because it was a relation between two sums instead of a closed form. But now we can regard (5.19) as an identity between hypergeometric series. If we differentiate it n times with respect to y and then replace k by $m - n - k$, we get

$$\begin{aligned}&\sum_{k\geqslant 0} \binom{m+r}{m-n-k}\binom{n+k}{n} x^{m-n-k} y^k\\ &\quad = \sum_{k\geqslant 0} \binom{-r}{m-n-k}\binom{n+k}{n} (-x)^{m-n-k}(x+y)^k\,.\end{aligned}$$

This yields the following hypergeometric transformation:

$$F\left({a, -n \atop c}\middle|\, z\right) = \frac{(a-c)^{\underline{n}}}{(-c)^{\underline{n}}}\, F\left({a, -n \atop 1-n+a-c}\middle|\, 1-z\right), \quad \begin{matrix}\text{integer}\\ n \geqslant 0.\end{matrix} \qquad (5.105)$$

Notice that when $z = 1$ this reduces to Vandermonde's convolution, (5.93).

Differentiation seems to be useful, if this example is any indication; we also found it helpful in Chapter 2, when summing $x + 2x^2 + \cdots + nx^n$. Let's see what happens when a general hypergeometric series is differentiated with respect to z:

$$\begin{aligned}
\frac{d}{dz} F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n} \middle| z\right) &= \sum_{k \geqslant 1} \frac{a_1^{\overline{k}} \ldots a_m^{\overline{k}} \, z^{k-1}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}} \, (k-1)!} \\
&= \sum_{k+1 \geqslant 1} \frac{a_1^{\overline{k+1}} \ldots a_m^{\overline{k+1}} \, z^{k}}{b_1^{\overline{k+1}} \ldots b_n^{\overline{k+1}} \, k!} \\
&= \sum_{k \geqslant 0} \frac{a_1 (a_1{+}1)^{\overline{k}} \ldots a_m (a_m{+}1)^{\overline{k}} \, z^{k}}{b_1 (b_1{+}1)^{\overline{k}} \ldots b_n (b_n{+}1)^{\overline{k}} \, k!} \\
&= \frac{a_1 \ldots a_m}{b_1 \ldots b_n} F\left({a_1{+}1, \ldots, a_m{+}1 \atop b_1{+}1, \ldots, b_n{+}1} \middle| z\right). \qquad (5.106)
\end{aligned}$$

The parameters move out and shift up.

It's also possible to use differentiation to tweak just one of the parameters while holding the rest of them fixed. For this we use the operator

How do you pronounce ϑ?
(Dunno, but TeX calls it 'vartheta'.)

$$\vartheta = z \frac{d}{dz},$$

which acts on a function by differentiating it and then multiplying by z. This operator gives

$$\vartheta F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n} \middle| z\right) = z \sum_{k \geqslant 1} \frac{a_1^{\overline{k}} \ldots a_m^{\overline{k}} \, z^{k-1}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}} \, (k-1)!} = \sum_{k \geqslant 0} \frac{k \, a_1^{\overline{k}} \ldots a_m^{\overline{k}} \, z^{k}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}} \, k!},$$

which by itself isn't too useful. But if we multiply F by one of its upper parameters, say a_1, and add it to ϑF, we get

$$\begin{aligned}
(\vartheta + a_1) F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n} \middle| z\right) &= \sum_{k \geqslant 0} \frac{(k + a_1) a_1^{\overline{k}} \ldots a_m^{\overline{k}} \, z^{k}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}} \, k!}, \\
&= \sum_{k \geqslant 0} \frac{a_1 (a_1{+}1)^{\overline{k}} \, a_2^{\overline{k}} \ldots a_m^{\overline{k}} \, z^{k}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}} \, k!} \\
&= a_1 F\left({a_1{+}1, \, a_2, \, \ldots, \, a_m \atop b_1, \, \ldots, \, b_n} \middle| z\right).
\end{aligned}$$

Only one parameter has been shifted.

A similar trick works with lower parameters, but in this case things shift down instead of up:

$$\begin{aligned}(\vartheta + b_1 - 1)\,F\left({a_1,\dots,a_m \atop b_1,\dots,b_n}\,\Big|\,z\right) &= \sum_{k\geqslant 0}\frac{(k+b_1-1)a_1^{\overline{k}}\dots a_m^{\overline{k}}\,z^k}{b_1^{\overline{k}}\dots b_n^{\overline{k}}\,k!}\,,\\ &= \sum_{k\geqslant 0}\frac{(b_1-1)\,a_1^{\overline{k}}\dots a_m^{\overline{k}}\,z^k}{(b_1-1)^{\overline{k}}\,b_2^{\overline{k}}\dots b_n^{\overline{k}}\,k!}\\ &= (b_1{-}1)\,F\left({a_1,\ \dots,\ a_m \atop b_1{-}1,\ b_2,\ \dots,\ b_n}\,\Big|\,z\right).\end{aligned}$$

We can now combine all these operations and make a mathematical "pun" by expressing the same quantity in two different ways. Namely, we have

Ever hear the one about the brothers who named their cattle ranch Focus, because it's where the sons raise meat?

$$(\vartheta + a_1)\dots(\vartheta + a_m)F \;=\; a_1\dots a_m\;F\left({a_1{+}1,\ \dots,\ a_m{+}1 \atop b_1,\ \dots,\ b_n}\,\Big|\,z\right),$$

and

$$\begin{aligned}&(\vartheta + b_1 - 1)\dots(\vartheta + b_n - 1)F\\ &\qquad = (b_1{-}1)\dots(b_n{-}1)\,F\left({a_1,\ \dots,\ a_m \atop b_1{-}1,\ \dots,\ b_n{-}1}\,\Big|\,z\right),\end{aligned}$$

where $F = F(a_1,\dots,a_m;b_1,\dots,b_n;z)$. And (5.106) tells us that the top line is the derivative of the bottom line. Therefore the general hypergeometric function F satisfies the differential equation

$$D(\vartheta + b_1 - 1)\dots(\vartheta + b_n - 1)F \;=\; (\vartheta + a_1)\dots(\vartheta + a_m)F\,, \tag{5.107}$$

where D is the operator $\frac{d}{dz}$.

This cries out for an example. Let's find the differential equation satisfied by the standard 2-over-1 hypergeometric series $F(z) = F(a,b;c;z)$. According to (5.107), we have

$$D(\vartheta + c - 1)F \;=\; (\vartheta + a)(\vartheta + b)F\,.$$

What does this mean in ordinary notation? Well, $(\vartheta + c - 1)F$ is $zF'(z) + (c-1)F(z)$, and the derivative of this gives the left-hand side,

$$F'(z) + zF''(z) + (c-1)F'(z)\,.$$

On the right-hand side we have

$$\begin{aligned}(\vartheta+a)\big(zF'(z)+bF(z)\big) &= z\frac{d}{dz}\big(zF'(z)+bF(z)\big) + a\big(zF'(z)+bF(z)\big)\\ &= zF'(z)+z^2F''(z)+bzF'(z)+azF'(z)+abF(z)\,.\end{aligned}$$

Equating the two sides tells us that

$$z(1-z)F''(z) + \big(c - z(a+b+1)\big)F'(z) - abF(z) \;=\; 0\,. \tag{5.108}$$

This equation is equivalent to the factored form (5.107).

Conversely, we can go back from the differential equation to the power series. Let's assume that $F(z) = \sum_{k\geqslant 0} t_k z^k$ is a power series satisfying (5.107). A straightforward calculation shows that we must have

$$\frac{t_{k+1}}{t_k} \;=\; \frac{(k+a_1)\ldots(k+a_m)}{(k+b_1)\ldots(k+b_n)(k+1)}\,;$$

hence $F(z)$ must be $t_0\,F(a_1,\ldots,a_m;b_1,\ldots,b_n;z)$. We've proved that the hypergeometric series (5.76) is the only formal power series that satisfies the differential equation (5.107) and has the constant term 1.

It would be nice if hypergeometrics solved all the world's differential equations, but they don't quite. The right-hand side of (5.107) always expands into a sum of terms of the form $\alpha_k z^k F^{(k)}(z)$, where $F^{(k)}(z)$ is the kth derivative $D^kF(k)$; the left-hand side always expands into a sum of terms of the form $\beta_k z^{k-1}F^{(k)}(z)$ with $k>0$. So the differential equation (5.107) always takes the special form

$$z^{n-1}(\beta_n - z\alpha_n)F^{(n)}(z) + \cdots + (\beta_1 - z\alpha_1)F'(z) - \alpha_0 F(z) \;=\; 0\,.$$

Equation (5.108) illustrates this in the case $n=2$. Conversely, we will prove in exercise 6.13 that any differential equation of this form can be factored in terms of the ϑ operator, to give an equation like (5.107). So these are the differential equations whose solutions are power series with rational term ratios.

The function $F(z) = (1-z)^r$ satisfies $\vartheta F = z(\vartheta - r)F$. This gives another proof of the binomial theorem.

Multiplying both sides of (5.107) by z dispenses with the D operator and gives us an instructive all-ϑ form,

$$\vartheta(\vartheta+b_1-1)\ldots(\vartheta+b_n-1)F \;=\; z(\vartheta+a_1)\ldots(\vartheta+a_m)F\,. \tag{5.109}$$

The first factor $\vartheta = (\vartheta+1-1)$ on the left corresponds to the $(k+1)$ in the term ratio (5.81), which corresponds to the k! in the denominator of the kth term in a general hypergeometric series. The other factors $(\vartheta+b_j-1)$ correspond to the denominator factor $(k+b_j)$, which corresponds to $b_j^{\overline{k}}$ in (5.76). On the right, the z corresponds to z^k, and $(\vartheta+a_j)$ corresponds to $a_j^{\overline{k}}$.

One use of this differential theory is to find and prove new transformations. For example, we can readily verify that both of the hypergeometrics

$$F\left({2a,\,2b \atop a+b+\frac12}\,\middle|\,z\right) \quad\text{and}\quad F\left({a,\,b \atop a+b+\frac12}\,\middle|\,4z(1-z)\right)$$

satisfy the differential equation

$$z(1-z)F''(z) + (a+b+\tfrac12)(1-2z)F'(z) - 4abF(z) \;=\; 0\,;$$

hence *Gauss's identity* [143, equation 102]

$$F\left({2a,\,2b \atop a+b+\frac12}\,\middle|\,z\right) \;=\; F\left({a,\,b \atop a+b+\frac12}\,\middle|\,4z(1-z)\right) \tag{5.110}$$

must be true. In particular,

$$F\left({2a,\,2b \atop a+b+\frac12}\,\middle|\,\frac12\right) \;=\; F\left({a,\,b \atop a+b+\frac12}\,\middle|\,1\right), \tag{5.111}$$

whenever both infinite sums converge.

(Caution: We can't use (5.110) safely when $|z| > 1/2$, unless both sides are polynomials; see exercise 53.)

Every new identity for hypergeometrics has consequences for binomial coefficients, and this one is no exception. Let's consider the sum

$$\sum_{k\le m}\binom{m-k}{n}\binom{m+n+1}{k}\left(\frac{-1}{2}\right)^k, \qquad \text{integers } m \ge n \ge 0.$$

The terms are nonzero for $0 \le k \le m-n$, and with a little delicate limit-taking as before we can express this sum as the hypergeometric

$$\lim_{\epsilon\to0}\binom{m}{n}F\left({n-m,\;-n-m-1+\alpha\epsilon \atop -m+\epsilon}\,\middle|\,\frac12\right).$$

The value of α doesn't affect the limit, since the nonpositive upper parameter $n-m$ cuts the sum off early. We can set $\alpha = 2$, so that (5.111) applies. The limit can now be evaluated because the right-hand side is a special case of (5.92). The result can be expressed in simplified form,

$$\sum_{k\le m}\binom{m-k}{n}\binom{m+n+1}{k}\left(\frac{-1}{2}\right)^k = \binom{(m+n)/2}{n}2^{n-m}[m+n \text{ is even}]\,, \qquad \begin{matrix}\text{integers}\\ m\ge n\ge 0,\end{matrix} \tag{5.112}$$

as shown in exercise 54. For example, when $m = 5$ and $n = 2$ we get $\binom52\binom80 - \binom42\binom81/2 + \binom32\binom82/4 - \binom22\binom83/8 = 10-24+21-7 = 0$; when $m = 4$ and $n = 2$, both sides give $\frac34$.

We can also find cases where (5.110) gives binomial sums when $z = -1$, but these are really weird. If we set $a = \frac{1}{6} - \frac{n}{3}$ and $b = -n$, we get the monstrous formula

$$F\left({\frac{1}{3} - \frac{2}{3}n,\ -2n \atop \frac{2}{3} - \frac{4}{3}n} \,\Big|\, -1 \right) \;=\; F\left({\frac{1}{6} - \frac{1}{3}n,\ -n \atop \frac{2}{3} - \frac{4}{3}n} \,\Big|\, -8 \right).$$

These hypergeometrics are nondegenerate polynomials when $n \not\equiv 2 \pmod 3$; and the parameters have been cleverly chosen so that the left-hand side can be evaluated by (5.94). We are therefore led to a truly mind-boggling result,

$$\sum_k \binom{n}{k}\binom{\frac{1}{3}n - \frac{1}{6}}{k} 8^k \Big/ \binom{\frac{4}{3}n - \frac{2}{3}}{k}$$
$$= \binom{2n}{n}\Big/\binom{\frac{4}{3}n - \frac{2}{3}}{n}, \qquad \text{integer } n \geqslant 0,\ \ n \not\equiv 2 \pmod 3. \qquad (5.113)$$

This is the most startling identity in binomial coefficients that we've seen. Small cases of the identity aren't even easy to check by hand. (It turns out that both sides do give $\frac{81}{7}$ when $n = 3$.) But the identity is completely useless, of course; surely it will never arise in a practical problem.

The only use of (5.113) is to demonstrate the existence of incredibly useless identities.

So that's our hype for hypergeometrics. We've seen that hypergeometric series provide a high-level way to understand what's going on in binomial coefficient sums. A great deal of additional information can be found in the classic book by Bailey [18] and its sequel by Gasper and Rahman [141].

5.7 PARTIAL HYPERGEOMETRIC SUMS

Most of the sums we've evaluated in this chapter range over all indices $k \geqslant 0$, but sometimes we've been able to find a closed form that works over a general range $a \leqslant k < b$. For example, we know from (5.16) that

$$\sum_{k<m} \binom{n}{k}(-1)^k \;=\; (-1)^{m-1}\binom{n-1}{m-1}, \qquad \text{integer } m. \qquad (5.114)$$

The theory in Chapter 2 gives us a nice way to understand formulas like this: If $f(k) = \Delta g(k) = g(k+1) - g(k)$, then we've agreed to write $\sum f(k)\,\delta k = g(k) + C$, and

$$\textstyle\sum_a^b f(k)\,\delta k \;=\; g(k)\,\big|_a^b \;=\; g(b) - g(a).$$

Furthermore, when a and b are integers with $a \leqslant b$, we have

$$\textstyle\sum_a^b f(k)\,\delta k \;=\; \displaystyle\sum_{a\leqslant k<b} f(k) \;=\; g(b) - g(a).$$

Therefore identity (5.114) corresponds to the indefinite summation formula

$$\sum \binom{n}{k}(-1)^k \,\delta k \;=\; (-1)^{k-1}\binom{n-1}{k-1} + C\,,$$

and to the difference formula

$$\Delta\left((-1)^k\binom{n}{k}\right) \;=\; (-1)^{k+1}\binom{n+1}{k+1}.$$

It's easy to start with a function $g(k)$ and to compute $\Delta g(k) = f(k)$, a function whose sum will be $g(k)+C$. But it's much harder to start with $f(k)$ and to figure out its indefinite sum $\sum f(k)\,\delta k = g(k) + C$; this function g might not have a simple form. For example, there is apparently no simple form for $\sum \binom{n}{k}\,\delta k$; otherwise we could evaluate sums like $\sum_{k\leqslant n/3}\binom{n}{k}$, about which we're clueless. Yet maybe there is a simple form for $\sum \binom{n}{k}\,\delta k$ and we just haven't thought of it; how can we be sure?

In 1977, R. W. Gosper [154] discovered a beautiful way to find indefinite sums $\sum f(k)\,\delta k = g(k) + C$ whenever f and g belong to a general class of functions called hypergeometric terms. Let us write

$$F\left({a_1,\,\ldots,\,a_m \atop b_1,\,\ldots,\,b_n}\,\middle|\,z\right)_k \;=\; \frac{a_1^{\overline{k}}\ldots a_m^{\overline{k}}}{b_1^{\overline{k}}\ldots b_n^{\overline{k}}}\,\frac{z^k}{k!} \tag{5.115}$$

for the kth term of the hypergeometric series $F(a_1,\ldots,a_m;b_1,\ldots,b_n;z)$. We will regard $F(a_1,\ldots,a_m;b_1,\ldots,b_n;z)_k$ as a function of k, not of z. In many cases it turns out that there are parameters c, A_1, …, A_M, B_1, …, B_N, and Z such that

$$\sum F\left({a_1,\,\ldots,\,a_m \atop b_1,\,\ldots,\,b_n}\,\middle|\,z\right)_k \delta k \;=\; c\,F\left({A_1,\,\ldots,\,A_M \atop B_1,\,\ldots,\,B_N}\,\middle|\,Z\right)_k + C\,, \tag{5.116}$$

given a_1, …, a_m, b_1, …, b_n, and z. We will say that a given function $F(a_1,\ldots,a_m;b_1,\ldots,b_n;z)_k$ is *summable in hypergeometric terms* if such constants c, A_1, …, A_M, B_1, …, B_N, Z exist. Gosper's algorithm either finds the unknown constants or proves that no such constants exist.

In general, we say that $t(k)$ is a *hypergeometric term* if $t(k+1)/t(k)$ is a rational function of k, not identically zero. This means, in essence, that $t(k)$ is a constant multiple of a term like (5.115). (A technicality arises, however, with respect to zeros, because we want $t(k)$ to be meaningful when k is negative and when one or more of the b's in (5.115) is zero or a negative integer. Strictly speaking, we get the most general hypergeometric term by multiplying (5.115) by a nonzero constant times a power of 0, then cancelling zeros of the numerator with zeros of the denominator. The examples in exercise 12 help clarify this general rule.)

Suppose we want to find $\sum t(k)\,\delta k$, when $t(k)$ is a hypergeometric term. Gosper's algorithm proceeds in two steps, each of which is fairly straightforward. Step 1 is to express the term ratio in the special form

$$\frac{t(k+1)}{t(k)} = \frac{p(k+1)}{p(k)}\frac{q(k)}{r(k+1)}, \tag{5.117}$$

(Divisibility of polynomials is analogous to divisibility of integers. For example, $(k+\alpha)\backslash q(k)$ means that the quotient $q(k)/(k+\alpha)$ is a polynomial. It's easy to see that $(k+\alpha)\backslash q(k)$ if and only if $q(-\alpha)=0$.)

where p, q, and r are polynomials subject to the following condition:

$$\begin{aligned}&(k+\alpha)\backslash q(k) \quad\text{and}\quad (k+\beta)\backslash r(k)\\ &\qquad \Longrightarrow\ \alpha-\beta \text{ is not a positive integer.}\end{aligned} \tag{5.118}$$

This condition is easy to achieve: We start by provisionally setting $p(k)=1$, and we set $q(k)$ and $r(k+1)$ to the numerator and denominator of the term ratio, factoring them into linear factors. For example, if $t(k)$ has the form (5.115), we start with the factorizations $q(k) = (k+a_1)\ldots(k+a_m)z$ and $r(k) = (k+b_1-1)\ldots(k+b_n-1)k$. Then we check if (5.118) is violated. If q and r have factors $(k+\alpha)$ and $(k+\beta)$ where $\alpha-\beta = N > 0$, we divide them out of q and r and replace $p(k)$ by

$$p(k)(k+\alpha-1)^{\underline{N-1}} = p(k)(k+\alpha-1)(k+\alpha-2)\ldots(k+\beta+1). \tag{5.119}$$

The new p, q, and r still satisfy (5.117), and we can repeat this process until (5.118) holds. We'll see in a moment why (5.118) is important.

Step 2 of Gosper's algorithm is to finish the job—to find a hypergeometric term $T(k)$ such that

$$t(k) = T(k+1) - T(k), \tag{5.120}$$

whenever possible. But it's not obvious how to do this; we need to develop some theory before we know how to proceed. Gosper noticed, after studying a lot of special cases, that it is wise to write the unknown function $T(k)$ in the form

$$T(k) = \frac{r(k)\,s(k)\,t(k)}{p(k)}, \tag{5.121}$$

(Exercise 55 gives a clue about why we might want to make this magic substitution.)

where $s(k)$ is a secret function that must be discovered somehow. Plugging (5.121) into (5.120) and applying (5.117) gives

$$\begin{aligned} t(k) &= \frac{r(k+1)s(k+1)t(k+1)}{p(k+1)} - \frac{r(k)s(k)t(k)}{p(k)}\\ &= \frac{q(k)s(k+1)t(k)}{p(k)} - \frac{r(k)s(k)t(k)}{p(k)};\end{aligned}$$

so we need to have

$$p(k) \;=\; q(k)s(k+1) - r(k)s(k)\,. \tag{5.122}$$

If we can find $s(k)$ satisfying this fundamental recurrence relation, we've found $\sum t(k)\,\delta k$. If we can't, there's no T.

We're assuming that $T(k)$ is a hypergeometric term, which means that $T(k+1)/T(k)$ is a rational function of k. Therefore, by (5.121) and (5.120), $r(k)s(k)/p(k) = T(k)/\bigl(T(k+1) - T(k)\bigr)$ is a rational function of k, and $s(k)$ itself must be a quotient of polynomials:

$$s(k) \;=\; f(k)/g(k)\,.$$

But in fact we can prove that $s(k)$ is itself a polynomial. For if $g(k)$ is not constant, and if $f(k)$ and $g(k)$ have no common factors, let N be the largest integer such that $(k+\beta)$ and $(k+\beta+N-1)$ both occur as factors of $g(k)$ for some complex number β. The value of N is positive, since $N = 1$ always satisfies this condition. Equation (5.122) can be rewritten

$$p(k)g(k{+}1)g(k) \;=\; q(k)f(k{+}1)g(k) - r(k)g(k{+}1)f(k)\,,$$

and if we set $k = -\beta$ and $k = -\beta - N$ we get

$$r(-\beta)g(1{-}\beta)f(-\beta) \;=\; 0 \;=\; q(-\beta{-}N)f(1{-}\beta{-}N)g(-\beta{-}N)\,.$$

Now $f(-\beta) \neq 0$ and $f(1-\beta-N) \neq 0$, because f and g have no common roots. Also $g(1-\beta) \neq 0$ and $g(-\beta-N) \neq 0$, because $g(k)$ would otherwise contain the factor $(k+\beta-1)$ or $(k+\beta+N)$, contrary to the maximality of N. Therefore

$$r(-\beta) \;=\; q(-\beta-N) \;=\; 0\,.$$

But this contradicts condition (5.118). Hence $s(k)$ must be a polynomial.

I see: Gosper came up with condition (5.118) in order to make this proof go through.

Our task now boils down to finding a polynomial $s(k)$ that satisfies (5.122), when $p(k)$, $q(k)$, and $r(k)$ are given polynomials, or proving that no such polynomial exists. It's easy to do this when $s(k)$ has any particular degree d, since we can write

$$s(k) \;=\; \alpha_d k^d + \alpha_{d-1}k^{d-1} + \cdots + \alpha_0\,, \qquad \alpha_d \;\neq\; 0 \tag{5.123}$$

for unknown coefficients $(\alpha_d, \ldots, \alpha_0)$ and plug this expression into the fundamental recurrence (5.122). The polynomial $s(k)$ will satisfy the recurrence if and only if the α's satisfy the linear equations that result when we equate coefficients of each power of k in (5.122).

But how can we determine the degree of s? It turns out that there actually are at most two possibilities. We can rewrite (5.122) in the form

$$2p(k) \;=\; Q(k)\bigl(s(k+1)+s(k)\bigr)+R(k)\bigl(s(k+1)-s(k)\bigr)\,, \qquad (5.124)$$
$$\text{where}\quad Q(k) \;=\; q(k)-r(k) \quad\text{and}\quad R(k) \;=\; q(k)+r(k)\,.$$

If $s(k)$ has degree d, then the sum $s(k+1)+s(k)=2\alpha_d k^d+\cdots$ also has degree d, while the difference $s(k+1)-s(k)=\Delta s(k)=d\alpha_d k^{d-1}+\cdots$ has degree $d-1$. (The zero polynomial can be assumed to have degree -1.) Let's write $\deg(P)$ for the degree of a polynomial P. If $\deg(Q)\geqslant\deg(R)$, then the degree of the right-hand side of (5.124) is $\deg(Q)+d$, so we must have $d=\deg(p)-\deg(Q)$. On the other hand if $\deg(Q)<\deg(R)=d'$, we can write $Q(k)=\lambda' k^{d'-1}+\cdots$ and $R(k)=\lambda k^{d'}+\cdots$ where $\lambda\neq 0$; the right-hand side of (5.124) has the form

$$(2\lambda'\alpha_d+\lambda d\,\alpha_d)k^{d+d'-1}+\cdots\,.$$

Ergo, two possibilities: Either $2\lambda'+\lambda d\neq 0$, and $d=\deg(p)-\deg(R)+1$; or $2\lambda'+\lambda d=0$, and $d>\deg(p)-\deg(R)+1$. The second case needs to be examined only if $-2\lambda'/\lambda$ is an integer d greater than $\deg(p)-\deg(R)+1$.

OK, we now have enough facts to perform Step 2 of Gosper's two-step algorithm: By trying at most two values of d, we can discover $s(k)$, whenever equation (5.122) has a polynomial solution. If $s(k)$ exists, we can plug it into (5.121) and we have our T. If it doesn't, we've proved that $t(k)$ is not summable in hypergeometric terms.

Time for an example: Let's try the partial sum (5.114). Gosper's method should be able to deduce the value of

$$\sum\binom{n}{k}(-1)^k\,\delta k$$

for any fixed n, so we seek the sum of

$$t(k) \;=\; \binom{n}{k}(-1)^k \;=\; \frac{n!\,(-1)^k}{k!\,(n-k)!}\,.$$

Step 1 is to put the term ratio into the required form (5.117); we have

$$\frac{t(k+1)}{t(k)} \;=\; \frac{k-n}{k+1} \;=\; \frac{p(k+1)\,q(k)}{p(k)\,r(k+1)}$$

Why isn't it $r(k)=k+1$? *Oh, I see.*

so we simply take $p(k)=1$, $q(k)=k-n$, and $r(k)=k$. This choice of p, q, and r satisfies (5.118), unless n is a negative integer; let's suppose it isn't.

Now we do Step 2. According to (5.124), we should consider the polynomials $Q(k)=-n$ and $R(k)=2k-n$. Since R has larger degree than Q, we need to look at two cases. Either $d = \deg(p) - \deg(R) + 1$, which is 0; or $d = -2\lambda'/\lambda$ where $\lambda' = -n$ and $\lambda = 2$, hence $d = n$. The first case is nicer, because it doesn't require n to be a positive integer, so let's try it first; we'll need to try the other possibility for d only if the first case fails. Assuming that $d = 0$, the value of $s(k)$ is simply α_0, and equation (5.122) reduces to

$$1 = (k-n)\alpha_0 - k\alpha_0 .$$

Hence we choose $\alpha_0 = -1/n$. This satisfies the equation and gives

$$\begin{aligned} T(k) &= \frac{r(k)\,s(k)\,t(k)}{p(k)} \\ &= k \cdot \left(\frac{-1}{n}\right) \cdot \binom{n}{k}(-1)^k \\ &= \binom{n-1}{k-1}(-1)^{k-1}, \qquad \text{if } n \ne 0, \end{aligned}$$

precisely the answer we were hoping to confirm.

If we apply the same method to find the indefinite sum $\sum \binom{n}{k}\,\delta k$, without the $(-1)^k$, everything will be almost the same except that $q(k)$ will be $n-k$; hence $Q(k) = n - 2k$ will have greater degree than $R(k) = n$, and we will conclude that d has the impossible value $\deg(p) - \deg(Q) = -1$. (The polynomial $s(k)$ cannot have negative degree, because it cannot be zero.) Therefore the function $\binom{n}{k}$ is not summable in hypergeometric terms.

However, once we have eliminated the impossible, whatever remains—however improbable—must be the truth (according to S. Holmes [83]). When we defined p, q, and r in Step 1, we decided to ignore the possibility that n might be a negative integer. What if it is? Let's set $n = -N$, where N is positive. Then the term ratio for $\sum \binom{n}{k}\,\delta k$ is

$$\frac{t(k+1)}{t(k)} = \frac{-(k+N)}{(k+1)} = \frac{p(k+1)}{p(k)}\,\frac{q(k)}{r(k+1)}$$

and it should be represented by $p(k) = (k+1)^{\overline{N-1}}$, $q(k) = -1$, $r(k) = 1$, according to (5.119). Step 2 of Gosper's algorithm now tells us to look for a polynomial $s(k)$ of degree $d = N - 1$; maybe there's hope after all. For example, when $N = 2$ recurrence (5.122) says that we should solve

$$k+1 = -\bigl((k+1)\alpha_1 + \alpha_0\bigr) - (k\alpha_1 + \alpha_0) .$$

Equating coefficients of k and 1 tells us that

$$1 = -\alpha_1 - \alpha_1; \qquad 1 = -\alpha_1 - \alpha_0 - \alpha_0;$$

hence $s(k) = -\frac{1}{2}k - \frac{1}{4}$ is a solution, and

$$T(k) \;=\; \frac{1 \cdot \left(-\frac{1}{2}k - \frac{1}{4}\right) \cdot \binom{-2}{k}}{k+1} \;=\; (-1)^{k-1}\,\frac{2k+1}{4}\,.$$

"Excellent, Holmes!"
"Elementary, my dear Watson."

Can this be the desired sum? Yes, it checks out:

$$(-1)^k\,\frac{2k+3}{4} - (-1)^{k-1}\,\frac{2k+1}{4} \;=\; (-1)^k(k+1) \;=\; \binom{-2}{k}.$$

Incidentally, we can write this summation formula in another form, by attaching an upper limit:

$$\begin{aligned}\sum_{k<m}\binom{-2}{k} \;&=\; (-1)^{k-1}\,\frac{2k+1}{4}\,\Big|_0^m \\ &=\; \frac{(-1)^{m-1}}{2}\left(m + \frac{1-(-1)^m}{2}\right) \\ &=\; (-1)^{m-1}\left\lceil\frac{m}{2}\right\rceil, \qquad \text{integer } m \geqslant 0.\end{aligned}$$

This representation conceals the fact that $\binom{-2}{k}$ is summable in hypergeometric terms, because $\lceil m/2\rceil$ is not a hypergeometric term. (See exercise 12.)

A problem might arise in the denominator of (5.121) if $p(k) = 0$ for some integer k. Exercise 97 gives some insight into what can be done in such situations.

Notice that we need not bother to compile a catalog of indefinitely summable hypergeometric terms, analogous to the database of definite hypergeometric sums mentioned earlier in this chapter, because Gosper's algorithm provides a quick, uniform method that works in all summable cases.

Marko Petkovšek [291] has found a nice way to generalize Gosper's algorithm to more complicated inversion problems, by showing how to determine all hypergeometric terms $T(k)$ that satisfy the lth-order recurrence

$$t(k) \;=\; p_l(k)T(k+l) + \cdots + p_1(k)T(k+1) + p_0(k)T(k)\,, \tag{5.125}$$

given any hypergeometric term $t(k)$ and polynomials $p_l(k), \ldots, p_1(k), p_0(k)$.

5.8 MECHANICAL SUMMATION

Gosper's algorithm, beautiful as it is, finds a closed form for only a few of the binomial sums we meet in practice. But we need not stop there. Doron Zeilberger [383] showed how to extend Gosper's algorithm so that it becomes even more beautiful, making it succeed in vastly more cases. With

Zeilberger's extension we can handle summation over all k, not just partial sums, so we have an alternative to the hypergeometric methods of Sections 5.5 and 5.6. Moreover, as with Gosper's original method, the calculations can be done by computer, almost blindly; we need not rely on cleverness and luck.

The basic idea is to regard the term we want to sum as a function $t(n,k)$ of two variables n and k. (In Gosper's algorithm we wrote just $t(k)$.) When $t(n,k)$ does not turn out to be indefinitely summable in hypergeometric terms, with respect to k—and let's face it, relatively few functions are—Zeilberger noticed that we can often modify $t(n,k)$ in order to obtain another term that *is* indefinitely summable. For example, it often turns out in practice that $\beta_0(n)t(n,k)+\beta_1(n)t(n+1,k)$ is indefinitely summable with respect to k, for appropriate polynomials $\beta_0(n)$ and $\beta_1(n)$. And when we carry out the sum with respect to k, we obtain a recurrence in n that solves our problem.

Let's start with a simple case in order to get familiar with this general approach. Suppose we have forgotten the binomial theorem, and we want to evaluate $\sum_k \binom{n}{k} z^k$. How could we discover the answer, without clairvoyance or inspired guesswork? Earlier in this chapter, for example in Problem 3 of Section 5.2, we learned how to replace $\binom{n}{k}$ by $\binom{n-1}{k}+\binom{n-1}{k-1}$ and to fiddle around with the result. But there's a more systematic way to proceed.

Or without looking on page 174.

Let $t(n,k)=\binom{n}{k}z^k$ be the quantity we want to sum. Gosper's algorithm tells us that we can't evaluate the partial sums $\sum_{k\leqslant m} t(n,k)$ for arbitrary n in hypergeometric terms, except in the case $z=-1$. So let's consider a more general term

$$\hat{t}(n,k) \;=\; \beta_0(n)t(n,k)+\beta_1(n)t(n+1,k) \qquad (5.126)$$

instead. We'll look for values of $\beta_0(n)$ and $\beta_1(n)$ that make Gosper's algorithm succeed. First we want to simplify (5.126) by using the relation between $t(n+1,k)$ and $t(n,k)$ to eliminate $t(n+1,k)$ from the expression. Since

$$\begin{aligned}\frac{t(n+1,k)}{t(n,k)} \;&=\; \frac{(n+1)!\,z^k}{(n+1-k)!\,k!}\,\frac{(n-k)!\,k!}{n!\,z^k}\\ &=\; \frac{n+1}{n+1-k}\,,\end{aligned}$$

we have

$$\hat{t}(n,k) \;=\; p(n,k)\frac{t(n,k)}{n+1-k}\,,$$

where

$$p(n,k) \;=\; (n+1-k)\beta_0(n)+(n+1)\beta_1(n)\,.$$

We now apply Gosper's algorithm to $\hat{t}(n,k)$, with n held fixed, first writing

$$\frac{\hat{t}(n,k+1)}{\hat{t}(n,k)} = \frac{\hat{p}(n,k+1)}{\hat{p}(n,k)}\frac{q(n,k)}{r(n,k+1)} \tag{5.127}$$

as in (5.117). Gosper's method would find such a representation by starting with $\hat{p}(n,k) = 1$, but with Zeilberger's extension we are better off starting with $\hat{p}(n,k) = p(n,k)$. Notice that if we set $\bar{t}(n,k) = \hat{t}(n,k)/p(n,k)$ and $\bar{p}(n,k) = \hat{p}(n,k)/p(n,k)$, equation (5.127) is equivalent to

$$\frac{\bar{t}(n,k+1)}{\bar{t}(n,k)} = \frac{\bar{p}(n,k+1)}{\bar{p}(n,k)}\frac{q(n,k)}{r(n,k+1)}\,. \tag{5.128}$$

So we can find $\hat{p}$, q and r satisfying (5.127) by finding $\bar{p}$, q and r satisfying (5.128), starting with $\bar{p}(n,k) = 1$. This makes life easy, because $\bar{t}(n,k)$ does not involve the unknown quantities $\beta_0(n)$ and $\beta_1(n)$ that appear in $\hat{t}(n,k)$. In our case $\bar{t}(n,k) = t(n,k)/(n+1-k) = n!\,z^k/(n+1-k)!\,k!$, so we have

$$\frac{\bar{t}(n,k+1)}{\bar{t}(n,k)} = \frac{(n+1-k)z}{k+1}\,;$$

This time I remembered why $r(n,k)$ *isn't* $k+1$.

we may take $q(n,k) = (n+1-k)z$ and $r(n,k) = k$. These polynomials in k are supposed to satisfy condition (5.118). If they don't, we're supposed to remove factors from q and r and include corresponding factors (5.119) in $\bar{p}(n,k)$; but we should do this only when the quantity $\alpha - \beta$ in (5.118) is a positive integer constant, independent of n, because we want our calculations to be valid for arbitrary n. (The formulas we derive will, in fact, be valid even when n and k are not integers, using the generalized factorials (5.83).)

Our first choices of q and r do satisfy (5.118), in this sense, so we can move right on to Step 2 of Gosper's algorithm: We want to solve the analog of (5.122), using (5.127) in place of (5.117). So we want to solve

$$\hat{p}(n,k) = q(n,k)s(n,k+1) - r(n,k)s(n,k) \tag{5.129}$$

for the secret polynomial

$$s(n,k) = \alpha_d(n)k^d + \alpha_{d-1}(n)k^{d-1} + \cdots + \alpha_0(n)\,. \tag{5.130}$$

(The coefficients of s are considered to be functions of n, not just constants.) In our case equation (5.129) is

$$\begin{aligned}(n+1-k)\beta_0(n) &+ (n+1)\beta_1(n)\\ &= (n+1-k)zs(n,k+1) - ks(n,k)\,,\end{aligned}$$

and we regard this as a polynomial equation in k with coefficients that are functions of n. As before, we determine the degree d of s by considering

$Q(n,k) = q(n,k) - r(n,k)$ and $R(n,k) = q(n,k) + r(n,k)$. Since $\deg(Q) = \deg(R) = 1$ (assuming that $z \neq -1$), we have $d = \deg(\hat{p}) - \deg(Q) = 0$ and $s(n,k) = \alpha_0(n)$ is independent of k. Our equation becomes

The degree function $\deg(Q)$ refers here to the degree in k, treating n as constant.

$$(n+1-k)\beta_0(n) + (n+1)\beta_1(n) \;=\; (n+1-k)z\alpha_0(n) - k\alpha_0(n)\,;$$

and by equating powers of k we get the equivalent k-free equations

$$\begin{aligned}(n+1)\beta_0(n) + (n+1)\beta_1(n) - (n+1)z\alpha_0(n) &= 0\,,\\ -\beta_0(n) \qquad\qquad\qquad\qquad + (z+1)\alpha_0(n) &= 0\,.\end{aligned}$$

Hence we have a solution to (5.129) with

$$\beta_0(n) \;=\; z+1\,, \qquad \beta_1(n) \;=\; -1\,, \qquad \alpha_0(n) \;=\; s(n,k) \;=\; 1\,.$$

(By chance, n has dropped out.)

We have discovered, by a purely mechanical method, that the term $\hat{t}(n,k) = (z+1)t(n,k) - t(n+1,k)$ is summable in hypergeometric terms. In other words,

$$\hat{t}(n,k) \;=\; T(n,k+1) - T(n,k)\,, \tag{5.131}$$

where $T(n,k)$ is a hypergeometric term in k. What is this $T(n,k)$? According to (5.121) and (5.128), we have

$$T(n,k) \;=\; \frac{r(n,k)s(n,k)\hat{t}(n,k)}{\hat{p}(n,k)} \;=\; r(n,k)s(n,k)\bar{t}(n,k)\,, \tag{5.132}$$

because $\bar{p}(n,k) = 1$. (Indeed, $\bar{p}(n,k)$ almost always turns out to be 1 in practice.) Hence

$$T(n,k) \;=\; \frac{k}{n+1-k}\,t(n,k) \;=\; \frac{k}{n+1-k}\binom{n}{k}z^k \;=\; \binom{n}{k-1}z^k.$$

And sure enough, everything checks out — equation (5.131) is true:

$$(z+1)\binom{n}{k}z^k - \binom{n+1}{k}z^k \;=\; \binom{n}{k}z^{k+1} - \binom{n}{k-1}z^k\,.$$

But we don't actually need to know $T(n,k)$ precisely, because we are going to sum $t(n,k)$ over all integers k. All we need to know is that $T(n,k)$ is nonzero for only finitely many values of k, when n is any given nonnegative integer. Then the sum of $T(n,k+1) - T(n,k)$ over all k must telescope to 0.

Let $S_n = \sum_k t(n,k) = \sum_k \binom{n}{k}z^k$; this is the sum we started with, and we're now ready to compute it, because we now know a lot about $t(n,k)$. The

Gosper-Zeilberger procedure has deduced that

$$\sum_k \big((z+1)t(n,k) - t(n+1,k)\big) \;=\; 0\,.$$

In fact, $\lim_{k\to\infty} T(n,k) = 0$ *when* $|z| < 1$ *and* n *is any complex number. So (5.133) is true for all* n*, and in particular* $S_n = (z+1)^n$ *when* n *is a negative integer.*

But this sum is $(z+1)\sum_k t(n,k) - \sum_k t(n+1,k) = (z+1)S_n - S_{n+1}$. Therefore we have

$$S_{n+1} \;=\; (z+1)S_n\,. \tag{5.133}$$

Aha! This is a recurrence we know how to solve, provided that we know S_0. And obviously $S_0 = 1$. Hence we deduce that $S_n = (z+1)^n$, for all integers $n \geqslant 0$. QED.

Let's look back at this computation and summarize what we did, in a form that will apply also to other summands $t(n,k)$. The Gosper-Zeilberger algorithm can be formalated as follows, when $t(n,k)$ is given:

0 Set $l := 0$. (We'll seek recurrences in n of order l.)

1 Let $\hat{t}(n,k) = \beta_0(n)t(n,k)+\cdots+\beta_l(n)t(n+l,k)$, where $\beta_0(n), \ldots, \beta_l(n)$ are unknown functions. Use properties of $t(n,k)$ to find a linear combination $p(n,k)$ of $\beta_0(n), \ldots, \beta_l(n)$ with coefficients that are polynomials in n and k, so that $\hat{t}(n,k)$ can be written in the form $p(n,k)\bar{t}(n,k)$, where $\bar{t}(n,k)$ is a hypergeometric term in k. Find polynomials $\bar{p}(n,k)$, $q(n,k)$, $r(n,k)$ so that the term ratio of $\bar{t}(n,k)$ is expressed in the form (5.128), where $q(n,k)$ and $r(n,k)$ satisfy Gosper's condition (5.118). Set $\hat{p}(n,k) = p(n,k)\bar{p}(n,k)$.

2a Set $d_Q := \deg(q-r)$, $d_R := \deg(q+r)$, and

$$d \;:=\; \begin{cases} \deg(\hat{p}) - d_Q, & \text{if } d_Q \geqslant d_R;\\ \deg(\hat{p}) - d_R + 1, & \text{if } d_Q < d_R. \end{cases}$$

2b If $d \geqslant 0$, define $s(n,k)$ by (5.130), and consider the linear equations in $\alpha_0, \ldots, \alpha_d, \beta_0, \ldots \beta_l$ obtained by equating coefficients of powers of k in the fundamental equation (5.129). If these equations have a solution with $\beta_0, \ldots, \beta_l$ not all zero, go to Step 4. Otherwise, if $d_Q < d_R$ and if $-2\lambda'/\lambda$ is an integer greater than d, where λ is the coefficient of k^{d_R} in $q+r$ and λ' is the coefficient of k^{d_R-1} in $q-r$, set $d := -2\lambda'/\lambda$ and repeat Step 2b.

3 (The term $\hat{t}(n,k)$ isn't hypergeometrically summable.) Increase l by 1 and go back to Step 1.

4 (Success.) Set $T(n,k) := r(n,k)s(n,k)\bar{t}(n,k)/\bar{p}(n,k)$. The algorithm has discovered that $\hat{t}(n,k) = T(n,k+1) - T(n,k)$.

We'll prove later that this algorithm terminates successfully whenever $t(n,k)$ belongs to a large class of terms called proper terms.

The binomial theorem can be derived in many ways, so our first example of the Gosper-Zeilberger approach was more instructive than impressive. Let's tackle Vandermonde's convolution next. Can Gosper and Zeilberger deduce algorithmically that $\sum_k \binom{a}{k}\binom{b}{n-k}$ has a simple form? The algorithm starts with $l = 0$, which essentially reproduces Gosper's original algorithm, trying to see if $\binom{a}{k}\binom{b}{n-k}$ is summable in hypergeometric terms. Surprise: That term actually does turn out to be summable, if $a+b$ is a specific nonnegative integer (see exercise 94). But we are interested in general values of a and b, and the algorithm quickly discovers that the indefinite sum is not a hypergeometric term in general. So l is increased from 0 to 1, and the algorithm proceeds to try $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k)$ instead. The next step, as in our derivation of the binomial theorem, is to write $\hat{t}(n,k) = p(n,k)\bar{t}(n,k)$, where $p(n,k)$ is obtained by clearing fractions in $t(n+1,k)/t(n,k)$. In this case — the reader should please work along on a piece of scratch paper to check all these calculations — they aren't as hard as they look — everything goes through in an analogous fashion, but now with

$$\begin{aligned} p(n,k) &= (n+1-k)\beta_0(n) + (b-n+k)\beta_1(n) \;=\; \hat{p}(n,k)\,, \\ \bar{t}(n,k) &= t(n,k)/(n+1-k) \;=\; a!\,b!/(a-k)!\,k!\,(b-n+k)!\,(n+1-k)!\,, \\ q(n,k) &= (n+1-k)(a-k)\,, \\ r(n,k) &= (b-n+k)k\,. \end{aligned}$$

Step 2a finds $\deg(q-r) < \deg(q+r)$, and $d = \deg(\hat{p}) - \deg(q+r) + 1 = 0$, so $s(n,k)$ is again independent of k. Gosper's fundamental equation (5.129) is equivalent to two equations in three unknowns,

The crucial point is that the Gosper-Zeilberger method always leads to equations that are linear in the unknown α's and β's, because the left side of (5.129) is linear in the β's and the right side is linear in the α's.

$$\begin{aligned} (n+1)\beta_0(n) + (b-n)\beta_1(n) \quad - (n+1)a\,\alpha_0(n) &= 0\,, \\ -\beta_0(n) \qquad\qquad +\beta_1(n) + (a+b+1)\alpha_0(n) &= 0\,, \end{aligned}$$

which have the solution

$$\beta_0(n) = a+b-n\,, \qquad \beta_1(n) = -n-1\,, \qquad \alpha_0(n) = 1\,.$$

We conclude that $(a+b-n)t(n,k) - (n+1)t(n+1,k)$ is summable with respect to k; hence if $S_n = \sum_k \binom{a}{k}\binom{b}{n-k}$ the recurrence

$$S_{n+1} = \frac{a+b-n}{n+1}\,S_n$$

holds; thus $S_n = \binom{a+b}{n}$ since $S_0 = 1$. A piece of cake.

What about the Saalschützian triple-binomial identity in (5.28)? The proof of (5.28) in exercise 43 is interesting, but it requires inspiration. When we transform an art into a science, we aim to replace inspiration by perspiration; so let's see if the Gosper-Zeilberger approach to summation is able to discover and prove (5.28) in a purely mechanical way. For convenience we make the substitutions $m = b + d$, $n = a$, $r = a + b + c + d$, $s = a + b + c$, so that (5.28) takes the more symmetrical form

$$\sum_k \frac{(a+b+c+d+k)!}{(a-k)!\,(b-k)!\,(c+k)!\,(d+k)!\,k!} \\ = \frac{(a+b+c+d)!\,(a+b+c)!\,(a+b+d)!}{a!\,b!\,(a+c)!\,(a+d)!\,(b+c)!\,(b+d)!}\,. \tag{5.134}$$

To make the sum finite, we assume that either a or b is a nonnegative integer.

Deciding what parameter to call n is the only nonmechanical part.

Let $t(n,k) = (n+b+c+d+k)!/(n-k)!\,(b-k)!\,(c+k)!\,(d+k)!\,k!$ and $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k)$. Proceeding along a path that is beginning to become well worn, we set

$$\begin{aligned} p(n,k) &= (n+1-k)\beta_0(n) + (n+1+b+c+d+k)\beta_1(n) \;=\; \hat{p}(n,k)\,, \\ \bar{t}(n,k) &= \frac{t(n,k)}{n+1-k} \;=\; \frac{(n+b+c+d+k)!}{(n+1-k)!\,(b-k)!\,(c+k)!\,(d+k)!\,k!}\,, \\ q(n,k) &= (n+b+c+d+k+1)(n+1-k)(b-k)\,, \\ r(n,k) &= (c+k)(d+k)k\,, \end{aligned}$$

and we try to solve (5.129) for $s(n,k)$. Again $\deg(q-r) < \deg(q+r)$, but this time $\deg(\hat{p}) - \deg(q+r) + 1 = -1$ so it looks like we're stuck. However, Step 2b has an important second choice, $d = -2\lambda'/\lambda$, for the degree of s; we had better try it now before we give up. Here $R(n,k) = q(n,k) + r(n,k) = 2k^3 + \cdots$, so $\lambda = 2$, while the polynomial $Q(n,k) = q(n,k) - r(n,k)$ almost miraculously turns out to have degree 1 in k—the coefficient of k^2 vanishes! Therefore $\lambda' = 0$; Gosper allows us to take $d = 0$ and $s(n,k) = \alpha_0(n)$.

Notice that λ' is not the leading coefficient of Q, although λ is the leading coefficient of R. The number λ' is the coefficient of $k^{\deg(R)-1}$ in Q.

The equations to be solved are now

$$\begin{aligned} &(n+1)\beta_0(n) + (n+1+b+c+d)\beta_1(n) \\ &\qquad - (n+1)(n+1+b+c+d)b\alpha_0(n) \;=\; 0\,, \\ &-\beta_0(n) + \beta_1(n) \\ &\qquad - \big((n+1)b - (n+1+b)(n+1+b+c+d) - cd\big)\alpha_0(n) \;=\; 0\,; \end{aligned}$$

and we find

$$\begin{aligned} \beta_0(n) &= (n+1+b+c)(n+1+b+d)(n+1+b+c+d)\,, \\ \beta_1(n) &= -(n+1)(n+1+c)(n+1+d)\,, \\ \alpha_0(n) &= 2n+2+b+c+d\,, \end{aligned}$$

Perspiration flows, identity follows.

after only a modest amount of perspiration. The identity (5.134) follows immediately.

A similar proof of (5.134) can be obtained if we work with $n = d$ instead of $n = a$. (See exercise 99.)

The Gosper-Zeilberger approach helps us evaluate definite sums over a restricted range as well as sums over all k. For example, let's consider

$$S_n(z) \;=\; \sum_{k=0}^{n} \binom{n+k}{k} z^k\,. \tag{5.135}$$

When $z=\frac{1}{2}$ we obtained an "unexpected" result in (5.20); would Gosper and Zeilberger have expected it? Putting $t(n,k)=\binom{n+k}{k}z^k$ leads us to

$$\begin{aligned} p(n,k) \;&=\; (n+1)\beta_0(n) + (n+1+k)\beta_1(n) \;=\; \hat{p}(n,k)\,,\\ \bar{t}(n,k) \;&=\; t(n,k)/(n+1) \;=\; (n+k)!\,z^k/k!\,(n+1)!\,,\\ q(n,k) \;&=\; (n+1+k)z\,,\\ r(n,k) \;&=\; k\,, \end{aligned}$$

and $\deg(s)=\deg(\hat{p})-\deg(q-r)=0$. Equation (5.129) is solved by $\beta_0(n)=1$, $\beta_1(n)=z-1$, $s(n,k)=1$. Therefore we find

$$t(n,k) + (z-1)t(n+1,k) \;=\; T(n,k+1) - T(n,k)\,, \tag{5.136}$$

where $T(n,k)=r(n,k)s(n,k)\hat{t}(n,k)/\hat{p}(n,k)=\binom{n+k}{k-1}z^k$. We can now sum (5.136) for $0\leqslant k\leqslant n+1$, getting

$$\begin{aligned} S_n(z) + t(n,n+1) + (z-1)S_{n+1}(z) \;&=\; T(n,n+2) - T(n,0)\\ &=\; \binom{2n+2}{n+1} z^{n+2}\\ &=\; 2\binom{2n+1}{n} z^{n+2}\,. \end{aligned}$$

But $t(n,n+1)=\binom{2n+1}{n+1}z^{n+1}=\binom{2n+1}{n}z^{n+1}$, so

$$S_{n+1}(z) \;=\; \frac{1}{1-z}\left(S_n(z) + (1-2z)\binom{2n+1}{n} z^{n+1}\right). \tag{5.137}$$

We see immediately that the case $z=\frac{1}{2}$ is special, and that $S_{n+1}(\frac{1}{2})=2S_n(\frac{1}{2})$. Moreover, the recurrence (5.137) can be simplified by applying the summation factor $(1-z)^{n+1}$ to both sides; this yields the general identity

$$(1-z)^n \sum_{k=0}^{n}\binom{n+k}{k} z^k \;=\; 1 + \frac{1-2z}{2-2z}\sum_{k=1}^{n}\binom{2k}{k}\bigl(z(1-z)\bigr)^k\,, \tag{5.138}$$

which comparatively few people would have expected before Gosper and Zeilberger came along. Now the production of such identities is routine.

How about the similar sum

$$S_n(z) \;=\; \sum_{k=0}^{n} \binom{n-k}{k} z^k\,, \tag{5.139}$$

which we encountered in (5.74)? Flushed with confidence, we set $t(n,k) = \binom{n-k}{k} z^k$ and proceed to calculate

$$\begin{aligned} p(n,k) &= (n+1-2k)\beta_0(n) + (n+1-k)\beta_1(n) \;=\; \hat{p}(n,k)\,,\\ \bar{t}(n,k) &= t(n,k)/(n+1-2k) \;=\; (n-k)!\,z^k/k!\,(n+1-2k)!\,,\\ q(n,k) &= (n+1-2k)(n-2k)z\,,\\ r(n,k) &= (n+1-k)k\,. \end{aligned}$$

$S_n(-\frac{1}{4})$ *equals* $(n+1)/2^n$.

But whoa — there's no way to solve (5.129), if we assume that $z \neq -\frac{1}{4}$, because the degree of s would have to be $\deg(\hat{p}) - \deg(q-r) = -1$.

No problem. We simply add another parameter $\beta_2(n)$ and try $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k) + \beta_2(n)t(n+2,k)$ instead:

$$\begin{aligned} p(n,k) &= (n+1-2k)(n+2-2k)\beta_0(n)\\ &\qquad + (n+1-k)(n+2-2k)\beta_1(n)\\ &\qquad + (n+1-k)(n+2-k)\beta_2(n) \;=\; \hat{p}(n,k)\,,\\ \bar{t}(n,k) &= t(n,k)/(n{+}1{-}2k)(n{+}2{-}2k) \;=\; (n{-}k)!\,z^k/k!\,(n{+}2{-}2k)!\,,\\ q(n,k) &= (n+2-2k)(n+1-2k)z\,,\\ r(n,k) &= (n+1-k)k\,. \end{aligned}$$

Now we can try $s(n,k) = \alpha_0(n)$ and (5.129) does have a solution:

$$\beta_0(n) \;=\; z\,,\qquad \beta_1(n) \;=\; 1\,,\qquad \beta_2(n) \;=\; -1\,,\qquad \alpha_0(n) \;=\; 1\,.$$

We have discovered that

$$zt(n,k) + t(n+1,k) - t(n+2,k) \;=\; T(n,k+1) - T(n,k)\,,$$

where $T(n,k)$ equals $r(n,k)s(n,k)\hat{t}(n,k)/\hat{p}(n,k) \;=\; (n+1-k)k\bar{t}(n,k) \;=\; \binom{n+1-k}{k-1} z^k$. Summing from $k=0$ to $k=n$ gives

$$\begin{aligned} &zS_n(z) + \Bigl(S_{n+1}(z) - \tbinom{0}{n+1} z^{n+1}\Bigr) - \Bigl(S_{n+2}(z) - \tbinom{0}{n+2} z^{n+2} - \tbinom{1}{n+1} z^{n+1}\Bigr)\\ &\qquad = \; T(n,n+1) - T(n,0)\,. \end{aligned}$$

And $\binom{1}{n+1} = \binom{0}{n} z^{n+1} = T(n,n+1)$ for all $n \geqslant 0$, so we obtain

$$S_{n+2}(z) \;=\; S_{n+1}(z) + zS_n(z)\,,\qquad n \geqslant 0. \tag{5.140}$$

We will study the solution of such recurrences in Chapters 6 and 7; the methods of those chapters lead directly from (5.140) to the closed form (5.74), when $S_0(z) = S_1(z) = 1$.

One more example—a famous one—will complete the picture. The French mathematician Roger Apéry solved a long-standing problem in 1978 when he proved that the number $\zeta(3) = 1+2^{-3}+3^{-3}+4^{-3}+\cdots$ is irrational [14]. One of the main components of his proof involved the binomial sums

$$A_n \;=\; \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2, \tag{5.141}$$

for which he announced a recurrence that other mathematicians were unable to verify at the time. (The numbers A_n have since become known as Apéry numbers; we have $A_0 = 1$, $A_1 = 5$, $A_2 = 73$, $A_3 = 1445$, $A_4 = 33001$.) Finally [356] Don Zagier and Henri Cohen found a proof of Apéry's claim, and their proof for this special (but difficult) sum was one of the key clues that ultimately led Zeilberger to discover the general approach we are discussing.

By now, in fact, we have seen enough examples to make the sum in (5.141) almost trivial. Putting $t(n,k) = \binom{n}{k}^2\binom{n+k}{k}^2$ and $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k) + \beta_2(n)t(n+2,k)$, we try to solve (5.129) with

(First we try doing without β_2, but that attempt quickly peters out.)

$$\begin{aligned}
p(n,k) \;&=\; (n+1-k)^2(n+2-k)^2\beta_0(n) \\
&\qquad + (n+1+k)^2(n+2-k)^2\beta_1(n) \\
&\qquad + (n+1+k)^2(n+2+k)^2\beta_2(n) \;=\; \hat{p}(n,k)\,, \\
\bar{t}(n,k) \;&=\; t(n,k)/(n{+}1{-}k)^2(n{+}2{-}k)^2 \;=\; (n{+}k)!^2/k!^4(n{+}2{-}k)!^2\,, \\
q(n,k) \;&=\; (n+1+k)^2(n+2-k)^2\,, \\
r(n,k) \;&=\; k^4\,.
\end{aligned}$$

(We don't worry about the fact that q has the factor $(k+n+1)$ while r has the factor k; this does not violate (5.118), because we are regarding n as a variable parameter, not a fixed integer.) Since $q(n,k)-r(n,k) = -2k^3+\cdots$, we are allowed to set $\deg(s) = -2\lambda'/\lambda = 2$, so we take

$$s(n,k) \;=\; \alpha_2(n)k^2 + \alpha_1(n)k + \alpha_0(n)\,.$$

With this choice of s, the recurrence (5.129) boils down to five equations in the six unknown quantities $\beta_0(n)$, $\beta_1(n)$, $\beta_2(n)$, $\alpha_0(n)$, $\alpha_1(n)$, $\alpha_2(n)$. For example, the equation arising from the coefficients of k^0 simplifies to

$$\beta_0 + \beta_1 + \beta_2 - \alpha_0 - \alpha_1 - \alpha_2 \;=\; 0\,;$$

the equation arising from the coefficients of k^4 is

$$\beta_0 + \beta_1 + \beta_2 + \alpha_1 + (6+6n+2n^2)\alpha_2 \;=\; 0\,.$$

The other three equations are more complicated. But the main point is that these linear equations — like all the equations that arise when we come to this stage of the Gosper-Zeilberger algorithm — are *homogeneous* (their right-hand sides are 0). So they always have a nonzero solution when the number of unknowns exceeds the number of equations. A solution, in our case, turns out to be

$$\begin{aligned}
\beta_0(n) &= (n+1)^3\,,\\
\beta_1(n) &= -(2n+3)(17n^2+51n+39)\,,\\
\beta_2(n) &= (n+2)^3\,,\\
\alpha_0(n) &= -16(n+1)(n+2)(2n+3)\,,\\
\alpha_1(n) &= -12(2n+3)\,,\\
\alpha_2(n) &= 8(2n+3)\,.
\end{aligned}$$

Consequently

$$\begin{aligned}
(n+1)^3 t(n,k) - (2n+3)(17n^2+51n+39)t(n+1,k)&\\
+ (n+2)^3 t(n+2,k) &= T(n,k+1) - T(n,k)\,,
\end{aligned}$$

where $T(n,k) = k^4 s(n,k)\bar{t}(n,k) = (2n+3)(8k^2-12k-16(n+1)(n+2)) \times (n+k)!^2/(k-1)!^4(n+2-k)!^2$. Summing on k gives Apéry's once-incredible recurrence,

"Professor Littlewood, when he makes use of an algebraic identity, always saves himself the trouble of proving it; he maintains that an identity, if true, can be verified in a few lines by anybody obtuse enough to feel the need of verification. My object in the following pages is to confute this assertion."
— *F. J. Dyson [89]*

$$(n+1)^3 A_n + (n+2)^3 A_{n+2} = (2n+3)(17n^2+51n+39)A_{n+1}\,. \tag{5.142}$$

Does the Gosper-Zeilberger method work with all the sums we've encountered in this chapter? No. It doesn't apply when $t(n,k)$ is the summand $\binom{n}{k}(k+1)^{k-1}(n-k+1)^{n-k-1}$ in (5.65), because the term ratio $t(n,k+1)/t(n,k)$ is not a rational function of k. It also fails to handle cases like $t(n,k) = \binom{n}{k}n^k$, because the other term ratio $t(n+1,k)/t(n,k)$ is not a rational function of k. (We can do that one, however, by summing $\binom{n}{k}z^k$ and then setting $z = n$.) And it fails on a comparatively simple summand like $t(n,k) = 1/(nk+1)$, even though both $t(n,k+1)/t(n,k)$ and $t(n+1,k)/t(n,k)$ are rational functions of n and k; see exercise 107.

But the Gosper-Zeilberger algorithm is guaranteed to succeed in an enormous number of cases, namely whenever the summand $t(n,k)$ is a so-called *proper term* — a term that can be written in the form

$$t(n,k) = f(n,k)\frac{(a_1n+a_1'k+a_1'')!\,\ldots\,(a_pn+a_p'k+a_p'')!}{(b_1n+b_1'k+b_1'')!\,\ldots\,(b_qn+b_q'k+b_q'')!}\,w^n z^k\,. \tag{5.143}$$

Here $f(n,k)$ is a polynomial in n and k; the coefficients a_1, a_1', …, a_p, a_p', b_1, b_1', …, b_q, b_q' are specific integer constants; the parameters w and z

are nonzero; and the other quantities $a''_1, \ldots, a''_p, b''_1, \ldots, b''_q$ are arbitrary complex numbers. We will prove that whenever $t(n,k)$ is a proper term, there exist polynomials $\beta_0(n), \ldots, \beta_l(n)$, not all zero, and a proper term $T(n,k)$, such that

What happens if $t(n,k)$ is independent of n?

$$\beta_0(n)t(n,k) + \cdots + \beta_l(n)t(n+l,k) \;=\; T(n,k+1) - T(n,k)\,. \qquad (5.144)$$

The following proof is due to Wilf and Zeilberger [374].

Let N be the operator that increases n by 1, and let K be the operator that increases k by 1, so that, for example, $N^2K^3t(n,k) = t(n+2,k+3)$. We will study linear difference operators in N, K, and n, namely operator polynomials of the form

$$H(N,K,n) \;=\; \sum_{i=0}^{I}\sum_{j=0}^{J}\alpha_{i,j}(n)N^iK^j\,, \qquad (5.145)$$

where each $\alpha_{i,j}(n)$ is a polynomial in n. Our first observation is that, if $t(n,k)$ is any proper term and $H(N,K,n)$ is any linear difference operator, then $H(N,K,n)t(n,k)$ is a proper term. Suppose t and H are given respectively by (5.143) and (5.145); then we define a "base term"

$$\bar t(n,k)_{I,J} \;=\; \frac{\prod_{i=1}^{p}\bigl(a_in + a'_ik + a_iI[a_i<0] + a'_iJ[a'_i<0] + a''_i\bigr)!}{\prod_{i=1}^{q}\bigl(b_in + b'_ik + b_iI[b_i>0] + b'_iJ[b'_i>0] + b''_i\bigr)!}\,w^nz^k\,.$$

For example, if $t(n,k)$ is $\binom{n-2k}{k} = (n-2k)!/k!\,(n-3k)!$, the base term corresponding to a linear difference operator of degrees I and J is $\bar t(n,k)_{I,J} = (n-2k-2J)!/(k+J)!\,(n-3k+I)!$. The point is that $\alpha_{i,j}(n)N^iK^jt(n,k)$ is equal to $\bar t(n,k)_{I,J}$ times a polynomial in n and k, whenever $0 \leqslant i \leqslant I$ and $0 \leqslant j \leqslant J$. A finite sum of polynomials is a polynomial, so $H(N,K,n)t(n,k)$ has the required form (5.143).

The next step is to show that whenever $t(n,k)$ is a proper term, there is always a nonzero linear difference operator $H(N,K,n)$ such that

$$H(N,K,n)\,t(n,k) \;=\; 0\,.$$

If $0 \leqslant i \leqslant I$ and $0 \leqslant j \leqslant J$, the shifted term $N^iK^jt(n,k)$ is $\bar t(n,k)_{I,J}$ times a polynomial in n and k that has degree at most

$$\begin{aligned} D_{I,J} \;=\; \deg(f) &+ |a_1|I + |a'_1|J + \cdots + |a_p|I + |a'_p|J \\ &+ |b_1|I + |b'_1|J + \cdots + |b_q|I + |b'_q|J \end{aligned}$$

in the variable k. Hence the desired H exists if we can solve $D_{I,J}+1$ homogeneous linear equations in the $(I+1)(J+1)$ variables $\alpha_{i,j}(n)$, with coefficients that are polynomials in n. All we need to do is choose I and J large enough that $(I+1)(J+1) > D_{I,J}+1$. For example, we can take $I = 2A'+1$ and $J = 2A + \deg(f)$, where

$$\begin{aligned} A &= |a_1| + \cdots + |a_p| + |b_1| + \cdots + |b_q|\,; \\ A' &= |a_1'| + \cdots + |a_p'| + |b_1'| + \cdots + |b_q'|\,. \end{aligned}$$

The last step in the proof is to go from the equation $H(N,K,n)t(n,k) = 0$ to a solution of (5.144). Let H be chosen so that J is minimized, i.e., so that H has the smallest possible degree in K. We can write

The trick here is based on regarding H as a polynomial in K and then replacing K by $\Delta+1$.

$$H(N,K,n) \;=\; H(N,1,n) - (K-1)G(N,K,n)$$

for some linear difference operator $G(N,K,n)$. Let $H(N,1,n) = \beta_0(n) + \beta_1(n)N + \cdots + \beta_l(n)N^l$ and $T(n,k) = G(N,K,n)t(n,k)$. Then $T(n,k)$ is a proper term, and (5.144) holds.

The proof is almost complete; we still have to verify that $H(N,1,n)$ is not simply the zero operator. If it is, then $T(n,k)$ is independent of k. So there are polynomials $\beta_0(n)$ and $\beta_1(n)$ such that $\bigl(\beta_0(n)+\beta_1(n)N\bigr)T(n,k) = 0$. But then $\bigl(\beta_0(n)+\beta_1(n)N\bigr)G(N,K,n)$ is a nonzero linear difference operator of degree $J-1$ that annihilates $t(n,k)$; this contradicts the minimality of J, and our proof of (5.144) is complete.

Once we know that (5.144) holds, for some proper term T, we can be sure that Gosper's algorithm will succeed in finding T (or T plus a constant). Although we proved Gosper's algorithm only for the case of hypergeometric terms $t(k)$ in a single variable k, our proof can be extended to the two-variable case, as follows: There are infinitely many complex numbers n for which condition (5.118) holds when $q(n,k)$ and $r(n,k)$ are completely factored as polynomials in k, and for which the calculations of d in Step 2 agree with the calculations of Gosper's one-variable algorithm. For all such n, our previous proof shows that a suitable polynomial $s(n,k)$ in k exists; therefore a suitable polynomial $s(n,k)$ in n and k exists; QED.

We have proved that the Gosper-Zeilberger algorithm will discover a solution to (5.144), for some l, where l is as small as possible. That solution gives us a recurrence in n for evaluating the sum over k of any proper term $t(n,k)$, provided that $t(n,k)$ is nonzero for only finitely many k. And the roles of n and k can, of course, be reversed, because the definition of proper term in (5.143) is symmetrical in n and k.

Exercises 99–108 provide additional examples of the Gosper-Zeilberger algorithm, illustrating some of its versatility. Wilf and Zeilberger [374] have significantly extended these results to methods that handle generalized binomial coefficients and multiple indices of summation.

Exercises

Warmups

1 What is 11^4? Why is this number easy to compute, for a person who knows binomial coefficients?

2 For which value(s) of k is $\binom{n}{k}$ a maximum, when n is a given positive integer? Prove your answer.

3 Prove the hexagon property,

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

4 Evaluate $\binom{-1}{k}$ by negating (actually un-negating) its upper index.

5 Let p be prime. Show that $\binom{p}{k} \bmod p = 0$ for $0 < k < p$. What does this imply about the binomial coefficients $\binom{p-1}{k}$?

6 Fix up the text's derivation in Problem 6, Section 5.2, by correctly applying symmetry.

A case of mistaken identity.

7 Is (5.34) true also when $k < 0$?

8 Evaluate

$$\sum_k \binom{n}{k}(-1)^k(1-k/n)^n .$$

What is the approximate value of this sum, when n is very large? *Hint:* The sum is $\Delta^n f(0)$ for some function f.

9 Show that the generalized exponentials of (5.58) obey the law

$$\mathcal{E}_t(z) = \mathcal{E}(tz)^{1/t}, \qquad \text{if } t \neq 0,$$

where $\mathcal{E}(z)$ is an abbreviation for $\mathcal{E}_1(z)$.

10 Show that $-2\bigl(\ln(1-z)+z\bigr)/z^2$ is a hypergeometric function.

11 Express the two functions

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots$$
$$\arcsin z = z + \frac{1\cdot z^3}{2\cdot 3} + \frac{1\cdot 3\cdot z^5}{2\cdot 4\cdot 5} + \frac{1\cdot 3\cdot 5\cdot z^7}{2\cdot 4\cdot 6\cdot 7} + \cdots$$

in terms of hypergeometric series.

12 Which of the following functions of k is a hypergeometric term, as defined in Section 5.7? Explain why or why not.

a n^k.

b k^n.

c $\bigl(k! + (k+1)!\bigr)/2$.

d H_k, that is, $\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{k}$.

e $1/\binom{n}{k}$.

(Here t and T aren't necessarily related as in (5.120).)

f $t(k)T(k)$, when t and T are hypergeometric terms.

g $t(k) + T(k)$, when t and T are hypergeometric terms.

h $t(n-k)$, when t is a hypergeometric term.

i $a\,t(k) + b\,t(k{+}1) + c\,t(k{+}2)$, when t is a hypergeometric term.

j $\lceil k/2 \rceil$.

k $k\,[k>0]$.

Basics

13 Find relations between the superfactorial function $P_n = \prod_{k=1}^{n} k!$ of exercise 4.55, the hyperfactorial function $Q_n = \prod_{k=1}^{n} k^k$, and the product $R_n = \prod_{k=0}^{n} \binom{n}{k}$.

14 Prove identity (5.25) by negating the upper index in Vandermonde's convolution (5.22). Then show that another negation yields (5.26).

15 What is $\sum_k \binom{n}{k}^3 (-1)^k$? *Hint:* See (5.29).

16 Evaluate the sum

$$\sum_k \binom{2a}{a+k}\binom{2b}{b+k}\binom{2c}{c+k}(-1)^k$$

when a, b, c are nonnegative integers.

17 Find a simple relation between $\binom{2n-1/2}{n}$ and $\binom{2n-1/2}{2n}$.

18 Find an alternative form analogous to (5.35) for the product

$$\binom{r}{k}\binom{r-1/3}{k}\binom{r-2/3}{k}.$$

19 Show that the generalized binomials of (5.58) obey the law

$$\mathcal{B}_t(z) \;=\; \mathcal{B}_{1-t}(-z)^{-1}.$$

20 Define a "generalized bloopergeometric series" by the formula

$$G\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n} \,\middle|\, z\right) \;=\; \sum_{k\geq 0} \frac{a_1^{\underline{k}} \ldots a_m^{\underline{k}}}{b_1^{\underline{k}} \ldots b_n^{\underline{k}}} \, \frac{z^k}{k!},$$

using falling powers instead of the rising ones in (5.76). Explain how G is related to F.

21 Show that Euler's definition of factorials is consistent with the ordinary definition, by showing that the limit in (5.83) is $1/\big((m-1)\ldots(1)\big)$ when $z=m$ is a positive integer.

22 Use (5.83) to prove the *factorial duplication formula*:

$$x!\,(x-\tfrac{1}{2})! \;=\; (2x)!\,(-\tfrac{1}{2})!/2^{2x}\,.$$

By the way, $(-\frac{1}{2})! = \sqrt{\pi}$.

23 What is the value of $F(-n,1;;1)$?

24 Find $\sum_k \binom{n}{m+k}\binom{m+k}{2k}4^k$ by using hypergeometric series.

25 Show that

$$(a_1-b_1)\,F\left({a_1,\,a_2,\,\ldots,\,a_m \atop b_1{+}1,\,b_2,\,\ldots,\,b_n}\,\Big|\,z\right) = a_1\,F\left({a_1{+}1,\,a_2,\,\ldots,\,a_m \atop b_1{+}1,\,b_2,\,\ldots,\,b_n}\,\Big|\,z\right) - b_1\,F\left({a_1,\,a_2,\,\ldots,\,a_m \atop b_1,\,b_2,\,\ldots,\,b_n}\,\Big|\,z\right).$$

Find a similar relation between the hypergeometrics

$$F\left({a_1,\,a_2,\,a_3,\,\ldots,\,a_m \atop b_1,\,\ldots,\,b_n}\,\Big|\,z\right),$$

$$F\left({a_1{+}1,\,a_2,\,a_3,\,\ldots,\,a_m \atop b_1,\,\ldots,\,b_n}\,\Big|\,z\right), \qquad \text{and}$$

$$F\left({a_1,\,a_2{+}1,\,a_3,\,\ldots,\,a_m \atop b_1,\,\ldots,\,b_n}\,\Big|\,z\right).$$

26 Express the function $G(z)$ in the formula

$$F\left({a_1,\,\ldots,\,a_m \atop b_1,\,\ldots,\,b_n}\,\Big|\,z\right) \;=\; 1+G(z)$$

as a multiple of a hypergeometric series.

27 Prove that

$$F\left({a_1,\,a_1{+}\frac{1}{2},\,\ldots,\,a_m,\,a_m{+}\frac{1}{2} \atop b_1,\,b_1{+}\frac{1}{2},\,\ldots,\,b_n,\,b_n{+}\frac{1}{2},\,\frac{1}{2}}\,\Big|\,(2^{m-n-1}z)^2\right) = \frac{1}{2}\left(F\left({2a_1,\ldots,2a_m \atop 2b_1,\ldots,2b_n}\,\Big|\,z\right) + F\left({2a_1,\ldots,2a_m \atop 2b_1,\ldots,2b_n}\,\Big|\,{-z}\right)\right).$$

28 Prove *Euler's identity*

$$F\left({a,b \atop c}\,\Big|\,z\right) \;=\; (1-z)^{c-a-b}\,F\left({c{-}a,\,c{-}b \atop c}\,\Big|\,z\right)$$

by applying Pfaff's reflection law (5.101) twice.

29 Show that confluent hypergeometrics satisfy

$$e^z F\left({a \atop b}\middle| -z\right) = F\left({b-a \atop b}\middle| z\right).$$

30 What hypergeometric series F satisfies $zF'(z) + F(z) = 1/(1-z)$?

31 Show that if $f(k)$ is any function summable in hypergeometric terms, then f itself is a hypergeometric term. For example, if $\sum f(k)\,\delta k = cF(A_1, \ldots, A_M; B_1, \ldots, B_N; Z)_k + C$, then there are constants $a_1, \ldots, a_m$, $b_1, \ldots, b_n$, and z such that $f(k)$ is a multiple of (5.115).

32 Find $\sum k^2\,\delta k$ by Gosper's method.

33 Use Gosper's method to find $\sum \delta k/(k^2-1)$.

34 Show that a partial hypergeometric sum can always be represented as a limit of ordinary hypergeometrics:

$$\sum_{k\leqslant c} F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n}\middle| z\right)_k = \lim_{\epsilon\to 0} F\left({-c,\ a_1, \ldots, a_m \atop \epsilon-c,\ b_1, \ldots, b_n}\middle| z\right),$$

when c is a nonnegative integer. (See (5.115▷) Use this idea to evaluate $\sum_{k\leqslant m}\binom{n}{k}(-1)^k$.

Homework exercises

35 The notation $\sum_{k\leqslant n}\binom{n}{k}2^{k-n}$ is ambiguous without context. Evaluate it

a as a sum on k;

b as a sum on n.

36 Let p^k be the largest power of the prime p that divides $\binom{m+n}{m}$, when m and n are nonnegative integers. Prove that k is the number of carries that occur when m is added to n in the radix p number system. *Hint:* Exercise 4.24 helps here.

37 Show that an analog of the binomial theorem holds for factorial powers. That is, prove the identities

$$(x+y)^{\underline{n}} = \sum_k \binom{n}{k} x^{\underline{k}}\, y^{\underline{n-k}},$$

$$(x+y)^{\overline{n}} = \sum_k \binom{n}{k} x^{\overline{k}}\, y^{\overline{n-k}},$$

for all nonnegative integers n.

38 Show that all nonnegative integers n can be represented uniquely in the form $n = \binom{a}{1} + \binom{b}{2} + \binom{c}{3}$ where a, b, and c are integers with $0 \leqslant a < b < c$. (This is called the *binomial number system.*)

39 Show that if $xy = ax + by$ then

$$x^n y^n \;=\; \sum_{k=1}^{n} \binom{2n-1-k}{n-1} (a^n b^{n-k} x^k + a^{n-k} b^n y^k)$$

for all $n > 0$. Find a similar formula for the more general product $x^m y^n$. (These formulas give useful partial fraction expansions, for example when $x = 1/(z-c)$ and $y = 1/(z-d)$.)

40 Find a closed form for

$$\sum_{j=1}^{m} (-1)^{j+1} \binom{r}{j} \sum_{k=1}^{n} \binom{-j+rk+s}{m-j}, \qquad \text{integers } m, n \geqslant 0.$$

41 Evaluate $\sum_k \binom{n}{k} k!/(n+1+k)!$ when n is a nonnegative integer.

42 Find the indefinite sum $\sum \bigl((-1)^x / \binom{n}{x}\bigr)\,\delta x$, and use it to compute the sum $\sum_{k=0}^{n} (-1)^k / \binom{n}{k}$ in closed form.

43 Prove the triple-binomial identity (5.28). *Hint:* First replace $\binom{r+k}{m+n}$ by $\sum_j \binom{r}{m+n-j}\binom{k}{j}$.

44 Use identity (5.32) to find closed forms for the double sums

$$\sum_{j,k} (-1)^{j+k} \binom{j+k}{j} \binom{a}{j} \binom{b}{k} \binom{m+n-j-k}{m-j} \quad \text{and}$$

$$\sum_{j,k\geqslant 0} (-1)^{j+k} \binom{a}{j} \binom{m}{j} \binom{b}{k} \binom{n}{k} \Bigm/ \binom{m+n}{j+k},$$

given integers $m \geqslant a \geqslant 0$ and $n \geqslant b \geqslant 0$.

45 Find a closed form for $\sum_{k\leqslant n} \binom{2k}{k} 4^{-k}$.

46 Evaluate the following sum in closed form, when n is a positive integer:

$$\sum_k \binom{2k-1}{k} \binom{4n-2k-1}{2n-k} \frac{(-1)^{k-1}}{(2k-1)(4n-2k-1)}\,.$$

Hint: Generating functions win again.

47 The sum

$$\sum_k \binom{rk+s}{k} \binom{rn-rk-s}{n-k}$$

is a polynomial in r and s. Show that it doesn't depend on s.

48 The identity $\sum_{k\leqslant n}\binom{n+k}{n}2^{-k} = 2^n$ can be combined with the formula $\sum_{k\geqslant 0}\binom{n+k}{n}z^k = 1/(1-z)^{n+1}$ to yield

$$\sum_{k>n}\binom{n+k}{n}2^{-k} \;=\; 2^n\,.$$

What is the hypergeometric form of the latter identity?

49 Use the hypergeometric method to evaluate

$$\sum_k (-1)^k\binom{x}{k}\binom{x+n-k}{n-k}\frac{y}{y+n-k}\,.$$

50 Prove Pfaff's reflection law (5.101) by comparing the coefficients of z^n on both sides of the equation.

51 The derivation of (5.104) shows that

$$\lim_{\epsilon\to 0} F(-m,-2m-1+\epsilon;-2m+\epsilon;2) \;=\; 1\Big/\binom{-1/2}{m}\,.$$

In this exercise we will see that slightly different limiting processes lead to distinctly different answers for the degenerate hypergeometric series $F(-m,-2m-1;-2m;2)$.

a Show that $\lim_{\epsilon\to 0} F(-m+\epsilon,-2m-1;-2m+2\epsilon;2) = 0$, by using Pfaff's reflection law to prove the identity $F(a,-2m-1;2a;2) = 0$ for all integers $m \geqslant 0$.

b What is $\lim_{\epsilon\to 0} F(-m+\epsilon,-2m-1;-2m+\epsilon;2)$?

52 Prove that if N is a nonnegative integer,

$$\begin{aligned} &b_1^{\overline{N}}\ldots b_n^{\overline{N}}\,F\left(\begin{matrix}a_1,\ldots,a_m,-N\\ b_1,\ldots,b_n\end{matrix}\,\Big|\,z\right)\\ &\quad= a_1^{\overline{N}}\ldots a_m^{\overline{N}}(-z)^N\,F\left(\begin{matrix}1-b_1-N,\ldots,1-b_n-N,-N\\ 1-a_1-N,\ldots,1-a_m-N\end{matrix}\,\Big|\,\frac{(-1)^{m+n}}{z}\right).\end{aligned}$$

53 If we put $b = -\frac{1}{2}$ and $z = 1$ in Gauss's identity (5.110), the left side reduces to -1 while the right side is $+1$. Why doesn't this prove that $-1 = +1$?

54 Explain how the right-hand side of (5.112) was obtained.

55 If the hypergeometric terms $t(k) = F(a_1,\ldots,a_m;\, b_1,\ldots,b_n;\, z)_k$ and $T(k) = F(A_1,\ldots,A_M;B_1,\ldots,B_N;Z)_k$ satisfy $t(k) = c\big(T(k+1)-T(k)\big)$ for all $k \geqslant 0$, show that $z = Z$ and $m - n = M - N$.

56 Find a general formula for $\sum\binom{-3}{k}\,\delta k$ using Gosper's method. Show that $(-1)^{k-1}\big\lfloor\frac{k+1}{2}\big\rfloor\big\lfloor\frac{k+2}{2}\big\rfloor$ is also a solution.

57 Use Gosper's method to find a constant θ such that

$$\sum \binom{n}{k} z^k (k+\theta)\, \delta k$$

is summable in hypergeometric terms.

58 If m and n are integers with $0 \leqslant m \leqslant n$, let

$$T_{m,n} = \sum_{0 \leqslant k < n} \binom{k}{m} \frac{1}{n-k}.$$

Find a relation between $T_{m,n}$ and $T_{m-1,n-1}$, then solve your recurrence by applying a summation factor.

Exam problems

59 Find a closed form for

$$\sum_{k \geqslant 1} \binom{n}{\lfloor \log_m k \rfloor}$$

when m and n are positive integers.

60 Use Stirling's approximation (4.23) to estimate $\binom{m+n}{n}$ when m and n are both large. What does your formula reduce to when $m = n$?

61 Prove that when p is prime, we have

$$\binom{n}{m} \equiv \binom{\lfloor n/p \rfloor}{\lfloor m/p \rfloor} \binom{n \bmod p}{m \bmod p} \pmod{p},$$

for all nonnegative integers m and n.

62 Assuming that p is prime and that m and n are positive integers, determine the value of $\binom{np}{mp} \bmod p^2$. *Hint:* You may wish to use the following generalization of Vandermonde's convolution:

$$\sum_{k_1+k_2+\cdots+k_m=n} \binom{r_1}{k_1}\binom{r_2}{k_2}\ldots\binom{r_m}{k_m} = \binom{r_1+r_2+\cdots+r_m}{n}.$$

63 Find a closed form for

$$\sum_{k=0}^{n} (-4)^k \binom{n+k}{2k},$$

given an integer $n \geqslant 0$.

64 Evaluate $\sum_{k=0}^{n} \binom{n}{k} \Big/ \left\lceil \frac{k+1}{2} \right\rceil$, given an integer $n \geqslant 0$.

65 Prove that

$$\sum_k \binom{n-1}{k} n^{-k}(k+1)! \;=\; n\,.$$

66 Evaluate "Harry's double sum,"

$$\sum_{0 \leqslant j \leqslant k} \binom{-1}{j - \lfloor \sqrt{k-j} \rfloor} \binom{j}{m} \frac{1}{2^j}\,, \qquad \text{integer } m \geqslant 0,$$

as a function of m. (The sum is over both j and k.)

67 Find a closed form for

$$\sum_{k=0}^{n} \binom{\binom{k}{2}}{2} \binom{2n-k}{n}\,, \qquad \text{integer } n \geqslant 0.$$

68 Find a closed form for

$$\sum_k \binom{n}{k} \min(k, n-k)\,, \qquad \text{integer } n \geqslant 0.$$

69 Find a closed form for

$$\min_{\substack{k_1,\ldots,k_m \geqslant 0\\ k_1+\cdots+k_m=n}} \sum_{j=1}^{m} \binom{k_j}{2}$$

as a function of m and n.

70 Find a closed form for

$$\sum_k \binom{n}{k} \binom{2k}{k} \left(\frac{-1}{2}\right)^k\,, \qquad \text{integer } n \geqslant 0.$$

71 Let

$$S_n \;=\; \sum_{k \geqslant 0} \binom{n+k}{m+2k} a_k\,,$$

where m and n are nonnegative integers, and let $A(z) = \sum_{k \geqslant 0} a_k z^k$ be the generating function for the sequence $\langle a_0, a_1, a_2, \ldots \rangle$.

a Express the generating function $S(z) = \sum_{n \geqslant 0} S_n z^n$ in terms of $A(z)$.

b Use this technique to solve Problem 7 in Section 5.2.

72 Prove that, if m, n, and k are integers and $n > 0$,

$$\binom{m/n}{k} n^{2k-\nu(k)} \quad \text{is an integer,}$$

where $\nu(k)$ is the number of 1's in the binary representation of k.

73 Use the repertoire method to solve the recurrence

$$\begin{aligned} X_0 &= \alpha\,; \qquad X_1 = \beta\,; \\ X_n &= (n-1)(X_{n-1} + X_{n-2})\,, \qquad \text{for } n > 1. \end{aligned}$$

Hint: Both $n!$ and $n¡$ satisfy this recurrence.

74 This problem concerns a deviant version of Pascal's triangle in which the sides consist of the numbers 1, 2, 3, 4, ... instead of all 1's, although the interior numbers still satisfy the addition formula:

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 2 & & 2 & & & \\ & & 3 & & 4 & & 3 & & \\ & 4 & & 7 & & 7 & & 4 & \\ 5 & & 11 & & 14 & & 11 & & 5 \\ . & . & & . & & . & & . & . \end{array}$$

If $\left(\!\binom{n}{k}\!\right)$ denotes the kth number in row n, for $1 \leqslant k \leqslant n$, we have $\left(\!\binom{n}{1}\!\right) = \left(\!\binom{n}{n}\!\right) = n$, and $\left(\!\binom{n}{k}\!\right) = \left(\!\binom{n-1}{k}\!\right) + \left(\!\binom{n-1}{k-1}\!\right)$ for $1 < k < n$. Express the quantity $\left(\!\binom{n}{k}\!\right)$ in closed form.

75 Find a relation between the functions

$$\begin{aligned} S_0(n) &= \sum_k \binom{n}{3k}\,, \\ S_1(n) &= \sum_k \binom{n}{3k+1}\,, \\ S_2(n) &= \sum_k \binom{n}{3k+2} \end{aligned}$$

and the quantities $\lfloor 2^n/3 \rfloor$ and $\lceil 2^n/3 \rceil$.

76 Solve the following recurrence for $n, k \geqslant 0$:

$$\begin{aligned} Q_{n,0} &= 1\,; \qquad Q_{0,k} = [k=0]\,; \\ Q_{n,k} &= Q_{n-1,k} + Q_{n-1,k-1} + \binom{n}{k}\,, \qquad \text{for } n, k > 0. \end{aligned}$$

77 What is the value of

$$\sum_{0\leqslant k_1,\ldots,k_m\leqslant n}\ \prod_{1\leqslant j<m}\binom{k_{j+1}}{k_j}, \qquad \text{if } m>1?$$

78 Assuming that m is a positive integer, find a closed form for

$$\sum_{k=0}^{2m^2}\binom{k \bmod m}{(2k+1) \bmod (2m+1)}.$$

79 **a** What is the greatest common divisor of $\binom{2n}{1}$, $\binom{2n}{3}$, $\ldots$, $\binom{2n}{2n-1}$? *Hint:* Consider the sum of these n numbers.

b Show that the least common multiple of $\binom{n}{0}$, $\binom{n}{1}$, $\ldots$, $\binom{n}{n}$ is equal to $L(n+1)/(n+1)$, where $L(n)=\mathrm{lcm}(1,2,\ldots,n)$.

Handy to know.

80 Prove that $\binom{n}{k}\leqslant(en/k)^k$ for all integers $k,n\geqslant 0$.

81 If $0<\theta<1$ and $0\leqslant x\leqslant 1$, and if l,m,n are nonnegative integers with $m<n$, prove the inequality

$$(-1)^{n-m-1}\sum_k\binom{l}{k}\binom{m+\theta}{n+k}x^k \;>\; 0\,.$$

Hint: Consider taking the derivative with respect to x.

Bonus problems

82 Prove that Pascal's triangle has an even more surprising hexagon property than the one cited in the text:

$$\gcd\Bigl(\tbinom{n-1}{k-1},\tbinom{n}{k+1},\tbinom{n+1}{k}\Bigr) \;=\; \gcd\Bigl(\tbinom{n-1}{k},\tbinom{n+1}{k+1},\tbinom{n}{k-1}\Bigr),$$

if $0<k<n$. For example, $\gcd(56,36,210)=\gcd(28,120,126)=2$.

83 Prove the amazing five-parameter double-sum identity (5.32).

84 Show that the second pair of convolution formulas, (5.61), follows from the first pair, (5.60). *Hint:* Differentiate with respect to z.

85 Prove that

$$\sum_{m=1}^{n}(-1)^m\sum_{1\leqslant k_1<k_2<\cdots<k_m\leqslant n}\binom{k_1^3+k_2^3+\cdots+k_m^3+2^n}{n} \;=\; (-1)^n n!^3-\binom{2^n}{n}.$$

(The left side is a sum of 2^n-1 terms.) *Hint:* Much more is true.

86 Let $a_1, \ldots, a_n$ be nonnegative integers, and let $C(a_1, \ldots, a_n)$ be the coefficient of the constant term $z_1^0 \ldots z_n^0$ when the $n(n-1)$ factors

$$\prod_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}} \left(1-\frac{z_i}{z_j}\right)^{a_i}$$

are fully expanded into positive and negative powers of the complex variables $z_1, \ldots, z_n$.

a Prove that $C(a_1, \ldots, a_n)$ equals the left-hand side of (5.31).

b Prove that if $z_1, \ldots, z_n$ are distinct complex numbers, then the polynomial

$$f(z) = \sum_{k=1}^{n} \prod_{\substack{1\leqslant j\leqslant n\\ j\neq k}} \frac{z-z_j}{z_k-z_j}$$

is identically equal to 1.

c Multiply the original product of $n(n-1)$ factors by $f(0)$ and deduce that $C(a_1, a_2, \ldots, a_n)$ is equal to

$$C(a_1-1, a_2, \ldots, a_n) + C(a_1, a_2-1, \ldots, a_n) \\ + \cdots + C(a_1, a_2, \ldots, a_n-1)\,.$$

(This recurrence defines multinomial coefficients, so $C(a_1, \ldots, a_n)$ must equal the right-hand side of (5.31).)

87 Let m be a positive integer and let $\zeta = e^{\pi i/m}$. Show that

$$\sum_{k\leqslant n/m} \binom{n-mk}{k} z^{mk} \\ = \frac{\mathcal{B}_{-m}(z^m)^{n+1}}{(1+m)\mathcal{B}_{-m}(z^m)-m} \\ - \sum_{0\leqslant j<m} \frac{\bigl(\zeta^{2j+1} z\,\mathcal{B}_{1+1/m}(\zeta^{2j+1}z)^{1/m}\bigr)^{n+1}}{(m+1)\mathcal{B}_{1+1/m}(\zeta^{2j+1}z)^{-1}-1}\,.$$

(This reduces to (5.74) in the special case $m = 1$.)

88 Prove that the coefficients s_k in (5.47) are equal to

$$(-1)^k \int_0^\infty e^{-t}(1-e^{-t})^{k-1}\,\frac{dt}{t}\,,$$

for all $k > 1$; hence $|s_k| < 1/(k-1)$.

89 Prove that (5.19) has an infinite counterpart,

$$\sum_{k>m}\binom{m+r}{k}x^k y^{m-k} = \sum_{k>m}\binom{-r}{k}(-x)^k(x+y)^{m-k}, \quad \text{integer } m,$$

if $|x| < |y|$ and $|x| < |x+y|$. Differentiate this identity n times with respect to y and express it in terms of hypergeometrics; what relation do you get?

90 Problem 1 in Section 5.2 considers $\sum_{k\geqslant 0}\binom{r}{k}/\binom{s}{k}$ when r and s are integers with $s \geqslant r \geqslant 0$. What is the value of this sum if r and s aren't integers?

91 Prove *Whipple's identity*,

$$F\left({\tfrac{1}{2}a,\ \tfrac{1}{2}a+\tfrac{1}{2},\ 1+a-b-c \atop 1+a-b,\ 1+a-c}\,\middle|\, \frac{-4z}{(1-z)^2}\right) = (1-z)^a\, F\left({a,\ b,\ c \atop 1+a-b, 1+a-c}\,\middle|\, z\right),$$

by showing that both sides satisfy the same differential equation.

92 Prove *Clausen's product identities*

$$F\left({a,\ b \atop a+b+\tfrac{1}{2}}\,\middle|\, z\right)^2 = F\left({2a,\ a+b,\ 2b \atop 2a+2b,\ a+b+\tfrac{1}{2}}\,\middle|\, z\right);$$

$$F\left({\tfrac{1}{4}+a,\ \tfrac{1}{4}+b \atop 1+a+b}\,\middle|\, z\right) F\left({\tfrac{1}{4}-a,\ \tfrac{1}{4}-b \atop 1-a-b}\,\middle|\, z\right) = F\left({\tfrac{1}{2},\ \tfrac{1}{2}+a-b,\ \tfrac{1}{2}-a+b \atop 1+a+b,\ 1-a-b}\,\middle|\, z\right).$$

What identities result when the coefficients of z^n on both sides of these formulas are equated?

93 Show that the indefinite sum

$$\sum\left(\prod_{j=1}^{k-1}\bigl(f(j)+\alpha\bigr)\Big/\prod_{j=1}^{k} f(j)\right)\delta k$$

has a (fairly) simple form, given any function f and any constant α.

94 Find $\sum\binom{a}{k}\binom{-a}{n-k}\,\delta k$.

95 What conditions in addition to (5.118) will make the polynomials p, q, r of (5.117) uniquely determined?

96 Prove that if Gosper's algorithm finds no solution to (5.120), given a hypergeometric term $t(k)$, then there is no solution to the more general equation

$$t(k) = \bigl(T_1(k+1) + \cdots + T_m(k+1)\bigr) - \bigl(T_1(k) + \cdots + T_m(k)\bigr),$$

where $T_1(k), \ldots, T_m(k)$ are hypergeometric terms.

97 Find all complex numbers z such that $k!^2 / \prod_{j=1}^{k}(j^2 + jz + 1)$ is summable in hypergeometric terms.

98 What recurrence does the Gosper-Zeilberger method give for the sum $S_n = \sum_k \binom{n}{2k}$?

99 Use the Gosper-Zeilberger method to discover a closed form for $\sum_k t(n,k)$ when $t(n,k) = (n+a+b+c+k)!/(n-k)!\,(c+k)!\,(b-k)!\,(a-k)!\,k!$, assuming that a is a nonnegative integer.

100 Find a recurrence relation for the sum

$$S_n = \sum_{k=0}^{n} \frac{1}{\dbinom{n}{k}},$$

and use the recurrence to find another formula for S_n.

101 Find recurrence relations satisfied by the sums

Better use computer algebra for this one (and the next few).

a $\quad S_{m,n}(z) = \sum_k \binom{m}{k}\binom{n}{k} z^k;$

b $\quad S_n(z) = S_{n,n}(z) = \sum_k \binom{n}{k}^2 z^k.$

102 Use the Gosper-Zeilberger procedure to generalize the "useless" identity (5.113): Find additional values of a, b, and z such that

$$\sum_k \binom{n}{k}\binom{\frac{1}{3}n - a}{k} z^k \Big/ \binom{\frac{4}{3}n - b}{k}$$

has a simple closed form.

103 Let $t(n,k)$ be the proper term (5.143). What are the degrees of $\hat{p}(n,k)$, $q(n,k)$, and $r(n,k)$ in terms of the variable k, when the Gosper-Zeilberger procedure is applied to $\hat{t}(n,k) = \beta_0(n)t(n,k) + \cdots + \beta_l(n)t(n+l,k)$? (Ignore the rare, exceptional cases.)

104 Use the Gosper-Zeilberger procedure to verify the remarkable identity

$$\sum_k (-1)^k \binom{r-s-k}{k}\binom{r-2k}{n-k}\frac{1}{r-n-k+1} \;=\; \binom{s}{n}\frac{1}{r-2n+1}\,.$$

Explain why the simplest recurrence for this sum is not found.

105 Show that if $\omega = e^{2\pi i/3}$ we have

$$\sum_{k+l+m=3n} \binom{3n}{k,l,m}^2 \omega^{l-m} \;=\; \binom{4n}{n,n,2n}\,, \qquad \text{integer } n \geqslant 0.$$

106 Prove the amazing identity (5.32) by letting $t(r,j,k)$ be the summand divided by the right-hand side, then showing that there are functions $T(r,j,k)$ and $U(r,j,k)$ for which

$$\begin{aligned} t(r+1,j,k) - t(r,j,k) \;&=\; T(r,j+1,k) - T(r,j,k) \\ &\qquad + U(r,j,k+1) - U(r,j,k)\,. \end{aligned}$$

107 Prove that $1/(nk+1)$ is not a proper term.

108 Show that the Apéry numbers A_n of (5.141) are the diagonal elements $A_{n,n}$ of a matrix of numbers defined by

$$A_{m,n} \;=\; \sum_{j,k} \binom{m}{j}^2 \binom{m}{k}^2 \binom{2m+n-j-k}{2m}\,.$$

Prove, in fact, that this matrix is symmetric, and that

$$\begin{aligned} A_{m,n} \;&=\; \sum_k \binom{m+n-k}{k}^2 \binom{m+n-2k}{m-k}^2 \\ &=\; \sum_k \binom{m}{k}\binom{n}{k}\binom{m+k}{k}\binom{n+k}{k}\,. \end{aligned}$$

109 Prove that the Apéry numbers (5.141) satisfy

$$A_n \;\equiv\; A_{\lfloor n/p \rfloor} A_{n \bmod p} \qquad (\text{mod } p)$$

for all primes p and all integers $n \geqslant 0$.

Research problems

110 For what values of n is $\binom{2n}{n} \equiv (-1)^n \pmod{2n+1}$?

111 Let $q(n)$ be the smallest odd prime factor of the middle binomial coefficient $\binom{2n}{n}$. According to exercise 36, the odd primes p that do *not* divide $\binom{2n}{n}$ are those for which all digits in n's radix p representation are $(p-1)/2$ or less. Computer experiments have shown that $q(n) \leqslant 11$ for $1 < n < 10^{10000}$, except that $q(3160) = 13$.

a Is $q(n) \leqslant 11$ for all $n > 3160$?

b Is $q(n) = 11$ for infinitely many n?

A reward of $\$7 \cdot 11 \cdot 13$ is offered for a solution to either (a) or (b).

112 Is $\binom{2n}{n}$ divisible either by 4 or by 9, for all $n > 4$ except $n = 64$ and $n = 256$?

113 If $t(n+1,k)/t(n,k)$ and $t(n,k+1)/t(n,k)$ are rational functions of n and k, and if there is a nonzero linear difference operator $H(N,K,n)$ such that $H(N,K,n)t(n,k) = 0$, does it follow that $t(n,k)$ is a proper term?

114 Let m be a positive integer, and define the sequence $c_n^{(m)}$ by the recurrence

$$\sum_k \binom{n}{k}^m \binom{n+k}{k}^m = \sum_k \binom{n}{k}\binom{n+k}{k} c_k^{(m)} .$$

Are these numbers $c_n^{(m)}$ integers?