3. GENERATING FUNCTIONS

3.1. ORDINARY GENERATING FCNS

3.1.1. There are 18 distinguishable selections of six marbles from a pile consisting of three red marbles, four white marbles, and five blue marbles. We have a factor for each color, with one way for it to contribute each given number of marbles up to the bound for that color. Thus the generating function for distinguishable selections from the pile, indexed by number of marbles, is \((1 + x + x^2 + x^3)(1 + x + x^2 + x^3 + x^4)(1 + x + x^2 + x^3 + x^4 + x^5)\). We rewrite this as \((1 - x^6)(1 - x^9)(1 - x^{15})/(1 - x)^3\). The expansion of \((1 - x)^{-3}\) is \(\sum_{n=0}^{\infty} \binom{n+2}{2} x^n\), or \(1 + 3x + 6x^2 + 10x^3 + 15x^4 + 21x^5 + 28x^6 + \cdots\). We multiply by \(1 - x^4 - x^5 - x^6 + x^9 + \cdots\) and extract the coefficient of \(x^6\), which is \(28 - 6 - 3 - 1\), equal to 18.

Explicitly, the 11 allowed red+white+blue compositions are \(0 + 1 + 5\), \(1 + 0 + 5\), \(0 + 2 + 4\), \(1 + 1 + 4\), \(2 + 0 + 4\), \(0 + 3 + 3\), \(1 + 2 + 3\), \(2 + 1 + 3\), \(3 + 0 + 3\), \(4 + 1 + 2\), \(1 + 3 + 2\), \(2 + 2 + 2\), \(3 + 2 + 2\), \(1 + 4 + 1\), \(2 + 3 + 1\), \(3 + 2 + 1\), \(2 + 4 + 0\), \(3 + 3 + 0\).

3.1.2. Generating function for selections with restriction. Selecting from \(n\) types, we have the lower bound \(r_i\) and upper bound \(s_i\) on the usage of the \(i\)th type. Equivalently, let \(a_i\) be the number of integer solutions to \(e_1 + \cdots + e_n = k\) in which \(r_i \leq e_i \leq s_i\). Let \(A(x) = \sum_{k \geq 0} a_k x^k\). In the generating function, the factor for the \(i\)th type allows contributions of weights \(r_i\) through \(s_i\). Since there is only one way to select \(k\) objects from 1 type, each coefficient in each factor is 1. Hence \(a_k\) is the coefficient of \(x^k\) in the function \(A(x)\) given by

\[
A(x) = (x^{r_1} + \cdots + x^{s_1}) \cdots (x^{r_n} + \cdots + x^{s_n}) = \frac{\prod x^{c_i} (1 - x^{s_i - r_i + 1})}{(1 - x)^n}
\]

If any type is unlimited, then the corresponding factor is the infinite geometric series that sums to \(1/(1 - x)\) instead of the factor given above.

3.1.3. Enumerator by cost for selections from types of candy with costs 2, 1, 2, 5, respectively, is \(\frac{1}{(1-x^2)^2} \frac{1}{1-x} \frac{1}{1-x^5}\). With one factor for each type, the OGF is \(\sum_{n \geq 0} x^{2n} \sum_{n \geq 0} x^n \sum_{n \geq 0} x^{2n} \sum_{n \geq 0} x^{5n}\). The exponent accumulates the money spent for a particular selection of candies. The product equals the OGF given above.

3.1.4. The OGF for the number \(a_n\) of nonnegative integer solutions to \(4e_1 + 2e_2 + e_3 + 2e_4 = n\) is \(\left(\frac{1}{1-x}\right)^2 \frac{1}{1-x^2} \frac{1}{1-x^4}\). We have a factor for each coefficient, using that each \(e_i\) may be any nonnegative integer. Thus the OGF is \(\sum_{k \geq 0} x^{4k} \sum_{k \geq 0} x^{2k} \sum_{k \geq 0} x^k \sum_{k \geq 0} x^{2k}\). The exponent accumulates the contribution to the total sum. The product equals the OGF given above.

3.1.5. The number \(a_{n,k}\) of permutations of \([n]\) with \(k\) inversions satisfies \(a_{n,k} = \sum_{j=0}^{n-1} a_{n-1,k-j}\), with \(a_{n,k} = \delta_{k,0}\). There is one empty permutation, and it has no inversions. For \(n > 0\), we build a permutation by successively inserting elements 1 through \(n\). When element \(n\) is inserted, it creates \(j\) inversions if it goes into position \(n - j\), and this does not alter inversions in the permutation existing before inserting \(n\). Grouping the permutations by the number \(k\) of inversions in the final permutation yields the recurrence.

3.1.6. Proof of \(x^n = \sum_{k=0}^{n} A(n, k) \binom{x+k-1}{n}\) from Worpitzky’s Identity. The given identity is \(x^n = \sum_{k=0}^{n} A(n, k) \binom{x+k-1}{n}\). Letting \(j = n - k + 1\) rewrites it as \(x^n = \sum_{j=0}^{n} A(n, n - j + 1) \binom{x+j-1}{n}\). The reverse of a permutation with \(k\) runs has \(n - k + 1\) runs. (With \(k\) runs, there are \(k - 1\) descents, hence \(n - k\) ascents, hence \(n - k\) descents and \(n - k + 1\) runs in the reverse.) Hence \(A(n, n - j + 1) = A(n, j)\). Finally, \(A(n, n - 0 + 1) = 0\), so setting \(j = k\) turns the sum into the desired form.

3.1.7. The numbers \(A(4, k)\) of permutations of \([4]\) with \(k\) runs are 1 for \(k \in \{1, 4\}\) and 11 for \(k \in \{2, 3\}\). There is one permutation with one run and one (its reverse) with four runs. The permutations with two and three runs are reversals of each other, so they must split the remaining 22 permutations evenly.

The formula \(A(n, k) = \sum_{i=0}^{k} (-1)^i \binom{k-i}{n} k^n\) in Theorem 3.1.26 yields the same information more laboriously. We have

\[
\begin{align*}
A(4, 1) &= 1 \cdot 1^4 = 1 \\
A(4, 2) &= 1 \cdot 2^4 - 5 \cdot 1^4 = 16 - 5 = 11 \\
A(4, 3) &= 1 \cdot 3^4 - 5 \cdot 2^4 + 10 \cdot 1^4 = 81 - 80 + 10 = 11 \\
A(4, 4) &= 1 \cdot 4^4 - 5 \cdot 3^4 + 10 \cdot 2^4 - 10 \cdot 1^4 = 256 - 405 + 16 - 10 = 1
\end{align*}
\]

It is somewhat amazing from this computation that \(A(n, n - k + 1) = A(n, k)\), but that follows easily by reversing permutations.
The congruence class of $A(p - 1, k)$ modulo $p$ when $p$ is prime. In $A(p - 1, k) = \sum_{i=0}^{k}(-1)^i \binom{p}{i}(k - i)^{p-1}$, the binomial coefficient $\binom{p}{i}$ is divisible by $p$ except when $i \in \{0, p\}$. Also, the sum can be nonzero only when $k < p$. Hence $A(p - 1, k) \equiv k^{p-1} (mod p) \equiv 1 (mod p)$, by Fermat’s Little Theorem (Application 1.3.10). For $p = 5$, the values computed above are $1, 11, 11, 1, 1$ indeed congruent to 1 modulo 5.

3.1.8. For the number $b_{n,k}$ of $k$-subsets of $[n]$ with no consecutive integers, the OGF indexed by $n$ is $B_k(x) = \sum_{n \geq 0} b_{n,k} x^n = x^{2k}/(1-x)^{k+1}$. (Exception: for $k = 0$ it is $\frac{1}{1-x}$.) Given the elements in increasing order, a subset of $[n]$ is specified by the least element, the differences between successive elements, and the difference between the largest and $n$. These numbers sum to $n$. For a $k$-set, there are $k - 1$ differences between elements. With consecutive integers forbidden, the differences are at least 2.

Thus $b_{n,k}$ is the number of integer solutions to $\sum_{i=0}^{k} e_i = n$ such that $e_0 \geq 1, e_i \geq 2$ for $1 \leq i \leq k - 1$, and $e_k \geq 0$. This is the coefficient of $x^n$ in the product of $k + 1$ generating functions, where the factor for $e_0$ is $\frac{2}{x^2}$, the factor for $e_k$ is $\frac{1}{1-x}$, and the factors for the other variables are $\frac{x^{e_i}}{1-x}$. This computes $B_k(x)$ for $k > 0$.

The solution formula is $b_{n,k} = \binom{n-k+1}{k}$. To obtain $b_{n,k}$, we compute

$$[x^n] B_k(x) = [x^{n-2k+1}] (1-x)^{-k} = \binom{n-2k+1+k(k+1)-1}{k+1}/(1-x)^{k+1} = \binom{n-k+1}{k}.$$

3.1.9. Distribution of sums of two dice.

a) The four-sided case. The given example $\{(1, 2, 3, 3), (1, 3, 3, 5)\}$ has the same distribution as $\{(1, 2, 3, 4), (1, 2, 3, 4)\}$ because $\{x+2x^2+3x^3+4x^4\} = \{x+2x^2+x^3+4x^4+5x^5+6x^6+7x^7+x^8\}$. To consider other pairs with this distribution, we must have both dice with smallest face 1, since 2 is a possible sum. To obtain two way to score 3, we must then have two copies of 2, whether on the same die or one copy on each. To get three copies of 4, we must then have 3 on each die in the balanced case or get 1 + 3 in three ways in the unbalanced case, which completes $\{(1, 2, 2, 3)\}$ and puts $(1, 3, 3, 3)$ on the other die. Now the top value of 8 finishes the unbalanced case, while $4x^5$ requires that the balanced case be finished with 4 on each die.

b) The six-sided case: in addition to normal dice, one other pair has the same distribution. To match two normal dice, we then need distribution $x^2+2x^3+3x^4+4x^5+5x^6+6x^7+5x^8+4x^9+3x^{10}+2x^{11}+x^{12}$. The analysis through the effect of $x^4$ is almost the same as above. We have $\{(1, 2, 3, a, b, c), (1, 2, 3, x, y, z)\}$ (balanced case) or $\{(1, 2, 3, a, b, c), (1, v, w, x, y, z)\}$ with three copies of 3 added (unbalanced case).

Case 1: balanced case. So far $(1, 2, 3, a, b, c)$ and $(1, 2, 3, x, y, z)$. We already have two sums of 5 and need two more. Hence we need two copies of 4. If the 4s are on the same die, then we have three sums of 6 and need two more by adding two copies of 5. If the 5s are split, then we need a 7 to reach 12, yielding $(1, 2, 3, 4, 4, 5)$ and $(1, 2, 3, 5, y, 7)$, but then no choice for $y$ yields the right distribution. Otherwise, we have $(1, 2, 3, 4, 4, c)$ and $(1, 2, 3, 5, 5, z)$; to obtain six sums of 7, we need $c = z = 6$, but then there are only two ways to make 10.

Hence we must have $(1, 2, 3, 4, b, c)$ and $(1, 2, 3, 4, y, z)$. To have five sums of 6 we must add two copies of 5. If on the same die, then the two available positions on the other die cannot give us three sums of 10. Hence the copies of 5 are split, and we must finish with the normal dice.

Case 2: unbalanced case. So far we have $(1, 2, 2, a, b, c)$ and $(1, v, w, x, y, z)$. We must put three 3s on the dice to satisfy $3x^4$, since there is no $2 + 2$. However, $4x^5$ prevents having them all on the second die. Also, if we put no 3 on the second die, then $4x^5$ requires four copies of 4 there, which violates $5x^6$.

Hence we must split the three copies of 3. Suppose first that we put two on the second die, making $(1, 2, 2, 3, b, c)$ and $(1, 3, 3, x, y, z)$. We now have four sums of 5, so we cannot put 4 anywhere. Hence we can only sum to 6 by $3+3$ (twice) or $1+5$, so $5x^5$ requires three copies of 5. Putting two on the first die prevents getting six sums of 7. Putting one there requires $(1, 2, 2, 3, 5, 6)$ and $(1, 3, 3, 5, 5, 6)$, which violates $2x^{11}$. Putting them all on the second die requires $(1, 2, 2, 3, b, c)$ and $(1, 3, 3, 5, 5, 5)$. That needs the other two numbers to be 7 in order to satisfy $5x^8$, but then rolling a sum of 9 is impossible.

Hence we may assume $(1, 2, 2, 3, 3, c)$ and $(1, 3, w, x, y, z)$. We have already two sums of 5, so achieving $4x^5$ requires two copies of 4. If we put them both on the second die, then we already have six sums of 6, which is too many. Hence we must have $(1, 2, 2, 3, 4, 4)$ and $(1, 3, 4, x, y, z)$. We already have four sums of 6. We need one more, so the second die must have one 5. Now we have five sums of 7, so the second die must have one 6. To reach 12, the remaining number must be 8, so we have $(1, 2, 2, 3, 4, 5)$ and $(1, 3, 4, 5, 6, 8)$, which has the complete correct distribution!

3.1.10. When a coin is flipped 14 times, and three of the flips are tails, the probability is $1/7$ that no five consecutive flips are heads. With three tails, there are four places where heads can occur. To avoid five consecutive heads, the enumerator for the options for heads in each location is $1 + x + x^2 + x^3 + x^4$. Hence the number of successful lists is the coefficient of $x^{11}$ in $[1 - x^5]/(1-x)^4$. The generating function is $\sum_{k=0}^{4}(-1)^k\binom{4}{k}x^k = (\sum_{n \geq 0}0(n+3)x^n)$. The coefficient of $x^{11}$ is $\binom{14}{3} - 4\binom{9}{3} + 6\binom{4}{3}$.
which equals 52.

The denominator for the probability is the total number of strings of 14 flips with three tails, which is \((\binom{14}{3})\), equaling 364. Surprisingly, the fraction reduces to the simple value 1/7.

3.1.11. The generating function for distinguishable flippings of pennies, nickels, and dimes by number of coins is \((1 - x)^{-6}\).

### Algebraic argument.

We are given the generating function \((\sum_{k=0}^{\infty} (k + 1)x^k)^3\) from analyzing pennies/nickels/dimes separately. We notice that each factor is \((d/dx) \sum_{k=0}^{\infty} x^k = (d/dx)(1/(1 - x)) = 1/(1 - x)^2\). Thus the full generating function is \((1 - x)^{-6}\).

### Combinatorial argument.

Consider contributions of six types to building such an arrangement. An arrangement is determined by knowing the numbers of Head-pennies, Tail-pennies, Head-nickels, Tail-nickels, Head-dimes, and Tail-dimes. We choose some number of each type, each such choice being made exactly in one way. Thus the number of ways with \(k\) coins is the coefficient of \(x^k\) in \((\sum_{k=0}^{\infty} x^k)^6 = (1/(1 - x))^6\).

3.1.12. Committee delegates.

When each university sends five delegates, a university can contribute one delegate in five ways, and \((1 + 5x)\) enumerates the choices of at most one delegate from one university, weighted by the number of delegates contributed. Since there are 100 universities, the OGF for committees with at most one delegate from each university is \((1 + 5x)^{100}\). The coefficient of \(x^{25}\), which is \((\binom{100}{25})5^{25}\), is the number of ways to form such a committee of size 25.

When a university can contribute up to three delegates, the generating function for contributions from a university is \(1 + 5x + 10x^2 + 10x^3\). We seek the coefficient of \(x^{25}\) in \((1 + 5x + 10x^2 + 10x^3)^{100}\). To attack such a problem, we describe the contributions to the desired coefficient. We study nonnegative solutions to \(e_1 + \cdots + e_{100} = 25\); each solution corresponds to a choice of 25 delegates, listed by how many come from each university.

We can group these contributions according to the multiset of nonzero values. These correspond to partitions of the integer 25 using parts of size at most 3. Using exponents to denote multiplicities, the possibilities are \(3\,^1, 3\,^2, 3^2\,^1, 3^3\,^1, 3^2\,^2\,^1, 3^2\,^3\,^1, 3^3\,^2\,^1, 3^3\,^3\,^1, \ldots\). The number of each type is the coefficient of \(x^{25}\) in the expansion of \(\frac{1}{(1-x-3x)(1-x-2x)^3}\). Let \(S\) be the set of these partitions.

The number of times a particular term \(3^m\,^n\,^1\) arises is the number of ways to arrange \(e_3\,^3, e_2\,^2, e_1\,^1\), and \(100 - e_1 - e_2 - e_3\,^9\), which is \(\binom{100}{e_1|e_2|e_3|100-e_1-e_2-e_3}\). For each such occurrence, the number of ways to make the selections is \(5^{e_1}10^{e_2}10^{e_3}\).

Therefore, the formula for the number of possible committees is \(\sum_{e \in S} \frac{100!}{e_1!e_2!e_3!}5^{e_1}10^{e_2}10^{e_3}\).

3.1.13. From ten distinct pairs of socks in the laundry, exactly eight socks can survive in 6765 ways. We have multiplicity at most two for ten types of items, and we select a multiset of size eight. Hence we seek the coefficient of \(x^8\) in \((1 + x + x^2)^{10}\). Writing this as \((\frac{1+x}{1-x})^{10}\), we want the coefficient of \(x^8\) in the product of \(\sum_{k=0}^{10} (-1)^k \binom{10}{k} x^k \) and \(\sum_{k=0}^{\infty} \binom{k+9}{k} x^k\). Hence the answer is \(\binom{10}{0} \binom{10}{1} \binom{10}{2} \binom{10}{3} \binom{10}{4} \binom{10}{5} \binom{10}{6} \binom{10}{7} \binom{10}{8} \binom{10}{9} \binom{10}{10}\), which simplifies to \(17 \cdot 13 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5\) and then 6765.

Alternatively, using the multinomial coefficient to expand \((1 + x + x^2)^{10}\), we can get eight socks with \(k\) pairs and \(8 - 2k\) singletons, where \(0 \leq k \leq 4\). The coefficient for each such possibility is \(\binom{10}{k} \binom{10}{2k} \binom{10}{2k+2} \binom{10}{2k+4} \binom{10}{2k+6} \binom{10}{2k+8} \binom{10}{2k+10}\), since we make a word of \(k\) 2s, \(8 - 2k\) 1s and \(2k + 2\) 0s to determine how many socks are contributed by each pair. Summing these contributions yields \(\binom{10}{0} \binom{10}{2} \binom{10}{4} \binom{10}{6} \binom{10}{8} \binom{10}{10} + \binom{10}{1} \binom{10}{3} \binom{10}{5} \binom{10}{7} \binom{10}{9} \binom{10}{10} + \binom{10}{2} \binom{10}{4} \binom{10}{6} \binom{10}{8} \binom{10}{10}\) as the coefficient of \(x^8\). These computations are easier: we obtain \(45 + 840 + 3150 + 2520 + 210\), which equals 6765.

Intuitively, the distinguishable ways to have eight socks survive are not equally likely. Having both socks from one pair and none from other occurs physically in only one way, while having one sock from each pair occurs in four ways.

3.1.14. The generating function for repeated rolls of a six-sided die, indexed by total sum, is \(x/1-x-x^2-x^3-x^4-x^5-x^6\). When the die is rolled \(r\) times, the generating function is \((\sum_{i=1}^{6} x^i)^r\). We may choose any \(r \in N_0\), so we sum these options:

\[
\sum_{r=0}^{\infty} \left(\sum_{i=1}^{6} x^i\right)^r = \frac{1}{1-\frac{x}{1-x-\sum_{i=1}^{6} x^i}} = \frac{1}{1-x-\sum_{i=1}^{6} x^i}.
\]

Recurrence. Let \(a_n = [x^n]A(x)\). The denominator in the expression for \(A(x)\) yields \(a_n = 2a_{n-1} - a_{n-2}\) for \(n \geq 7\). Combinatorially, from a list of rolls summing to \(n - 1\), one can form two lists summing to \(n\), by appending a 1 or by increasing the last element by 1. This generates each valid list summing to \(n\) exactly once, plus some invalid lists. The invalid lists are those ending in 7. There are \(a_{n-7}\) of these, so subtracting \(a_{n-7}\) yields \(a_n\).

3.1.15. The generating function for words without consecutive consonants from an alphabet of a vowels and c consonants, indexed by length, is \((1+x)/1-x-\alpha x\). We build words. To keep consonants apart, we iteratively choose a one-letter vowel or a two-letter consonant/vowel pair. The factor for this
3.1.16. When \( a_{n,k} \) is the number of distinguishable ways to have a multi-set of \( n \) coins from \( t \) types on a table with \( k \) showing heads, \( \sum_{n,k} a_{n,k}x^n y^k = [(1 - x)(1 - xy)]^{-t} \). The choices for various types of coins are independent, so we have identical factors for each type. For each type of coin, some number \( n \) of coins is used, and then there is one way of showing \( k \) heads, for \( 0 \leq k \leq n \). To make the number of heads accumulate in the exponent of \( y \) and the number of coins accumulate in the exponent of \( x \), we use \((1 + y + \cdots + y^n)x^y\) to model using \( n \) coins from a given type. Thus to compute the factor for each type of coin, \( \sum_{n=0}^\infty (1 + y + \cdots + y^n)x^n = \sum_{n=0}^\infty 1/x^n = \frac{1}{1-x}\). Therefore, \( a_{n,k} = \sum_{n=0}^\infty \sum_{k=0}^n x^n y^k = \sum_{m=0}^\infty x^{m}y^{m}y^{m} = \frac{1}{1-x}\).

3.1.17. When \( a_{n,k} \) is the number of ways to tile a 3-by-\( n \) rectangle using unit squares and copies of \( n \), \( \sum_{n,k} a_{n,k}x^n y^k = 1/(1 - x - xy - x^3y^2) \). Let a break be a vertical line crossed by no tile. From a break to the next break there are four options: a column of unit squares, one \( L \) in the bottom two rows with three unit squares above it, one \( L \) in the top two rows plus three unit squares, and two \( Ls \) plus three unit squares. To build the generating function, we allow any number \( m \) of such steps. With the count of columns accumulating in the exponent on \( x \) and the count of \( Ls \) accumulating in the exponent on \( y \), we have \( \sum_{n,k} a_{n,k}x^n y^k = \sum_{m=0}^\infty (x + 2x^2 + x^3y^2)^m \).

3.1.18. The generating function for Delannoy paths from \((0, 0)\) to \((m, n)\) with \( k \) diagonal steps is \( 1/(1 - x - y - xz) \). The three types of steps are modeled by \( x, y, \) and \( xyz \) so that horizontal movement accumulates in the exponent on \( x \), vertical movement in the exponent on \( y \), and the number of diagonal steps (which cause both horizontal and vertical movement) in the exponent on \( z \). Hence one step is modeled by \((x+y+xyz)\). Any number of steps can be taken, so the generating function is \( \sum_{r=0}^\infty (x+y+xyz)^r \).

3.1.19. Let \( \hat{F}(x) \) be the OGF for the adjusted Fibonacci sequence (\( \hat{F} \)).

a) Derivation of \( \hat{F}(x) \) from the Fibonacci recurrence. Applying the generating function method to \( a_n = a_{n-1} + a_{n-2} \) for \( n \geq 2 \), with \( a_0 = a_1 = 1 \), we multiply by \( x^2 \) and then sum over \( n \geq 2 \). This yields \( \hat{F}(x) - 1 - x = x(\hat{F}(x)-1)+x^2 \hat{F}(x) \), or \( \hat{F}(x)(1-x-x^2) = 1 \). Therefore, \( \hat{F}(x) = (1-x-x^2)^{-1} \).

b) Combinatorial derivation of \( \hat{F}(x) \). The generating function enumerates 1, 2-lists, indexed by the sum. Each term contributes 1 or 2 to the sum, so the generating function for a single term is \( x + x^2 \); for a list with \( k \) terms it is \((x + x^2)^k \). The list may have any number of terms, so the generating function is \( \sum_{k=0}^\infty (x + x^2)^k \), which equals \((1-((x + x^2))^{-1})^{-1} \), as in part (a). The coefficient of \( x^n \) is the number of 1, 2-lists with sum \( n \).

c) \( F_n = \sum_k (n-k)^k \). Since the generating function is \( \sum_{k=0}^\infty (x + x^2)^k \), \( \hat{F}_n = [x^n] \sum_{k=0}^\infty (x + x^2)^k = \sum_{k=0}^\infty [x^{n-k}](x + x^2)^k = \sum_k (n-k) = \sum_k (n-k)^k \).

3.1.16. The enumerator of \( k \)-colored graphs by number of edges is \( \frac{1}{k!} \sum_{n_1, \ldots, n_k} (x + 1)^{k(n-k)}/2 \), where the sum is over choices of nonzero \( n_1, \ldots, n_k \) with sum \( n \). Temporarily index the sets of the partition as 1, \( \ldots, k \). The classes are distinguished by the vertices within them, since they are all nonempty, so we can remove the indexing at the end by dividing by \( k! \).

The choice of \( n_1, \ldots, n_k \) is the choice of sizes of the classes; we sum over this choice. Once the sizes are fixed, the multinomial coefficient counts the distributions of the vertices to the classes. Once the vertices are distributed, we allocate edges. Each edge that occurs contributes 1 to the exponent on \( x \), which accumulates the number of edges. Each allowed edge contributes or does not, so we have the factor \((1 + x)\) for each allowed edge. On \( n \) vertices, there are \( \binom{n}{2} \) possible edges, but the edges within color classes are not allowed, so we subtract \( \sum \binom{n}{2} \). In the exponent \( \sum n_i/2 \) cancels \( n/2 \).

3.1.18. The number \( a_{n,k} \) of acyclic directed graphs with vertex set \([n]\) that have exactly \( k \) sources satisfies

\[ a_{n,k} = \binom{n}{k} \sum_{j=1}^{n-k} a_{n-k,j} (2^k - 1)^j 2^{(n-k-j)} \]

Note \( a_{n,n} \), which allows the inductive computation to proceed. For \( n > k \), we first choose \( k \) vertices to be the set \( S \) of sources of the digraph \( G \) we are building. Since \( G \) has no cycle, the digraph \( G' \) obtained by deleting \( S \) from \( G \) also has no cycle and thus has some nonempty set \( S' \) of sources (this point is discussed formally in Chapter 5). We sum over the size \( j \) of \( S' \); it ranges from 1 to \( n - k \). Having chosen \( G' \), we must add edges so that each source in \( G' \) is not a source in \( G \). For each vertex \( v \in S' \), we specify a nonempty subset of \( S \) to have edges pointing to \( v \); this yields the factor \((2^k - 1)^j \). Finally, for each \((u, v) \) with \( u \in S \) and \( v \in V(G') - S' \), the edge \( uv \) may or may not be present; this explains the factor \( 2^{(n-k-j)} \).
To obtain from this a generating function indexed by number of edges, we want to contribute \( 1 \) to the exponent of \( x \) for each edge created. Thus we replace each 2 in the recurrence by \( 1 + x \). Also we replace \( a_{n-k,j} \) with the generating function by number of edges for acyclic digraphs with vertex set \([n - k]\) having \( j \) sources.

The reason for restricting the problem to acyclic directed graphs is that this argument does not provide a way to start building a digraph that has no sources.

3.1.22. Over all \( k \)-sets \( S \subseteq [n] \) occurring with equal probability, the expected number of inversions in \( \pi(S) \) is \( k(n-k)/2 \), where \( \pi(S) \) is the permutation consisting of \( S \) in increasing order followed by \([n] - S \) in increasing order. Fix \( k \) and \( n \) with \( 0 \leq k \leq n \). The formula is trivial when \( k = 0 \) or \( k = n \), because then there is only one subset \( S \), and \( X(S) = 0 \). Thus the formula holds for \( n = 1 \), and we use induction on \( n \).

For \( n > 1 \), we may assume \( 0 < k < n \). Consider separately the cases where \( n \notin S \) and \( n \in S \). If \( n \notin S \), then the inversions are precisely the inversions that occur when \( S \) is chosen from \([n-1] \), so we get the resulting expectation \((n-1)^k\) times. When \( n \in S \), we choose only \( k - 1 \) elements from \([n-1] \) to complete \( S \), but the element \( n \) at the end of \( S \) contributes \( n - k \) additional inversions. Thus

\[
\sum X(S) = \left(\frac{n-1}{k}\right)\frac{k(n-1-k)}{2} + \left(\frac{n-1}{k+1}\right)\frac{(k-1)(n-k)}{2} + (n-k)\]

and hence \( E(X) = k(n-k)/2 \).

3.1.23. The expected number of cycles in a random permutation of \([n]\) is \( \sum_{k=1}^{n} 1/k \) (approximately \( \ln n \)).

(a) By generating functions. The OGF \( C(x) \) for permutations of \([n]\) by number of cycles is \( x^n \). The probability that a random permutation has \( k \) cycles is \( 1/k! C(x) \), and the expected value is \( 1/k! C'(1) \). The product rule for differentiation yields \( C'(x) = C(x) \sum_{i=1}^{n} 1/(x+i) \). Thus \( C'(1) = n! \sum_{i=1}^{n} 1/i \), and the expected number of cycles is \( \sum_{i=1}^{n} 1/i \).

(b) By linearity of expectation. Let \( X \) count the cycles in a random permutation.

Proof 1 For \( i \in [n] \), let \( X_i \) be the reciprocal of the length of the cycle containing \( i \). Now each cycle of length \( k \) corresponds to \( k \) variables with values \( 1/k \), so \( X = \sum_{i=1}^{n} X_i \). All of \( X_1, \ldots, X_n \) have the same distribution, so \( E(X) = nE(X_1) \). From Proposition 3.1.19, \( P(X_1 = 1) = 1/n \). Thus \( E(X_1) = \sum_{k=1}^{n} \frac{1/n}{k} = \frac{1/n}{n} \sum_{k=1}^{n} 1/k \). We conclude \( E(X) = \sum_{k=1}^{n} 1/k \).

Section 3.1: Ordinary Generating Functions

Proof 2 For each possible cycle \( C \), let \( X_C \) be 1 if \( C \) is a cycle in the permutation, 0 if it is not. Thus \( X = \sum C X_C \) and \( E(X) = \sum P(X_C = 1) \).

A cycle of length \( k \) occurs with probability \( (k-1)/n^k \), since the other elements can be permuted arbitrarily. There are \( \binom{n}{k} \) possible cycles with \( k \) elements. Thus \( E(X) = \sum_{k=1}^{n} \binom{n}{k} (k-1)! (n-k)! = \sum_{k=1}^{n} \frac{1}{k} \).

3.1.24. The probability that a random permutation of \([n]\) has exactly two cycles is \( 1/n^2 \sum_{k=1}^{n-1} 1/k \).

(a) Using the generating function. We seek \([x^2] x^n \). Terms in \( \prod_{i=1}^{n} (x + i - 1) \) that contribute to this coefficient use two factors of \( x \) and \( n-2 \) constant factors. The nonzero terms select \( n-2 \) of the \( n-1 \) constant factors \( 1, \ldots, n-1 \); summing them yields the formula claimed.

(b) Using canonical cycle representations. A permutation has two cycles if and only if its canonical cycle representation has 1 after the first position and has the least of the numbers before 1 in the first position. For \( 0 \leq k \leq n-1 \), there are \( (n-1)! \) permutations whose canonical cycle representation has 1 in position \( k + 1 \). For each choice of \( k \) elements in the first \( k \) positions, \( 1/k \) of the representations have the least of them in the first position. Thus \( (n-1)!/k \) permutations with two cycles have a canonical cycle representation with 1 in position \( k + 1 \).

3.1.25. In a random permutation of \([n]\), the probability is \( 1/r \) that any \( r \) specified elements are in the same cycle. The answer is the same for any \( r \) elements, so consider \( \{1, \ldots, r\} \). Since 1 is least in its cycle, its cycle consists of 1 and all later elements in the canonical cycle representation. Hence \( 1, \ldots, r \) are in the same cycle if and only if 1 is first among these in the canonical cycle representation. There are the same number of permutations with 1 in each position among \( 1, \ldots, r \), so the probability that 1 is first among them is \( 1/r \). Sums to count the good permutations yield the same result less simply.

3.1.26. Equivalence of two random vectors. Let \( m_0 = n \). Starting with \( i = 0 \), iteratively choose \( x_i \in [m_i] \) at random, subtract \( x_i \) from \( m_i \) to obtain \( m_{i+1} \), and continue (stopping if \( m_{i+1} = 0 \)). Let \( Y_k \) be the number of times that \( k \) is the randomly chosen integer. The distribution of the random vector \( (Y_1, \ldots, Y_n) \) is the same as the distribution of the random vector \((X_1, \ldots, X_n)\), where \( X_k \) is the number of \( k \)-cycles in a random permutation of \([n]\).

We show that the cycle lengths in the canonical cycle representation, from back to front, are produced by the same process as \((Y_1, \ldots, Y_n)\). The last cycle starts wherever element 1 appears; its length has the uniform distribution over \([n]\), just like \( x_0 \) in the other process. Let its value be \( k \). No matter which \( k-1 \) other elements are chosen to be in that cycle, the
remainder of the canonical cycle representation is generated in the same way: it follows the same process on \( n - k \) elements, with the next cycle (from the end) started by the smallest of those \( n - k \) remaining elements. Thus the next length generated is uniform over \([n - k]\), just as the next integer generated in the other process is uniform over \([m_1]\). Step by step, both lists of numbers are generated under the same process.

3.1.27. The sets \( A_n \) and \( B_n \) have the same size, where \( A_n = \{ \pi \in S_n : \pi_i \neq i + 1 \text{ for } 1 \leq i \leq n - 1 \} \) and \( B_n = \{ \sigma \in S_n : \sigma_{j+1} \neq \sigma_j + 1 \text{ for } 1 \leq j \leq n - 1 \} \). Consider \( \pi \in S_n \). If \( \pi_i > i \), then \( \pi_i \) is not the least element in its cycle. Thus if \( \pi_i > i \), then \( \pi_i \) immediately follows \( i \) in the canonical cycle representation of \( \pi \). This means that if the canonical cycle representation of \( \pi \) is in \( B_n \), then \( \pi \) is in \( A_n \). Conversely, if \( i + 1 \) immediately follows \( i \) in the canonical cycle representation of \( \pi \), then \( \pi_i = i + 1 \); therefore, if \( \pi \in A_n \), then the canonical cycle representation of \( \pi \) is in \( B_n \).

Thus the number of permutations whose canonical cycle representations are in \( B_n \) is \( |A_n| \). Since canonical cycle representation defines a bijection on \( S_n \), this yields \( |B_n| = |A_n| \).

3.1.28. When the keys to \( n \) boxes are randomly permuted into the boxes, the probability of being able to open all by breaking \( k \) boxes is \( k/n \). Letting \( \sigma(i) \) be the index of the key in box \( i \) defines a permutation from \([n]\) to \([n]\). When a box is broken, one can open all boxes whose indices are in the cycle of \( \sigma \) containing the index of the broken box.

To compute the probability that every element lies in a cycle with at least one of the \( k \) selected indices, it suffices by symmetry to assume that the \( k \) boxes broken are 1 through \( k \).

In a permutation, every element lies in a cycle with an element of \([k]\) if and only if the first element in the canonical cycle representation is at most \( k \). By Lemma 3.1.18, generating a random permutation is equivalent to generating a random canonical cycle representation. Of the \( n! \) permutations in word form, the fraction with first entry in \([k]\) is \( k/n \).

3.1.29. The number of fixed points of the permutation on \( S_n \) that takes a permutation to its reverse canonical cycle representation (with the parentheses dropped) is the adjusted Fibonacci number \( \hat{F}_n \). The permutations consist of fixed points and disjoint transpositions of two consecutive integers. Let \( \psi_n \) be the specified mapping, and let \( a_n \) be the number of fixed points of \( \psi_n \). Note that \( a_0 = a_1 = 1 \).

If \( \pi \) is a fixed point of \( \psi_n \), such that \( \pi(n) = n \), where \( n \geq 1 \), then the restriction of \( \pi \) to \([n - 1]\) must be a fixed point of \( \psi_{n-1} \), since the computation of its reverse canonical representation is unaffected by the presence of \( n \). Conversely, every fixed point \( \pi \) of \( \psi_{n-1} \) extends to a fixed point of \( \psi_n \) by setting \( \pi(n) = n \). Hence there are \( a_{n-1} \) fixed points \( \pi \) of \( \psi_n \) such that \( \pi(n) = n \).

Let \( \pi \) be a fixed point of \( \psi_n \), such that \( \pi(n) < n \), where \( n \geq 2 \). Since \( n \) is the largest element of its cycle, this cycle appears in the reverse canonical representation as \( (n, \pi(n), \ldots) \). Thus the word form of \( \psi_n(\pi) \) ends with \( n, \pi(n), \ldots \). Since the word form of \( \pi \) ends with \( \pi(n) \), to be a fixed point of \( \psi_n \) it must end with \( n, \pi(n) \). Thus \( \pi(n - 1) = n \), and the cycle of \( \pi \) containing \( n \) has only the two elements \( n \) and \( n - 1 \). Now as above the restriction of \( \pi \) to \([n - 2]\) must be a fixed point of \( \psi_{n-2} \), and such fixed points extend to fixed points of \( \psi_n \), by adding the cycle \((n, n - 1)\).

Thus \( a_n = a_{n-1} + a_{n-2} \) for \( n \geq 2 \), with \( a_0 = a_1 = 1 \), so \( a_n = \hat{F}_n \). Since each fixed point ends has the cycle \((n)\) or \((n, n - 1)\), inductively the fixed points of \( \psi_n \) consist of fixed points and disjoint transpositions of two consecutive integers.

3.1.30. \(|E_{2n}| = |O_{2n}| \), where elements of \( S_{2n} \) are in \( E_{2n} \) or \( O_{2n} \), if their cycle lengths are all even or all odd, respectively. Note that all elements of \( O_{2n} \) have an even number of cycles, though elements of \( E_{2n} \) need not.

Consider the canonical cycle representation of \( \pi \in O_{2n} \); let the cycles in order be \( \pi_1, \ldots, \pi_{2k} \). Form a permutation \( \pi' \) as follows. For \( 1 \leq i \leq k \), move the second element of \( \pi_{2i} \) to the end of \( \pi_{2k} \) unless \( \pi_{2k-1} \) is a single element (fixed point), in which case move that element itself to the end of \( \pi_{2k} \). Because the first element of \( \pi_{2k} \) is less than the first element of \( \pi_{2k-1} \), the permutation \( \pi' \) given by cycles is also already in canonical form. Furthermore, its cycles all have even length.

To show that this map is a bijection, consider \( \pi' \in E_{2n} \), with cycles \( \pi_1, \ldots, \pi_j \) in order in canonical form. Let \( a \) denote the last element of \( \pi_j \), and let \( b \) denote the first element of \( \pi'_j \) (if \( j > 1 \)). If \( a \) exceeds \( b \), then move \( a \) from \( \pi_j \) so that it immediately follows \( b \) in \( \pi'_{j-1} \) and begin the process anew at \( \pi'_{j-1} \) (if present). If \( b \) exceeds \( a \) or if \( j = 1 \), then make \( a \) a singleton placed immediately before \( \pi'_{j-1} \) and begin the process anew at \( \pi'_{j-2} \). This undoes in reverse the steps performed to obtain \( \pi' \) from \( \pi \).

3.1.31. Let \( c(n, k) \) be the number of permutations of \([n]\) with \( k \) cycles.

a) \( c(n + 1, m + 1) = \sum_{k=m}^{n} c(n, k) \binom{k}{m} \), bijectively. The left side counts permutations of \([n + 1]\) with \( m + 1 \) cycles. We associate each with a unique permutation of \([n]\) having at least \( m \) cycles, with \( m \) cycles marked; there are \( \sum_{k=m}^{n} c(n, k) \binom{k}{m} \) of these.

A permutation \( \sigma \) of \([n + 1]\) (with \( m + 1 \) cycles) has \( n + 1 \) in one cycle. Mark the \( m \) other cycles. Rotate the unmarked cycle so that \( n + 1 \) is at the end and view the rest of it as a linear list \( \tau \). From \( \tau \) form cycles that begin at the left-to-right minima. The result is a permutation of \([n]\), described by cycles, having at least \( m \) cycles, with \( m \) of them marked.
To reverse the process, combine the unmarked cycles into one cycle by writing them with least element first, in descending order of least elements. Add + 1 at the end to complete the \((m + 1)\)th cycle.

b)\( \sum_{k=1}^{n} c(n, k)x^k = x^{(n)}, \) by induction on \(n.\) Let \( A_n(x) = \sum_{k=1}^{n} c(n, k)x^k. \) Note that \( A_0(x) = 1 = x^{(0)} \) (the empty permutation has no cycles). For the induction step, multiplying part (a) by \( x^n \) and summing over \( m \geq 0 \) yields the following computation, where we interchange the order of summation and use the induction hypothesis to introduce \( A_n. \)

\[
\frac{d}{dx}A_{n+1}(x) = \sum_{m \geq 0} c(n + 1, m + 1)x^m = \sum_{m \geq 1} \sum_{k=m}^{n} c(n, k)\binom{k}{m}x^m
\]

\[
= \sum_{k=1}^{n} c(n, k)\sum_{m=1}^{k} \binom{k}{m}x^m = \sum_{k=1}^{n} c(n, k)(1 + x)^k
\]

\[
= A_n(1 + x) = \prod_{i=1}^{n}(1 + x + i - 1) = \prod_{i=1}^{n}(x + i).
\]

Multiplying by \( x \) completes the induction step.

c) OGF from recurrence. We have \( c(n, k) = (n - 1)c(n - 1, k) + c(n - 1, k - 1) \) for \( n, k \geq 1, \) with \( c(0, k) = \delta_{0,k} \) and \( c(n, 0) = \delta_{n,0}. \) For \( n \geq 1, \) multiply by \( x^k \) and sum over \( k \geq 1 \) to obtain

\[
\sum_{k=1}^{n} c(n, k)x^k = (n - 1)\sum_{k=1}^{n} c(n - 1, k)x^k + x\sum_{k=1}^{n} c(n - 1, k - 1)x^{k-1},
\]

or \( A_n(x) = (n - 1)A_{n-1}(x) + xA_{n-1}(x). \) With \( A_0(x) = 1, \) from \( A_n(x) = (x + n - 1)A_{n-1}(x) \) we obtain \( A_n(x) = x^n. \)

3.1.32. Addition method for generating factorials. Every entry in row 0 is 1. To obtain row \( j + 1 \) from row \( j, \) delete the first element of row \( j, \) then the second after that, then the third after that, and so on. Take partial sums of the rest to form row \( j + 1, \) with blanks under deleted entries.

The process forms triangular wedges: wedge \( n \) has columns indexed \( 0, \ldots, n, \) with one less column occupied in each succeeding row. Let \( a_{n,m} \) be the bottom entry in column \( m \) (it is in row \( n - m \)). We prove \( a_{n,m} = c(n + 1, m + 1) \), by induction on \( n; \) then \( a_{n,0} = c(n + 1, 1) = n! \), as desired. The basis is \( a_{0,0} = 1 = c(1, 1). \) By Exercise 3.1.31(a), \( c(n + 1, m + 1) = \sum_{k=m}^{n} c(n, k)\binom{k}{m}. \) Thus it suffices to show \( a_{n,m} = \sum_{k=m}^{n} c(n, k)\binom{k}{m}a_{n-1,k-1} \) for \( n \geq 1. \)

\[
\begin{array}{cccccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
2 & 6 & 11 & 18 & 26 & 35 & 4 \times 5 & 6 & 50 & 24 \\
1 & 5 & 10 & 15 & 20 & 25 & 30 & 35 & 40 & 45 \\
\end{array}
\]

Each entry is the sum of the entries above and to its left. For \( 0 \leq m \leq n, \) the number of times \( a_{n-1,k-1} \) contributes to \( a_{n,m} \) is the number of paths from row \( n - k \) of column \( k - 1 \) in wedge \( n - 1 \) to row \( n - m \) of column \( m \) in wedge \( n, \) moving right or down at each step. Each such path steps first to row \( n - k \) of column 0 of wedge \( n. \) It then takes \( m \) rightward steps and \( k - m \) downward steps in some order to reach its goal. Hence the number of paths is the binomial coefficient \( \binom{k}{m}, \) as desired.

This completes the proof; setting \( m = 0 \) yields \( a_{n,0} = n! \), as desired.

3.1.33. If \( F_n(z) = \sum_{k \geq 0} c(k, n)z^k/k!, \) where \( c(k, n) \) is the number of permutations of \( [k] \) with \( n \) cycles, then \( F_n(z) = [\ln(1 - z)]^n/n! \). The generating function for the permutations of \( [k] \) by number of cycles is \( \sum_{n \geq 0} c(k, n)x^n = \prod_{i=0}^{k-1} (x + i) \) (Theorem 3.1.20). In the generating function \( \sum_{n \geq 0} F_n(x) \), we interchange the order of summation to compute

\[
\sum_{n \geq 0} F_n(x) = \sum_{k \geq 0} \frac{z^k}{k!} \sum_{n \geq 0} c(k, n)x^n = \sum_{k \geq 0} \frac{z^k}{k!} \prod_{i=0}^{k-1} (x + i).
\]

\[
= \sum_{k \geq 0} \frac{z^k}{k!} (1) = \sum_{k \geq 0} \left( \frac{z}{k} \right) (-1)^k = \left( \frac{-x}{k} \right) (-z)^k = (1 - z)^{-x}
\]

\[
= e^{-x\ln(1-z)} = \sum_{n \geq 0} \frac{[-x\ln(1-z)]^n}{n!} = \sum_{n \geq 0} \frac{[-\ln(1-z)]^n}{n!} x^n
\]

Finally, \( F_n(z) \) must be the coefficient of \( x^n \) in the OGF.

3.1.34. The Eulerian number \( A(n, k) \) is given by

\[
A(n, k) = \sum_{i=0}^{k} (-1)^i \binom{n+1}{i} (k-i)^n.
\]

Into the desired sum we substitute Woritzky’s Identity, stated with \( x \) being \( k-i \); we then collect the coefficient of each \( A(n, j). \) Thus

\[
\sum_{i=0}^{k} (-1)^i \binom{n+1}{i} (k-i)^n = \sum_{i=0}^{k} (-1)^i \binom{n+1}{i} \sum_{j=1}^{n} A(n, j) \binom{k-i+n-j}{n}
\]

\[
= \sum_{j=1}^{k} A(n, j) \sum_{i=0}^{k} (-1)^i \binom{n+1}{i} \binom{n+k-j-i}{n}
\]

The last factor is 0 when \( i > k-j, \) so we can stop the inner sum at \( k-j. \)

To complete the proof, we want the inner sum to be 0 when \( j < k \) and 1 when \( j = k. \) For this we use the negative binomial identity and Vandermonde Convolution from Section 1.2. A noted there, these are polynomial identities. For all \( u, x, y \in \mathbb{R} \) and \( r \in \mathbb{N}, \) we have \((u)_r = (-1)^r (u+r-1)\) and \( \sum_{i} \binom{r}{i} (y)^i = (x+y)^r. \) Letting \( s = k-j, \) we compute
The sum on the left counts all runs in all permutations. A permutation and its reverse together have \( n - 1 \) descents, so together they have \( n + 1 \) runs. With \( n + 1 \) runs for every pair of permutations of \([n]\), there are altogether \((n + 1)!/2\) runs.

**Proof 2** (descents). The sum on the left counts all runs in all permutations. A permutation and its reverse together have \( n - 1 \) descents, so together they have \( n + 1 \) runs. With \( n + 1 \) runs for every pair of permutations of \([n]\), there are altogether \((n + 1)!/2\) runs.

**3.1.37.** \( B(n, k) = A(n, k) \), where \( B(n, k) \) counts the lists \((b_1, \ldots, b_n) \in \mathbb{N}^n\) such that \( b_i \leq i \) for all \( i \) and \( k - 1 \) elements of \([n]\) are omitted. It suffices to show that \( B(n, k) \) satisfies the same recurrence as the Eulerian numbers. By Proposition 3.1.24, \( A(n, k) = kA(n - 1, k) + (n - k + 1)A(n - 1, k - 1) \) for \( n, k \in \mathbb{N} \). For \( n = 0 \), the initial conditions are the same: \( A(0, k) = B(0, k) = \delta_{0,k} \).

For \( B(n, k) = kB(n - 1, k) + (n - k + 1)B(n - 1, k - 1) \), when extending a list of length \( n - 1 \), the new element may already be present or not. The number of missing elements increases by 1 or remains unchanged, respectively. Hence lists counted by \( B(n, k) \) arise from each list counted by \( B(n - 1, k - 1) \) in \((n - 1) - (k - 2)\) ways \( (n \) enters the set of missing values) and from each list counted by \( B(n - 1, k) \) in \( k \) ways \((b_0 = \text{a previously missing element or is } n)\). These two cases generate all lists, so \( B(n, k) = kB(n - 1, k) + (n - k + 1)B(n - 1, k - 1) \).

**3.1.38.** \( \sum_{k \geq 0} x^k k^2 \) is an integer when \( n \in \mathbb{N} \). Let \( C(x) = \sum_{k \geq 0} x^k k^2 \). In Theorem 3.1.26, it is proved that \( C(x) = A(x)(1 - x)^{-(n + 1)} \), where \( A(x) = \sum_{k = 0}^{n} A(n, k) x^k \). Setting \( x = 1/2 \), we compute \( \sum_{k \geq 0} x^k k^2 = 2^{n - 1} \sum_{k = 0}^{n} A(n, k) 2^k = \sum_{k = 0}^{n} A(n, k) 2^{n + 1 - k} \). We have expressed the desired series as a finite sum of integers.

**3.1.39.** Games adding dollars. Two games begin with players A and B each having $1. On each play, A or B wins $1. When A has $a$ and B has $b$, the probability that A gets the next dollar is $a/(a + b)$ in Game 1, but in Game 2 it is $b/(a + b)$. For $1 \leq k \leq n$, the probability that A has $\$k$ when the total is $\$n + 1$ is $1/n$ in Game 1 and is $A(n, k)!/n!$ in Game 2.

Let \( A_n \) be the amount held by A when the total reaches \( n + 1 \), so \( 1 \leq A_n \leq n \). We have \( A_1 = 1 \) with probability 1, so the claims hold for \( n = 1 \). Let \( X_n \) record the recipient of the dollar in the round that reaches total \( n + 1 \); the outcome determines \( A_n \). In either game,

\[
P(A_n = k) = P(X_n = B \mid A_{n-1} = k) P(A_{n-1} = k) + P(X_n = A \mid A_{n-1} = k - 1) P(A_{n-1} = k - 1).
\]

**Game 1: Uniform distribution.** The inductive computation is

\[
P(A_n = k) = \frac{n - k}{n} \frac{1}{n - 1} + \frac{k - 1}{n} \frac{1}{n - 1} = \frac{1}{n}.
\]

**Game 2: Eulerian distribution.** The inductive computation is

\[
P(A_n = k) = \frac{n - k}{n} \frac{1}{n - 1} + \frac{k - 1}{n} \frac{1}{n - 1} = \frac{1}{n}.
\]

**Proof 1** (marking runs). The sum on the left counts permutations with a marked run, by the number of runs. To mark a run, insert \( n + 1 \) at its end. Permutations of \([n + 1]\) starting with \( n + 1 \) don’t arise, while those ending with \( n + 1 \) do. Other permutations have a string \( \ldots, a, n + 1, b, \ldots \); these arise if and only if \( a > b \), because that is when \( a \) ends a run in the permutation obtained by deleting \( n + 1 \). Hence reversal exchanges good and bad permutations of \([n + 1]\).
Here we use the recurrence for $A(n, k)$ from Proposition 3.1.24.

### 3.1.40. Let $f(\sigma)$ map a permutation $\sigma$ of $[n]$ to the graph on $[n] \cup \{0\}$ in which each $i$ has an edge to the last element of $\sigma$ before $i$ that is less than $i$, where $\sigma$ has been augmented by putting $0$ at the beginning. Let $T_n$ be the set of increasing trees with vertex set $[n] \cup \{0\}$.

- $f$ is a bijection from the set of permutations of $[n]$ to $T_n$. 

  By induction on $n$, we claim that after adding edge $ij$ with $j < i$, the graph built thus far is a tree with vertex set $[i] \cup \{0\}$. This is true at the start. If it is true after $i-1$ steps, then adding $ij$ adds an edge to a new vertex, so the graph remains connected and without cycles.

  Therefore, after $n$ steps, we have a tree with vertex set $[n] \cup \{0\}$. Furthermore, since $i$ when it is added is the largest vertex in the current tree, all paths from the root remain increasing at all times. Hence $f(\sigma) \in T_n$.

  To complete the proof that $f$ is a bijection, it suffices to show that for each $T \in T_n$, there is exactly one solution to $f(\sigma) = T$. We use induction on $n$. For $n = 1$, the one tree and one permutation match up. For $n > 1$, the element $n$ is a leaf of $T$, since $T$ is increasing. Let $T' = T - n$. Deleting a leaf from a tree yields a smaller tree (proved in the Introduction), so $T'$ is a tree. Therefore, the induction hypothesis yields a unique permutation $\sigma'$ such that $f(\sigma') = T'$.

  Since $n$ exceeds all earlier values, the location of $n$ is irrelevant during the first $n-1$ steps of the algorithm. That is, every permutation that produces $T$ after $n$ steps produces $T'$ after $n-1$, and thus the word form of it when element $n$ is suppressed must be the same as $\sigma'$. Finally, since all elements preceding $n$ are less than $n$, in such a permutation $n$ must immediately follow the element that is its neighbor in $T$. Hence every such $\sigma$ is obtained from $\sigma'$ by inserting $n$ in $\sigma'$ immediately following the neighbor of $n$ in $T$. Since $\sigma'$ is uniquely determined, the resulting $\sigma$ is also uniquely determined.

- **b)** The number of elements of $T_n$ whose root has degree $k$ is the number $c_{n,k}$ of permutations of $[n]$ with $k$ cycles. In forming $f(\sigma)$, an element $i$ becomes a neighbor of the root 0 if and only if all the elements between $i$ and the prepended 0 in $\sigma$ exceed $i$. That is, if and only if $i$ is a left-to-right minimum in $\sigma$. By Lemma 3.1.18, the number of permutations of $[n]$ with $k$ left-to-right minima is the number of permutations of $[n]$ with $k$ cycles.

- **c)** The number of elements of $T_n$ with $k$ leaves is the Eulerian number $A(n, k)$. It suffices to show that $x$ is a leaf in $f(\sigma)$ if and only if $x$ is the end of a run in $\sigma$.

  If $x$ is a leaf, then the element $i$ following $x$ must be less than $x$, since otherwise we add the edge $ix$, giving $x$ a second neighbor.

  If $x$ is not a leaf, then a second incident edge $ix$ is added after the step when $x$ is first reached. This requires (1) $i$ is later than $x$ in $\sigma$, (2) $x < i$, and (3) all elements between $x$ and $i$ exceed $x$. Thus the element immediately following $x$ exceeds $x$, so $x$ does not end its run.

### 3.1.41. A permutation $\sigma$ has an excedance at $i$ if and only if the permutation $\delta$ obtained by dropping parentheses in the canonical cycle representation of $\sigma$ has an ascent at $i$, where an excedance at $i$ means $\sigma(i) > i$ and an ascent at $i$ means $\delta(i+1) > \delta(i)$.

- If $\sigma(i) > i$, then $\sigma(i)$ is not the least element in its cycle. Thus if $\sigma(i) > i$, then $\sigma(i)$ immediately follows $i$ in $\delta$; excedance at $i$ in $\sigma$ implies ascent at $i$ in $\delta$. Conversely, if $j$ immediately follows $i$ in $\delta$ and $j > i$, then $\sigma(j) = j$; ascent at $i$ in $\delta$ implies excedance at $i$ in $\sigma$.

### 3.1.42. The number of permutations of $[n]$ having $k$ weak excedances is $A(n, k)$. The reverse complement $\sigma^*$ of a permutation $\sigma$ is defined by $\sigma^*(i) = n + 1 - \sigma(n + 1 - i)$, or $\sigma^*(n+1-i) = n+1-\sigma(i)$. Thus $\sigma(i) \geq i$ if and only if $\sigma^*(n+1-i) \leq n+1-i$, which in turn is equivalent to saying that $\sigma^*$ does not have an excedance at $n+1-i$. Hence the number of positions where $\sigma$ has a weak excedance equals the number of positions where $\sigma^*$ does not have an excedance. That is, $k$ has $k$ weak excedances if and only if $\sigma^*$ has $n-k$ excedances. By Exercise 3.1.41, the number of such permutations is $A(n, k)$.

### 3.1.43. Smith College Diploma Problem.

- **a)** Cyclically shifting the non-fixed points reduces the number of excedances by 1. For $\sigma \in S_n$, with $s_1, \ldots, s_m$ being the non-fixed elements in increasing order, let $\sigma^*(s_i) = \sigma(s_{i+1})$ if $i < m$ and $\sigma^*(s_m) = \sigma(s_1)$. The fixed points of $\sigma$ are also fixed points of $\sigma^*$. We show that if $\sigma$ has $k$ excedances, then $\sigma^*$ has $k-1$ excedances. Specifically, $\sigma^*$ has an excedance at $s_i$ if and only if $\sigma$ has an excedance at $s_{i+1}$, for $1 \leq i < m$. Neither has an excedance at $s_m$, so the excedance that $\sigma$ has at $s_1$ disappears.

  Since $s_1 < \cdots < s_m$, for $i \leq m$ we have $\sigma^*(s_i) = \sigma(s_{i+1}) > s_{i+1} > s_i$ if $\sigma$ has an excedance at $s_{i+1}$. Hence $\sigma^*$ has an excedance at $s_i$ if $\sigma$ has an excedance at $s_{i+1}$. The other possibility, when $\sigma$ does not have an excedance at $s_{i+1}$, is $\sigma(s_{i+1}) < s_{i+1}$, since we move only the elements that are not fixed by $\sigma$. Since $s_i$ is the next element less than $s_{i+1}$ that is a possible value for $\sigma(s_{i+1})$, we have $\sigma^*(s_i) \leq s_i$, and $\sigma^*$ does not have an excedance at $s_i$.

- **b)** The number of permutations of the $n$ diplomas that result in $k$ rounds for sorting is $A(n, k)$. The only permutation with no excedances is the
identity permutation, which as described takes one round, and \( A(n, 1) = 1 \). For \( k > 1 \), the shifting process studied in part (a) models one round of the sorting; it reduces the number of excedances by 1. Hence the number of permutations taking \( k \) rounds to sort is the number of permutations having \( k - 1 \) excedances.

In Exercise 3.1.41 it is observed that \( A(n, k) \) is the number of permutations having \( n - k \) excedances. However, \( A(n, k) = A(n, n + 1 - k) \), by reversal of permutations. Hence the number of permutations with \( n - k \) excedances also equals \( A(n, n + 1 - k) \). Letting \( j = n + 1 - k \), this statement becomes that the number of permutations with \( j - 1 \) excedances is \( A(n, j) \). That is, \( A(n, k) \) is the number of permutations whose sorting takes \( k \) rounds.

### 3.2. COEFFICIENTS & APPLICATIONS

#### 3.2.1. Computing coefficients.

- **a)** \( [x^{10}] (1 + x)^3 (1 - x)^{-3} = 402 \). Note that \( (1 - x)^{-3} = \sum_{n \geq 0} (n+2)x^n \).

  Hence the coefficient of \( x^{10} \) is \( \sum_{k=0}^{3} \binom{10}{2k} (2k+2) \), which is \( \binom{12}{2} + 3 \binom{11}{2} + 3 \binom{10}{2} \).

- **b)** \( [x^{10}] (x^2 + x^3 + x^4)^4 = 10 \). Factoring out \( (x^2)^4 \) shows that the answer is \( \binom{n}{2}(1 + x + x^2)^3 \). To obtain contributions to \( x^2 \), one factor can contribute \( x^2 \) and the rest \( x^0 \), or two factors can contribute \( x^4 \) and the other two \( x^0 \). There are four ways in the first case and \( \binom{3}{2} \) in the second.

- **c)** \( [x^{10}] (1 - 2x)^{-3} = 67584 \). We have \( (1 - 2x)^{-3} = \sum_{n \geq 0} \binom{n+3}{3}(2x)^n \).

  Hence the coefficient of \( x^{10} \) is \( \binom{12}{2} 2^{10} \).

#### 3.2.2. The generating function \( \sum_{k \geq 0} a_k x^k \), where \( a_k = \binom{n}{k/2} \). Consider the even and odd terms separately.

\[
\sum_{k \geq 0} a_k x^k = \sum_{j \geq 0} \binom{n}{2j} x^{2j} + \sum_{j \geq 0} \binom{n}{2j+1} x^{2j+1} = (1 + x) \sum_{j \geq 0} \binom{n}{j} x^j = (1 + x(1 + x^2)) = (1 + x)(1 + x^2)^n.
\]

#### 3.2.3. If \( a_n = \sum_{k \geq 1} \frac{1}{kx^k} \) for \( n \geq 1 \), then \( \sum_{n \geq 1} a_n x^n = -\ln(1-x)/(1-\frac{x}{2}) \). The sum is a convolution. It is the coefficient of \( x^n \) in the product of the generating functions \( \sum_{k \geq 1} \frac{1}{kx^k} \) and \( \sum_{k \geq 0} \frac{1}{2x^k} \). These generating functions are \( -\ln(1-x) \) and \( (1 - \frac{x}{2})^{-1} \).

#### 3.2.4. Summation by convolution.

- **a)** \( \sum_{k \geq 0} (-1)^k \binom{k}{r} \binom{n}{r-k} \). The convolution is the coefficient of \( x^r \) in the product of \( \sum_{k \geq 0} (-1)^k \binom{k}{r} x^k \) and \( \sum_{k \geq 0} \binom{n}{k} x^k \). We compute \( [x^r](1-x)^n(1+x)^n = [x^r](1-x^2)^n \). This is \( (-1)^{r/2} \binom{n}{r/2} \) when \( r \) is even and is 0 when \( r \) is odd.

- **b)** \( \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} 2^{n-k} = n \). We can add the term \( k = 0 \). Differentiating the Binomial Theorem yields \( x n (1 + x)^{n-1} = \sum_{k=0}^{n} \binom{n}{k} x^k \). Setting \( x = -1/2 \) and multiplying the sum by \( -2^n \) produces the desired sum, and the formula on the other side then simplifies to \( n \).

The sum can also be evaluated using ordinary convolution and OGFs. There is a short proof using the concepts of exponential generating functions and binomial convolution developed in Section 3.3.

#### 3.2.5. More summations.

- **a)** \( \sum_{k=0}^{2n} \binom{n}{k} = 2^{2n-2} \) if \( n > 1 \); otherwise 0.

  **Proof** 1 (generating functions). Let \( f(x) = \sum_{k \geq 0} \binom{n}{k} x^k \). We seek the sum of the even coefficients in \( f(x) \), which equals \( \frac{1}{2} (f(x) + f(-x)) \) evaluated at \( x = 1 \). From \( (1 + x)^n = \sum_{k \geq 0} \binom{n}{k} x^k \), we obtain \( f(x) = x \frac{d}{dx} (1 + x)^n \) = \( nx(1 + x)^{n-1} \). To compute the desired value,

\[
\frac{1}{2} (f(x) + f(-x)) \bigg|_{x=1} = \frac{1}{2} \left[ 2n^{n-1} - n \cdot 0^{n-1} \right].
\]

  **Proof** 2 (combinatorial argument). The sum counts the even-sized chaired committees from a set of \( n \) people. For \( n \geq 1 \), we can form all such committees by picking the chair first and then picking an odd-sized subset of the remaining people to complete the committee. If \( n - 1 \) is positive, then half the subsets of the remaining people have odd size, yielding \( 2^{n-2} \) as the answer. For \( n < 2 \), there is no such committee.

- **b)** \( \sum_{k=0}^{n} \binom{n}{k}^2 = \frac{1}{2} (2^n) \).

  **Proof** 1 (algebraic manipulation). Since \( \binom{n}{k} = \binom{n-k}{k} \), this sum is the coefficient of \( x^n \) in the product of \( \sum_{k \geq 0} \binom{n-k}{k} x^k \) and \( \sum_{k \geq 0} \binom{n}{k} x^k \). These sums are the generating functions \( x n (1 + x)^{n-1} \) and \( (1 + x)^n \). We compute

\[
[x^n] x n (1 + x)^{n-1} = nx^{n-1}(1 + x)^{2n-1} = n \left( \frac{2n-1}{n-1} \right) = n \left( \frac{2n}{n} \right).
\]

  **Proof** 2 (counting two ways). Taking the product of the numbers of ways to do the first portion and the last portion, \( \binom{n}{k}^2 \) counts the lattice paths from \( (0, 0) \) to \( (n, n) \) that visit \( k, n - k \). Multiplying by \( k \) corresponds to marking one horizontal step among the first \( n \) steps. Summing over \( k \) counts the paths from \( (0, 0) \) to \( (n, n) \) with a horizontal step marked before the middle.
On the other hand, \( \binom{2n}{n} \) counts marked paths with any horizontal step marked. Exactly half of these objects have the mark in the first half of the path, because we can pair such marked paths with marked paths having the horizontal mark in the last half by rotating the picture by 180° around the point \((n/2, n/2)\).

**Alternative combinatorial argument.** Since \( \binom{n}{k} = \binom{n}{n-k} \), the sum counts the chaired committees from \([2n]\) such that the committee has size \(n\) and the chair is from \([n]\), grouped by how many committee members, denoted by \(k\), belong to \([n]\). To form such committees, we can select the chair first, in \(n\) ways, and then select the remaining \(n-1\) members from the remaining \(2n-1\) people.

**3.2.6. A formal power series** \( \sum_{n \geq 0} a_n x^n \) **has a multiplicative inverse if and only if** \(a_0 \neq 0\). A multiplicative inverse is a formal power series \( \sum_{n \geq 0} b_n x^n \) such that \((\sum_{n \geq 0} a_n x^n)(\sum_{n \geq 0} b_n x^n) = 1\). The coefficient \(c_n \) of \(x^n\) in the product is \(\sum_{k=0}^{n} a_k b_{n-k}\). If \(a_0 = 0\), then \(a_0 b_0 = 0\), so \(c_0\) is always 0, and the product cannot equal 1. If \(a_0 \neq 0\), then an inverse requires \(b_0 = 1/a_0\).

We continue with the existence and uniqueness of the formula for the inverse, computing \(b_n\) in terms of earlier values, for \(n > 1\). With \(a_0, \ldots, a_n\) and \(b_0, \ldots, b_{n-1}\) already known, \(\sum_{k=0}^{n} a_k b_{n-k} = c_k = 0\) requires \(b_n = -(\sum_{k=1}^{n} a_k b_{n-k})/a_0\).

**3.2.7. Binomial inversion formulae:** The following four statements (each over all \(k \in \mathbb{N}\)) are equivalent.

\[
\begin{align*}
(a) \quad b_k &= \sum_{i=0}^{k} \binom{n}{i} a_{k-i} \\
(b) \quad b_k &= \sum_{i=0}^{k} \binom{n}{i} a_i \\
(c) \quad a_k &= \sum_{i=0}^{k} (-1)^i \binom{n}{i} b_{k-i} \\
(d) \quad a_k &= \sum_{i=0}^{k} (-1)^i \binom{n}{k-i} b_i
\end{align*}
\]

\((a)\) and \((b)\) are equivalent by substituting \(k - i\) for \(i\) in the sum. The same is true for the equivalence of \((c)\) and \((d)\). For the equivalence of \((a)\) and \((c)\), let \(A(x)\) and \(B(x)\) be the OGFs for \((a)\) and \((c)\). Statement \((a)\) is \(B(x) = (1 + x)^n A(x)\). Statement \((b)\) is \(A(x) = (1 + x)^{-n} B(x)\). These are clearly equivalent.

**3.2.8. Binomial inversion formulae:** The following four statements (each over all \(k \in \mathbb{N}\)) are equivalent.

\[
\begin{align*}
(a) \quad b_k &= \sum_{i=0}^{k} \binom{n+i}{i} a_{k-i} \\
(b) \quad b_k &= \sum_{i=0}^{k} \binom{k+i}{n} a_i \\
(c) \quad a_k &= \sum_{i=0}^{k} (-1)^i \binom{n+i}{i} b_{k-i} \\
(d) \quad a_k &= \sum_{i=0}^{k} (-1)^i \binom{k+i}{k-i} b_i
\end{align*}
\]

\((a)\) and \((b)\) are equivalent by substituting \(k - i\) for \(i\) in the sum. The same is true for the equivalence of \((c)\) and \((d)\). For the equivalence of \((a)\) and \((c)\), let \(A(x)\) and \(B(x)\) be the OGFs for \((a)\) and \((c)\). Statement \((a)\) is \(B(x) = (1 - x)^{-n+1} A(x)\). Statement \((b)\) is \(A(x) = (1 - x)^{n+1} B(x)\). These are clearly equivalent.

**3.2.9.** When \((a)\) and \((b)\) with respective generating functions \(f(x)\) and \(g(x)\) are related by \(b_n = \sum_{k \geq 0} a_k \delta_k\) for \(n \geq 0\), the generating functions are related by \(g(x) = f(1/x - 1/f(1))\). We compute

\[
(1 - x)g(x) = (1 - x) \sum_{n \geq 0} \sum_{k \geq n} a_k x^n = f(1 - a_0 + \sum_{n \geq 1} \sum_{k \geq n} a_k - \sum_{k \geq n-1} a_k) x^n = f(1) - \sum_{n \geq 0} a_n x^n = f(1) - f(x).
\]

**3.2.10. Restricted multisets.** Let \(n\) be even.

- **a)** The OGF by size for multisets with odd numbers restricted to odd multiplicity and even numbers restricted to even multiplicity is \(x^n/(1 - x^n)\). The factor for unrestricted multiplicity is \((1 - x)^{-1}\). To keep only the odd terms, the factor is \(\frac{1}{2}[(1 - x)^{-1} - (1 + x)^{-1}]\). To keep only the even terms, the factor is \(\frac{1}{2}[(1 - x)^{-1} + (1 + x)^{-1}]\). The expressions can be simplified to get \(x/(1 - x^2)\) for odd multiplicity and \(1/(1 - x^2)\) for even multiplicity, which can be argued directly since the members of the multiset get added in pairs. We have \(n/2\) factors of each type.

- **b)** The number of such multisets of size \(k\) is \((k + 2n\mathcal{A} - 4n)/4\). Expanding the generating function in part \(a)\) yields \(\sum_{j \geq 0} (j + n - 1)x^{j+n-2}/j!)\). To obtain the coefficient of \(k,\) set \(k = n/2 + 2j\). Note that the parity of \(k\) must agree with the parity of \(i/2\); otherwise there is no such multiset. Setting \(j = k/2 - n/4\) yields \((k^2 + 3n^2)/4\). This checks with the value \((k + 1)/2\) when \(n = 2\); multisets are then determined by the number of 2s, which may be any even number from 0 through \(k - 1\).

**3.2.11.** If \(\frac{3 - 3x}{1 - 3x}\) is the generating function for the sequence \((a)\), then \(\sum_{n \geq 0} a_n = 2^{n+1} + (-1)^n\). Let \(b_n = \sum_{k \geq 0} a_k\). If \(f(x)\) is the generating function for \((a)\), then \(f(x)/(1 - x)\) is the generating function for \((b)\). Thus we expand \(3(1 - x - 2x^2)\) as a power series in \(x\). The denominator factors as \((1 + x)(1 - 2x)\), so the generating function equals \(\frac{A}{1 + x} + \frac{B}{1 - 2x}\). Clearing fractions yields \(3 = A(1 - 2x) + B(1 + x)\). At \(x = -1\) and \(x = 1/2\), we obtain \(3 = 3A + 3 = 2B\), so \(A = 1\) and \(B = 2\). The conclusion, using the geometric series, is \(b_n = 2^{n+1} + (-1)^n\).

**3.2.12.** \(\sum_{n \geq 0} (-1)^n \binom{n}{k} = \binom{3}{k}\), by generating functions. We show that each side is the coefficient of \(x^k\) in the same generating function. The right side is the coefficient of \(x^k\) in \((1 + x)^n\).

The left side is a convolution. We claim it is the coefficient of \(x^k\) in
the product of \((1-x^2)^n(1-x)^{-n}\), which also equals \((1+x)^n\). The coefficient of \(x^i\) in \((1-x^2)^n\) is \((-1)^i\binom{n}{i}\), but the coefficient of \(x^{2i+1}\) is 0. The coefficient of \(x^j\) in \((1-x)^{-n}\) is \(\binom{-n}{j}\). Thus we get nonzero contributions to the convolution for \(a_kb_{n-j}\) only when \(j\) is even. Setting \(j = 2t\), the contribution is \((-1)^t\binom{n}{2t+2n-1}\).

### 3.2.13. Alternative computation of \(\sum_{k=1}^{n} k^2\).

- **a)** \(k^2 = 2\binom{k+2}{3} - 3\binom{k+1}{2} + 1\). We compute \(k^2 + (k+1) - 3(k+1) + 1 = (k-1)^2 + 1 = k^2\).
- **b)** \(\sum_{k=1}^{n} k^2 = \binom{n+1}{3}2(n+1)\). Note that \(\sum_{k\geq0} (k^{n+1})x^k = (1-x)^{-n}\). Part (a) thus yields \(\sum_{k\geq0} k^2x^k = 2(1-x)^{-3} - 3(1-x)^{-2} + 1(1-x)^{-1}\). To find \(\sum_{k=0}^{n} k^2\), we multiply by \((1-x)^{-1}\) and take the coefficient of \(x^n\). From the expansion for \((1-x)^{-1}\),

\[
\sum_{k=0}^{n} k^2 = \left[ x^n \right] \left( \frac{1}{(1-x)^3} - \frac{3}{(1-x)^2} + \frac{1}{1-x} \right) = 2\binom{n+3}{3} - 3\binom{n+2}{2} + 1\binom{n+1}{1} = \frac{(n+1)n(2n+1)}{6}.
\]

### 3.2.14. \([x^n](1-x)^{-1/2} = \begin{cases} \binom{n}{n/2}2^{-n} & \text{if } n \text{ even} \\ 0 & \text{if } n \text{ odd} \end{cases}\)

Always \([x^n] f(x^2) = [x^{n/2}] f(x)\) when \(n\) is even, and \([x^n] f(x^2) = 0\) when \(n\) is odd, which completes the odd case. For the even case, we compute \([x^{n/2}] A(x)\), where \(A(x) = (1-x)^{-1/2}\). By the Extended Binomial Theorem (Theorem 1.2.5), we compute

\[
[x^m](1-x)^{-1/2} = \binom{-1/2}{m}\binom{m}{0} = \frac{1}{m!}\frac{1}{2}...\frac{m-1}{2} = \frac{(2m)!}{m!2^m m!} = \frac{2m}{m} = 2^{2m}.
\]

Note: We previously computed the expansion \((1-4y)^{-1/2} = \sum_{m\geq0} \binom{2m}{m} y^m\).

### 3.2.15. \(\sum_{k=1}^{n} \binom{2k}{k} \binom{2n-2k}{n-k} = C_n + 1\)

#### Proof 1 (generating functions).

The sum is \(\sum_{k=0}^{n} C_k C_{n-k}\), where \(C_n\) is the \(n\)th Catalan number. Let \(C(x) = (1-\sqrt{1-4x})/2x\). Since \(C(x)\) is the OGF for the Catalan numbers, the desired sum is \([x^n] (C(x))^2\). By direct computation,

\[
\left( \frac{1-\sqrt{1-4x}}{2x} \right)^2 = \frac{1}{x} + \frac{1}{x} - \frac{2\sqrt{1-4x}}{2x} = -1 + \frac{C(x)}{x}.
\]

Since \(C(0) = 1\), the resulting coefficient of \(x^{-1}\) is 0, as it must be. For \(n \geq 0\), we have \([x^n] (C(x))^2 = [x^{n+1}] C(x) = C_{n+1}\).

### 3.2.16. Fibonacci numbers of fixed parity.

Let \(A(x) = \sum_{n\geq0} F_n x^n\), where \(\langle F\rangle\) is the adjusted Fibonacci sequence. Recall that \(A(x) = 1/(1-x-x^2)\). If \(A(x) = \sum_{n\geq0} a_n x^n\) and \(b_n = 2a_{2n}\), then \(\sum_{n\geq0} b_n y^n = \frac{1}{2} [A(y^{1/2}) + A(-y^{1/2})]\), since \(\frac{1}{2} (A(x) + A(-x))\) cancels the odd terms, and then setting \(y = x^2\) divides the exponents by 2. Similarly, \(\sum_{n\geq0} c_n y^n = \frac{1}{2} [A(y^{1/2}) - A(-y^{1/2})]\) when \(c_n = a_{2n+1}\); we compute \(\frac{1}{2} [A(x) - A(-x)]\) to cancel the even terms, then divide by \(x\), then set \(y = x^2\).

With \(A(x) = \frac{1}{1-x-x^2}\), for (b) we obtain \(B(x^2) = \frac{1-x^2}{1-3x^2+x^4}\) and then \(B(y) = \frac{1-y}{1-3y^2+y^4}\). For (c), we obtain \(C(x^2) = \frac{1}{1-3x^2+x^4}\) and then \(C(y) = \frac{1}{1-3y^2+y^4}\).

The denominators in the rational generating function give the same recurrence \(b_n = 3b_{n-1} - b_{n-2}\) and \(c_n = 3c_{n-1} - c_{n-2}\) for \(n \geq 2\), but the initial conditions are different: \(b_0 = \hat{F}_0 = 1\), \(b_1 = \hat{F}_2 = 2\), \(c_0 = F_1 = 1\), \(c_1 = F_3 = 3\).

### 3.2.17. \(\sum_{n\geq0} a_n x^n = \frac{1}{2} (1-4x)^{-1/2}\), where \(a_n = \binom{2n-1}{n}\).

Using the Complementation and Committee-Chair Identities, \(\binom{2n-1}{n} = \binom{2n-1}{n+1}\). Thus this OGF is half the OGF \((1-4x)^{-1/2}\) for the central binomial coefficients (Example 3.2.13).

### 3.2.18. Convolutions versus combinatorial arguments.

- **a)** \(\sum_{k=0}^{n} \binom{n+k-j-1}{k-j} \binom{m+j-1}{j} = \binom{k+m+n-1}{k}\). The convolution is the coefficient of \(x^k\) in the product of the generating functions \(\sum_{r\geq0} \binom{r+m-1}{m-1} x^r\) and
$\sum_{r \geq 0} \binom{r+n-1}{n-1} x^r$, which are $(1-x)^{-m}$ and $(1-x)^{-n}$. Hence the value is $[x^k](1-x)^{-m+n}$, which equals $\binom{k+m+n-1}{m+n-1}$.

Both sides count multisets of size $k$ from $m+n$ types. Every such selection chooses $j$ elements from the first $m$ types and $k-j$ elements from the remainder, for some $j$. Summing over $j$ completes the proof.

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

**Proof 1** (generating functions). The convolution is the coefficient of $x^n$ in the product of the generating functions $\sum \frac{1}{k+1}(\frac{2k}{k}) x^k$ and $\sum (\frac{2k}{k}) x^k$. The first is the Catalan generating function $(1 - \sqrt{1 - 4x})/2x$. The second is $1/\sqrt{1 - 4x}$ (Example 3.2.13). The product is $1/(1 - 4x)^{-1/2} - 1$. The $-1$ cancels the constant term in $(1 - 4x)^{-1/2}$, so the coefficient of $x^1$ is 0, as desired. For $n \geq 0$, we seek $[x^n+1] \frac{1}{2}(1 - 4x)^{-1/2}$, which is $\frac{1}{2}(-1/2)(-4)^{n+1}$.

By the Extended Binomial Formula, $\binom{-1/2}{n+1} = \frac{(-1/2)(-3/2)\cdots(-n-1/2)}{n!} \binom{n+1}{n}$. Hence $\frac{1}{2}(-1/2)(-4)^{n+1} = \frac{2^{n+1} (-1)^n}{2(n+1)n!} \frac{n!}{(2i+1)!} = \binom{2n+1}{n+1}$.

**Proof 2** (combinatorial proof) The value $\binom{2n+1}{n+1}$ is the number of standard lattice walks from (0, 0) to $(n, n)$. Each rises above the line $y = x$ for the first time at some point $(k, k+1)$. Before that point, it follows a ballot path of length $2k$. After the vertical step to $(k, n+1)$, it may follow any lattice path of length $2n-2k$ with the same number of horizontal and vertical steps. By the rules of sum and product, each such path is counted once in $\sum_{k=0}^{n} \frac{1}{k+1}(\frac{2k}{k})^{n-k}$.

### 3.2.19. The number of monotone lattice paths with endpoints in $[n] \times [n]$ is $\binom{2n+1}{n+1}$. To determine the endpoints, it suffices to choose a horizontal interval $I$ and a vertical interval $J$ for the coordinates. The path then extends from the lower left corner of the rectangle $I \times J$ to its upper right corner. The number of lattice paths with a given pair of endpoints depends only on the horizontal and vertical distances between the endpoints. For $0 \leq i \leq n-1$, there are $n-i$ ways to choose $I$ with length $i$; similarly $n-j$ ways to choose $J$ with length $j$ for $0 \leq j \leq n-1$. Having chosen such intervals, the number of paths with the specified endpoints is $\binom{i+j}{i}$. Extending the sums to $n$ without change, the total number of paths is $\sum_{i=0}^{n} \sum_{j=0}^{n} \binom{n-i}{i} \binom{n-j}{j} = \sum_{i=0}^{n} \sum_{j=0}^{n} \binom{n-i}{i} \binom{n-j}{j}$.

To simplify, we first evaluate the inner sum. For fixed $i$ and $n$, the sum $\sum_{j=0}^{n} \binom{n-j}{j} = \binom{n}{i}$. The sum now simplifies to $\binom{n}{i} \frac{1}{(1-x)^{i+1}} \sum_{i=0}^{n} \frac{1}{(1-x)^{i+1}} = \binom{n}{i} \frac{(i+1+n)}{(i+2)}$. The sum now simplifies to $\sum_{i=0}^{n} \binom{n}{i} \binom{n+1}{n-1}$, which is the coefficient of $x^n$ in the product of $\sum_{i=0}^{n} \binom{n-1}{i} x^i$ and $\sum_{k=0}^{n} k x^k$. For the first of these generating functions, we shift indices to compute $\sum_{i=0}^{n} \binom{n-1}{i} x^{i+1} = \frac{1}{(1-x)^n} - 1 - nx$.

With $\sum_{k=0}^{n} k x^k = x/(1-x)^2$ as before, we compute $\sum_{i=0}^{n} \sum_{j=0}^{n} \binom{n-i}{i} \binom{n-j}{j} = \frac{1}{(1-x)^n} - 1 - nx$.

**3.2.20.** $\sum_{n=0}^{\infty} \binom{2n+k}{n} x^n = \frac{C(x)^k}{\sqrt{1-4x}}$. The coefficient of $x^n$ is the number of lattice paths from the origin to $(n, n+k)$. To build such a path, we must visit each line $y = x + i$, for $1 \leq i \leq k$. Before visiting $y = x + 1$ for the first time, we follow a ballot path and then a vertical step. From the first visit to $y = x + i$ to the first visit to $y = x + i + 1$, we follow a ballot path and then a vertical step. After we first reach $y = x + k$, we follow any lattice path that has an equal number of steps up and right.

Hence each path being counted is a concatenation of $k$ ballot paths, separated each from the next by one vertical step, and an equal up/right lattice path at the end. The index is the number of horizontal steps. The factor for the choice of each ballot path is $C(x)$, the factor for each special vertical step is 1 (no horizontal contribution), and the factor for the path chosen at the end is $\sum \binom{2n}{n} x^n$, which equals $(1-4x)^{-1/2}$ (Example 3.2.13).

**Comment:** An alternative proof by induction on $k$ is really the same idea. For $k = 0$, we have $\sum \binom{2n}{n} x^n = (1-4x)^{-1/2}$. For $k > 0$, we concatenate a ballot path with $r$ horizontal steps, a step up (reach $y = x + 1$ for the first time), and a path with $n-r$ horizontal steps from $(r, r+1)$ to $(n, n+k)$. The second factor in this convolution is given by the induction hypothesis, so the desired generating function is $C(x)C(x)^{k-1}/\sqrt{1-4x}$.

### 3.2.21. The identity $\sum \binom{\frac{n-i}{k+1}}{k+1} = \frac{2^{n+1}}{k+1}$, where the sum is over $k, l \in \mathbb{N}$ with $k + l = n$.

a) $B(x)B'(x) = \frac{1}{2}B'(x) - 1$, where $B(x) = \sum_{k=0}^{n} C_k x^{k+1} = \sum_{k \geq 0} \binom{2k}{k} x^{k+1}$. Begin with the usual Catalan recurrence $C_n = \sum_{k=1}^{n-1} C_{k-1} C_{n-k}$ for $n \geq 1$,
with $C_0 = 1$. Multiplying by $x^{r+1}$ and summing over $n$ yields $B(x) - x = B(x)^2$, since $\sum_{k=1}^{n} C_{k-1} C_{n-k} x^{k+1} = \sum_{k=0}^{n-1} (C_{k} x^{k+1}) (C_{n-k-1} x^{-k})$. Differentiating both sides then yields $B'(x) - 1 = 2B(x)B'(x)$.

The summand on the left side of the identity is the product of the coefficients of $x^{k+1}$ in $B(x)$ and $x^{k+1}$ in $B'(x)$, summed over $(k+1) + (l+1) = n+2$, but lacking the term for $l = -1$, which would invoke the term for $x^0$ in $B'(x)$. Since the constant term in $B'(x)$ is 1, the sum is the coefficient of $x^{n+2}$ in $B(x)(B'(x) - 1)$. We compute

$$B(x)(B'(x) - 1) = \frac{1}{2} B'(x) - B(x) = \frac{1}{2} \sum_{k=1}^{\infty} \left( \frac{2k}{k} \right) x^k - \sum_{k=0}^{\infty} \left( \frac{2k}{k} \right) x^{k+1}.$$  

The coefficient of $x^{n+2}$ is $\frac{1}{2} \left( \frac{2n+4}{n+2} \right) - \frac{1}{n+2} \left( \frac{2n+2}{n+1} \right)$. We compute

$$\frac{1}{2} \left( \frac{2n+4}{n+2} \right) - \frac{1}{n+2} \left( \frac{2n+2}{n+1} \right) = \frac{2n+3}{n+1} \left( \frac{2n+2}{n} \right) - \frac{1}{n+1} \left( \frac{2n+2}{n} \right) = 2 \left( \frac{2n+2}{n} \right).$$

b) Combinatorial proof. The number of lattice paths from $(0, 0)$ to $(n, n)$ not rising above the line $y = x$ is the Catalan number $C_n$, equal to $\frac{1}{n+1} \binom{2n}{n}$.

The total number of lattice paths from $(0, -1)$ to $(n+1, n+1)$ is $\binom{2n+3}{n+1}$. Among these, the number that first meet the line $y = x$ at the point $(k, k)$ is $\frac{1}{k+1} \left( \frac{2k}{n+1-k} \right)$, since the initial portion does not rise above the line $y = x-1$. Therefore, the left side of the identity counts all paths from $(0, -1)$ to $(n+1, n+1)$ except those that first meet the line $y = x$ at $(n+1, n+1)$, since the term for $k = n+1$ is missing from the sum. Thus the value of the sum is $(\binom{2n+3}{n+1} - C_n + 1$. Since $C_{n+1} = \frac{1}{n+2} \binom{2n+2}{n+1} = \frac{(2n+2)}{n+1} \binom{2n}{n}$, to evaluate the sum we compute

$$\left( \frac{2n+3}{n+1} \right) - C_n = \left( \left( \frac{2n+2}{n+1} + \frac{2n+2}{n} \right) \right) - \left( \left( \frac{2n+2}{n+1} + \frac{2n+2}{n} \right) \right) = 2 \left( \frac{2n+2}{n} \right).$$

3.2.22. $\sum_{k=0}^{n-1} 4^{n-k} \frac{1}{1+x} \binom{2k}{k} = 2 \left[ 4^n - \binom{2n}{n} \right]$. The convolution is the coefficient of $x^{n-1}$ in the product of $\sum_{k=0}^{1} \frac{1}{2} \binom{2k}{k} x^k$ and $\sum_{k=0}^{\infty} 4^{k+1} x^k$. The first is the Catalan generating function $(1 - \sqrt{1-4x})/(2x)$; the second is $4(1-4x)^{-1}$. We compute

$$[x^{n-1}] \frac{1}{2x} \left( \frac{1}{1-4x} - \frac{1}{\sqrt{1-4x}} \right) = 2 [x^n] \left( \frac{1}{1-4x} - \frac{1}{\sqrt{1-4x}} \right).$$

Since $[x^n] \frac{1}{\sqrt{1-4x}} = \binom{2n}{n}$, the value of the sum is $2 \left[ 4^n - \binom{2n}{n} \right]$.  

3.2.23. Evaluation of $\sum_{k=0}^{n} (-1)^k \binom{n-k}{k} x^k$. Let $A_n(x)$ be the specified sum. Using Pascal’s Formula, we obtain the recurrence $A_n = A_{n-1} - x A_{n-2}$, with $A_0(x) = A_1(x) = 1$. The characteristic equation is $x^2 - x + x = 0$, with roots $(1 \pm \sqrt{1-4x})/2$ (the generating function method can also be used). Using the initial conditions $A_0(x) = A_1(x) = 1$, we obtain

$$A_n(x) = \frac{1}{\sqrt{1-4x}} \left[ \left( \frac{1 + \sqrt{1-4x}}{2} \right)^{n+1} - \left( \frac{1 - \sqrt{1-4x}}{2} \right)^{n+1} \right].$$

Setting $x = -1$ in $A_n(x)$ yields $\sum_{k=0}^{n} (-k) \binom{n-k}{k} x^k = \binom{n}{n} (\frac{1}{\sqrt{3}})^{n+1} - \binom{n}{n} (\frac{-1}{\sqrt{3}})^{n+1}$. This final expression is the value of $\hat{P}_n$ found in Example 2.2.4.

3.2.24. $\sum_{i=0}^{n} \binom{n}{i} (2i-1)! (2n-2i-1)! = 2^n n!$. Since $(2k-1)! = (2k)!/(2^{k-1})$, we compute

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

3.2.25. Given $P(m, n, r) = \sum_{k=0}^{m} (-1)^k \binom{m-k}{k} \binom{n+k}{n}$, if $0 \leq r \leq n \leq m$ and $n > (m+1)/2$, then $P(m, n, r) > 0$ and $\sum_{r=0}^{m} P(m, n, r) = \binom{m+2}{n+2}$. Let $F_{n,r}(x) = \sum_{m \geq 0} P(m, n, r) x^m$. Note that nonzero contributions for a given $k$ require $m \geq n + 2k$, so we can start the sum over $m$ there. Since $\sum_{j=0}^{n} \binom{n-j}{n} x^j = (1-x)^{-(n+1)}$, we have

$$F_{n,r}(x) = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \sum_{m \geq n + 2k} \binom{n + m - (n + 2k)}{n} x^{m-(n+2k)}$$

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

With $(1-x^2) = (1-x)(1+x)$ and $r \leq n$, we have

$$P(m, n, r) = [x^{m-n}] (1+x)^r (1-x)^{-(n-r+1)} = \sum_{k=0}^{m-n} \binom{r}{k} \frac{x^{m-r-k}}{n-r}.$$  

With $m \geq n \geq r$, the summands and $P(m, n, r)$ are positive. Returning to the expression for $F_{n,r}(x)$,
\[
\sum_{r=0}^{n} P(m, n, r) = \frac{x^m}{1 - x} \frac{1}{(1-x)^{n+1}} \sum_{r=0}^{n} (1-x^2)^r = \frac{x^m}{1 - x} \frac{1}{(1-x)^{n+1}} \frac{1 - (1-x^2)^{n+1}}{x^2}
\]
\[
= \frac{x^{m-n+2}}{1-x} \frac{1 - (1-x)^{(n+1)}}{x^2} - \frac{x^{m-2n+2}}{1-x} \frac{1 + x^{n+1}}{x^2}
\]
\[
= \left( m - n + 2 + n \right) \frac{1}{n} + 0 = \left( m + 2 \right).
\]

The contribution from the second term is 0 since \( m - n + 2 < n + 1 \).

**3.2.26.** \([x^m y^n](1 - x - y + 2xy)^{-1} = \sum_{k \geq 0} (-1)^k \binom{m}{k} \binom{j}{k} \). Let \( a_{m,n} = [x^m y^n](1 - x - y + 2xy)^{-1} \). Consider the formal power series expansion

\[
\frac{1}{1 - x - y + 2xy} = \frac{1}{(1-x)(1-y) 1 + \frac{2x}{(1-x)(1-y)}} = \sum_{k=0}^{\infty} \frac{(-1)^k x^k y^k}{(1-x)^{k+1}(1-y)^{k+1}}
\]
\[
= \sum_{i,j,k=0}^{\infty} (-1)^k \binom{k+i}{k} \binom{k+j}{k} x^i y^j = \sum_{m,n} \sum_{k=0}^{\infty} (-1)^k \binom{m}{k} \binom{n}{k}.
\]

Thus the coefficient is \( \sum_{k \geq 0} (-1)^k \binom{m}{k} \binom{j}{k} \). Writing \( \binom{m}{k} \) as \( \binom{m-k}{k} \), the convolution is the coefficient of \( x^n \) in \((1+x)^m(1-x)^n\). With \( m = 2j \) and \( n = 2j + 2 \), this generating function becomes \((1-x^2)^{2j}(1-x)^2\), and we compute

\[a_{2j,2j+2} = [x^{2j}] \left( \binom{2j}{j} + \binom{2j}{j} \right) = \binom{2j}{j} \frac{x^{2j}}{2^{2j}}.\]

**3.2.27.** Let \( f(n) \) be the number of binary words of \( n \) length in which the numbers of occurrences of consecutive 00 and consecutive 01 are the same. Show that

\[
\sum_{n=0}^{\infty} f(n) t^n = \frac{1}{2} \left( \frac{1}{1-t} + \frac{1+2t}{\sqrt{1-2t}(1+2t+2t^2)} \right). \quad \text{(Stanley [2011])}
\]

Let \( a_{k,n} \) count the \( n \)-words with \( k \) occurrences of 00 and \( n-1 \) that end with 0, and let \( b_{k,n} \) count those that end with 1. With \( k \) occurrences of 01, words counted by \( a_{k,n} \) have \( k + 1 \) runs of 0s. They arise from the alternating word 010⋯10 of length \( 2k + 1 \) by \( k \) insertions of 0 at the beginnings of 0-runs and \( n - 3k - 1 \) insertions of 1 before the beginnings of 0-runs. Since the number of available locations is \( k + 1 \) and the resulting words are determined by how many insertions go into each location, the formula for multisets from \( k + 1 \) types yields \( a_{k,n} = \binom{2k}{k} \binom{n-2k}{k} \) (with \( \binom{k}{k} = 0 \) when \( u < k \)). Thus

\[
\sum_{k=0}^{\infty} \sum_{m=0}^{\infty} a_{k,m+3k+1} t^{m+3k+1} = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{2k}{k} \binom{m+k}{k} t^{m+3k+1} = \sum_{k=0}^{\infty} \frac{2k}{k} \frac{t^{2k+1}}{(1-t)^{k+1}}.
\]

Similarly, a word counted by \( b_{k,n} \) arises from the alternating word 010⋯01 of length \( 2k \) by \( k \) insertions of 0 and then \( n - 3k \) insertions of 1, again at the starts of 0-runs. Hence \( b_{n,k} = \frac{(2k-1)}{k} t^{n-2k} \) for \( k > 0 \), so

\[
\sum_{n=1}^{\infty} b_{n,k} t^n = \sum_{k=1}^{\infty} \sum_{m=0}^{\infty} b_{k,m+3k} t^{m+3k} = \sum_{k=1}^{\infty} \sum_{m=0}^{\infty} \frac{2k-1}{k} \binom{m+k}{k} t^{m+3k} = \sum_{k=1}^{\infty} \frac{2k-1}{k} \frac{t^{3k}}{(1-t)^{k+1}}.
\]

Since \( b_{n,k} = 1 \) for \( n > 0 \) and \( f(0) = 1 \), the desired generating function is obtained by summing the two displayed equations and \( 1/(1-t) \). Using \( \sum_{k=0}^{\infty} = \frac{2k}{k} \frac{t^{2k}}{(1-t)^{k+1}} \) with \( x = t^2/(1-t) \), we have

\[
\sum_{n=0}^{\infty} f(n) t^n = \frac{1}{1-t} \left[ 1 + \frac{t}{\sqrt{1-4t^2/(1-t)}} + \frac{1}{2} \left( \frac{1}{\sqrt{1-4t^2/(1-t)}} - 1 \right) \right],
\]

which equals the desired expression.

**3.2.28.** A gambler who has probability \( p \) of winning any single game, independently, insists on playing until he is ahead by one game.

a) The probability \( \alpha \) that the match ends is \( 1 \) if \( p \geq 1/2 \), less than \( 1 \) if \( p < 1/2 \). The probability \( \alpha \) is the sum, over all game sequences that end the match, of the probability that that sequence occurs. A sequence of length \( 2n+1 \) that ends the match occurs with probability \( p^{n+1}(1-p) \). Such a sequence puts the gambler ahead only at the end; before that it is a ballot list for the opponent. Hence the number of winning sequences of length \( 2n+1 \) is the Catalan number \( C_n \). Using the Catalan generating function, \( C(x) = (1 - \sqrt{1 - 4x})/(2x) \), we have

\[
\alpha = \sum_{n \geq 0} p^{n+1}(1-p)^n C_n = p \sum_{n \geq 0} C_n (p(1-p))^n = p \frac{1 - \sqrt{1 - 4p(1-p)}}{2p(1-p)}.
\]

Setting \( p = 1/2 \) yields \( \alpha = 1 \). When \( p < 1/2 \), we have \( \alpha = \frac{1 - \sqrt{1 - 4p(1-p)}}{2p(1-p)} < \frac{1}{2p(1-p)} < 1 \). When \( p > 1/2 \), the gambler cannot become less likely to win that at \( p = 1/2 \).

b) Expected number of games, by generating functions. In part (a), summing the probabilities of all the winning sequences yielded \( \alpha = \sum_{n \geq 0} C_n (p(1-p))^n \). Weighting the probability of each game sequence
by its length yields $p \sum_{(2 \leq j \leq n+1)} C_n(p(1-p)^n$ for the expectation. Thus the expectation should be $2pC'(x) + pC(x)$, evaluated at $x = p(1-p)$.

e) For $p > 1/2$, the expected number of matches in the match is $1/(2p-1)$.

Let $t$ be the expected number of games in the match. The first game ends the match with probability $p$. With probability $1-p$, the match continues. Furthermore, to win the overall match, the gambler must now twice win a match under the same conditions; for each the expected number of games is $t$. Using linearity of expectation (see Chapter 0 or Chapter 9), we have $t = 1 + (1-p)2t$, which simplifies to $t = 1/(2p-1)$. The expectation tends to 1 as $p \to 1$ and to $\infty$ as $p \to 1/2$.

3.2.29. For nonnegative integers $m$ and $n$, $\sum_{k=0}^{\infty} 2^k \binom{2m-k}{m+n} = 4^n - \sum_{j=1}^{n} \binom{2m+1}{m+j}$.

**Proof 1** (generating functions). Let $P(x) = \sum_{k=0}^{\infty} 2^k(1+x)^{2m-k}$. Since $(\binom{2m-k}{m+n}) = 0$ when $k > m$, the left side is $[x^{m+n}]P(x)$. Since

$$P(x) = \frac{(1 + x)^{2m+1}}{(1 + x)} \frac{(1 + x)^{m+1} - 1}{1 - x} = \frac{2^{m+1}(1 + x)^{m+1}}{1 - x} - \frac{(1 + x)^{2m+1}}{1 - x},$$

we extract the coefficient of $x^{m+n}$ by

$$[x^{m+n}]P(x) = 2^{m+1} \sum_{k=0}^{m} \binom{m}{k} - \sum_{k=0}^{m} \binom{2m+1}{k} = 2^{m+1} - \sum_{k=0}^{m+n} \binom{2m+1}{k}.$$  

Setting $j = k - m$ in the last sum yields $4^n - \sum_{j=1}^{n} \binom{2m+1}{m+j}$.

**Proof 2** (counting two ways). Both sides count the subsets of $[2m+1]$ with size at least $m+n+1$. For the right side, exactly half of the $2 \cdot 4^n$ subsets of $[2m+1]$ have size at least $m+n+1$. From these, we eliminate those with sizes from $m+1$ to $m+n$ by subtracting the given sum.

For the left side, such subsets are constructed by picking an element $z$ to be the $(m+n+1)$-th smallest element chosen, picking $m+n$ elements below $z$, and picking any subset of the elements above $z$. With $k = 2m + 1 - z$, the value of $k$ indexes the choices for $z$. Since there are $2m-k$ elements below $z$ from which $m+n$ are picked, and $k$ elements above $z$ from which any are picked, the left side counts the specified subsets.

3.2.30. Differentiation of formal power series.
the coefficient of odd powers being 0. Also \( \binom{m-2k+n-1}{n-1} \) is the coefficient of \( x^{m-2k} \) in \( (1-x)^n \). Thus the sum is \( [x^{m}](1-x)^n \), which equals \( [x^{m}](1+x)^n \).

3.2.34. Converting sums to integrals.

a) \( \int_0^1 x^r(1-x)^s dx = \frac{1}{r+s+1} \binom{r+s}{s}^{-1} \) for \( r, s \in \mathbb{N} \). We use induction on \( s \).

When \( s = 0 \), we have \( \int_0^1 x^r dx = \frac{1}{r+1} \) for all \( r \).

For \( s > 0 \), we use integration by parts and the induction hypothesis:

\[
\int_0^1 x^r(1-x)^s dx = \frac{x^{r+1}(1-x)^{s+1}}{r+1} \bigg|_0^1 + \int_0^1 x^{r+1} \frac{s}{r+1} (1-x)^{s-1} dx \\
= \frac{s}{r+1} \binom{r+s}{s}^{-1} = \frac{1}{r+s+1} \binom{r+s}{s}^{-1}.
\]

b) \( \sum_{k=0}^n (-1)^k \binom{n}{k} = \frac{1}{m+n} \binom{m+n}{n}^{-1} \) for \( n, m \in \mathbb{N} \). Using part (a), we compute

\[
\sum_{k=0}^n (-1)^k \binom{n}{k} \frac{1}{m+k} = \sum_{k=0}^n (-1)^k \binom{n}{k} \int_0^1 x^{m-1} x^k dx = \int_0^1 x^{m-1} \sum_{k=0}^n (-1)^k \binom{n}{k} x^k dx = \int_0^1 x^{m-1}(1-x)^n dx = \frac{1}{m+n} \binom{m+n}{n}^{-1}.
\]

c) \( \sum_{k=1}^n (-1)^k \binom{n}{k} = -H_n \) for \( n \in \mathbb{N} \), where \( H_n = \sum_{i=1}^n \frac{1}{i} \). We use induction on \( n \). For \( n = 1 \), the only term in the sum is \(-1\), which equals \(-H_1\). For \( n > 1 \), we use the induction hypothesis (without part (a)) to compute

\[
\sum_{k=1}^n (-1)^k \binom{n-1}{k} = \sum_{k=1}^{n-1} (-1)^k \binom{n-1}{k} \frac{1}{k} + \sum_{k=1}^{n-1} (-1)^k \binom{n-1}{k-1} \frac{1}{k} \\
= -H_{n-1} - \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{1}{k+1} = -H_{n-1} - \int_0^1 \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} x^k dx \\
= -H_{n-1} - \int_0^1 x^{n-1} dx = -H_{n-1} + \frac{1}{n} = -H_n
\]

3.2.35. If \( b, r, s, t \) are nonnegative integers such that \( r + s + t = r b > 0 \), then

\[
\sum_{k=0}^{|s|/b} (-1)^k \binom{|s|/b}{k} \left( \frac{s + r - 1 - bk}{b} \right) = \sum_{k=0}^{|t|/b} (-1)^k \binom{|t|/b}{k} \left( \frac{t + r - 1 - bk}{b} \right).
\]

Proof 1 (generating functions). If we divide both sides by \( b^r \), then the left side gives the probability of rolling the sum \( r + s \), and the right side gives the probability of rolling the sum \( r + t \), with \( r \) rolls of a \( b \)-sided die, as verified below. By symmetry, the probability of rolling \( r + s \) is the same as the probability of rolling \( r b - s = r + t \). For example, with two 6-sided dice, the probability is symmetric around 7.

For the probability computation, let \( X_i \) be the outcome of the \( i \)th roll, and let \( Z = \sum_{i=1}^r X_i \). The probability generating function for \( X_i \) is \( \phi(x) = b^{-1} \sum_{k=0}^b x^k = b^{-1}(1-x^b)/(1-x) \). Hence the generating function for \( Z \) is \( \psi(x) = b^{-1} \sum_{i=0}^r (-1)^i \binom{r}{i} x^i b^k \sum_{n=0}^\infty \binom{n+r-1}{r} x^n \). The probability of rolling \( r + s \) is the coefficient of \( x^{r+s} \) in \( \psi(x) \). Summing over all \( k \) with \( bk + n = s \) in the convolution of the two sums completes the proof.

Proof 2 (inclusion-exclusion; see Chapter 4). By inclusion-exclusion, the number of integer solutions to the equation \( \sum_{i=1}^r x_i = s \) that have \( 0 \leq x_i < b \) is given by the formula on the left side. By the transformation \( y_i = b - 1 - x_i \), we obtain a one-to-one correspondence between these and the integer solutions to the equation \( \sum_{i=1}^r y_i = rb - r - s = t \) that have \( 0 \leq y_i < b \), but the number of these is given by the right side.

3.2.36. \( \sum_{k \geq 0} \binom{k}{n-k} 2^k = Aa^n + B\beta^n \), where \( a = 1 + \sqrt{3} \), \( \beta = 1 - \sqrt{3} \), \( A = \frac{a}{a-\beta} \), and \( B = \frac{\beta}{a-\beta} \). Let \( a_n \) denote the sum. Using Snake Oil,

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

To extract the coefficient of \( x^n \), we expand by partial fractions: \( 1 - 2x - 2x^2 = (1 - ax)/(1 - bx) \), where \( a = 1 + \sqrt{3} \) and \( \beta = 1 - \sqrt{3} \). Thus \( \frac{1}{1-2x-2x^2} = \frac{A}{1-ax} + \frac{B}{1-bx} \), where \( A = a/(\alpha - \beta) = (3 + \sqrt{3})/6 \) and \( B = \beta/(\beta - \alpha) = (3 - \sqrt{3})/6 \). With \( A \) and \( B \) so specified, we have \( a_n = Aa^n + B\beta^n \).

3.2.37. \( \sum_{k \geq 0} \binom{n+k}{2k} 2^{n-k} = \frac{1}{3}(1 + 2^{n+1}) \), which is always an integer since \( 4^n \equiv 1 \mod 3 \). Let \( a_n \) be the desired sum. Note that the nonzero terms in the sum run from \( k = 0 \) to \( k = n \). We use Snake Oil to find the generating function for \( a_n \).
Chapter 3: Generating Functions

Section 3.2: Coefficients and Applications

3.2.38. For $m, n \in \mathbb{N}_0$, prove that

$$
\sum \binom{n+k}{m+2k} \frac{(-1)^k}{k+1} = \binom{n-1}{m-1}.
$$

We apply Snake Oil and the Catalan generating function. First multiply by $x^n$ and sum over $n$:

$$
\sum_n \sum_k \binom{n+k}{m+2k} \frac{(-1)^k}{k+1} x^n = \sum_k \binom{2k}{k} \frac{(-1)^k}{k+1} \sum_n \binom{n+k}{m+2k} x^{n+k}.
$$

For $m \geq k$, we have

$$
\binom{n+k}{m+2k} \frac{(-1)^k}{k+1} = \frac{\binom{k}{k} (-1)^k}{k+1} \sum_n \binom{n+k}{m+2k} x^{n+k}.
$$

We apply Snake Oil and the Catalan generating function. First multiply by $x^n$ and sum over $n$:

$$
\sum_n \sum_k \binom{n+k}{m+2k} \frac{(-1)^k}{k+1} x^n = \sum_k \binom{2k}{k} \frac{(-1)^k}{k+1} \sum_n \binom{n+k}{m+2k} x^{n+k}.
$$

For $m \geq k$, we have

$$
\binom{n+k}{m+2k} \frac{(-1)^k}{k+1} = \frac{\binom{k}{k} (-1)^k}{k+1} \sum_n \binom{n+k}{m+2k} x^{n+k}.
$$

Note that

$$
\frac{1}{1-x} \sum_k \binom{2k}{k} \frac{1}{k+1} \left(\frac{-x}{1-x}\right)^k = \frac{1}{1-x} - \frac{\sqrt{1+4x/(1-x)^2}}{2x/(1-x)^2}.
$$

To extract the value of the sum, we compute

$$
[x^n] \left(\frac{x}{1-x}\right)^m = [x^{n-m}] (1-x)^{-m} = \binom{n-m + m - 1}{m-1}.
$$

3.2.39. Let $A_n(y) = \sum_k \binom{n+k}{k} \frac{(-1)^k}{k+1} y^k$ for $n \in \mathbb{N}_0$, and let $A(x, y) = \sum_{n \geq 0} A_n(y)x^n$.

a) $A(x, y) = \frac{1}{1-x} \left(1 - x - 4xy\right)^{-1/2}$. Letting $z = xy/(1-x)$ and using the generating function for the central binomial coefficients, we compute

$$
A(x, y) = \sum_k \binom{n+k}{k} \frac{(-1)^k}{k+1} y^k x^n = \sum_k \binom{2k}{k} y^k \sum_n \binom{n+k}{n} x^n
$$

We set $A_n = \binom{n+k}{n}$. Note that

$$
A_n = \binom{n+k}{k} \frac{2n}{(1-x)^2} = \binom{2n}{n} \frac{2n}{(1-x)^2}
$$

b) $\sum_k \binom{2k}{k} \frac{(-1)^k}{k+1} x^n = \frac{1}{2n!} \binom{n}{2n}$. We set $y = -1/4$ in the generating function and then extract $[x^n]$. Note that $1 - x - 4xy = 1$ when $y = -1/4$.

$$
\sum_k \binom{2k}{k} \frac{(-1)^k}{k+1} x^n = \frac{1}{2n!} \binom{n}{2n}.
$$

c) $\sum_k \binom{2k}{k} \frac{(-1)^k}{k+1} x^n = \binom{n+1}{2n}$. We set $y = -1/4$ in the generating function and then extract $[x^n]$. Note that $1 - x - 4xy = 1$ when $y = -1/4$.

$$
\sum_k \binom{2k}{k} \frac{(-1)^k}{k+1} x^n = \binom{n}{2n}.
$$

The desired sum is $[x^n] \sum_{m \geq 0} \frac{\binom{2m}{m}}{2^{2m} m!} x^{2m}$. Hence it is 0 when $n$ is odd, and when $n$ is even it is $\frac{1}{2^{n(n/2)}}$.

3.2.40. Let $c(n+1, k+1) \binom{k}{m} (-1)^{k-m} x^m = \sum_k (-1)^k c(n+1, k+1) \binom{k}{m} (-x)^m

\sum_k \sum_m c(n+1, k+1) \binom{k}{m} (-1)^{k-m} x^m = \sum_k (-1)^k c(n+1, k+1) \binom{k}{m} (-x)^m

We introduce the generating function, interchange order of summation, and twice apply $\sum_r c(n, r) = x^n$.

$$
\sum_m c(n+1, k+1) \binom{k}{m} (-1)^{k-m} x^m = \sum_k (-1)^k c(n+1, k+1) \binom{k}{m} (-x)^m
$$

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

$$
= \frac{1}{x-1} (x-1)^{(n+1)} = x^{(n)} = \sum_m c(n, m) x^m.
$$
Thus the value of the sum is \( c(n, m) \).

3.2.41. The central Delannoy number \( d_{n,n} \).

\( \textbf{a)} \) \( d_{n,n} = \sum_k \binom{n+k}{2k} \). If a Delannoy path to \((n, n)\) takes \( h \) horizontal steps, then it also takes \( n-k \) diagonal steps and \( k \) vertical steps, for \( n+k \) steps in total. Such a path is formed by choosing \( 2k \) positions for the nondiagonal steps, and then from those positions choose positions for the \( k \) horizontal steps. Every path has exactly one number of horizontal steps, so summing over \( k \) counts each path exactly once.

\( \textbf{b)} \) \( \sum_{n=0}^{\infty} d_{n,n} x^n = (1 - 6x + x^2)^{-1/2} \), by \textit{Snake Oil}. We have a finite sum over \( k \) for the coefficient of \( x^n \), with only one appearance of \( n \). After forming the generating function, we interchange the order of summation and perform the inner sum. Note that terms with \( k \geq n \) in the original sum are 0. Later we substitute \( y = x/(1-x)^2 \) and invoke the generating function for the central binomial coefficients.

\[
\sum_{n=0}^{\infty} d_{n,n} x^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n+k}{2k} \frac{(2k)}{k} x^n = \sum_{k=0}^{\infty} \left( \frac{2k}{k} \right) x^k \sum_{n=0}^{\infty} \binom{n+k}{2k} \frac{(2k)}{k} x^n
\]

\[
= \sum_{k=0}^{\infty} \left( \frac{2k}{k} \right) x^k \frac{1}{(1-x)^{2k+1}} = \frac{1}{1-x} \sum_{k=0}^{\infty} \left( \frac{2k}{k} \right) y^k = \frac{1}{1-x} \frac{1}{\sqrt{1-4y}}
\]

3.2.42. \( \sum_{m,k} a_{m,k} x^m y^k = \frac{1}{\sqrt{(1-x)^2(1-4x^2)}} \), where \( a_{m,k} = \sum_r \binom{r}{(k-r)} \).

To obtain the generating function, we compute

\[
\sum_{m,k,r} \binom{r}{(k-r)} \binom{m}{r} x^m y^k = \sum_{k,r} \binom{r}{(k-r)} y^k x^r \sum_{m=0}^{\infty} \binom{m}{r} x^m
\]

\[
= \sum_{k,r} \binom{r}{(k-r)} y^k x^r (1-x)^{r+1} = \sum_{r} \frac{x^r}{(1-x)^{r+1}} y^k \sum_{k} \binom{r}{(k-r)} y^{-r}
\]

\[
= \frac{1}{1-x} \sum_{r} \left( \frac{xy(1+y)}{1-x} \right)^r
\]

\[
= \frac{1}{1-x} \frac{1}{1-(xy(1+y)^2/x)} = \frac{1}{1-x(1+y+y^2)}.
\]

When we view the generating function as \( \sum_{m=0}^{\infty} [x(1+y+y^2)]^m \), we can interpret it as choosing a value \( m \) and making a string of \( m \) elements that are in \{0, 1, 2\}. The exponent on \( x \) is the number of values chosen, and the exponent on \( y \) accumulates the sum of the values. Thus the coefficient of \( x^m y^n \) counts the strings of \( m \) numbers in in \{0, 1, 2\} that sum to \( k \).

Therefore, to complete a combinatorial proof of the derivation, it suffices to show that the sum counts these strings. Let \( r \) be the number of nonzero entries. Given \( r \), we select the positions for the nonzero entries in \( \binom{m}{r} \) ways and then fill them with a 1, 2-list having sum \( k \). To reach \( k \) in \( r \) values, the list must have \( k-r \) 2s and \( 2r-k \) 1s; there are \( \binom{r}{k} \) ways to choose the positions of the 2s. Summing over \( r \) counts all the strings.

3.2.43. Two proofs of \( \sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{n-k}{m} = \binom{m}{n} \).

\textbf{Proof 1 (convolutions).} Since \( \binom{m+n+k}{n-k} = \binom{n-k}{m} \), the sum is the coefficient of \( x^n \) in the product of \( (1-x)^n \) and \( (1-x)^{(m+1)} \). If \( n > m \), then \( x^n(1-x)^{n-m} = 0 \), because the generating function is a polynomial of degree less than \( n \); this agrees with \( \binom{m}{n} \). If \( n \leq m \), then the exponent is negative, and \( \binom{m}{n} = \binom{m-n}{n} = \binom{m}{n} \).

\textbf{Proof 2 (Snake Oil).} Since \( m \) appears only once, we treat the sum as the coefficient of \( x^m \) in a generating function. Then we interchange the order of summation.

\[
\sum_{m,k} (-1)^k \binom{n}{k} \binom{m-n-k}{m} x^m = \sum_{k=0}^{\infty} (-1)^k \binom{n}{k} \sum_{m=0}^{\infty} \binom{m-n-k}{m} x^m
\]

\[
= \sum_{k=0}^{\infty} (-1)^k \binom{n}{k} \left( \frac{1}{1-x} \sum_{m=0}^{\infty} \binom{m}{r} (-1)^k \left( \frac{1}{1-x} \right) \right)^n-k
\]

\[
= \frac{1}{1-x} \left[ \frac{1}{1-x} - 1 \right] = \frac{x^n}{(1-x)^{n+1}}.
\]

Since \( \sum_{m=0}^{\infty} \binom{m}{n} x^m = \frac{x^n}{(1-x)^{n+1}} \) (Remark 3.2.3), the coefficient of \( x^m \) is \( \binom{m}{n} \), as desired. Alternatively, \( \binom{n-k}{n} = \binom{n}{n-k} = \binom{n-k}{m} \).

3.2.44. \textit{Snake Oil proof that} \( \sum_{k=0}^{\infty} \binom{n}{k} 2^k = \sum_{k=0}^{\infty} \binom{n}{2k} 2^{2k} \).

\textit{We show that both sides have the same generating function, indexed by \( n \).}

On the left, with \( y = x/(1-x) \), we compute

\[
\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \binom{n}{2k} x^n = \sum_{k=0}^{\infty} \binom{n}{2k} \sum_{n=0}^{\infty} \binom{n}{k} x^n = \sum_{k=0}^{\infty} \binom{2k}{k} \frac{x^k}{(1-x)^{k+1}}
\]

\[
= \frac{1}{1-x} \sum_{k=0}^{\infty} \binom{n}{k} \left( \frac{x}{1-x} \right)^k = \frac{1}{1-x} \frac{1}{\sqrt{4y}} = \frac{1}{\sqrt{(1-x)^2 - 4y(1-x)}}
\]

On the right, with \( z = x^2/(1-3x)^2 \), we compute
\[
\sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{2k} (2k)! \binom{2n}{k} 3^{n-2k} x^n = \sum_{k=0}^{\infty} \binom{2k}{k} 3^{2k} \binom{n}{2n} (3x)^n = \sum_{k=0}^{\infty} \binom{2k}{k} 3^{2k} \frac{(3x)^{2k}}{(1-3x)^{2k+1}}
\]

\[
= \frac{1}{1-3x} \sum_{k=0}^{\infty} \binom{2k}{k} \left( \frac{x^2}{1-(3x)^2} \right)^k = \frac{1}{1-3x} \frac{1}{\sqrt{1-4x}} = \frac{1}{\sqrt{(1-3x)^2-4x^2}}
\]

Both generating functions equal \((1-6x + 5x^2)^{-1/2}\).

3.2.45. \[
\sum_{j=0}^{n} \binom{n-j}{k-j} = \binom{n-k}{k-1}.\]

When \(k\) and \(j\) are fixed, \(\sum_{n=0}^{\infty} \binom{n-j}{k-j} x^n = x^k \sum_{n=0}^{\infty} \binom{n-k}{k-j} x^{n-k} = x^k (1-x)^{-(k-j-1)}.\) Therefore,

\[
\sum_{j=0}^{n} (-1)^j \binom{n-j}{k-j} = \sum_{j=0}^{n} (-1)^j \binom{k-1}{j} (k-j)^{x^k} \frac{1}{(1-x)^{k-j+1}} = \frac{x^k}{(1-x)^{k+1}} \sum_{j=0}^{n-1} \binom{k-1}{j} (1-x)^j = \frac{x^k}{(1-x)^{k+1}} [1-(1-x)]^{k-1}
\]

Since \(1-(1-x) = x\), the last expression simplifies to \(x^{2k-1}(1-x)^{(k-1)}\). The desired sum is the coefficient of \(x^n\) in this generating function.

\[
[x^n] \frac{x^{2k-1}}{(1-x)^{k+1}} = [x^{n-k+1}] \frac{x^k}{(1-x)^{k+1}} = \binom{n-k+1}{k}
\]

3.2.46. \[
\sum_{n=0}^{\infty} \binom{n}{k} \binom{2n}{2k} 2^{n-k} = \binom{2n+1}{k}.\]

The left side counts the positive lattice walks of length \(n\), but that is not needed to evaluate the sum.

**a)** Computation of \(\sum_{k=0}^{n} b_k y^k\), where \(b_k = \binom{k}{k/2}\). We obtain OGFs for the even terms and the odd terms separately, with the substitution \(x = y^2\) to spread the terms. From Example 3.2.13, \(\sum_{n=0}^{\infty} \binom{n}{2n} x^n = (1-4x)^{-1/2}\). Hence \(\sum_{n=0}^{\infty} \binom{2n}{n} y^{2n} = (1-4y^2)^{-1/2}\). To this we add \(\sum_{n=0}^{\infty} \binom{2n-1}{n-1} y^{2n-1}\). Since \(\binom{2n}{n} = \frac{1}{n} \binom{2n}{n-1}\), we compute \(\sum_{n \geq 1} \frac{1}{n} \binom{2n}{n-1} y^{2n-1} = \frac{1}{2y} (1-4y^2)^{-1/2} - \frac{1}{2y}\). Thus

\[
\sum_{k=0}^{\infty} b_k y^k = \frac{1+2y}{1-4y^2} = \frac{1}{2y} \frac{1+2y}{1-4y^2} = \frac{1}{2y}.
\]

**b)** Evaluation of \(\sum_{k=0}^{\infty} \binom{k}{k/2} 2^{n-k} y^n\) by Snake Oil. Multiply by \(x^n\) and sum over \(n \geq 0\) to form a generating function. Interchange the order of summation and evaluate the inner sum. To invoke part (a), let \(y = \frac{x}{1-2x}\); note then that \((1-2x)2y = 2x\) and \(\frac{1+2y}{2y} = \frac{1}{1-4x}\).

\[
\sum_{k \geq 0} \binom{k}{k/2} x^k \sum_{n \geq 0} \binom{n}{k} 2^{n-k} x^{n-k} = \sum_{k \geq 0} \binom{k}{k/2} x^k \sum_{m \geq 0} \binom{m+k}{k} (2x)^m
\]

\[
= \sum_{k \geq 0} \binom{k}{k/2} x^k \frac{1}{(1-2x)^{k+1}} = \frac{1}{2x} \sum_{k \geq 0} \binom{k}{k/2} y^k
\]

\[
= \frac{1}{2x} \left( \frac{1}{2y \sqrt{1-2y}} - 1 \right) = \left[ x^{n+1} \frac{1}{2} \frac{1}{\sqrt{1-4x}} \right] = \frac{1}{2} \frac{2n+2}{n+1} = \binom{2n+1}{n+1}
\]

The desired sum is the coefficient of \(x^n\). For \(n \geq 0\), we extract

\[
[x^n] \frac{1}{2x} \left( \frac{1}{\sqrt{1-4x}} - 1 \right) = [x^{n+1}] \frac{1}{2} \frac{1}{\sqrt{1-4x}} = \binom{2n+2}{n+1} = \binom{2n+1}{n+1}
\]

3.2.47. \[
\sum_{r=0}^{\infty} \binom{n/2}{r} \binom{r}{k} = \binom{n-k}{2n-k} 2^{n-k} \text{ for } n \geq 2k.\]

Fix \(k\), and let \(a_n\) denote the left side. Applying Snake oil, we form the generating function for \(\langle a \rangle\), switch the order of summation, and use \(\sum_{n=0}^{\infty} \binom{n+k}{k} y^n = (1-y)^{(k+1)}\) and \(\sum_{r=0}^{\infty} \binom{r}{k} z^r = z^r (1-z)^{(k+1)}\) to compute

\[
\sum_{n=2k}^{\infty} a_n x^n = \sum_{n=2k}^{\infty} \sum_{r=0}^{\infty} \binom{n-1}{2r+1} \binom{r}{k} x^n = \sum_{r=0}^{\infty} \sum_{n=2r+1}^{\infty} \binom{n-1}{2r+1} \binom{r}{k} x^n
\]

\[
= \sum_{r=0}^{\infty} \binom{r}{k} \sum_{n=2r+1}^{\infty} \binom{n-1}{2r+1} x^n = \sum_{r=0}^{\infty} \binom{r}{k} x^{2r+2} \sum_{m=0}^{\infty} \binom{m+2r+1}{2r+1} x^{m+2r}
\]

\[
= \sum_{r=0}^{\infty} \binom{r}{k} \frac{1}{(1-x)^{2r+2}} = \sum_{m=0}^{\infty} \binom{m+k}{k} 2^m x^{m+2k} = \sum_{n=2k}^{\infty} \binom{n-k}{k} 2^{n-2k} x^n
\]

Comparing coefficients of \(x^n\) yields \(a_n = \binom{n-k}{k} 2^{n-k}\).

3.2.48. \[
\sum_{k=0}^{\infty} \binom{n-m}{k} \binom{m-k}{k} 2^{k+1} = \binom{m+n}{n}.\]

Let \(a_m, n\) be the desired sum, and \(A(x, y) = \sum_{m, n \geq 0} a_m, n x^m y^n\). Interchange the order of summation, moving \(k\) to the outside and \(m, n\) to the inside. Now

\[
A(x, y) = \sum_{k} x^{2k} y^k \sum_{m=0}^{\infty} \binom{m-k}{k} x^{m-2k} \sum_{n=0}^{\infty} \binom{n+k}{2k} y^{n-k}
\]

\[
+ \sum_{k} x^{2k+1} y^{k+1} \sum_{m=0}^{\infty} \binom{m-k-1}{k} x^{m-2k-1} \sum_{n=0}^{\infty} \binom{n+k}{2k+1} y^{n-k}.
\]

Using \(\sum_{r=0}^{\infty} \binom{r+k}{k} z^r = (1-z)^{-(k+1)}\),
Chapter 3: Generating Functions

3.3. Exponential Generating Functions

\[ A(x, y) = \sum_{k \geq 0} x^{2k} y^k (1-x)^{-(k+1)}(1-y)^{-(2k+2)} + \sum_{k \geq 0} x^{2k+1} y^{k+1} (1-x)^{-(k+1)}(1-y)^{-(2k+3)} \]

\[ = \frac{1+xy}{(1-x)(1-y)} \sum_{k \geq 0} \left( \frac{x^2y}{(1-x)(1-y)} \right)^k = \frac{1 - y + xy}{(1-x)(1-y)^2} (1-x)(1-y)^2 - x^2y \]

the last step being verified by \((1-y+xy)(1-x-y) = 1 - x - 2y + 2xy + y^2 - x^2y - xy^2\). Hence \(a_{m,n}\) is the coefficient of \(x^m y^n\) in \((1 - (x+y))^{-1}\), which by the geometric series and binomial theorem is \((m+n)^n\).

Once we expect the sum to equal \((\frac{m+n}{n})\), there is a short combinatorial proof, because \((\frac{m+n}{n})\) counts the lattice paths from \((0,0)\) to \((m,n)\). The summation counts these paths by their last intersection with the set \(S = \{(m-2i,i)\} \cup \{(m-2i-1,i)\}\). There are \((\frac{m-i}{i})\) walks for which the last intersection is \((m-2i,i)\). If the last intersection is \((m-2i-1,i)\), then the next step must be to \((m-2i-1,i+1)\), and the number of walks of this form is \((m-i)\).

3.3.49. For \(n \geq 1\), the value of \(\sum_{0 \leq k \leq n/3} 2^k \frac{n-k}{n} = 2^{n-1}\) when \(n\) is odd and \(2^{n-1} - 1\) when \(n\) is even. Let \(s(n)\) be the desired sum, and form the generating function \(A(x) = \sum_{n \geq 0} s(n) x^n\), specifying \(s(0) = 1\). Using \(\frac{k}{n} = 1 + \frac{k}{n}\), express \(s(n)\) as \(s(n) = u(n) + v(n)\), where

\[ u(n) = \sum_{0 \leq k \leq n/3} 2^k \binom{n-k}{2k}, \quad v(n) = \sum_{1 \leq k \leq n/3} 2^{k-1} \binom{n-k-1}{2k-1}. \]

Applying Snake Oil, we interchange the order of summation and use \((1-x)^{-s} = \sum_{m \geq 0} (\frac{m+s-1}{s-1}) x^n\),

\[ 1 + \sum_{n \geq 1} u(n) x^n = \sum_{k \geq 0} 2^k \sum_{n \geq 3k} \binom{n-k}{2k} x^n = \sum_{k \geq 0} 2^k \frac{x^{3k}}{(1-x)^{2k+1}}, \]

\[ \sum_{n \geq 1} v(n) x^n = \sum_{k \geq 1} 2^{k-1} \sum_{n \geq 3k} \binom{n-k-1}{2k-1} x^n = \sum_{k \geq 1} 2^{k-1} \frac{x^{3k}}{(1-x)^{2k}}. \]

Thus \(A(x)\) is the sum of two geometric series, the second missing the first term. Simplifying and using partial fractions, we compute

\[ A(x) = \frac{1}{1-x} \sum_{k \geq 0} \left[ \frac{2x^3}{(1-x)^2} \right]^k + \frac{1}{2} \sum_{k \geq 1} \left[ \frac{2x^3}{(1-x)^2} \right]^k \]

\[ = \frac{1}{1-x} \cdot \frac{(1-x)^2}{(1-x)^2 - 2x^3} + \frac{1}{2} \cdot \frac{(1-x)^2}{(1-x)^2 - 2x^3} = \frac{1-x + x^3}{(1-x)^2 - 2x^3} = \frac{1-x + x^3}{(1-x)(1+x^2)} \]

\[ = \frac{1}{2} + \frac{1}{1-2x} + \frac{1}{1+x^2} = \frac{1}{2} + \sum_{n \geq 0} 2^{n-1} x^n + \sum_{n \geq 0} (-1)^n x^{2n}. \]

3.3. EXPONENTIAL GENERATING FCNS

3.3.1. For even \(n\), the exponential generating function \(B_n(x)\) for lists from \([n]\) such that each odd number is used an odd number of times and each even number is used an even number of times is \(2^{-n}(e^{2x} - e^{-2x})^{n/2}\). There is one way to select an element a particular number of times, but the selections are labeled by their positions in the list. Hence we use EGFs. The EGF for elements used an odd number of times is \(\sum_{k \geq 0} x^{2k+1}/(2k+1)!\); for those used an even number of times it is \(\sum_{k \geq 0} x^{2k}/(2k)!\). Using the standard trick for extracting odd terms or even terms from a power series, these EGFs are \((e^x - e^{-x})/2\) and \((e^x + e^{-x})/2\). We need the product of \(n/2\) factors of each type.

3.3.2. The number of ways to distribute 10 people into three rooms so that each room has at least one person is 55,980. The EGF for distributions of distinguishable people (labels) into one nonempty room is \(e^{-x} - 1\). Hence we seek \([x^{10}/10!] (e^x - 1)^3\). Since \((e^x - 1)^3 = e^{3x} - 3e^{2x} + 3e^x - 1\), we extract \(3^{10} - 3 \cdot 2^{10} + 3 \cdot 1^{10} - 0^{10}\).

3.3.3. The EGF for rows of cards from four each of 13 values, indexed by length, is \((1 + x + x^2/2 + x^3/6 + x^4/24)^{13}\). The labels are the positions. Each rank of four cards provides a factor. It can contribute up to four cards, so the EGF for each type of card is \(1 + x + x^2/2 + x^3/6 + x^4/24\). Note that the exponent accumulates the number of cards used. There are 13 types of cards. (-) Cards are dealt in a row from a standard 52-card deck, and the values are recorded (suits are ignored). Build an EGF for these lists, indexed by length.

3.3.4. The EGF for the ways to put distinct objects into \(k\) distinct boxes with at least \(m\) objects in each box, indexed by the number of objects, is
Objects are being assigned to boxes. Each object is used only once, while boxes can be used repeatedly, and the index is the number of objects. Hence the objects are the labels. We have a factor for each box. Given the objects going into a box, there is only one way to put them in. Hence the factor for each box is \( \sum_{n \geq m} x^n/n! \).

Modification for indistinguishable boxes. If the boxes are not distinguishable and \( m \geq 1 \), then each arrangement occurs \( k! \) times, and we simply divide by \( k! \). However, if \( m = 0 \), then boxes can be empty, and the problem is to partition a set of size \( n \) into at most \( k \) parts. In this case the coefficient of \( x^n/n! \) is \( \sum_{j=0}^k S(n, j) \), where \( S(n, j) \) is the Stirling number of the second kind.

3.3.5. Arrangements of people in a club, partitioned into a set \( S \), a queue \( T \), and the remainder. Let \( a_n \) count the arrangements. Each person is used, so the label set is \( [n] \); we seek \( \sum_{n \geq 0} a_n x^n/n! \). The EGF for an unordered set (\( S \)) is \( e^x \). The EGF for arranging people in a queue (index by number of people) is \( 1/(1-x) \). The EGF for the remainder (people sent home) is \( e^x \). By the product rule, the answer is \( e^{2x}/(1-x) \).

3.3.6. \( \sum_{k=0}^n k^n = n 2^{n-1} \). Using binomial convolution, the sum is the coefficient of \( x^n/n! \) in the product of the EGFs \( \sum_{k \geq 0} k x^k/k! \) and \( \sum_{k \geq 0} x^k/k! \). The latter is \( e^x \). The former is \( x \frac{d}{dx} e^x \), which is \( xe^x \). Since \( xe^x = \sum_{n \geq 0} 2^n n^{n+1}/n! = \sum_{n \geq 0} (n+1)! \), we have \( [x^n/n!] xe^{2x} = n 2^{n-1} \).

3.3.7. Binomial Inversion on Stirling numbers. Fixing \( n \), the formula is \( k! S(n, k) = \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n \). Write this as \( b_k = \sum_{i=0}^k (-1)^i \binom{k}{i} a_{k-i} \), where \( b_k = k! S(n, k) \) and \( a_k = k^n \). Binomial Inversion makes this equivalent to \( b_k = \sum_{i=0}^k (-1)^i \binom{k}{i} S(n, k) \). Combinatorially, to form \( n \)-words from \( [k] \), choose \( i \) letters to omit, partition the positions into \( k-i \) blocks, and assign them to the \( k-i \) letters used.

3.3.8. The statements \( \sum_{k=1}^n s(n, k) b_k \) for all \( n \in \mathbb{N} \) and \( \sum_{k=1}^n S(n, k) a_k \) for all \( n \in \mathbb{N} \) are equivalent. If the first statement holds, then by Corollary 3.3.14 we compute

\[
\sum_{k=1}^n S(n, k) a_k = \sum_{k=1}^n \sum_{m=1}^k S(n, k) s(k, m) b_m
\]

\[
= \sum_{m=1}^n \sum_{k=m}^n S(n, k) s(k, m) b_m = \sum_{m=1}^n \delta_{n,m} b_m = b_n
\]

If the second statement holds, then since the transpose of the identity matrix is the identity matrix, we similarly compute

\[
\sum_{k=1}^n s(n, k) b_k = \sum_{k=1}^n \sum_{m=1}^n s(n, k) S(k, m) a_m = \sum_{m=1}^n \delta_{n,m} a_m = a_n.
\]

3.3.9. The EGF for arrangements of distinct children in teams with captains, indexed by number of children, is \( e^{e^x} \). The children are the labels, and we have no arrangements of teams. Thus a team is a component structure. With \( n \) children on a team, there are \( n \) ways to designate the captain. Note also that the existence of a team requires at least one child (the captain). Hence the EGF for one team is \( \sum_{n \geq 1} x^n/(n-1)! = x e^x = xe^x \). By the Exponential Formula, the generating function with any number of teams allowed is \( e^{e^x} \).

3.3.10. If \( a_n = a_{n-1} + n! \) for \( n \geq 1 \), with \( a_0 = 1 \), then \( a_n = (n+1)! \), by the generating function method. Let \( A(x) = \sum_{n \geq 0} a_n x^n/n! \). Multiply the recurrence by \( x^n/n! \) and sum over \( n \geq 1 \) to obtain \( A(x) - a_0 = x A(x) + \sum_{n \geq 1} x^n \), which simplifies to \( A(x) = \frac{1}{1-x} + \frac{1}{(1-x)^2} = \frac{1}{1-x} \). The coefficient of \( x^n/n! \) in \( 1/(1-x)^2 \) is \( n!/(n+1)! \). Hence \( a_n = (n+1)! \).

3.3.11. Coefficient of \( x^3 \) in \((1+x)^a\), by composing \( e^{e^x} \) and \( \ln(1+x) \). By the Extended Binomial Theorem, \([x^3](1+x)^a = (a-1)(a-2)/6 \). To compute this from \( e^{e^x} \), we consider all contributions to the coefficient of \( x^3 \). We have \( \ln(1+x) = x - x^2/2 + x^3/3 + \cdots \). We multiply by \( \alpha \) and substitute into the series for \( e^x \). Since \( \ln(1+x) \) has no constant term, we need only consider up to \( y^3/6 \).

The constant term is \( 1 \). We have

\[
y = \alpha \ln(1+x) = \alpha x - \alpha^2 x^2/2 + \alpha^3 x^3/3 + \cdots
\]

\[
y^2 = \alpha^2(\ln(1+x))^2 = \alpha^2 x^2 - \alpha^2 2 x^3/2 + \cdots
\]

\[
y^3 = \alpha^3(\ln(1+x))^3 = \alpha^3 x^3 + \cdots
\]

To find \([x^3](1+y+y^2/2+y^3/6+\cdots \), we just take the contribution from each term, computing

\[
0 + \alpha/3 - \alpha^2/2 + \alpha^3/6 = (2-\alpha)\alpha/2 + \alpha^3/6 = (\alpha-1)(\alpha-2)/6.
\]

3.3.12. If \( A^1, \ldots, A^n \) and \( C \) are symmetric families of labeled structures such that objects in \( C[k] \) correspond bijectively to distributions of the label set \([k]\) into \( n \) sets \( S_1, \ldots, S_n \) and choices of elements from \( A_{S_i} \) for \( 1 \leq i \leq n \).
k, then the EGF for the counting sequence of C is the product of those for $A^1, \ldots, A^n$. Let $a_i^k = |A^k_i|$ for $1 \leq i \leq n$ and $c_k = |C_k|$. The description of objects in $C_k$ yields $c_k = \sum (r_1, \ldots, r_n) \prod_{i=1}^n a_{r_i}^k$, where the sum is over all solutions to $\sum_{i=1}^n r_i = k$ in nonnegative integers.

By induction on $n$, this formula is the product of the EGFs for the counting sequences of $A^1, \ldots, A^n$. The statement is trivial for $n = 1$. For $n \geq 2$, group the contributions by the value of $r_n$ to compute

$$\sum (r_1, \ldots, r_n) \prod_{i=1}^n a_{r_i}^k = \sum_{r_n=0}^k \left( \sum (r_1, \ldots, r_{n-1}) \prod_{i=1}^{n-1} a_{r_i}^k \right).$$

Thus the EGF for $\langle c \rangle$ is the product of those for $\langle a^n \rangle$ and $\langle b \rangle$, where by the induction hypothesis the EGF of $\langle b \rangle$ is the product of those for the counting sequences of $A^1, \ldots, A^{n-1}$.

3.3.13. $4^{n-1}$ is the number of words of length $n$ from $\{w, x, y, z\}$ such that the usage of $w$ is even and the usage of $y$ is odd.

Building the EGF. We make independent choices for usage of the four letters. The labels are the positions in words, and words are built by allocating labels to the four types, so the EGF for $\langle a \rangle$ is the product of the EGFs for words using one type of letter with the specified restrictions on usage. Since $w$ and $z$ are unrestricted, the factor for each is $e^w$. Since $x$ and $y$ must be used an even number and an odd number of times, respectively, the factors for them are $(e^x + e^{-x})/2$ and $(e^y + e^{-y})/2$, respectively. Hence the EGF for $\langle a \rangle$ is $e^{2w}(e^{2x} - e^{-2y})/4$, which equals $(e^{4x} - 1)/4$. The coefficient of $x^n/n!$ in the expansion is 0 when $n = 0$ and $4^{n-1}$ when $n > 0$.

Combinatorial explanation.

Proof 1. Since $y$ must be used an odd number of times, $a_0 = 0$. For $n > 0$, we form binary words by thinking of $w$ and $z$ both as 0 and $x$ and $y$ both as 1. Thus we want binary words with an odd number of $1$s. There are $2^{n-1}$ of these, since we put an arbitrary binary $(n-1)$-tuple first and choose the $n$th bit to make the number of $1$s odd. Now we expand each bit in two ways: 0s turn into $w$ or $z$, and 1s turn into $x$ or $y$, except that the last 1 becomes an x if so far we have an odd number of $x$s, and it becomes a $y$ if so far we have an even number of $x$s. Hence each binary word expands in $2^{n-1}$ ways, yielding $4^{n-1}$ words in total.

Proof 2. Start with an arbitrary word of length $n - 1$ in the alphabet $\{x, y, z, w\}$. There are $4^{n-1}$ such words, and there are two ways to extend each to obtain a word of length $n$ in which the total number of $x$s and $y$s is odd. In such words, the numbers of $x$s and $y$s have opposite parity. Exactly half have an even number of $x$s; we establish a one-to-one correspondence between the two types by changing all $x$s into $ys$ and all $ys$ into $x$s. This brings the number of desired words back to $4^{n-1}$.

3.3.14. $\sum_{k=0}^m \binom{n}{k} \binom{n-k}{m-k} = \binom{n}{m} 2^m$. Observe first that the upper limit of the sum $S_{n,m}$ can be changed to $n$ without changing its value, since the terms added or deleted are 0. Next replace $\binom{n-k}{m-k}$ with $\binom{n-k}{n-m+k}$. Now $S_{n,m}$ has the form $\sum_{k=0}^n \binom{n}{k} b_{k,n-k}$, where $a_k = 1$ and $b_k = \binom{n-k}{n-m+k}$. Thus $S_{n,m}$ is the coefficient of $x^n/n!$ in the product of $e^x$ and $B(x) = \sum_{r \geq 0} (r^m/n!)$, which we can change to summation over $r \geq n-m$ since the earlier coefficients are 0. Thus $B(x) = (x^{n-m}/(n-m)! \sum_{r \geq n-m} x^{r-n+m}/(r-n+m)! = e^x x^{n-m}/(n-m)!$. Now $S_{n,m}$ is the coefficient of $x^n/n!$ in $e^x x^{n-m}/(n-m)!$. The coefficient is the coefficient of $x^n/n!$ in $e^x x^{n-m}/(n-m)!$. This is $\binom{2(n-m)}{n-m}$.

This value $\binom{n}{2^k}$ counts the number of ternary $n$-tuples with $m$ 0’s and 1’s by picking the positions for 0’s and 1’s and choosing some subset to have 0’s. The sum counts the same set by first picking some number of positions ($k$) to have 0’s and then choosing positions for 1’s from the remaining positions.

Comment. In terms of ordinary generating functions, the sum is the coefficient of $x^n$ in $(1 + x)^n x^{n-m}/(1-x)^{n-m+1}$. It is not easy to extract directly from this a formula for the coefficient of $x^n$.

3.3.15. Surjective distributions of flags to poles. Let $b_{r,n}$ be the number of distinguishable ways to place $n$ distinct flags onto $r$ distinct flagpoles with each pole having at least one flag.

a) $\sum_{r \geq 1} b_{r,n} x^r = x^r(1-x)^{-r}$. The labels, used once each, are the flags, and the EGF is indexed by the number of flags. There are $n!$ ways to put $n$ flags on one pole, except that $n = 0$ is not allowed. Hence the EGF for one pole is $x^r/(1-x)$. With $r$ poles, we allocate the flags to the poles and then place them on the poles, so the EGF is $[x^r/(1-x)]^r$.

b) $b_{r,n} = n! \left(\frac{1}{r-1}\right)$, from the EGF. We compute

$$b_{r,n} = \left[ x^n/n! x^r(1-x)^{-r} = n! [x^{n-r}](1-x)^{-r} = n! \binom{n-r+r-1}{r-1} \right].$$

c) $b_{r,n} = n! \left(\frac{1}{r-1}\right)$, directly. List the flags in order from pole 1, then pole 2, up to pole $r$. This yields a permutation of $[n!)$, with locations marked where we move from one pole to the next. Since each pole is nonempty, we form such a distribution from any permutation of $[n!)$ by choosing $r - 1$ distinct places between elements to move from one pole to the next.

3.3.16. Recurrence and EGF for $\langle a \rangle$, where $a_n$ counts the involutions on $[n]$. For $n \geq 2$, element $n$ may occur by itself or with one other element, so $a_n = a_{n-1} + (n-1) a_{n-2}$, with $a_0 = 1$ and $a_1 = 1$. Let $A(x) = \sum_{n \geq 0} a_n x^n/n!$. Multiply by $x^{n-1}/(n-1)!$ and sum over $n \geq 2$. Now the left side is $A'(-x) - a_1$, and the right side is $A(x) - a_0 + x A(x)$. Thus $A'(x)/A(x) = 1 + x$, using both ini-
3.3.17. Rankings. Let \( a_n \) denote the number of rankings of \( n \) distinct candidates.

\[ \text{a) The EGF for } (a) \text{ is } 1/(2 - e^x). \]

A ranking on a set is just an ordered partition of the set, with any number of blocks, so \( a_n = \sum_{k \geq 0} k! S(n, k) \).

When there are \( k \) nonempty (ranked) blocks, the EGF is \( (e^x - 1)^k \), as in the discussion of Stirling numbers. However, we cannot use any number of blocks, so \( \sum_{k \geq 0} a_n/k! = \sum_{k \geq 0} (e^x - 1)^k \).

\( \) (There is one ranking of no candidates, achieved by using no blocks.) Evaluating the geometric sum yields \( 1/(2 - e^x) \).

\[ b \sum_{k \geq 0} k^n/2^k \text{ is an integer. The coefficient of } x^n/n! \text{ in } 1/(2 - e^x) \text{ is the integer } a_n, \text{ so twice } a_n \text{ is also an integer. Considering twice the EGF,} \]

\[ \frac{1}{1 - e^x} = \sum_{k \geq 0} \frac{e^{2k} x^n}{2^k} = \sum_{n \geq 0} \left( \sum_{k \geq 0} \frac{k^n}{2^k} \right) \frac{x^n}{n!}. \]

Section 3.3.3: Exponential Generating Functions

3.3.18. Among trees with vertex set \( [n] \), there are \((n!/k)!S(n-2, n-k)\) with \( k \) leaves. In the Pr"ufer code for a tree, the label \( d(i) \) appears \( d(i) - 1 \) times, \( d(i) \) is the degree of vertex \( i \) in the tree. Thus the leaves are the labels that do not appear in the code. After drawing the remaining \( n-k \) labels must appear, and thus they partition the \( n-2 \) positions of the Pr"ufer code into \( n-k \) blocks. Each such block is named by the label occurring in those positions. Thus the number of these partitions, when we have chosen the labels for leaves, is \((n-k)!S(n-2, n-k)\). Since we can choose the labels in \( \binom{n}{k} \) ways, the total number of Pr"ufer codes with \( n-k \) labels appearing is \((n!/k)!S(n-2, n-k)\), and thus this is the desired number of trees.

3.3.19. There are \( S(n-1, k-1) \) ways to partition \([n]\) into \( k \) sets so that no two consecutive values are in the same group. Let \( a_{n,k} \) be this value. We use induction on \( n \); note that \( a_{1,k} = \delta_{1,k} = S(0, k-1) \). For \( n \geq 2 \), element \( n \) can be in a group alone or in a group not containing \( n-1 \). There are \( a_{n-1,k-1} \) partitions of the first type and \((k-1)a_{n-1,k-1} \) of the second type. Hence \( a_{n,k} = (k-1)a_{n-1,k-1} + a_{n-1,k-1} \). By the induction hypothesis, \( a_{n,k} = (k-1)S(n-2, k-1) + S(n-2, k-2) = S(n-1, k-1) \), using the recurrence for the Stirling numbers (valid since \( n-1 \geq 1 \)).

3.3.20. \( \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} i^k = 0 \) when \( 0 \leq k < n \).

\[ \text{a) Induction on } k. \text{ When } k = 0, \text{ always } i^k = 1 \text{ for } i \in \mathbb{N}_0. \text{ Thus the sum is } \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} = 0 \text{ when } n \geq 1 \text{ (the numbers of even and odd subsets of } [n] \text{ are equal when } n \geq 1). \text{ We may thus assume } 1 \leq k < n. \text{ The Committee-Chair Identity } \binom{n}{i} = \frac{n}{i} \binom{n-1}{i-1} \]

expresses the sum as a linear combination of sums that equal 0 by the induction hypothesis.

\[ \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} i^k = \sum_{i=1}^{n} (-1)^{i} \binom{n}{i-1} i^k = n \sum_{i=1}^{n} (-1)^{i} \binom{n}{i-1} i^{k-1} = -n \sum_{j=0}^{n-1} \binom{n-1}{j} (j+1)^{k-1} = -n \sum_{j=0}^{n-1} \binom{n}{j} \sum_{r=0}^{k-1} \binom{k-1}{r} \left( \frac{n-1}{r} \right)_{j} = 0. \]

\[ \text{b) Proof using OGFs.} \]

\[ \text{Proof 1 (manipulating binomial expansion). The sum } \sum_{i=0}^{n} i^k \binom{n}{i} \text{ arises from } (1 + x)^n \text{ by } k \text{ applications of } x \frac{d}{dx} \text{, followed by setting } x = -1. \text{ Application of the product rule for differentiation yields, after } r \text{ steps, the product of } (1 + x)^{n-r} \text{ with a polynomial in } x. \text{ Since } k < n, \text{ after step } k \text{ the summation retains } 1 + x \text{ as a factor. Setting } x = -1 \text{ thus yields } 0. \]

\[ \text{Proof 2 (convolution). Written as } (-1)^n \sum_{i=0}^{n} i^k (-1)^{n-i} \binom{n}{i}, \text{ the sum is } (-1)^n [x^n] C(x) B(x), \text{ where } C(x) = \sum_{i=0}^{\infty} i^k x^i \text{ and } B(x) = \sum_{i=0}^{\infty} (-1)^{i} \binom{n}{i} x^i. \]

Here \( B(x) = (1 - x)^n \), but \( C(x) \) is hard to compute. We only need to know that \( C(x) = P(x)(1-x)^{-k+1} \text{ for some polynomial } P \text{ of degree at least } k. \text{ We then have } C(x) B(x) = P(x)(1-x)^{n-k-1}. \text{ Since } k < n, \text{ (1-x)^{n-k-1} \text{ is a polynomial. Hence } C(x) B(x) \text{ is a polynomial in } x \text{ with degree at most } n-1, \text{ and } [x^n] C(x) B(x) = 0. \]

Worpitzky’s Identity is \( x^k = \sum_{k=0}^{n} A(n, k)(x-k+n) \), where \( A(k, j) \) is the Eulerian number. Rename the parameters to get \( i^k = \sum_{j=0}^{k} A(k, j)(-i+j)^k \). The convolution says \( C(x) = P(x)(1-x)^{-k+1} \), where \( P(x) = \sum_{i=0}^{\infty} A(k, j)x^i \); this is a polynomial of degree \( k. \)

\[ \text{c) Proof using Stirling numbers. Letting } j = n-i, \text{ the sum becomes } \sum_{i=0}^{n} (-1)^n (\binom{n}{i} (n-j)^k. \text{ This sum is } (-1)^n \text{ times the sum in Theorem 3.3.11 for } n! S(k, n), \text{ where } S(k, n) \text{ counts partitions of } [k] \text{ with } n \text{ blocks. For } k < n, \text{ none exists, so the value is } 0. \]

3.3.21. Properties of Stirling numbers.

\[ \text{a) } S(m + n, n) = \sum_{a \in P(m, n)} (n + 1)(n + 2) \cdots (m + n)(a_{n-1...a_1}) \prod_{i=2}^{m+n} \frac{1}{a_i}, \text{ where } P(m, n) \text{ is the set of nonnegative integer vectors } (a_1, \ldots, a_{m+1}) \text{ such that } \sum a_i = n \text{ and } \sum i a_i = m + n. \text{ Membership in } P(m, n) \text{ is necessary and sufficient for the existence of a partition of } m + n \text{ elements into } n \text{ sets consisting of } a_i \text{ sets of size } i \text{ for } 1 \leq i \leq m + 1; \text{ note that no subset can have more than } m + 1 \text{ elements. For fixed } a \text{ the number of these partitions is the product of the number of ways of partitioning the } m + n \text{ elements into labeled sets of sizes } 1a_1, \ldots, ma_m \text{ and the numbers of ways of partitioning the resulting } i \text{th set into } a_i \text{ unlabeled } i \text{-sets, for} \]

...
Section 3.3: Exponential Generating Functions

3.3.24. Stirling numbers and Eulerian numbers.

a) \( k!S(n, k) = \sum_{i=0}^{k} \binom{k-i}{i} A(n, i) \). The left side counts the ordered partitions of \([n]\) into \(k\) blocks. As in Gessel’s proof of Worpitzky’s Identity, these partitions become barred permutations by writing the elements in increasing order in each successive bin, followed by a bar. All these bins are nonempty, so there are no consecutive bars. The right side counts the barred permutations with \(k\) nonconsecutive bars by the number of runs, \(i\). Each run must be followed by a bar, and then \(k-i\) additional bars are inserted into distinct locations among the remaining \(n-i\) positions following elements. (These inserted bars correspond to the smallest element of one block being greater than the largest element of the preceding block.) Thus there

\[
\sum_{i=0}^{k} k!S(n, k)x^{n-i} = \sum_{i=0}^{n} A(n, i)(x+1)^{n-i}.
\]

Multiply the equation of part (a) by \(x^{n-k}\) and then sum over \(k\). Using \(\sum_{k=0}^{n} (n-i)x^{n-k} = \sum_{k=0}^{n-i} (n-i-k) = (1+x)^{n-i}\), we compute

\[
\sum_{k=0}^{n} k!S(n, k)x^{n-k} = \sum_{i=0}^{n} \binom{n-i}{i} A(n, i)x^{n-k} = \sum_{i=0}^{n} A(n, i)\binom{n-i}{i}x^{n-k} = \sum_{i=0}^{n} A(n, i)(1+x)^{n-i}.
\]

When \(n=0\), both sides equal 1. For \(n > 0\), we can start the sums at 1. Setting \(x = 1\) on the left counts all the ordered partitions of \([n]\). On the right this produces \(\sum_{i=0}^{n} A(n, i)2^{n-i}\). Setting \(j = n+1-i\) converts this to \(\sum_{j=0}^{n} A(n, n+1-j)2^{j-1}\), but \(A(n, n+1-j) = A(n, j)\).

3.3.25. The two statements below are equivalent for \(0 \leq k \leq n\).

a) \(b_k = \sum_{i=0}^{k} (-1)^{k-i}\binom{k-i}{i}a_i\) for \(0 \leq k \leq n\).

b) \(a_k = \sum_{i=0}^{k} (-1)^{k-i}\binom{k-i}{i}b_i\) for \(0 \leq k \leq n\).

Given (a), we substitute the expression for \(b_i\) from (a) into the right side of (b) and compute

\[
\sum_{i=0}^{k} (-1)^{k-i}\binom{k-i}{i}b_i = \sum_{i=0}^{k} (-1)^{k-i}\binom{n-i}{i-j}a_j = \sum_{j=0}^{k} a_j \sum_{i=j}^{k} (-1)^{i-j}\binom{n-j+l}{k-j-l} = \sum_{j=0}^{k} a_j \sum_{l=0}^{k-j} (-1)^{k-j-l}\binom{k-j-1}{k-j} = a_k.
\]
We have used the extended binomial coefficients and Vandermonde’s convolution. Not only Vandermonde in the usual way, but in addition its extension to a polynomial identity in two variables, allowing us to use a negative value for one argument. Furthermore, in the last step we use \((\frac{1}{m}) = 1\) while \((\frac{m-1}{m}) = 0\) for all \(m > 0\). This also requires the interpretation of the binomial coefficient as a polynomial.

There does not seem to be a nice inversion by generating functions.

With \(b_k = k!S(n, k)\) and \(a_k = A(n, k)\), part (a) of Exercise 3.3.24 gives statement (a) above. Hence we conclude statement (b): \(A(n, k) = \sum_{i=0}^{k} (-1)^{k-i} (k-i)! S(n, i)\).

3.3.26. \(n^m S(n, m) \geq m^n \binom{n}{m}\).

**Proof 1** (Thomas Horine). To partition the set \([n]\) into \(m\) nonempty unlabelled sets, first choose \(m\) elements and place one in each set, which can be done in \(\binom{n}{m}\) ways. The remaining \(n-m\) elements can be assigned to those \(m\) sets in \(m^{n-m}\) ways to complete a partition. However, a partition with part sizes \(s_1, \ldots, s_m\) has been counted \(\prod_{i=1}^{m} s_i\) times, since each part can be initiated by any of its \(s_i\) elements. Since always \(\sum_{i=1}^{m} s_i = n\), the arithmetic-geometric mean inequality yields \(\prod_{i=1}^{m} s_i \leq \left(\frac{n}{m}\right)^m\) for each partition.

Thus \(S(n, m) \geq \frac{\binom{n}{m} m^{n-m}}{\left(\frac{n}{m}\right)^m}\).

**Proof 2** (Mark Wildon). We interpret the desired inequality as comparing the sizes of certain sets.

The left side, \(n^m S(n, m)\), is the number of ways to form a partition \(P\) of \([n]\) into \(m\) parts and assign to each part a number in \([n]\) via a function \(f\). The right side, \(m^n \binom{n}{m}\), is the number of ways to choose a set \(Z\) of \(m\) elements in \([n]\) and assign to each member of \([n]\) one element of \(Z\) via a function \(g\).

Let \(\mathcal{L}\) be the set of such pairs \((P, f)\) such that exactly \(r\) elements of \([n]\) are used in the image of \(f\), and let \(\mathcal{R}\) be the set of such pairs \((Z, g)\) such that exactly \(r\) elements of \([n]\) are used in the image of \(g\). It suffices to prove \(|\mathcal{L}| \geq |\mathcal{R}|\) for all \(r\). Since \(m \leq n\), we may assume \(r \leq m\).

For \((Z, g) \in \mathcal{R}\), the preimages of elements of \([n]\) under \(g\) form a partition \(Q\) of \([n]\) into \(r\) parts. Since the image of \(g\) is contained in \(Z\), we can build the pair \((Z, g)\) by first choosing a \(r\)-element subset \(W\) of \([n]\) for the image of \(g\), then \(m-r\) additional members of \([n]\) to complete \(Z\), then the partition \(Q\), and finally the bijective assignment of elements of \(Z\) to parts of \(Q\). Thus \(|\mathcal{R}| = \binom{n}{r} (\binom{n-r}{m-r}) S(n, r)!\).

To count \(\mathcal{L}\), note that for any partition \(P\) of \([n]\), we can build the function \(f\) by first choosing an \(r\)-element subset \(X\) of \([n]\) to be the image of \(f\), then a partition of the \(m\) parts of \(P\) into \(r\) nonempty blocks, and finally the bijective assignment of elements of \(X\) to these blocks. Thus \(|\mathcal{L}| = S(n, m) \binom{n}{r} S(m, r)!\).

To prove \(|\mathcal{L}| \geq |\mathcal{R}|\), it thus suffices to prove \(S(n, m) S(m, r) \geq \binom{n-r}{m-r} S(n, r)\).

Given a partition \(Q\) of \([n]\) with \(r\) parts, let \(M(Q)\) be the \(r\)-subset of \([n]\) consisting of the largest element of each part. Choose also a set \(T\) of \(m-r\) elements from \([n] - M(Q)\). The right side is the number of such pairs \((Q, T)\).

Given such a pair \((Q, T)\), we define a partition \(P\) of \([n]\) into \(m\) parts such that \(P\) refines \(Q\). Simply extract each element of \(T\) from its part in \(Q\) and make it a new singleton part.

Next consider all the pairs \((Q, P)\) of partitions of \([n]\) such that \(Q\) has \(r\) parts, \(P\) has \(m\) parts, and \(P\) is a refinement of \(Q\). We can build such pairs by first choosing \(P\) and then grouping the parts of \(P\) into a partition with \(r\) parts. Hence there are \(S(n, m) S(m, r)\) such pairs. It therefore suffices to show that our map obtaining \((Q, P)\) from \((Q, T)\) is injective.

If \((Q, T)\) arises from \((Q, T)\) by this map, then \(P\) has at least \(m-r\) singleton parts. The element \(x\) of a singleton part lies in \(T\) if and only if \(x \notin M(Q)\). Thus we can reconstruct \((Q, T)\) from \((Q, P)\), and the map is injective, as desired.

*Comment:* Filip Nikšić posted this question at mathoverflow.net/questions/268544, noting that when the inequality is rewritten as

\[ \frac{S(n, m)!}{m^n} \geq \frac{n!}{m^n}, \]

the left side is the probability that a uniformly chosen random function \([n] \rightarrow [m]\) is surjective, while the right side is the probability that a uniformly chosen random function \([m] \rightarrow [n]\) is injective. Thus functions from \([n]\) to \([m]\) are more likely to be surjective than functions from \([m]\) to \([n]\) are to be injective.

3.3.27. \(\sum_{j=1}^{\min(n, N)} \binom{N}{j} j! S(n, j)(N-j)^m = \sum_{i=1}^{\min(m, N)} \binom{N}{i} i! S(m, i)(N-i)^n\).

The Stirling number \(S(n, m)\) counts the partition of an \(n\) element set into \(j\) nonempty parts. We interpret the \(j! S(n, j)\) ordered partitions as functions from \([n]\) onto \([j]\).

The two expressions are equal because both sides count the pairs of functions \((f, g)\) such that \(f: [n] \rightarrow [N], g: [m] \rightarrow [N]\), and the images
of $f$ and $g$ are disjoint. The first sum counts these pairs according to the size of the image of $f$: for a fixed size $j$, there are $\binom{N}{j}$ ways to select $j$ elements to form the image, $j!S(n, j)$ ways to select $f$ with that image, and then $(N - j)^m$ ways to select an arbitrary function $g$ with range disjoint from that of $f$. Similarly, the second sum counts the pairs according to the size of the range of $g$.

3.3.28. If $p_n(x) = \sum_{k=0}^{n-1} S(n, k)x^k$ for $n \in \mathbb{N}_0$, then $p_n(x + y)$ is the same polynomial as $\sum_{k=0}^{n} \binom{n}{k} p_k(x) p_{n-k}(y)$. For integer $z$, the value $p_n(z)$ is the number of pairs consisting of a partition of $[n]$ and a mapping of each block to an element of $[z]$. When $x = y$, some subset $A$ of $[n]$ winds up in blocks that get mapped to elements of $x$, and the remaining set $B$ goes into blocks that get mapped into the some of the remaining $y$ elements. If $|A| = k$ then we choose those $k$ elements, partition them into blocks, and map the blocks into $[x]$; there are $\binom{n}{k}$ $p_k(x)$ ways to do that. We must also partition $B$ and map those blocks into the remaining $y$ elements; there are $p_{n-k}(y)$ ways to do that. Summing over $k$ counts the same set of structures counted on the left side. Since the identity holds for all integer $x$ and $y$, by the Polynomial Principle it is a polynomial identity.

3.3.29. Special Stirling permutations. A Stirling permutation is an arrangement of two copies of the elements of $[n]$ such that for all $i$, the entries between the two copies of $i$ exceed $i$. A skyline is a Stirling permutation with the additional property that no strictly increasing triple has its last two entries consecutive in the arrangement. Let $a_n$ be the number of skylines of length $2n$, and let $A(x) = \sum_{n=0}^{\infty} a_n x^n/n!$.

a) $A'(x) = e^{2x}A(x)$, and $A(x) = e^{(e^{2x}-1)/2}$. In a skyline, the two copies of a number greater than 1 cannot have a 1 between them, so the 1st partition the arrangement into three segments. The first segment must satisfy both skyline properties. Since the other two segments follow a 1, they have no ascents and are weakly decreasing. Conversely, any distribution of the values of $[n] - 1$ into three segments that satisfy these constraints forms a skyline. With the multinomial coefficient to count the distributions when the multiplicities are fixed, we have $a_n = \sum_{i+j+k=n-1} \binom{n-1}{i,j,k} a_i$. Shifting the index and using the multiplication rule for EGFs yields $\sum_{n=0}^{\infty} a_{n+1} x^n/n! = A(x) \cdot e^x \cdot e^x$, and hence $A'(x) = e^{2x}A(x)$. With the convention that there is one skyline of length 0, we want $A(0) = 1$, and hence the solution of the differential equation is $A(x) = e^{(e^{2x}-1)/2}$.

b) $a_n = \sum_{k=0}^{n} 2^{n-k} S(n, k)$, from part (a).

Section 3.3: Exponential Generating Functions

\[ A(x) = e^{(e^{2x}-1)/2} = \sum_{k \geq 0} [(e^{2x} - 1)/2]^k / k! \]
\[ = \sum_{k \geq 0} 2^{-k} \frac{k!}{k!} \sum_{i=0}^{k} (-1)^i \left( \begin{array}{c} k \\ i \end{array} \right) e^{2x(k-i)} = \sum_{k \geq 0} 2^{-k} \frac{k!}{i!} \sum_{i=0}^{k} (-1)^i \left( \begin{array}{c} k \\ i \end{array} \right) \frac{2^n e^n (k-1)^n}{n!} \]
\[ = \sum_{n \geq 0} \frac{x^n}{n!} \sum_{k \geq 0} 2^{-k} \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \left( \begin{array}{c} k \\ i \end{array} \right) (k-1)^n = \sum_{n \geq 0} \frac{x^n}{n!} \sum_{k \geq 0} 2^{-k} S(n, k) \]

e) Combinatorial proof of part (b). Let $B_n$ be the set of objects consisting of a partition of $[n]$ with a subset of the non-minimal elements of each block marked. There are $2^{n-k} S(n, k)$ such objects in which the partition has $k$ blocks. We map $B_n$ bijectively into the set of skylines of length $2n$.

Note first that a skyline can be broken into segments starting at its left-to-right minima $m_1, \ldots, m_k$ so that the ith such segment has the form $m_i\alpha_i m_i\beta_i$, since the second copy of $m_i$ cannot come earlier and cannot follow the next left-to-right minimum. Furthermore, the second copy of each element in $\alpha_i$ must also be in $\alpha_i$, and similarly for $\beta_i$, to avoid surrounding a smaller element. Finally, since $m_i$ is the least entry in this segment, $\alpha_i$ and $\beta_i$ are nonincreasing.

Given a marked partition of $[n]$, index the blocks in decreasing order of their least elements. From the ith block, form the ith segment of a skyline by using the same least element, letting $\alpha_i$ consist of both copies of the marked elements of the block, and letting $\beta_i$ consist of both copies of the unmarked elements of the block. The steps of the construction are reversible, so the map is a bijection.

3.3.30. Bell numbers ($B_n$) is the number of partitions of $[n]$.

a) $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$ for $n \geq 0$. In each partition of $[n+1]$, the element $n+1$ lies in the same block with some subset of $[n]$, and the remainder of the partition is a partition of the remaining subset of $[n]$. Since $n-k$ elements can be chosen to lie with the element $n+1$ in $\binom{n}{k}$ ways, the desired formula follows, with $k$ representing the number of elements in blocks not containing the element $n+1$.

b) The exponential generating function for the Bell numbers is $e^{e^x-1}$. Let $B(x)$ be the EGF. The convolution on the right in (a) is the coefficient of $x^n/n!$ in $B(x) \cdot e^x$, viewed as a product of EGFs. The number $B_{n+1}$ is the coefficient of $x^n/n!$ in $B'(x)$. Since this holds for all $n \geq 0$, we have $B'(x) = B(x) e^x$. Integrating both sides of $B'(x)/B(x) = e^x$ yields $\ln B(x) = e^x + c$, and thus $B(x) = e^{e^x+c}$. We choose the constant $c = 0$ to agree with $B_0 = 1$.

3.3.31. Bell numbers.

a) $\sum_{k=1}^{\infty} \frac{k^n}{k!} x^k = e^x \sum_{k=0}^{\infty} S(n, k)x^k$. Begin with $\sum_{j=1}^{n} S(n, j)k_j = k^n$, a relabeling of part of Theorem 3.3.13. Divide both sides by $k!$, multiply by
Thus making the adjustments before dividing by 2 yields \((B_n + B_{n-1} + B_{n-2})/2\) as the number of distinguishable partitions.

3.3.34. A principal submatrix is a submatrix obtained by extracting the same set of columns as rows. A symmetric matrix is positive semidefinite if all its principal submatrices have nonnegative determinant.

a) The number of partial partitions of \([n]\) is \(B_{n+1} - 1\), where a partial partition of a set \(X\) is a partition of a nonempty subset of \(X\). To describe a partial partition, we can put the elements not used into a special block with the element 0. Thus partial partitions of \([n]\) correspond to arbitrary partitions of \([n+1]\), except that we do not allow the partition of \([n+1]\) in which all elements are in the block with 0.

b) The number of positive semidefinite 0,1-matrices of order \(n\) is the Bell number \(B_{n+1}\). Every all-1 square matrix is positive semidefinite; its 1-by-1 principal submatrices have determinant 1 and the others have determinant 0. Thus \(A\) is positive semidefinite if there is a simultaneous permutation of the rows and columns that transforms the matrix into a matrix with square all-1 blocks along the diagonal. Such a matrix is specified by a partial partition of \([n]\) specifying the row/column indices to be together in such blocks. We do also allow the 0-matrix, so the number of such matrices is \(B_{n+1}\).

We claim that these are the only positive semidefinite 0,1-matrices. Given a positive semidefinite 0,1-matrix \(A\), let \(J = \{i: a_{ii} = 1\}\). For \(i \notin J\), the 2 by 2 submatrices involving \(a_{ii} = 0\) show that all entries in the \(i\)th row and \(i\)th column are 0. Define a relation \(\sim\) on \(J\) by setting \(i \sim j\) if and only if \(a_{ij} = 1\). Reflexivity and symmetry are immediate for \(\sim\). If \(\sim\) is not transitive, there exist \(i, j, k \in J\) such that \(a_{ij} = a_{jk} = 1\) but \(a_{ik} = 0\). This yields a 3 by 3 principal submatrix \(B\) that is all 1’s except for one off-diagonal pair of 0’s. The determinant of such a matrix is \(-1\), so this is forbidden. We conclude that \(\sim\) is an equivalence relation, which expresses \(A\) in the form described above.

\[
\sum_{i=0}^{n} \binom{n}{i} (-1)^i (2^i) \sum_{m=0}^{\frac{n}{2}} \frac{(-1)^m}{m!} = \sum_{i=0}^{n} (-1)^i \left(\frac{n}{2}\right)^{2n-i(n+1)!}. \tag{2}\]

The factor \((2!)^i \frac{(-1)^m}{m!}\) is the well-known derangement number \(D_{2i}\), the number of ways to return hats to 2l people so that no hat returns to its owner. When hats are returned to \(n\) couples, let \(d(n, l)\) be the number of outcomes in which the hats for the people in \(l\) couples are deranged, while the remaining hats all go to their owners. The left side is \(\sum_{i=0}^{n} (-1)^i d(n, l)\), counting such outcomes positively or negatively depending on the parity of the number of couples in the derangement.

Now consider the right side. The expression \(\binom{n}{i} 2^{n-i} (n+l)!\) is the number of ways to return the hats so that one specified person in each of \(n-l\) couples receives the right hat (others may also receive the right hat).
These terms are weighted by the parity of the number of couples with no member specified. We claim that each permutation is counted with total weight 1, −1, or 0, depending on whether the set of those who get the wrong hat forms an even number of couples, an odd number of couples, or includes some “isolated” individuals whose mates get the right hat.

Consider an assignment of hats in which the correct hats are received by exactly $i$ couples and $j$ isolated individuals. This outcome is counted for each of the $2^r \binom{i}{r} \binom{j}{s}$ choices of one partner from each of $r$ of the couples and $s$ of the individuals; such a choice contributes weight $(-1)^{n-(r+s)}$. The total weight for this outcome is therefore

$$\sum_{r \leq i} \sum_{s \leq j} (-1)^{n-(r+s)} 2^r \binom{i}{r} \binom{j}{s} = (-1)^n \sum_{r \leq i} (-2)^r \binom{i}{r} \sum_{s \leq j} (-1)^s \binom{j}{s}$$

$$= (-1)^n (1-2)^i(1-1)^j = (-1)^n i^j,$$

with $0^0 = 1$. This gives ±1 or 0 in precisely the cases we need.

3.3.36. Derivation of the Eulerian polynomial from its EGF. The $n$th Eulerian polynomial $A_n$ is defined by $A_n(t) = \sum_{k=0}^{n} A(n, k) t^k$. We start with a recurrence for the Eulerian numbers.

a) $A(n, k) = A(n-1, k-1) + \sum_{m=1}^{n-1} \sum_{j=1}^{k} (n-1)_m A(m,j) A(n-1-m, k-j)$ for $n, k \geq 1$. Consider the position of element $n$ in a permutation of $[n]$ with $k$ runs. Some subset $S$ of $[n-1]$ is chosen to precede $n$; let $m = |S|$. If $m = 0$, then $n$ is a run by itself, and what follows is a permutation of $[n-1]$ with $k-1$ runs; these permutations are counted by the term $A(n-1, k-1)$. If $m > 0$, then $n$ extends the last run in the permutation of $S$ preceding it. Once $S$ is chosen, the permutation preceding $n$ has some number $j$ of runs, and the permutation of $[n-1] - S$ that follows $n$ then must have $k-j$ runs. Hence with each nonzero possibility for $m$ and $j$, the number of permutations contributing to $A(n, k)$ is $(n-1)_m A(m,j) A(n-1-m, k-j)$. Summing over $m$ and $j$ completes the count.

b) $A_n(t) = (t-1)A_{n-1}(t) + \sum_{m=0}^{n-1} (n-1)_m A_m(t) A_{n-1-m}(t)$ for $n \geq 1$, with $A_0(t) = 1$. We multiply the recurrence in part (a) by $t^k$ and sum over $k \geq 1$ (note that $A(n,0) = \delta_{n,0}$). We want to extend the summation to start at $m = 0$ and $j = 0$ to express the sum as a convolution. This adds $A(0,0) A(n-1, k)$ to the sum, so we must also subtract $A(n-1, k)$. For $n \geq 1$, we now have

$$A_n(t) = \sum_{k=0}^{n} A(n, k) t^k$$

$$= t \sum_{k=1}^{n} A(n-1, k) t^{k-1} - \sum_{k=1}^{n} A(n-1, k) t^k$$

$$+ \sum_{m=0}^{n-1} (n-1)_m \sum_{k=0}^{n} (n-1)_m A(m,j) t^j A(n-1-m, k-j)$$

$$= (t-1)A_{n-1}(t) + \sum_{m=0}^{n-1} (n-1)_m A_m(t) A_{n-1-m}(t).$$


section 3.3: Exponential Generating Functions

$$A_n(t) = \frac{t^n}{n!} \frac{\partial^n}{\partial t^n} \left( \frac{1}{1-\frac{t}{1-e^{t-1}}n} \right)$$

$$= \frac{t^n}{n!} \frac{\partial^n}{\partial t^n} \left( \frac{1}{1-\frac{t}{1-e^{t-1}}} \right).$$

The power series $A(x)$ is the EGF for the Eulerian polynomials. We introduce the EGF by multiplying the recurrence in part (b) by $x^{n-1}/(n-1)!$ and summing over $n \geq 1$. From the left side, we have $\sum_{n \geq 1} A_n(t)x^{n-1}/(n-1)! = A'(x)$. On the right, $(t-1)SGEn1A_{n-1}(t)x^{n-1}/(n-1)! = (t-1)A(x)$. The binomial convolution in the summation is the coefficient of $x^{n-1}/(n-1)!$ in the product of $A(x)$ with itself. Thus $A'(x) = (t-1)A(x) + A^2(x)$.

d) $A(x) = \frac{1-x}{1-te^{1-x}}$. We verify that this function satisfies the differential equation. On the left, $\frac{d}{dx} \frac{1-x}{1-te^{1-x}} = \frac{(1-x)(1-te^{1-x})}{(1-te^{1-x})^2} (-te^{1-x}) (1-x) = \frac{1-x}{1-te^{1-x}}$.

On the right, $(t-1)\frac{d}{dx} \frac{1-x}{1-te^{1-x}} + \left( \frac{1-x}{1-te^{1-x}} \right)^2 = (t-1)^2 \left( \frac{1-x-te^{1-x}}{(1-te^{1-x})^2} \right) = \frac{1-x}{1-te^{1-x}}$.

e) $A_n(t) = (t-1)^{n+1} \sum_{i \geq 0} i^n t^i$. We expand the EGF $\sum_{n \geq 0} A_m(t)x^n/n!$ and extract the coefficient of $x^n/n!$.

$$\frac{1-t}{1-te^{1-x}} = (t-1) \sum_{i \geq 0} t^i e^{1-x}$$

$$= (t-1) \sum_{n \geq 0} \sum_{i \geq 0} i^n t^i x^n/n!$$

$$= (t-1)^{n+1} \sum_{i \geq 0} i^n t^i x^n/n!. $$

Thus the coefficient of $x^n/n!$, which by definition is $\sum_{k \geq 0} A(n,k) t^k$, equals $(1-t)^{n+1} \sum_{i \geq 0} i^n t^i$.

3.3.37. The EGF for even graphs with vertex set $[n]$, indexed by $n$, is $1 + \int_{0}^{x} G(\lambda) d\lambda$, where $G(x)$ is the EGF for all graphs. The even graphs with vertex set $[n]$ correspond bijectively to all graphs with vertex set $[n-1]$; adding vertex $n$ to a graph $G$ with vertex set $[n-1]$ and making $n$ adjacent to all vertices having odd degree in $G$ creates an even graph, and all even graphs with vertex set $[n]$ arise in this way.
Chapter 3: Generating Functions

Section 3.3: Exponential Generating Functions

Given \( G(x) = \sum_{n \geq 0} G_n x^n/n! \) for all graphs, we want the coefficient of \( x^n/n! \) in the new generating function to be \( G_{n-1} \). We achieve this by integrating: \( \int_0^x G(t) dt = \sum_{n \geq 0} G_n \int_0^x t^{n-1} = \sum_{n \geq 0} G_n x^n/n! \). We add 1 for the constant term because there remains one graph with no vertices, and all its degrees are even.

Connected even graphs. Every component of an even graph is a connected even graph, so the relationship between general and component structures that is governed by the Exponential Formula holds here. Thus the EGF for connected even graphs is the natural logarithm of the EGF found above.

3.3.38. Unordered binary trees with \( n \) labeled leaves. Let \( b_n \) be the number of these, with \( b_0 = 0 \) and \( B(x) = \sum_{n \geq 0} b_n x^n/n! \).

a) \( B(x) = x + \frac{1}{2} B(x)^2 \). Note that \( b_1 = 1 \). For \( n \geq 2 \), the labels for the leaves must be partitioned into two sets, with an unordered binary tree placed on each. Picking a set or its complement as the “first” set yields the same object, since the subtrees are unordered. Thus \( b_n = \frac{1}{2} \sum_{k=1}^{n} \binom{n}{k} b_k b_{n-k} \). With \( b_0 = 0 \), the sum extends to \( 0 \leq k \leq n \). Multiplying by \( x^n/n! \) and summing over \( n \geq 2 \) introduces the EGF as \( B(x) - 0 = x + \frac{1}{2} (B(x)^2 - 0) \), as desired.

b) \( b_n = \prod_{i=0}^{n-1} (2i - 1) \). The equation for the EGF is \( \frac{1}{2} B^2 - B + x = 0 \), or \( B(x) = 1 + \sqrt{1 - 2x} \). Since the constant term must be 0, only \( 1 - \sqrt{1 - 2x} \) is allowed. For \( n \geq 1 \), we have \( b_n = \lfloor x^n/n! \rfloor (1 - 2x)^{1/2} = n! \binom{n}{1/2} (-2)^n = (-1)^{n-1} \prod_{i=1}^{n-1} (1 - 2i) \).

3.3.39. Arrangements of children. Here \( c_n \) is the number of ways that \( n \) children can arrange themselves into circles (holding hands) with one child standing alone inside each circle. Each “piece” of such an arrangement is a circle with a child inside. A piece with \( n \) children requires \( n \geq 2 \) and can be formed in \( n(n-2)! \) ways; we pick the inside child and form a cycle from the remaining \( n - 1 \) children. The EGF for the pieces is \( \sum_{n=3}^{\infty} n(n-2)!/n! x^n/n! \), which simplifies to \( \sum_{n \geq 2} x^n/(n-1) \) and then \( -x \ln (1 - x) \).

General structures are formed by allocating the children to such pieces. By the Exponential Formula, \( \sum c_n x^n/n! = e^{-x \ln (1 - x)} = (1 - x)^{-x} \).

3.3.40. The EGF \( A(x) \) for the sequence \( a_n \) in which \( a_n \) is the number of distinct matrices expressible as the sum of an \( n \)-by-\( n \) permutation matrix and its inverse is \( (1 - x)^{-1/2} e^{x/2 + x^2/4} \). Given a permutation \( \sigma \), we have \( \sigma^{-1}(j) = i \) if and only if \( \sigma(i) = j \). Hence the matrix for \( \sigma^{-1} \) is the transpose of the matrix for \( \sigma \), and every matrix to be counted is symmetric. Diagonal entries correspond to fixed points under \( \sigma \) and equal 2. Entries corresponding to a 2-cycle in \( \sigma \) yield a single symmetric pair of 1s, which are the only nonzero entries in their row and column. A cycle of length at least 3 can be followed forward or backward from each element and hence corresponds to a submatrix with two 1s in each row and column. Thus the desired \( n \)-by-\( n \) matrices correspond to undirected graphs with vertex set \([n]\) in which every component is an isolated vertex, an edge, or a cycle.

Proof 1 (Exponential Formula, directly). To form such a graph, we allocate the vertices to components (allocate elements of \([n]\) to cycles in \( \sigma \) and place a component structure on each component. The EGF \( C(x) \) for component structures, indexed by the number of elements in the component, is \( x + x^2/2 + \sum_{n=3}^{\infty} \frac{1}{2}(n-1)!/n! \). Here by convention we have one graph with no vertices, and it is not connected. We have \( \frac{1}{2}(n-1)! \) cycles with \( n \) vertices because taking \( \sigma \) and \( \sigma^{-1} \) together in the matrix makes the cycles undirected.

Since \( -\ln(1 - x) = \sum_{n=1}^{\infty} x^n/n! \), we have \( C(x) = x/2 + x^2/4 - \frac{1}{2} \ln(1 - x) \). By the Exponential Formula, \( A(x) = e^{C(x)} = (1 - x)^{-1/2} e^{x/2 + x^2/4} \).

Proof 2 (pairs of graphs). The square of the desired EGF is \( (1 - x)^{-1} e^{x + x^2/2} \), which is the product of the EGFs for all permutations and for involutions. Note that \( x^n/n! A(x)^2 \) is the number of ordered pairs \((G_1, G_2)\) of graphs of the desired form, whose union has vertex set \([n]\). We want each such ordered pair to correspond to a general permutation \( \sigma_1 \) and an involution \( \sigma_2 \), again with the union of the elements in the two permutations being \([n]\).

View \( \sigma_1 \) and \( \sigma_2 \) as functional digraphs. The cycles of lengths 1 and 2 in \( \sigma_1 \) become the isolated vertices and edges in \( G_1 \). This leaves cycles of length at least 3 in \( \sigma_1 \). Those in which the successor of the least element is greater than its predecessor become cycles in \( G_1 \); those in which the successor of the least element is less than its predecessor become cycles in \( G_2 \). The map is reversible, because whether a cycle comes from \( G_1 \) or \( G_2 \) determines which way to orient it in \( \sigma_1 \).

3.3.41. EGFs \( G(x) \) for 2-regular graphs by number of vertices and \( F(x) \) for permutations without cycles of length at most 2 by number of elements.

a) \( G(x) = (e^{-x^2/2 - x^2/4})/(\sqrt{1 - x}) \) and \( F(x) = (e^{-x^2/2})(1 - x) \). The 2-regular graphs are general structures whose component structures are cycles. With vertex set \([n]\), there are \((n-1)!/2\) possible cycles; cyclically order the vertices, but the two directions on the cycle yield the same graph. Since cycles have length at least 3, the EGF \( C(x) \) for component structures is \( \sum_{n \geq 3} \frac{1}{2}(n-1)!x^n/n! \), which equals \( \frac{1}{2} \sum_{n \geq 3} x^n/n! \). The series for \(-\ln(1 - x)\) yields \( C(x) = x/2 - x^2/4 - (1/2) \ln(1 - x) \). Since 2-regular graphs are formed by partitioning labels and forming a cycle on each set, the Exponential Formula applies. Thus \( G(x) = e^{C(x)} = (e^{-x^2/2 - x^2/4})/(\sqrt{1 - x}) \).
For permutations with cycles of length at least 3, the components are cycles; there are \((n - 1)!\) on a fixed set of \(n\) labels. Eliminating cycles of length at most 2, the EGF for the component structures is \(-x - x^2/2 - \ln(1 - x)\), and the Exponential Formula yields the expression for \(F(x)\).

b) Combinatorial proof for \(F(x) = G(x)^2\). Let \(A\) be the set of ordered pairs \((G_1, G_2)\) such that \(G_1\) and \(G_2\) are disjoint 2-regular graphs and \(V(G_1) \cup V(G_2) = [n]\). Let \(B\) be the set of permutations of \([n]\) with no cycles of length at most 2. For each \((G_1, G_2) \in A\), we form a permutation \(\sigma \in B\) by specifying its cycles.

For each cycle \(C\) in \(G_1\), put a cycle in \(\sigma\) starting with the least element of \(C\) and moving in the direction to its smaller neighbor, continuing thereafter around the cycle. For each cycle \(C\) in \(G_2\), put a cycle in \(\sigma\) starting with the least element of \(C\) and moving in the direction to its larger neighbor, continuing thereafter around the cycle. For each permutation \(\sigma\) in \(B\), retrieve the unique ordered pair mapped to it by placing each cycle of \(\sigma\) in \(G_1\) or \(G_2\) according to whether the successor of the least element on it is larger or smaller than the least element.

The size of \(B\) is \(f_n\), the coefficient of \(x^n/n!\) in \(F(x)\). Since labels must be allocated to the two graphs in the ordered pair, the size of \(A\) is \(\sum_{k=0}^{n} \binom{n}{k} g_k g_{n-k}\) where \(g_n\) is the coefficient of \(x^n/n!\) in \(G(x)\). The bijection yields \(|B| = |A|\), and hence \(F(x) = G(x)^2\).

3.3.42. Permutations with all cycles having odd length.

a) The EGFs \(A(x)\) and \(B(x)\) for permutations whose cycles have odd length and for such permutations with an even number of cycles are given by \(A(x) = \left(\frac{1+x}{2}\right)^{1/2}\) and \(B(x) = (1 - x^2)^{-1/2}\). A permutation of \([n]\) has some number of cycles. If \(C(x)\) is the EGF for allowed cycles, then the Exponential Formula yields \(A(x) = e^{C(x)}\).

We have \((n - 1)!\) allowed cycles when \(n\) is odd and 0 when \(n\) is even. Hence \(C(x)\) keeps just the odd terms in \(\sum_{n\geq 0} \binom{n}{n-1} x^n/n!\), which is \(-\ln(1 - x)\). To keep only the odd terms, \(C(x) = \frac{-\ln(1 - x) + \ln(1 + x)}{2} = \frac{1}{2} \ln \left(\frac{1 + x}{1 - x}\right)\).

By the Exponential Formula, \(A(x) = \left(\frac{1+x}{2}\right)^{1/2}\).

For \(B(x)\), we include only the general structures having an even number of components. Hence instead of \(e^{C(x)}\), we use \(\frac{1}{2}(e^{C(x)} + e^{-C(x)})\). Thus

\[
B(x) = \frac{1}{2} \left( \sqrt{1 + \frac{x}{1 - x}} + \sqrt{1 - \frac{x}{1 + x}} \right) = \frac{(1 + x) + (1 - x)}{2 \sqrt{1 - x^2}} = (1 - x^2)^{-1/2}.
\]

b) The probability that a random permutation of \([n]\) has an even number of cycles, all of odd length, equals the probability that a string of \(n\) coin tosses has exactly \(n/2\) heads. Both are 0 when \(n\) is odd. When \(n\) is even, the latter is \(2^{-n} \binom{n}{n/2}\), and the former is \([x^n] B(x)\). We can use the Extended Binomial Theorem to obtain \([x^n] B(x) = (-1)^{n/2} \binom{n}{-n/2}\) and multiply out the Extended Binomial Coefficient as usual, or use the known expansion \((1 - 4y)^{-1/2} = \sum_{k \geq 0} \frac{(2k)!}{k!}, k\). Setting \(y = x^2/4\) yields

\[
[x^n] B(x) = [x^n] \sum_{k \geq 0} \frac{2^n}{k!} \frac{x^{2k}}{4^{k}} = 2^{-n} \binom{n}{n/2}.
\]

3.3.43. Exponential generating functions for (labeled) bipartite graphs, indexed by number of vertices. Let \(F(x)\) and \(G(x)\) be the EGFs for 2-colored bipartite graphs and for bipartite graphs, respectively.

a) \(G(x) = \sqrt{F(x)}\). Let \(C(x)\) be the exponential enumerator for connected bipartite graphs. Since there are two ways to assign \(X\) and \(Y\) to the partite sets of a connected bipartite graph, the EGF for connected 2-colored bipartite graphs is \(2C(x)\). By the Exponential Formula, \(G(x) = e^{C(x)}\) and \(F(x) = e^{2C(x)}\), so \(G(x) = \sqrt{F(x)}\).

b) Bijective proof that \(F(x) = G(x)^2\). Under the rule for multiplying EGFs, it suffices to show that 2-colored bipartite graphs are constructed by allocating labels to two sets and forming an ordinary labeled bipartite graph on each set. For any 2-colored bipartite graph, the components can be grouped by whether the least-indexed label is in \(X\) or in \(Y\). In general, we form such a graph by deciding for each label whether the least label in its component will be in \(X\) or in \(Y\). Having allocated the labels in this way, we form any bipartite graph on each of the two sets of labels. The type-\(X\) graph is then 2-colored so that \(X\) contains the least label in each component; the type-\(Y\) graph is 2-colored the other way.

3.3.44. The difference between the numbers of even and odd derangements of \([n]\) is \(n - 1\), where the parity of a permutation of \([n]\) with \(k\) cycles is the parity of \(n - k\). Let \(c_k\) be the number of permutations of \([k]\) that are cycles, so \(c_k = (k - 1)!\). Let \(C(x)\) be the EGF for nontrivial cycles (without fixed points), indexed by length: \(C(x) = \sum_{k \geq 2} c_k x^k/k! = -\ln(1 - x) - x\). By the Exponential Formula, the EGF for derangements is \(e^{C(x)}\), obtained as \(\sum_{m \geq 0} C(x)^m/m!\).

In this computation, the term for \(m\) enumerates derangements with \(m\) cycles. We want to count such contributions positively when \(m\) is even, negatively when \(m\) is odd. If we accumulate \((-C(x))^m/m!\) for the derangements with \(m\) cycles and sum over \(m\), then for any fixed \(n\), the absolute value of the coefficient of \(x^n/n!\) will be the difference between the numbers of even and odd derangements. We compute

\[
e^{-C(x)} = (1 - x)e^x - xe^x = \sum_{n \geq 0} \frac{x^n}{n!} - \sum_{n \geq 0} \frac{x^{n+1}}{n!} = \sum_{n \geq 0} (1 - n) \frac{x^n}{n!}.
\]
3.3.45. The EGFs for blocks and connected graphs with vertex set \([n]\) are related by \(\ln C'(x) = B'(xC'(x))\). Here \(B(x) = \sum_{n=2} b_n x^n / n!\) and \(C(x) = \sum_{n\geq 1} c_n x^n / n!\), where \(b_n\) and \(c_n\) are the numbers of blocks and connected graphs with vertex set \([n]\), respectively. A graph with one vertex is not considered a block, since deleting the vertex leaves a graph with no vertices, which we consider disconnected. Let \(R(x)\) be the EGF for rooted connected graphs; note \(R(x) = \sum_{n\geq 1} nc_n x^n / n! = x C'(x)\).

a) \(R_k(x)/x = (R_k(x)/x)^{k}/k!,\) where \(R_k(x)\) is the EGF for rooted connected graphs in which the root appears in \(k\) blocks. We seek the number of rooted connected graphs with \(n+1\) vertices in which the root is not labeled, the labels of the other vertices are \(1\) through \(n\), and only one block contains the root. These correspond to rooted connected graphs with vertex set \([n+1]\) having vertex \(n+1\) as the root. Shifting the index and dividing by \(n+1\) shows that the EGF for this sequence is \(R_1(x)/x\). That is,

\[
\sum_{n\geq 0} \frac{r_{1,n+1} x^n}{n+1 \cdot n!} = \sum_{n\geq 1} \frac{x^{n-1}}{n!} = \frac{1}{x} R_1(x).
\]

Similarly, \(R_1(x)/x\) is the EGF for rooted connected graphs with unlabeled root in which the root appears in \(k\) blocks. To relate the two EGFs, we generate a \(k\)-set of the type in which the unlabeled root appears in one block, divide by \(k!\) since the order does not matter (the graphs are distinguished by the labels on the nonroot vertices), and then merge the unlabeled roots into one unlabeled root vertex.

b) \(R_1(x)/x = \sum_{j\geq 2} b_j R(x)^{j-1}/(j-1)!\). Recall that the coefficient of \(x^n/n!\) in \(R_1(x)/x\) is the number of rooted connected graphs having unlabeled root that appears in only one block, with the other vertices being \([n]\). Including the unlabeled root, the number of vertices in the block containing the root is some number \(j\), greater than \(1\).

To form such a graph, we first form \(j-1\) rooted connected graphs using \([n]\) as the full set of vertices. There is no order on these graphs, but they are distinguished by the names of their roots, so the EGF for this process is \(R(x)^{j-1}/(j-1)!\). Note that since \(R(0) = 0\), nonzero coefficients in this EGF require \(n\geq j-1\).

Next, the names of the roots have already been specified. At this point, we can view the unlabeled root as also having a name, such as “unlabeled”. We complete the structure by placing any block with \(j\) named vertices on the \(j\) vertices we have reserved for it. No matter what \(n\) is, this stage can be done in \(b_j\) ways. Summing over \(j\) completes the proof.

c) \(\ln C'(x) = B'(xC'(x))\). Note that \(R_0(x) = x\) and \(R(x) = \sum_{k\geq 2} R_k(x)\). Hence summing over \(k\) in part (a) yields \(R(x)/x = e^{R_1(x)/x}\), or \(R_1(x)/x = \ln(R(x)/x)\). We also noted earlier that \(R(x)/x = C'(x)\). Applying (b) yields

\[
\ln C'(x) = \frac{R_1(x)}{x} = \sum_{j\geq 1} \frac{b_j x^j}{j!} = B'(R(x)) = B'(xC'(x)).
\]

3.3.46. The number \(a_n\) of ordered \(k\)-ary trees with \(n\) vertices is \(\frac{1}{n} \binom{n}{(n-1)/k}\). Let \(y(x)\) be the OGF for \(\langle a \rangle\).

a) \(y = x(1 + y^k)\). The vertices do not have names, but one is distinguished as the root. To form an ordered \(k\)-ary tree, decide first whether the root has degree 0 or \(k\). The OGF for the first case is \(x\); there is exactly one vertex. In the second case, choose an ordered \(k\)-ary tree to root at each of the \(k\) children, in order. Since the total number of vertices used is additive over the subtrees, the coefficient of \(x^{n-1}\) in \(y^k\) is the number of ways to form the \(k\) subtrees. The root brings an extra factor of \(x\) for the \(n\)th vertex. Hence the generating function in this case is \(x y^k\), and we obtain \(y = x + x y^k\).

b) \([x^n] y = \frac{1}{n} \binom{n}{(n-1)/k}\). Part (a) yields \(x = y/\phi(y), \) where \(\phi(y) = 1 + y^k\). Since \(\phi(0) = 1,\) we can apply Lagrange Inversion. We compute \([x^n] y = \frac{1}{k} [y^{n-1}] (1 + y^k)^n\). The coefficient of \(y^{n-1}\) in \((1 + y^k)^n\) is nonzero only if \(k\) divides \(n - 1\), and in that case it is \((n-1)/k\). Comment: The divisibility condition agrees with the fact, proved by induction on \(m\), that a \(k\)-ary tree with \(m\) non-leaf vertices has \(km + 1\) vertices in total.

c) The Catalan number \(C_m\) equals \(\frac{1}{m+1} \binom{2m+1}{m}\). An ordered binary tree with \(m+1\) leaves has \(2m + 1\) vertices in total and \(2m\) edges. Section 1.3 shows that these trees are counted by the Catalan number \(C_m\). Setting \(k = 2\) and \(n = 2m + 1\) in part (b) thus yields

\[
C_m = \frac{1}{2m + 1} \binom{2m + 1}{m} = \frac{2m + 1}{2m + 1} \binom{2m + 1}{m+1} = \frac{1}{m+1} \binom{2m}{m}.
\]

3.3.47. If \(f\) is a formal Laurent series in \(y\), and \(g\) be a formal power series in \(x\) such that \([x^0] g(x) = \cdots = [x^{m-1}] g(x) = 0\) and \([x^m] g(x) \neq 0\), then

\[
[y^{-1}] f(y) = \frac{1}{m} [x^{-1}] f(g(x)) g'(x).
\]

We give a straightforward generalization of the proof of Lemma 3.3.33. By linearity, it suffices to consider the case \(f(y) = y^k\) for \(k \in \mathbb{Z}\). If \(k \neq -1\), then \([y^{-1}] f(y) = 0\). Since \(f(y) = y^k\), we have

\[
f(g(x)) g'(x) = g(x)^k g'(x) = \frac{1}{k+1} \frac{d}{dx} g(x)^{k+1}.
\]

Since \(g\) is a formal power series in \(x\), so are \(g^{k+1}\) and its derivative, and hence the coefficient of \(x^{-1}\) is 0, as desired.
If $k = -1$, then $[y^{-1}]f(y) = 1$. In this case, $f(g(x))g'(x) = g'(x)/g(x)$. The hypotheses on $g$ lead to $g(x) = x^m h(x)$, where $h(x)$ is a formal power series with nonzero constant term. Thus $h(x)$ has a multiplicative inverse, and we can write $h'(x)/h(x)$ as a formal power series.

By the product rule for differentiation, $g'(x) = mx^{m-1}h(x) + x^m h'(x)$. Thus
\[
\frac{g'(x)}{g(x)} = \frac{m}{x} + \frac{h'(x)}{h(x)}.
\]

Since $h'(x)/h(x)$ is a formal power series, it does not contribute to the coefficient of $x^{-1}$ in $\frac{m}{x}$. Thus $[x^{-1}]f(g(x))g'(x) = m$, and the desired conclusion holds.

**3.3.48. The general Lagrange Inversion Formula:** If $\phi(y)$ and $h(y)$ are formal power series in $y$ with $\phi(0) = 1$, and $x = y/\phi(y)$, then
\[
[x^n]h(y(x)) = \frac{1}{n} [y^{-1-n}] h'(y(y)) \phi(y)^n.
\]

We follow the proof of Theorem 3.3.34. To apply Lemma 3.3.33, we shift the series so that the desired coefficient is the residue. This uses $[z^{n-1}] z f(z) = [z^{-1}] f(z)$, twice.
\[
\frac{1}{n} [y^{-1-n}] h'(y(y)) \phi(y)^n = \frac{1}{n} [y^{-1}] h'(y) \left( \frac{y}{\phi(y)} \right)^n
\]
definition of $\phi$
\[
= \frac{1}{n} [y^{-1}] h'(y) \left( \frac{y}{\phi(y)} \right)^n
\]
shift of series
\[
= \frac{1}{n} [x^{-1}] h(y(x)) y'(x)
\]
Lemma 3.3.33 (see below)
\[
= \frac{1}{n} [x^{-1}] h'(y(x)) y'(x) \phi(y)^n
\]
shift of series
\[
= \frac{1}{n} [x^n] h(y(x))
\]
def. of differentiation
\[
= [x^n] h(y(x))
\]
In the crucial step, Lemma 3.3.33 yields
\[
[y^{-1}] f(y) = [x^{-1}] f(y(x)) y'(x)
\]
when (1) $y(x)$ is a formal power series with $[x^0] y(x) = 0$ and $[x^1] y(x) \neq 0$, and (2) $f(y)$ is a formal Laurent series in $y$. We apply this with $f(y) = h'(y)/x(y)^n$. Since $\phi(y)$ is a formal power series in $y$, we know that also $x(y)$ is a formal power series in $y$, and $f(y)$ is a formal Laurent series.

On the right side, $f(y(x))$ becomes $h'(y(x))y'(x)/x(y(x))^n$, expressed as a formal Laurent series in $x$. Since the composition $x(y(x))$ is the identity, the right side becomes $[x^{-1}] h'(y(x)) y'(x)/x^n$, as desired.

**3.3.49. If $y = xe^y$, then $e^{\alpha y} = \sum_{n \geq 0} a(n + 1)^n x^n/n!$. We apply the general form of the Lagrange Inversion Formula with $h(y) = e^{\alpha y}$ and $\phi(y) = e^y$. We compute**
\[
[x^n] e^{\alpha y} = [x^n] h(y(x)) = \frac{1}{n} [y^{n-1}] h'(y) \phi(y)^n
\]
\[
= \frac{1}{n} [y^{n-1}] a e^{\alpha y} e^{\alpha y} = \frac{a(a + n)^{n-1}}{n(n - 1)!}
\]

**3.3.50. Generalization of Cayley's Formula:** $a_{n,k} = kn^{n-k-1}$, where $a_{n,k}$ is the number of rooted forests with vertex set $[n]$ and root set $[k]$. Recall that $y = xe^y$ when $y(x)$ is the EGF for rooted labeled trees, using the Exponential Formula, since we choose one root and form any number of rooted trees on the remaining labels (see Theorem 3.3.29). When we require exactly $k$ rooted forests components, specified as components 1 through $k$, the EGF is $y^k$. We obtain one of the desired forests by exchanging the root labels $r_1, \ldots, r_k$ with the labels 1 through $k$ in order. Each resulting desired forest arises $n_{(k)}$ times in this process, since 1 through $k$ could have been any labels. Hence the desired answer $a_{n,k} = \frac{1}{n_{(k)}} [x^n / n!] y^k$.

Given $x = y/\phi(y)$, we can compute $[x^n] y^k$ using Theorem 3.3.35 with $h(y) = y^k$. Theorem 3.3.35 states $[x^n] h(y(x)) = \frac{1}{n} [y^{n-1}] h'(y) \phi(y)^n$; here $\phi(y) = e^y$. Thus
\[
[x^n] y^k = \frac{1}{n} [y^{n-1}] k y^{k-1} e^{\alpha y} = \frac{k}{n} [y^{n-k}] e^{\alpha y}
\]
\[
= \frac{k}{n} [y^{n-k}] \sum_{j \geq 0} \frac{(ny)^j}{j!} = \frac{k}{n} \sum_{j \geq 0} \frac{(ny)^j}{j!} = \frac{k}{n} \frac{n^{n-k}}{(n-k)!} = \frac{kn^{n-k}}{(n-k)!}
\]

Since $n_{(k)}(n-k)! = n!$, we have $a_{n,k} = \frac{1}{n_{(k)}} [x^n / n!] y^k = kn^{n-k-1}$.

**3.3.51. Given $C(x) - 1 = x C(x)^2$ for the Catalan generating function $C(x)$,**
\[
[x^n] C(x)^k = \frac{k}{2n + k} \binom{2n + k}{n}
\]

Let $y = x C(x)$. The given equation then becomes $C - 1 = x C^2 = y C$, which we write as $(1 - y) C = 1$. Multiplying by $x$ yields $(1 - y) x C = x$, or $y(1 - y) = x$. Thus $x = y/\phi(y)$, where $\phi(y) = 1/(1 - y)$. Note that $\phi(0) = 1$, so Lagrange Inversion applies. With $C = y/x$ and $h(y) = y^k$, we compute
Section 3.4: Partitions of Integers

3.4.1. The number of partitions of 30 into 1s and 3s using an odd number of 3s is 5. The answer is the number of odd multiples of 3 that are at most 30, since we choose an odd multiple of 3 to be the sum of the 3s used in the partition, and the remainder is achieved using 1s. There are five such multiples: 3, 9, 15, 21, 27.

Using generating functions, we seek the coefficient of $x^{30}$ in the OGF for partitions into 1s and an odd number of 3s. The factor enumerating the contribution by 1s is $\frac{1}{1-x}$. Partitions into 3s correspond to picking one 3 and some number of 6s, so the OGF is $\frac{1}{1-x^3}$. Hence we have $5 = [x^{30}]\frac{1}{1-x} \frac{1}{1-x^3}$.

3.4.2. Rolls of five dice summing to 20. For each die, the factor enumerating the outcomes is $\frac{1-x^2}{1-x}$. When the dice are different colors, we can tell which number comes from each die, so we seek $[x^{20}]\left(\frac{1-x^2}{1-x}\right)^5$. When the dice are identical, we can only tell how many times each number appears. So, we seek partitions into parts that are at most 6, and each part can appear at most five times. Thus we seek $[x^{20}]\prod_{i=1}^{6}\frac{1-x^i}{1-x^i}$. When the dice are distinguishable, we are enumerating compositions; there are more of these.
3.4.3. Positive integer solutions to \( \sum_{i=1}^{n} e_i = k \) such that \( e_1 \geq e_2 \geq \cdots \). The desired objects are the partitions of \( k \) with exactly \( n \) parts. By conjugation, these correspond bijectively to the partitions of \( k \) with largest part \( n \). Hence the OGF, indexed by the sum of the partitions, is \( \prod_{k=0}^{\infty} \frac{1}{1-x^{2^k}} \).

3.4.4. Every nonnegative integer has a unique binary expansion. A binary expansion of \( n \) is a partition of \( n \) into powers of 2, using each power at most once. The generating function for such partitions, indexed by the sum, is \( \prod_{k=0}^{\infty} \frac{1}{1-x^{2^k}} \). We compute

\[
\prod_{k=0}^{\infty} (1 + x^{2^k}) = \prod_{k=0}^{\infty} \frac{1-x^{2^{k+1}}}{1-x^{2^k}} = \frac{1}{1-x}.
\]

The final product telescopes; terms cancel, leaving only the denominator for \( k = 0 \). Since \( [x^n](1 - x)^{-1} = 1 \), there is one such partition for each \( n \).

3.4.5. The number of partitions of \( n+1 \) having no 1 is at most the number of partitions of \( n \) in which the number of copies of 1 is positive and is less than the size of the smallest other part (if such a part exists). For a partition \( \lambda \) of \( n+1 \) having no 1, let \( \phi(\rho) \) be the partition of \( n \) obtained by replacing a smallest part \( k \) by \( k-1 \) copies of 1. Note that \( \phi \) is injective; \( \lambda \) can be recovered from \( \phi(\lambda) \) by combining all \( k \) copies of 1 into one copy of \( k \), which will be a smallest part in \( \lambda \).

3.4.6. Distributions of bridge hands. The 16 distributions of bridge hands without voids are the partitions of 13 with four parts: \( 10 + 1 + 1 + 1, 9 + 2 + 1 + 1, 8 + 3 + 1 + 1, 8 + 2 + 2 + 1, 7 + 4 + 1 + 1, 7 + 3 + 2 + 1, 7 + 2 + 2 + 2, 6 + 5 + 1 + 1, 6 + 4 + 2 + 1, 6 + 3 + 3 + 1, 6 + 3 + 2 + 2, 5 + 5 + 2 + 1, 5 + 4 + 3 + 1, 5 + 4 + 2 + 2, 5 + 3 + 3 + 2, 4 + 3 + 3 + 3 \).

To form a hand with distribution 4333, we select the suit with four cards and then choose the specified number of cards from each suit: \( 4 \binom{13}{4} \binom{13}{3}^3 \). For 5431, there are 24 ways to specify how many cards come from each suit, so the number of hands is \( 24 \binom{13}{5} \binom{13}{4} \binom{13}{3} \binom{13}{1} \). After canceling common factors, the ratio of the first number to the second is 22/27, so the 5431 distribution is more common. Intuitively, this is due to the flexibility in ordering the suits.

3.4.7. With \( a_n \) being the number of ways to toss \( n \) indistinguishable 6-sided dice and obtain an even sum, the generating function \( \sum a_n x^n \) is \( \frac{1}{2}(1-x)^{-3}(1-x)^{-3} + (1+x)^{-3} \). Since the dice are indistinguishable, an outcome is how many of each number is rolled. An even number does not change the parity of the sum, so the usage of 2s, 4s, and 6s is unrestricted. The number of dice used is the index, and there is one way to have \( k \) copies of a particular roll, so the factors encoding the options for 2s, 4s, and 6s all equal \( 1/(1-x) \).

For odd numbers, the total number of odd values in the outcome must be even. The number of ways to get \( 2k \) odd dice altogether is the number of solutions to \( x_1 + x_2 + x_3 = 2k \), where these variables are the numbers of 1s, 3s, and 5s. We eliminate from \( (1-x)^{-3} \) the nonzero coefficients for odd powers. Hence the desired factor is \( \frac{1}{2}\left( (1-x)^{-3} + (1+x)^{-3} \right) \), and the product of this with \( (1-x)^{-3} \) is the generating function.

3.4.8. An arch with \( n \) blocks. Left and right pillars have the same total height and are built from blocks of height 1 or 2, with no blocks of height 2 above blocks of height 1. Let \( a_n \) count the arches using a total of \( n \) blocks.

a) OGF for \( \langle a \rangle \). Since no 2-block is above a 1-block, each design consists of some number of “one 2-block in each pillar” (two blocks total), then some number of “one 2-block opposite two 1-blocks” (left-right or right-left but not both, three blocks total), and then some number of “one 1-block in each pillar” (two blocks total). We choose each type with some multiplicity. We can view contributions of the first and last types as red and blue 2s. Contributions of the middle type are 3 reds or blue 3s, with only one color used. The factors for the first and last types are both \( 1/(1-x^2) \). The middle type has two ways to make any nonempty contribution; hence this factor is \( -1 + 1/2/(1-x^3) \). Thus \( A(x) = \left( \frac{1}{1-x} \right)^2 \left( \frac{2}{1-x^3} - 1 \right) = \frac{1+x^3}{1-x^2(1-x^3)} \) for the generating function.

b) The asymptotic behavior of \( a_n \) is \( a_n = n^2/12 + O(n) \). Using linear factors, the denominator of \( A(x) = (1-x)^3(1+x)^2(1-\omega x)(1+\omega x) \), where \( \pm\omega \) are the zeros of \( 1 + x + x^2 \) (cube roots of 1). After obtaining the partial fraction expansion, a term whose denominator is \( (1-ax)^k \) contributes a multiple of \( \frac{1}{(n-j-1)!} \) \( a^n \) to the coefficient of \( x^n \). Since \( \omega \) has magnitude 1, the term \( B/(1-x)^3 \) contributes a quadratic polynomial in \( n \), and the other terms have growth rates that are polynomials of lower degree. Since its denominator has the highest exponent, in the expansion

\[
\frac{1 + x^3}{1-x^2(1-x^3)} = \frac{B}{(1-x)^3} + \frac{C}{(1+x)^2} + \frac{D}{1-\omega x} + \frac{E}{1+\omega x}
\]

we care only about \( B \). Multiplying by \( (1-x)^3(1+x)^2(1+x+x^2) \) gives all terms on the right \( 1-x \) as a factor except the term involving \( B \). Setting \( x = 1 \) then yields \( 2 = B \cdot 2^2 \cdot 3 \), so \( B = 1/6 \). Since \( \binom{n+2}{2} = n^2/2 + O(n) \), we have \( a_n = n^2/12 + O(n) \).

3.4.9. Combinatorial arguments about partitions of integers. Let \( p_{n,k} \) be the number of partitions of \( n \) into \( k \) parts.
a) $p_{2r+k,r+k}$ is independent of $k$.

Proof 1. The partitions of $n$ with $m$ parts are in 1-1 correspondence with the partitions of $n - m$ with at most $m$ parts by removing the first column of dots in the Ferrers diagram. Thus $p_{2r+k,r+k}$ equals the number of partitions of $r$ with at most $r + k$ parts. As long as $k \geq 0$, this is independent of $r$, since no partitions of $r$ have more than $r$ parts.

Proof 2. Any partition of $2r + k$ into $r + k$ parts must have at least $k$ parts of size 1. Deleting $k$ such parts establishes a bijection from the partitions of $2r + k$ with $r + k$ parts to the partitions of $2r$ with $r$ parts.

b) $p_{r+k,k}$ equals the number of partitions of $r$ into parts of size at most $k$.

Every partition of $r + k$ into $k$ parts becomes a partition of $r$ into at most $k$ parts by removal of the first column of the Ferrers diagram. Distinct partitions have distinct images, and the map is reversible by adding a first column of length $k$, so the map is a bijection.

c) $p_{r+k,k}$ equals the number of partitions of $r + \left(\frac{k+1}{2}\right)$ into $k$ distinct parts. Any partition $\lambda$ of $r + k$ with $k$ parts can be converted into a partition of $r + k + \left(\frac{k}{2}\right)$ with $k$ distinct parts by adding a leading triangle of $\left(\frac{k}{2}\right)$ dots to the Ferrers diagram, changing $\lambda_i$ to $\lambda_i + k - i$. Distinct partitions have distinct images. The map is reversible, since in a partition with $k$ distinct parts, the $i$th part is larger than $i - k$. Hence the map is a bijection.

3.4.10. Let $p_{n,k}$ be the number of partitions of $n$ into $k$ parts.

a) $p_{n,k} = p_{n-1,k-1} + p_{n-k,k}$. In a partition $\lambda$ of $n$ into $k$ parts, the smallest part is 1 or is at least 2. Those of the first type correspond bijectively to partitions of $n - 1$ with $k - 1$ parts, by removing the last part. Those of the second type correspond bijectively to partitions of $n - k$ with $k$ parts, by removing the first column of the Ferrers diagram. Hence there are $p_{n-1,k-1}$ partitions of the first type and $p_{n-k,k}$ of the second type.

The argument for the recurrence is valid when $n \geq k > 0$, given the initial conditions $p_{n,0} = \delta_{n,0}$ and $p_{n,k} = 0$ when $n < k$.

b) $A_k(x) = \frac{1}{1-x}A_{k-1}(x)$ for $k \geq 1$, with $A_0(x) = 1$, where $A_k(x) = \sum_{k=0}^{\infty} p_{n,k}x^n$. Since $p_{n,0} = \delta_{n,0}$, we have $A_0(x) = 1$.

Multiply the recurrence in (a) by $x^n$ and sum over $n \geq k$. Since the coefficients of earlier terms in $A_k(x)$ are 0 when $k > 0$, introducing the generating functions yields $A_k(x) = xA_{k-1}(x) + x^kA_k(x)$ when $k \geq 1$, with no missing initial terms. Collecting like terms yields $A_k(x) = \frac{1}{1-x}A_{k-1}(x)$.

$A_k(x) = x^k\prod_{i=1}^{k}(1-x^{-i})^{-1}$. Starting with $A_0(x) = 1$, we accumulate a factor of $x$ in the numerator and $1 - x^i$ in the denominator for $1 \leq i \leq k$ when iterating the recurrence. Using the recurrence, we have obtained the generating function for partitions with $k$ parts. Notice that it is also the generating function for partitions with largest part $k$.

Section 3.4: Partitions of Integers

3.4.11. a) If $p_{j,k}(n)$ is the number of partitions of $n$ that have $j$ parts and have $k$ as the largest part, then $\sum_{n=0}^{\infty} p_{j,k}(n) = \left(\frac{i+j-2}{i-1}\right)$.

Proof 1 (Ferrers diagrams). With the upper-left dot in the Ferrers diagram placed at $(0,0)$, the bottom of the first column is $(0,-j+1)$, and the right end of the first row is $(k-1,0)$. Thus the outside of the diagram traces a lattice path from $(0,-j+1)$ to $(k-1,0)$. Every such lattice path arises from a partition in this way, so we have a bijection. These lattice paths move $k-1$ steps rightward and $j-1$ steps upward, so the number of them is $\left(\frac{i+j-2}{i-1}\right)$.

Proof 2a (multisets). Since we sum over $n$, we can only that we use $j$ parts altogether and that the largest one is $k$; the sum of the parts is unimportant. Thus it suffices to count multisets of size $j$ from $k$ types, with the last type used at least once. These correspond to multisets of size $j - 1$ from $k$ types, and the formula is $\left(\frac{i+j-2}{i-1}\right)$.

Proof 2b (generating functions). We build an OGF for partitions with largest part $k$, indexed by the number of parts (not the sum). The options for choosing parts equal to $r$ are modeled by $\sum_{i \geq 0} x^i$ for $r < k$ and by $\sum_{i \geq 1} x^i$ for $r = k$. The generating function is $\frac{x}{(1-x)^r}$, and $[x^j]\frac{x}{(1-x)^r} = \frac{[x^{j-1}]}{(1-x)^r} = (r-1)!$.

3.4.12. The number of partitions (of any integer) whose Ferrers diagram fits in an $i$-by-$(n-i)$ rectangle is one more than the number of permutations of $[n]$ having exactly one descent, at position $i$. The boundary of such a Ferrers diagram corresponds to a lattice path from $(0,-i)$ to $(n-i,0)$, where the path moves up the vertical axis to reach the first nonzero row (if the partition has fewer than $i$ parts) and along the horizontal axis to $(n-i,0)$ (if the largest part is less than $n-i$). There are $\binom{n}{i}$ such paths.

A permutation with one descent at position $i$ has two runs, of lengths $i$ and $n-i$. Such a permutation is determined by choosing $i$ elements of $[n]$ to form the first run (each run is listed in increasing order). However, the permutation in which the $i$ smallest elements are chosen for the first run must be excluded, since then the two runs combine into one run.

3.4.13. Algebraic proof of $\prod_{i=1}^{\infty}(1+x^i) = \prod_{i=1}^{\infty}(1-x^{2i-1})^{-1}$. Multiplying top and bottom by $\frac{1}{1-x^i}$ for each $i$ and then canceling the factors with even exponents, we compute

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

3.4.14. The number of partitions of $n$ having $k$ even parts is the same as the number of partitions of $n$ in which the largest repeated part is $k$. Fixing $k$, we find both generating functions, indexed by $n$. 

A partition with \( k \) as the largest repeated part can have parts less than \( k \) with any multiplicity, \( k \) at least twice, and parts larger than \( k \) at most once. Hence the generating function is

\[
\left( \prod_{i=1}^{k-1} \frac{1}{1 - x^i} \right) \frac{x^{2k}}{1 - x^{2k}} \prod_{i=k+1}^{\infty} (1 + x^i).
\]

For a partition with exactly \( k \) even parts, consider the even and odd parts separately. In the conjugate of the partition using the even parts, each part occurs an even number of times, and the largest part is \( k \) (occurring at least twice). Hence we can view this conjugate as a partition into even parts at most \( 2k \) with at least one occurrence of \( 2k \). The remaining parts form a partition into odd parts with unrestricted multiplicity. Hence the generating function is

\[
\left( \prod_{i=1}^{k-1} \frac{1}{1 - x^{2i}} \right) \frac{x^{2k}}{1 - x^{2k}} \prod_{i=k+1}^{\infty} 1 - x^{2i-1}.
\]

Collecting factors shows that both OGFs equal \( x^{2k} \prod_{i=1}^{\infty} \frac{1 - x^{2k+i}}{1 - x^i} \).

**Comment:** A bijective proof is obtained by specializing the bijective proof in part (b) of Exercise 3.4.15 to \( d = 2 \).

**3.4.15. Glaisher’s Theorem and generalization.**

a) For \( d \in \mathbb{N} \), the number of partitions of \( n \) with no part divisible by \( d \) equals the number of partitions of \( n \) in which no part appears at least \( d \) times. We map a partition with no part divisible by \( d \) into a partition in which no part appears at least \( d \) times by iteratively combining \( d \) equal parts into one part until no instance of \( d \) identical parts remains.

Each part in the resulting partition is a power of \( d \) times a number not divisible by \( d \), say \( d^k j \), obtained by combining \( d^k \) copies of \( j \) from the original partition, because no part in the original partition was divisible by \( d \). Since every positive integer is expressible as \( d^k j \) in a unique way, for each partition with no part appearing \( d \) times there is exactly one partition into parts not divisible by \( d \) that maps to it in this way; break each part of the form \( d^k j \) into \( d^k \) copies of \( j \). Hence the map is a bijection.

b) The number of partitions of \( n \) having exactly \( k \) parts divisible by \( d \) is the same as the number of partitions of \( n \) in which \( k \) is the largest part that occurs at least \( d \) times.

**Proof 1** (bijection). When \( n = 0 \), the claim is trivial, so assume \( n > 0 \). We construct a bijection. Let \( \lambda \) be a partition of \( n \) having exactly \( k \) parts divisible by \( d \). Let \( A \) consist of all the parts in \( \lambda \) that are divisible by \( d \), and let \( B \) consist of the other parts (\( A \) is empty when \( k = 0 \)). Map \( A \) to its conjugate partition \( A^* \), in which the largest part is \( k \) and occurs at least \( d \) times. The bijection in part (a) maps \( B \) (containing no part divisible by \( d \)) to a partition \( B' \) in which no part occurs at least \( d \) times. Combine \( A^* \) and \( B' \) to form the image of \( \lambda \).

To invert the map, we split a partition \( \mu \) in which \( k \) is the largest part occurring at least \( d \) times into \( A^* \) and \( B' \). For each part \( i \) with multiplicity \( m_i \) times in \( \mu \), put \( d \left\lfloor \frac{m_i}{d} \right\rfloor \) of the copies of \( i \) into \( A^* \). Put the remaining copies into \( B' \); no part occurs \( d \) or more times in \( B' \). This is the only way to split \( \mu \) into two partitions in the specified families. Inverting the two maps separately and recombining the output yields the only partition \( \lambda \) that maps to \( \mu \) as described above.

**Proof 2** (generating functions). Fixing \( k \) and \( d \), we find the generating functions of the two quantities, indexed by \( n \). When \( k \) is the largest part occurring at least \( d \) times, each smaller part is unrestricted, \( k \) appears at least \( d \) times, and larger parts appear fewer than \( d \) times. Hence the generating function is

\[
\left( \prod_{i=1}^{k-1} \frac{1}{1 - x^i} \right) \frac{x^{dk}}{1 - x^{dk}} \prod_{i=k+1}^{\infty} (1 + x^i + \cdots + x^{(d-1)i}).
\]

When exactly \( k \) parts are divisible by \( d \), consider the parts divisible by \( d \) and the others separately. In the conjugate of the partition \( \lambda \) consisting of the parts divisible by \( d \), each part occurs a multiple of \( d \) times, and the largest part is \( k \) (occurring at least \( d \) times). We view it as using parts that are multiples of \( d \), the largest one being \( dk \). For the remaining parts, not divisible by \( d \), there is no restriction on usage. Hence the full generating function is

\[
\left( \prod_{i=1}^{k-1} \frac{1}{1 - x^{di}} \right) \frac{x^{dk}}{1 - x^{dk}} \prod_{i=k+1}^{\infty} \frac{1 - x^{di}}{1 - x^i}.
\]

Collecting factors shows that both OGFs equal \( x^{dk} \prod_{i=1}^{\infty} \frac{1 - x^{dk+i}}{1 - x^i} \).

**3.4.16. For** \( n \geq 2 \), **exactly half of the partitions of** \( n \) **into powers of 2 have an even number of parts.** Call a partition into powers of 2 a 2-power partition, and let its parity be the parity of the number of parts.

**Proof 1** (induction on \( n \)). The claim holds by inspection for \( n = 2 \). For \( n > 2 \), group the 2-power partitions of \( n \) into two sets depending on whether 1 is a part. Those in which 1 is a part correspond bijectively to 2-power partitions of \( n - 1 \), with opposite parity. By the induction hypothesis, the number of 2-power partitions with each parity is the same.

Similarly, 2-power partitions not using \( 2^0 \) correspond bijectively to 2-power partitions of \( n/2 \) (by halving each part), since half a positive power
of 2 is a power of 2. In this case the number of parts is preserved by the transformation, and again the induction hypothesis applies.

**Proof 2** (bijection). Let $k$ be the largest part in a 2-power partition $\lambda$ of $n$. Define $\tau(\lambda)$ by splitting $k$ into two copies $k/2$ if $k$ appears only once in $\lambda$, or combining two copies of $k$ into one copy of $2k$ if $k$ appears more than once. If $n \geq 2$, then $\tau$ is well-defined and produces another 2-power partition. Because the next largest part is at most $k/2$, the resulting $\tau(\lambda)$ changes the property of whether the largest part is repeated or not. Hence $\tau$ pairs 2-power partitions having unique largest part with 2-power partitions having repeated largest part. Also $\tau$ changes the parity of the number of parts. Since $\tau$ has no fixed points, we obtain the desired equality.

**Proof 3** (generating functions). The generating function for 2-power partitions is $\prod_k (1-x^{2^k})^{-1}$. To compare parity, we make each 2-power partition of $n$ contribute $+1$ or $-1$ to the coefficient of $x^n$ according to its parity. To do this, change the factor for parts of size $2^k$ to $(1+x^{2^k})^{-1}$. The expansion is $\sum_j (-1)^j x^{2^j}$. When we make such a choice from each factor, the product of the coefficients of the terms chosen is $(-1)^m$, where altogether $m$ parts were chosen. Hence each partition makes the desired contribution, and the coefficient of $x^n$ is the number of even 2-power partitions of $n$ minus the number of odd 2-power partitions of $n$.

In $\prod_k (1+x^{2^k})^{-1}$, we have a term for each set of powers of two. Because every nonnegative integer has a unique expression as a sum of distinct powers of two, $\prod_k (1+x^{2^k}) = (1+x+x^2+\cdots)$. Taking reciprocals yields $\prod_k (1+x^{2^k})^{-1} = 1-x$. For $n \geq 2$, the coefficient of $x^n$ is 0, and hence the numbers of odd and even 2-power partitions of $n$ are the same.

**3.4.17.** $a_n = b_n$ for $n \geq 1$, where $a_n$ is the number of partitions of $n$ in which each part is at least as big as the sum of all subsequent parts, and $b_n$ is the number of partitions of $n$ into powers of 2. Since $a(1) = b(1) = 1$, it suffices to show that $\langle a \rangle$ and $\langle b \rangle$ satisfy the same recurrence.

When $n$ is odd, the first part in an acceptable partition of $n$ is more than $n/2$, and decreasing it by 1 gives an acceptable partition of $n-1$. Thus $a_{2m+1} = a_{2m}$ for $m \geq 1$. If $n$ is even, then either the first part is more than $n/2$, and decreasing it by 1 gives an acceptable partition of $n-1$, or the first part equals $n/2$, and the rest is an acceptable partition of $n/2$. Thus $a(2m) = a(2m-1) + a(m)$ for $m \geq 1$.

Call a partition into powers of 2 a 2-power partition. When $n$ is odd, a 2-power partition of $n$ must have 1 as a part, and the rest is a 2-power partition of $n-1$. Thus $b(2m+1) = b(2m)$ for $m \geq 1$. If $n$ is even, then a 2-power partition of $n$ either contains 1 plus a 2-power partition of $n-1$, or all the parts are even and the partition is a doubling of a 2-power partition of $n/2$. Thus $b(2m) = b(2m-1) + b(m)$ for $m \geq 1$.

**3.4.18.** The number $a_n$ of congruence classes of triangles with integer-length sides and perimeter $n$.

a) For $k \in \mathbb{N}_0$, $a_{2k}$ is the number of partitions of $k$ into three parts. The side-lengths of a triangle with perimeter $2k$ partition $2k$ into three parts satisfying the strict triangle inequality. Equivalently, the largest of the three parts is less than $k$. Hence we count partitions of $2k$ into three parts of size less than $k$. We establish a bijection to show that this is also the number of unrestricted partitions of $k$ into three parts.

Given a partition of $k$ into $x, y, z$, all positive, the numbers $x+y, x+z, y+z$ partition $2k$, with each part less than $k$. For any such partition $a, b, c$ of $2k$, we can retrieve the unique partition that maps to it by solving the system $x+y = a, x+z = b, y+z = c$. The coefficient matrix is nonsingular, and the solution is $(x, y, z) = (k-c, k-b, k-a)$.

b) For $k \geq 2$, $a_{2k-3} = a_{2k}$. Every triangle with perimeter $2k-3$ becomes a triangle with perimeter $2k$ when 1 is added to each side-length. Also, every triangle with perimeter $2k$ arises (exactly once) in this way, because an integer triangle with even perimeter cannot have a side of length 1. If the other lengths were $a$ and $b$, with $a \geq b$, then the triangle inequality requires $b+1 > a$, and hence $b \geq a$. Now $b \geq a$, which contradicts that the perimeter is even.

c) The OGF $A(x)$ for $\langle a \rangle$ is $x^2/[1-x^2(1-x^3)(1-x^4)]$. The generating function for partitions into three parts is given by $B(y) = y^3/[1-y(1-y^2)(1-y^3)]$. By part (a), we want $[x^{2k}]A(x) = [y^k]B(y)$, so setting $y = x^2$ in $B(y)$ will give us the even terms of $A(x)$. By part (b), the generating function for the odd terms is $x^{-3}B(x^2)$. Therefore,

$$A(x) = (1+x^{-3})B(x^2) = \frac{x^6+x^3}{(1-x^2)(1-x^4)(1-x^6)} = \frac{x^3}{(1-x^3)(1-x^4)(1-x^6)}.$$
Chapter 3: Generating Functions

exactly 1 more than the third don’t arise from a “triangular” partition of \( n - 3 \) by adding one to each part. The third part must be \((n-1)/2\), and there are \( \lfloor (n+1)/4 \rfloor \) ways to choose the other two parts summing to \((n+1)/2\). Hence \( a_n = a_{n-3} + \lfloor (n+1)/4 \rfloor \) when \( n \) is odd. The recurrence has degree 3, so we also specify \( a_0 = a_1 = a_2 = 0 \).

b) For \( n \geq 3 \), \( a_n = a_{n-3} + \frac{1}{n+1}(n+i+1) \). We first transform part (a) into a single recurrence as \( a_n = a_{n-3} + \frac{3}{2}(1 - (-1)^n) \lfloor n+1/2 \rfloor \). We are concerned about describing \( \lfloor n+1/2 \rfloor \) without the floor function, but only when \( n \) is odd; when \( n \) is even whatever we write with be multiplied by 0. We have \( \frac{3}{2} = \frac{n+1}{2} \) when \( n \equiv 3 \pmod{4} \), and \( \frac{n+1}{2} = \frac{n+1}{2} \) when \( n \equiv 1 \pmod{4} \). An exponential that is +1 when \( n \equiv 3 \pmod{4} \) and -1 when \( n \equiv 1 \pmod{4} \) is \( i^{n+1} \). Thus \( n + i^{n+1} \) is the desired factor in the numerator, and we have rewritten the recurrence in the desired form.

c) The generating function is \( x^3/[(1 - x^2)(1 - x^3)(1 - x^4)] \). Given the recurrence in (b), we multiply by \( x^n \), sum over \( n \geq 3 \), and adjust for the missing three terms (all zero) to obtain

\[
A(x) = x^3 A(x) + \frac{1}{4} \left( \sum_{n \geq 3} \frac{1 - (-1)^n}{2}(n+i+1)x^n \right).
\]

The factor \( \frac{1-(-1)^n}{2} \) cancels the terms for even \( n \) and keeps those for odd \( n \). Furthermore, the term for \( n = 1 \) is 0. Thus for the summation it suffices to evaluate \( B(y) = \sum_{n \geq 0} (n+i+1) y^n \) and take \( (B(x) - B(-x))/2 \).

By differentiating the geometric series, we have \( \sum_{n \geq 0} ny^n = \frac{y}{(1-y)^2} \). Thus \( B(y) = \frac{y}{(1-y)^2} + \frac{1}{1-iy} \). Now

\[
4(1 - x^3)A(x) = \frac{1}{2} \left( \frac{x}{(1-x)^2} - \frac{-x}{(1+x)^2} \right) + \frac{1}{2} \left( \frac{i}{1-ix} - \frac{i}{1+ix} \right).
\]

After putting the terms on the right over a common denominator, the right side simplifies to \( 4x^3/[(1 - x^2)(1 - x^4)] \), which yields \( A(x) = x^3/[(1 - x^3)(1 - x^3)(1 - x^4)] \).

3.4.20. The number \( a_n \) of congruence classes of triangles with integer-length sides and perimeter \( n \). We want to count expressions of \( n \) as a sum of three numbers \( x + y + z \) satisfying the strict triangle inequality, \( y + z > x \). Triangles with \( y \neq z \) arise by adding 1 to the two smaller parts of a triangle on \( n - 2 \). There are \( \lfloor n/3 \rfloor - \lfloor n/4 \rfloor \) triangles with \( y = z \), since there is one such triangle for each \( y \) with \( n/4 < y \leq n/3 \). Hence \( a_n = a_{n-2} + \lfloor n/3 \rfloor - \lfloor n/4 \rfloor \).

To solve this recurrence, we need the OGF for \( \lfloor n/k \rfloor \). With \( n = jk+i \),

\[
\sum_{n \geq 0} \left\lfloor \frac{n}{k} \right\rfloor x^n = \sum_{j \geq 0} jx^{jk} \sum_{i=0}^{k-1} x^i = \frac{x^k}{(1-x)^2} \sum_{i=0}^{k-1} x^i = \frac{x^k}{(1-x)(1-x^k)}.
\]

We get \( A(x) = x^3 A(x) + \sum_{n \geq 0} \lfloor n/3 \rfloor x^n - \sum_{n \geq 0} \lfloor n/4 \rfloor x^n \), and \( (1 - x^2)A(x) = \frac{x^5}{(1-x^2)(1-x^7)} - \frac{x^5}{(1-x^2)(1-x^7)}, \) which simplifies to \( A(x) = \frac{x^5}{(1-x^2)(1-x^7)(1-x^7)} \).

3.4.21. The identity

\[
\sum_{n=1}^{\infty} \frac{x^{n(n+1)/2}}{1-x^n} = \sum_{n=1}^{\infty} \frac{x^n}{1-x^{2n}}.
\]

The right side is the generating function \( \sum a_m x^m \), where \( a_m \) is the number of partitions of \( m \) into equal odd parts; the summand for \( n \) enumerates such partitions where \( n \) equal odd parts are used. The left side is the generating function \( \sum b_m x^m \), where \( b_m \) is the number of partitions of \( m \) into consecutive parts; the summand for \( n \) counts those where \( n \) parts are used.

Viewed directly, \( a_m \) is the number of odd divisors of \( m \). On the other hand, the sum of \( n \) consecutive integers starting with \( j \) is \( jn + (n-1)/2 \), which equals \( n(n+2j-1)/2 \). The value \( b_m \) is the number of ways to obtain \( m \) as a sum. When the sum is \( m \), there is exactly one choice of \( n \) and \( j \) such that \( n(n+2j-1) = 2m \) for each expression of \( 2m \) as a product of factors of opposite parity. There is one such product for each odd divisor of \( m \), and hence \( a_m = b_m \).

3.4.22. If \( A(x) = \sum_{n=0}^{\infty} a_n x^n/n! \) and \( B(x) = \sum_{n=1}^{\infty} b_n x^n/n! \), then

\[
A(B(x)) = a_0 + \sum_{n=1}^{\infty} \left( \sum_{k=1}^{n} a_k \cdot \sum_{i=1}^{\lfloor n/i \rfloor} \prod_{j=1}^{\lfloor n/i \rfloor} b_{i^j} \right) \frac{x^n}{n!},
\]

where the inner sum is over partitions \( \lambda \) of \( n \) with \( k \) parts, and \( e_i \) is the number of parts of size \( i \) in \( \lambda \). We want to expand the composition of two EGFs and collect the contributions to the coefficient of \( x^n/n! \). For \( n = 0 \), the coefficient is \( a_0 \), since \( B(0) = 0 \). For \( n \geq 1 \), contributions can come from plugging \( B(x) \) in for \( y \) in \( a_k y^k/k! \) when \( 1 \leq k \leq n \). The \( k \) factors of \( y \) contribute powers of \( x \) summing to \( n \).

When summing the values of \( n \) each \( \lambda \) has, the product of the coefficients of the powers of \( x \) is \( \prod \frac{b_{i^j}}{\Pi_{i^j}} \); when \( e_i \) parts equal \( i \), the factorials in the denominator are grouped as \( \Pi(\lfloor n/i \rfloor)! \). Also, the number of ways to assign the powers to the \( k \) factors is the number of arrangements of \( k \) things consisting of \( e_i \) things of type \( i \). This number is the multinomial coefficient \( \binom{n}{e_1 \ldots e_k} \), which equals \( \frac{\Pi_{i^j}}{\Pi_{i^j}} \). Finally, because we seek the coefficient of \( x^n/n! \) rather than the coefficient of \( x^n \), we must multiply by \( n! \).

3.4.23. The total number of 2s in all partitions of \( n \) equals the total number of nonrepeated parts over all partitions of \( n - 1 \). The number of partitions of \( n \) with exactly \( k \) copies of 2 is the coefficient of \( x^n \) in \( x^{2k}(1 - x^2)f(x) \),
where \( P(x) = \prod_{i \geq 0} \frac{1}{1 - x^i} \). Multiplying by \( k \) and summing over \( k \) yields the OGF for the total number of 2s, which we simplify as follows:

\[
\sum_{k=1}^{\infty} k x^{2k}(1 - x^2)P(x) = \frac{x^2}{(1 - x^2)^2} (1 - x^2)P(x) = \frac{x^2}{(1 - x^2)} P(x).
\]

On the other hand, the number of partitions of \( n - 1 \) containing the singleton \( k \) is the coefficient of \( x^{n-1} \) in \( x^k(1 - x^k)P(x) \). Summing over \( k \) and multiplying by \( x \) to obtain the OGF, we compute

\[
x \sum_{k=1}^{\infty} x^k(1 - x^k)P(x) = x \left( \frac{x}{1-x} - \frac{x^2}{1-x^2} \right) P(x) = \frac{x^2}{(1-x^2)} P(x).
\]

With equal generating functions, the two sequences are equal.

### 3.4.24. \( \sum_{k=0}^{\infty} x^{2k+1} \prod_{j=k+1}^{\infty} (1 + x^{2j-1}) = \frac{1}{1-x} \). As a formal power series, this is the statement that every positive integer has a unique expression as a sum of distinct numbers that are one less than a positive power of 2, except that the smallest number used (expressed as \( 2k+1 \)) can be repeated. We establish this by partitioning the positive integers into blocks of the form \([2^k-1, 2^{k+1}-2]\) for \( k \geq 1 \) and using induction on \( k \).

For \( k = 1 \), the block is \([1, 2]\), and the partitions are 1 and 1 + 1. For the block \([2^k-1, 2^{k+1}-2]\), one obtains such partitions by using \( 2^k - 1 \) and \( 2^{k+1} - 1 \) for the two extreme elements and adding the element \( 2^k - 1 \) to the partitions already found for 1 through \( 2^k - 2 \). For uniqueness, note that the largest number that can be so partitioned without using a number at least \( 2^k - 1 \) is in fact \( 2^k - 2 \). Since also \( 2^{k+1} - 1 \) is too big for numbers in the block \([2^k-1, 2^{k+1}-2]\), partitions of numbers in this block must use one copy of \( 2^k - 1 \), and then uniqueness follows inductively.

### 3.4.25. \( 2^{n-1} = \sum_1^{\lambda_j} \prod_{j \geq 1} (\lambda_j / \lambda_{j+1}) \), where the sum is over all partitions of \( n \), with trailing 0s irrelevant. The left side is the number of compositions of \( n \), obtained by placing dividers in a subset of the spaces in a row of \( n \) dots.

For the right side, view a composition \((a_1, \ldots, a_k)\) of \( n \) as columns of boxes sitting on the horizontal axis, with \( a_i \) dots in column \( i \). Counting the dots by rows yields a partition of \( n \). To form a composition associated with a partition \( \lambda \), place \( \lambda_1 \) boxes in the bottom row, left-justified. For \( j \geq 1 \), the boxes in row \( j+1 \) can be placed in any \( \lambda_1+j \) of the \( \lambda_j \) columns having boxes in row \( j \). Hence there are \((\lambda_j / \lambda_{j+1})\) ways to place the boxes in row \( j + 1 \), no matter how the earlier boxes were placed.

### 3.4.26. \( \sum_{m=0}^{\infty} \prod_{i=0}^{m} (\lambda_i / \lambda_{i+1}) = 2^n - F_n \), where \( m(\lambda) \) and \( m_i(\lambda) \) denote the total number of parts and number of parts equal to \( i \) in a partition \( \lambda \), the sum is over all partitions of \( n + 1 \) having at least one 1, and \( F_n \) is the \( n \)th classical Fibonacci number.

**Proof 1** (counting arguments). For each partition, the summand counts the ways to permute the parts, so the sum is the number of compositions of \( n + 1 \) having at least one 1. The number of compositions of \( n + 1 \) is \( 2^n \), so it suffices to prove that the number \( a_n \) of compositions of \( n + 1 \) with no 1 is \( F_n \). This is clear for \( n = 0 \) and \( n = 1 \). When \( n \geq 2 \), these compositions have last part 2 or greater than 2. Deleting the last part shows that there are \( a_{n-2} \) of the first type, and subtracting 1 from the last part shows that there are \( a_{n-1} \) of the second type. By induction, \( a_n = a_{n-1} + a_{n-2} = F_n \).

**Proof 2** (generating functions). We generate the partitions by number of parts. By eliminating those having no 1 from the series for all partitions, the sum is the coefficient of \( x^{n+1} \) in \( f(x) \), where

\[
f(x) = \sum_{m=0}^{\infty} ((x + x^2 + \cdots)^m - (x^2 + x^3 + \cdots)^m) = \sum_{m=0}^{\infty} \left( \left( \frac{x}{1-x} \right)^m - \left( \frac{x^2}{1-x^2} \right)^m \right) \]
\[
= \frac{1-x}{1-2x} - \frac{1-x}{1-x-x^2} = \frac{x(1-x)^2}{(1-2x)(1-x-x^2)}.
\]

Hence we seek the coefficient of \( x^n \) in

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

The coefficient subtracted in the second term is the number of \( 1, 2 \)-lists with sum \( n-1 \), which is \( F_n \), so the answer is \( 2^n - F_n \).

**3.4.27. Enumerating the parts over all partitions.**

a) With \( p(n, m) \) denoting the number of partitions of \( n \) with \( m \) parts, and \( H(x, t) = \sum_{m=0}^{\infty} p(n, m)x^mt^m \), we have \( H(x, t) = \prod_{k=1}^{\infty} \frac{1}{1-x^k t} \). To form a generating function where the exponent on \( t \) records the number of parts, we introduce a factor for each part-size. Each time \( k \) is used as a part, we contribute \( k \) to the total in the partition and 1 to the number of parts. The choice of the number of times \( k \) is used is made for each \( k \); thus \( H(x, t) = \prod_{k=1}^{\infty} \frac{1}{1-x^k t} \), with each partition of each integer contributing to the proper coefficient.

b) With \( f(n) \) denoting the total number of parts over all partitions of \( n \), and \( F(x) = \sum_{n=0}^{\infty} f(n)x^n \), we have \( F(x) = \sum_{k \geq 1} \frac{x^k}{1-x-x^k} P(x) \), where \( P(x) \) is the
OGF for all partitions (that is, $P(x) = \prod_{k=1}^{\infty} \frac{1}{1-x^k}$). With $H$ defined as in part (a), note that $H(x, 1) = P(x)$. Since each partition of $n$ with $m$ parts contributes $m$ to the coefficient of $x^m$ in $P_2(x)$, we have
\[ F(x) = \sum_{m,n} mp(n, m)x^n = \frac{\partial}{\partial t} H(x, t)|_{t=1} = \sum_{k=1}^{\infty} \frac{1}{1-x^k} \sum_{k=1}^{\infty} \frac{1}{1-x^r}. \]

3.4.28. The OGF $A_k(x)$ for double-partitions $(\lambda, \mu)$ with $\lambda$ and $\mu$ each having at most $k$ parts is given by $A_k(x) = \prod_{i=1}^{k} (1 + x^{2i-1}) / \prod_{i=1}^{2k} (1 - x^i)$. A double-partition of $n$ is an ordered pair $(\lambda, \mu)$ such that $\lambda$ and $\mu$ are partitions such that $\min(\lambda_i, \mu_i) \geq \max(\lambda_{i+1}, \mu_{i+1})$ for all $i$ (trailing zeros added as needed) and $\sum \lambda_i + \sum \mu_i = n$. Note that the numbers of nonzero parts in $\lambda$ and $\mu$ must differ by at most 1.

Let $a_{k,n}$ be the number of double-partitions of $n$ in which $\lambda$ and $\mu$ each have at most $k$ parts, so $A_k(x) = \sum_{n=0}^{\infty} a_{k,n} x^n$. Note that $a_{0,n} = \delta_{0,n}$, so $A_0(x) = 1$, giving a basis for induction on $k$. Let $b_{k,n}$ be the number of double-partitions in which $\lambda$ has $k$ parts and $\mu$ has $k-1$ parts. These numbers satisfy the following system of recurrences for $k > 0$:
\[ a_{k,n} = a_{k-1,n} + a_{k,n-2k} + 2b_{k,n}, \]
\[ b_{k,n} = b_{k,n-2k+1} + a_{k-1,n-2k+1}. \]

In the first recurrence, the contributions to $a_{k,n}$ count the double-partitions where both partitions have at most $k$ parts, where both partitions have exactly $k$ parts, and where one has $k$ parts and the other has $k-1$ parts, respectively. In the second recurrence, the contributions to $b_{k,n}$ count those with $\lambda_k > 1$ and with $\lambda_k = 1$, respectively, since $\lambda_k > 1$ implies that all parts of both partitions are greater than 1.

In addition to $a_{0,n} = \delta_{0,n}$, the boundary conditions are that $a_{n,k} = b_{n,k} = 0$ when $n$ or $k$ is negative. As noted above, $A_0(x) = 1$; also $B_0(x) = 0$, where $B_k(x) = \sum b_{k,n} x^n$.

Using the generating function method, multiply each recurrence by $x^n$ and sum over $n \geq 0$. The recurrences become
\[ (1 - x^{2k}) A_k(x) = A_{k-1}(x) + 2B_k(x), \]
\[ (1 - x^{2k-1}) B_k(x) = x^{2k-1} A_{k-1}(x). \]

Eliminating $B_k(x)$ yields
\[ A_k(x) = \frac{1 + x^{2k-1}}{1 - x^{2k}} A_{k-1}(x) = \frac{1 + x^{2k-1}}{1 - x^{2k}} A_{k-1}(x), \]
which yields the claim immediately by induction on $k$.

Comment. This solution is a bit disappointing. The generating function is the product of the generating functions for partitions into at most $2k$ parts and partitions into distinct odd parts that are less than $2k$. Hence one should be able to prove the result by a direct bijection converting a double-partition of $n$ into a pair of partitions with total sum $n$ having the two types described here.

3.4.29. If $\lambda$ and $\mu$ are integer partitions with conjugates $\lambda^*$ and $\mu^*$, then $\sum_{i,j} \min(\lambda_i, \mu_j) = \sum_k \lambda_k \mu_k^*$. Consider the Ferrers diagrams for $\lambda$ and $\mu$. The right side of the identity counts the pairs of dots selected from the same column in the two diagrams. The left side counts the same set, grouped by the ordered pair of rows that the two dots lie in. That is, if $\min(\lambda_i, \mu_j) = r$, then there are exactly $r$ pairs of same-column dots that come from row $i$ in $\lambda$ and row $j$ in $\mu$ (their columns are 1 through $r$).

Alternative phrasing. For a more geometric phrasing, consider an arrangement of unit cubes, where the height of the column above $(i, j)$ in the plane is $\min(\lambda_i, \mu_j)$. The left side counts the total number of cubes by column. The right side also counts the cubes, grouped by horizontal layers. Each layer is a rectangle of cubes. The length of one side of the $k$th layer is the number of parts of $\lambda$ that are at least $k$; this is exactly $\lambda_k$. Similarly, the length of the other side is $\mu_k$. Hence the $k$th layer has $\lambda_k \mu_k^*$ cubes, and the sum counts them all.

Comment: The special case $\lambda = \mu = (n, n-1, \ldots, 1)$ is the identity $\sum_{i=1}^{n} \sum_{j=1}^{n} \min(i, j) = \sum_{k=1}^{n} k^2$ of Exercise 1.2.13.

3.4.30. If $n, k \in \mathbb{N}$ and $q_k = \lfloor n/k \rfloor$, then $f(k) \geq f(k+1)$, where $f(k) = \left(\frac{q_k}{2}\right) k + r_k q_k$ and $r_k = n - k q_k$. We can partition $n$ into $k$ parts, where $r_k$ parts have size $q_k + 1$ and the rest have size $q_k$. Form the Ferrers diagram of this partition. The value of $f(k)$ is the number of pairs of dots in this diagram that lie in the same row.

To obtain the diagram for $k+1$ from the diagram for $k$, successively move one dot from the last column to a new last row, until the length of the new row catches up and there are again at most two distinct column lengths. Each such movement shifts a dot from its current row to a shorter row. Thus the number of pairs lost is at least the number of pairs gained in the new position. Reaching $f(k+1)$ is a combination of such shifts, so $f(k+1) \leq f(k)$.

3.4.31. Proof of Euler’s second identity.

$$(1 + x^2)(1 + x^4)(1 + x^6) \cdots = 1 + \sum_{k \geq 1} \frac{x^{k(k+1)}}{(1 - x^2)(1 - x^4)(1 - x^6) \cdots (1 - x^{2k})}$$

Proof. Using Euler’s first identity. The left side enumerates partitions into distinct even parts. Partitions of $n$ with $k$ distinct even parts
correspond to partitions of \( n - k \) into \( k \) distinct odd parts, by subtracting 1 from each part. Theorem 3.4.14 shows that these correspond to self-conjugate partitions of \( n - k \) with Durfee square of size \( k \). These are enumerated by the coefficient of \( x^{n-k} \) in \( x^{k^2}/\prod_{i=1}^{k}(1-x^{2i}) \), so we multiply by an additional \( x^k \) to make them contribute to the coefficient of \( x^n \) as desired.

**Proof.** (mimicking Euler’s first identity). Bend each even part into an L-shape with one more point in the row than in the column. Instead of the largest square in the upper right, look at the largest \( r \)-by-\((r+1) \) rectangle in the Ferrers diagram. Now the partitions with \( k \) distinct even parts bend into partitions with “Durfee rectangle” of order \( k \), and the dots in columns to the right of the rectangle pair with dots below the rectangle to make even parts of size at most \( k \). Thus \( x^{k(k+1)}/\prod_{i=1}^{k}(1-x^{2i}) \) is the generating function for “almost-conjugate” partitions of \( n \) with “Durfee rectangle” of order \( k \).

**Proof. 3.** (Tsu) Partitions of \( n \) into \( k \) distinct even parts are transformed into partitions of \( n \) into \( k \) even parts by subtracting \( 2k - 2i \) from the \( i \)th part. If we view the resulting partition of \( n - k(k-1) \) by columns, and take two columns at a time, we again get even parts. The first has size \( 2k \), from the \( k \) even parts of the original partition, and the subsequent parts have size at most \( 2k \). Hence partitions of \( n \) into \( k \) distinct even parts are counted by the coefficient of \( x^{n-k(k-1)} \) in \( x^{k^2}/\prod_{i=1}^{k}(1-x^{2i}) \), and multiplying by \( x^{k(k-1)} \) gives the desired generating function.

**3.4.32.**

\[
\prod_{i=1}^{\infty} \frac{1}{1-x^i} = 1 + \sum_{k \geq 1} \frac{x^{2k}}{(1-x^2)(1-x^2)\cdots(1-x^k)^2}.
\]

On the left side, the coefficient of \( x^n \) is the number of partitions of \( n \). Each partition (of any integer) has a unique Durfee square. When the Durfee square has \( k^2 \) dots, the partition is completed by having a partition with at most \( k \) parts (or parts that are at most \( k \)) to the right of the Durfee square and another such partition hanging below the Durfee square.

Hence partitions with Durfee square of order \( k \), indexed by the integer being partitioned, are enumerated by the product of \( x^{2i} \) with the square of the OGF for partitions into at most \( k \) parts. Summing over \( k \) (plus 1 for the empty partition of 0) enumerates all partitions.

**3.4.33.** Variation of Euler’s Identity for partitions into distinct parts.

\[
\prod_{i=1}^{\infty} (1 + x^i) = 1 + \sum_{k \geq 1} \frac{x^{k(k+1)/2}y^k}{(1-x)(1-x^2)\cdots(1-x^k)}.
\]

The coefficient of \( x^ny^k \) on the left side is the number of partitions of \( n \) into \( k \) distinct parts; we choose the term \( x^iy \) in the factors for \( k \) distinct values of \( i \) summing to \( n \).

On the right, there is one partition with no parts, the partition of 0. When there are \( k \) distinct parts, with \( k \geq 1 \), subtracting \( i \) from the \( i \)th smallest part (for all \( i \)) yields an unrestricted partition with at most \( k \) parts. These are enumerated by \( \prod_{i=1}^{k}(1-x^i)^{-1} \), and the numerator adjust the exponent so they contribute to the correct term.

**3.4.34.** For \( n \in \mathbb{N} \) the sum of \( f(\lambda) \) over all partitions of \( n \) equals 1, where a partition \( \lambda \) is described as having \( m_i \) parts equal to \( i \), and \( f(\lambda) = 1/\prod_{i \geq 1} m_i!^{l_m} \). It suffices to show that \( \sum n! f(\lambda) = n! \). We do this by showing that \( n! f(\lambda) \) is the number of permutations of \([n]\) that have \( m_i \) cycles of length \( i \), for all \( i \). Call such a permutation a \( \lambda \)-permutation.

With \( \lambda \) expressed as its parts \( \lambda_1, \ldots, \lambda_k \) in canonical nonincreasing order, write a \( \lambda \)-permutation as a listing of \([n]\) in order, with the first \( \lambda_1 \) elements forming the first cycle in order, then the next \( \lambda_2 \) elements the next cycle, and so on. The \( n! \) ways of writing the elements in order yield the same \( \lambda \)-permutation in many ways. First, each cycle of length \( i \) is the same cycle under its \( i \) rotations. Also, cycles of the same length can be permuted arbitrarily without changing the resulting permutation. Hence a single \( \lambda \)-permutation arises exactly \( \prod_{i \geq 1} m_i!^{l_m} \) times among the \( n! \) listings of \([n]\).

We conclude that the number of \( \lambda \)-permutations is \( n! f(\lambda) \). Hence \( \sum \lambda n! f(\lambda) = n! \), as desired.

**3.4.35.** For \( n > 1 \), at least half of all Rogers–Ramanujan partitions of \( n \) do not use 1 as a part, where a Rogers–Ramanujan partition is a partition into distinct parts that does not have two consecutive integers as parts. An RR partition of \( n \) that has 1 as a part does not have 2 as a part. Therefore, subtracting 1 from all parts other than the largest and adding the total subtracted amount to the largest part produces another RR partition of \( n \) that does not have 1 as a part. Furthermore, this transformation is injective, so the desired inequality holds.

**3.4.36.** Combinatorial problems with solution \( P(n) \), where \( P(n) = \sum_{i=0}^{n-1} p(i) \) and \( p(i) \) is the number of partitions of the integer \( i \).

a) If \( S \) consists of one black marble and \( n - 1 \) white marbles, then the number of distinguishable partitions of \( S \) is \( P(n) \). The number of distinguishable partitions in which there are \( k \) white marbles in the same block with the black marble is \( p(n-1-k) \), since for the blocks not containing the black marble, it matters only how many there are of each size. Setting \( i = n-1-k \) and summing over \( i \) yields the claim.
b) If $Q(n)$ is the sum, over all partitions of $n$, of the number of distinct parts in the partition, then $Q(n) = P(n)$.

**Proof 1.** For $k \in \mathbb{N}$, the number of partitions of $n$ using $k$ is $p(n - k)$. Summing over $k$ counts each number once for each partition in which it is used at least once, and the sum is $P(n)$. Note that $Q(0) = 0 = P(0)$.

**Proof 2.** Represent a partition of $n$ as a partition of $n$ white marbles. Mark a partition by adding a black marble to one part. The number of distinguishable ways to mark a partition is obtained by inserting copies of $\lambda$.

3.4.38. Euler’s Pentagonal Number Theorem.

a) $p_r(n) = p_s(n)$, where these denote the numbers of partitions of $n$ into an even number and an odd number of distinct parts, respectively, except that $|p_r(n) - p_s(n)| = 1$ when $n$ is of the form $\sum_{i=1}^{m}(m + i)$ or $\sum_{i=0}^{m-1}(m + i)$.

Given a partition $\lambda$ of $n$ into distinct parts, compare the smallest part $k$ with the length $l$ of the diagonal of the Ferrers diagram starting at the end of the first row ($l = 3$ in the example above). Define a new partition $\tau(\lambda)$. When $k \leq l$, move the smallest part to become a new farthest diagonal. When $k > l$, move the diagonal to become a new (smaller) smallest part. If the smallest part is moved, then $\tau(\lambda)$ has a larger smallest part than $\lambda$, since $\lambda$ has distinct parts. Hence $\tau(\lambda) = \lambda$. Similarly, if the diagonal is moved, then the diagonal in $\tau(\lambda)$ is at least as long as in $\lambda$, so again $\tau(\lambda) = \lambda$. The numbers of parts in $\tau(\lambda)$ and $\lambda$ differ by exactly 1.

We conclude that $\tau$, when it can be applied, pairs partitions counted by $p_r(n)$ and $p_s(n)$. It applies if and only if the last row and farthest diagonal are disjoint. If they share a dot, then the operation does not produce a valid Ferrers diagram. The only way to have such an intersection is for the smallest part to equal the number of parts ($k = l$) or one more than the number of parts ($k = l + 1$). In the first case, the common dot is on the main diagonal, and the parts are $m + i$ for $0 \leq i \leq m - 1$, where $m$ is the number of parts. In the second case, the parts are $m + i$ for $1 \leq i \leq m$.

Since the sums are $(3m^2 - m)/2$ and $(3m^2 + m)/2$, respectively, and $(3m^2 - m_1)/2 = (3m^2 + m_2)/2$ requires $m_1 - m_2 = 1/3$, the two cases cannot both occur for one value of $n$. Hence when $n$ has one of these forms, $p_r(n) = p_s(n) = 1$, and otherwise equality holds.

b) $\prod_{k=1}^{\infty}(1 - x^k) = 1 + \sum_{m \geq 1}(-1)^m(x^{\omega(m)} + x^{\omega(-m)})$, where $\omega(m) = (3m^2 - m)/2$. The left side is the generating function for partitions into distinct parts, except that each use of a part of size $k$ is counted with negative sign. Hence a partition of $n$ with an even or an odd number of distinct parts contributes $+1$ or $-1$ to the coefficient of $x^n$. Therefore,
the generating function equals \( \sum_{n=0}^{\infty} (p_n(n) - p_0(n)) \). By part (a), the coefficients are 0 except when \( n = \omega(m) \) or \( n = \omega(-m) \). For each \( m \), the two special partitions we found have sum \( \omega(m) \) and \( \omega(-m) \) and have \( m \) parts. Hence the factor \((-1)^m\) assigns the correct sign to \( p_n(n) - p_0(n) \) for those terms.

### 3.4.39. Pentagonal numbers, continued.

Let \( p(n) \) be the number of partitions of \( n \), and let \( \omega(m) = (3m^2 - m)/2 \).

\( \text{a) } p(n) = \sum_{m \geq 1} (-1)^{m-1} \left[ p(n - \omega(m)) + p(n - \omega(-m)) \right] \) when \( n \geq 1 \).

For \( n \geq 0 \), let \( a_n \) be \( p(n) \) minus the right side of the identity. Since \( p(0) = 1 \) and the right side equals 0 when \( n = 0 \), it suffices to prove \( \sum_{n \geq 0} a_n x^n = 1 \).

Using \( x^n p(n-t) = x^{n-t} p(n-t) x^t \), we group the instances of the partition function according to their argument. We apply Euler’s Pentagonal Number Theorem to the resulting sum. Thus

\[
\sum_{n \geq 0} a_n x^n = \sum_{n \geq 0} \left[ p(n)x^n + x^n \sum_{m \geq 1} (-1)^m \left[ p(n-\omega(m)) + p(n-\omega(-m)) \right] \right] \\
= \sum_{n \geq 0} p(n)x^n \left( 1 + \sum_{m \geq 1} (-1)^m [x^{\omega(m)} + x^{\omega(-m)}] \right) \\
= \sum_{n \geq 0} p(n)x^n \sum_{k \geq 1} (1 - x^k) = 1
\]

\( \text{b) } \sum_{j \geq 1} (-1)^j q_j (n - 3j(j-1)/2) = 0 \), where \( q_j(n) \) is the number of the partitions of \( n \) not having \( j \) or \( 2j \) as a part. For \( j \geq 1 \), accounting for appearances of \( j \) and \( 2j \) yields \( q_j(k) = p(k) - p(k-j) - p(k-2j) + p(k-3j) \).

Thus

\[
\sum_{j \geq 1} (-1)^j q_j \left( n - \frac{3(j-1)}{2} \right) = \sum_{j \geq 1} (-1)^j \left( p \left( n - \frac{3(j-1)}{2} \right) + p \left( n - \frac{3j+1}{2} \right) \right) \\
+ \sum_{j \geq 1} (-1)^{j-1} \left( p \left( n - \frac{3(j-1)}{2} \right) + p \left( n - \frac{3j+1}{2} \right) \right)
\]

The first sum telescopes to \(-p(n)\), and part (a) shows that the second sum equals \( p(n) \).

### 3.4.40. For \( n \in \mathbb{N} \), with each partition \( \lambda \) of \( n \) being described as having \( m_k \) parts equal to \( k \) (for all \( k \)), we have \( \prod_{k \geq 1} \left( \prod_{j \geq 1} m_j \right)^k = \prod_{k \geq 1} \left( \prod_{j \geq 1} m_j \right) \). Taking logarithms of both sides converts the claim to \( \sum_{k \geq 1} \sum_{j \geq 1} \log k = \sum_{k \geq 1} m_k \log k \). It suffices to show that for each \( k \), the number of times \( \log k \) appears in the sum is the same on both sides.

On the right side, the number of times \( \log k \) appears is the total number of copies of \( k \) over all partitions of \( n \). In order to have each partition of \( n \) contribute its number of copies of \( k \) to the coefficient of \( x^n \) instead of contributing 1, in the generating function \( P(x) \) for all partitions we replace the factor \( 1 + x^k + x^{2k} + \cdots \) with the factor \( 0 + x^k + 2x^{2k} + \cdots \). Note that \( \sum_{j \geq 0} x^j y^{j/j} = \frac{1}{1-x^j} \). Thus we want to replace the factor \( \frac{1}{1-x} \) with the factor \( \frac{1}{1-xy} \). Therefore, the number of appearances of \( k \) is

\[
[x^n] \frac{x^k}{1-xy} P(x).
\]

On the left side, for each partition \( \lambda \), we obtain one copy of \( \log k \) for each \( j \) such that the multiplicity of \( j \) in \( \lambda \) is at least \( k \). Interchanging the order of summation over \( \lambda \) and \( j \), for each \( j \) we obtain one copy of \( \log k \) for each partition \( \lambda \) of \( n \) such that \( j \) has multiplicity at least \( k \). Thus, in the generating function for partitions, in the factor for \( j \) we want to keep only the terms for using at least \( k \) copies of \( j \). Thus we replace \( 1 + x^j + x^{2j} + \cdots \) with \( x^j + x^{j+1} + x^{j+2} + \cdots \).

Since we do this for each \( j \), the desired value is \( [x^n] P(x) \sum_{j \geq 1} x^j \).

Since \( \sum_{j \geq 1} x^j = \frac{x}{1-x} \), the two sides are equal.

### 3.4.41. A corner of a partition is a dot in its Ferrers diagram that is the last dot in its row and in its column. Let \( \gamma(\lambda) \) denote the number of corners in a partition \( \lambda \), and let \( \mathcal{P}(n) \) denote the set of all partitions of \( n \).

\( \text{a) } \sum_{\lambda \in \mathcal{P}(n)} \gamma(\lambda) = |\mathcal{P}(n-1)| + \sum_{\lambda' \in \mathcal{P}(n-1)} \gamma(\lambda') \). Removing a corner from a partition \( \lambda \) of \( n \) yields a partition of \( n - 1 \). Indeed, this is the only way to obtain a partition of \( n - 1 \) whose Ferrers diagram is contained in that of \( \lambda \). Hence \( \sum_{\lambda \in \mathcal{P}(n)} \gamma(\lambda) \) is the number of pairs \((\lambda, \lambda')\) such that \( \lambda \in \mathcal{P}(n) \), \( \lambda' \in \mathcal{P}(n-1) \), and the diagram for \( \lambda' \) is contained in the diagram for \( \lambda \).

For \( \lambda' \in \mathcal{P}(n-1) \), the number of such pairs containing \( \lambda' \) is the number of places where one dot can be added to \( \lambda' \) to obtain the diagram of a partition of \( n \). This number of places is \( \gamma(n) + 1 \). Hence when we group the pairs by \( \lambda' \) we obtain \(|\mathcal{P}(n-1)| + \sum_{\lambda' \in \mathcal{P}(n-1)} \gamma(\lambda') \). This proves the desired equality by counting two ways.

\( \text{b) } \sum_{n \geq 0} \lambda_{k \in \mathcal{P}(n)} \gamma(\lambda) x^n = \frac{x}{1-x} \prod_{j \geq 1} \frac{1}{1-x^j} \). From part (a), inductively

\[
\sum_{\lambda \in \mathcal{P}(n)} \gamma(\lambda) x^n = \frac{x}{1-x} \prod_{j \geq 1} \frac{1}{1-x^j}.
\]

### 3.4.42. The number of partitions of \( n \) is the same as the number of partitions into distinct parts whose odd-indexed parts sum to \( n \). Given a partition \( \mu \) of \( n \), we build from the Ferrers diagram of \( \mu \) a partition \( \phi(\mu) \) whose odd-indexed parts sum to \( n \). The diagram consists of positions \((i, j)\) such that \( 1 \leq j \leq \mu_i \). The hook \( H_{i,j} \) of size \( h_{i,j} \) consists of position \((i, j)\) and the positions in the diagram to its right in row \( i \) and below it in column
Section 3.4: Partitions of Integers

Let $\mu$ be a partition of $n$. The weight of the edge from a parent labeled $k$ to its leaf child labeled $p$ is $(k-1)/p^2$, and the weight of the edge to its non-leaf child labeled $p$ is $(k-1)/(p^2 - p)$. The sum of the weights on all edges to its children is

$$
\sum_{i=1}^{m} \frac{(k-1)}{i^2} + \frac{k-1}{i^2(i-1)} = (k-1) \sum_{i=1}^{m} \left( \frac{1}{i} - \frac{1}{i-1} \right) = (k-1) \frac{1}{k-1} = 1.
$$

For the path reaching nodes labeled $d_1 \leq \cdots \leq d_m$, the last being a leaf, let $n = \prod_{i=1}^{m} d_i$. The probability of the path is

$$
\frac{1}{d_1^2(d_1-1)} \cdot \frac{d_1-1}{d_2^2(d_2-1)} \cdots \frac{d_m-2}{d_{m-1}^2(d_{m-1}-1)} \cdot \frac{d_m-1}{d_m^2} = \frac{1}{d_1^2 d_2^2 \cdots d_m^2} = \frac{1}{n^2}.
$$

3.4.44. Reducing a matrix to all-0. A move subtracts 1 from an entry and adds 1 to the next entry to its right or to the next entry below it or to no entry. For an $m$-by-$n$ integer matrix $A$ and $n \geq \lambda_1 \geq \cdots \geq \lambda_m \geq 0$, let $f_A(\lambda) = \sum_{\lambda_1}^{\lambda_m} \lambda_{\lambda_1}^{\lambda_2} \cdots \lambda_{\lambda_m}^{\lambda_{\lambda_m}}$.

Some sequence of moves reduces $A$ to the zero matrix if and only if $f_A(\lambda) \geq 0$ for every partition $\lambda$. A move cannot increase $f_A(\lambda)$ for any $\lambda$. Hence the condition is necessary. For sufficiency, note that $A$ can be reduced to 0 when $A$ has no negative entries, by always reducing entries without making additions. Hence it suffices to show that when $A$ has negative entries the absolute sum of those entries can be reduced.

Let $(r, s)$ be a position minimizing $r + s$ such that $a_{r,s}$ is negative. By the assumed condition, $\sum_{r' \leq r} \sum_{s' \leq s} a_{r',s'} = 0$. Hence there is some position $(r', s')$ with $r' \leq r$ and $s' \leq s$ such that $a_{r',s'} > 0$. Using $r-r'+s-s'$ moves, we can shift one unit from position $(r', s')$ to position $(r, s)$, accomplishing the needed change.

3.4.45. The set $S^+$ consisting of the elements of $[n]$ whose difference from $n$ is even is the eventual image of every subset of $[n]$ under the function $f$ defined by letting $f(S) = S \cup \{1\}$ if $1 \notin S$ and $f(S) = [n] - (k-1: k \in S)$ if $1 \in S$. When applying $f$ to $S^+$, the elements of $[n]$ having odd difference from $n$ are deleted from $[n]$, leaving those having even distance, which again is $S^+$. Hence $S^+$ is a fixed point.

Call each iteration when 1 is added a unit step, and call the other iterations sliding steps. Each unit step is followed by a sliding step. After the first sliding step, $n$ is present; thereafter it remains. The next sliding...
step eliminates \( n - 1 \) (if present), after which it cannot reappear, because the presence of \( n \) always deletes it. Continuing, after a sliding step that eliminates \( n - 2i - 1 \), the next sliding step introduces \( n - 2i - 2 \), and the status of the values from there up to \( n \) implies that the status of those values (including \( n - 2i - 1 \) and \( n - 2i - 2 \)) never changes in the future.

### 3.4.6. The partitions with parts 1 through \( k \) each appearing once are the only fixed points in Bulgarian solitaire.

These are fixed point; the parts 2 through \( k \) turn into 1 through \( k - 1 \), and new part is \( k \).

Let \( \lambda \) be a fixed point, and let \( k \) be the number of parts in \( \lambda \). Since one new part is produced (in fact, it is \( k \)), exactly one must disappear, so \( \lambda \) has one 1. Iteratively, for \( 2 \leq i \leq k - 1 \), the fact that \( \lambda \) has exactly one copy of \( i - 1 \) forces \( \lambda \) to have exactly one copy of \( i \).

Also the one copy of \( k - 1 \) forces \( \lambda \) to have exactly one copy of \( k \); it is regenerated by the new pile. Since we originally let \( k \) be the number of piles, we now have the full partition, since we obtained one copy each of 1 through \( k \).

#### 3.4.47. In Bulgarian solitaire, the number of moves from a single pile of size \( n \) to a position on a cycle is \( n - k \), where \( k \) is the greatest integer such that \( \binom{k}{2} \) \( < \) \( n \).

We prove that the largest pile shrinks by 1 with each move until it reaches size \( k \) and that then the position is on a cycle.

For \( m \in \mathbb{N} \), let \( \{m\} = \{1, \ldots, m\} \). We claim that after \( n - \ell \) moves, for \( \ell \leq k \), the position has the form \( [j] - (i) \cup \{\ell \} \) or \( [j] \cup \{\ell \} \) for some \( i, j, \ell \in \mathbb{N} \) with \( i < j < \ell < \ell - 1 \). The initial position has this form, with \( j = 0 \) and \( \ell = n \). From such a position, the next will be \( [j] - (i - 1) \cup \{\ell - 1 \} \) or \( [j + 1] \cup \{\ell - 1 \} \), respectively: all old piles shrink by 1, and the new pile has size \( j \) in the first case and size \( j + 1 \) in the second. Note also that \( j \) never decreases, and \( j \) increases only via \( [j] - (i) \cup \{\ell \} \) to \( [j] \cup \{\ell - 1 \} \) to \( [j + 1] - (i - 1) \cup \{\ell - 2 \} \).

Each move with \( \ell > j + 1 \) reduces the largest size by 1, so \( n - \ell \) moves shrink the large pile to size \( k \) if the desired form is maintained. For this we need \( \ell > j \) while \( \ell \geq k + 1 \). The number of coins outside the large pile is at least \( j(j + 1)/2 - j \) and equals \( n - \ell \), the number of moves that have been made. Thus \( n - \ell \geq j(j - 1)/2 \). While \( \ell \geq k + 1 \), this yields \( n - k + 1 \geq j(j - 1)/2 \). Since \( n \leq k(k + 1)/2 \), we obtain \( k(k - 1)/2 > j(j - 1)/2 \), so \( j < k < \ell \).

Thus the first \( n - k \) moves reach the position \( [k] - (h) \) for some \( h \in \{k - 1 \cup \{0\} \). When \( h = 0 \), we have \( n = \binom{k + 1}{2} \) and a fixed point. Otherwise, let \( p_1, \ldots, p_{k-1} \) denote the pile sizes in nondecreasing order; we have \( i \leq p_i \leq i + 1 \) for all \( i \). Showing that this position lies on a cycle in which every position has largest pile size \( k \) or \( k - 1 \) implies also that no earlier position was on a cycle.

More generally, we consider all positions with sizes \( p_0, \ldots, p_{k-1} \) such that \( i \leq p_i \leq i + 1 \) for \( 0 \leq i \leq k - 1 \). There are \( k \) piles when \( p_0 = 1 \) and \( k - 1 \) piles when \( p_0 = 0 \), such as after the first \( n - k \) moves above. We show that such positions form cycles. Such a position is specified by its augmented set \( \{i: p_i = i + 1\} \), viewed as congruence classes modulo \( k \). We show that for a position \( p \) of this form with augmented set \( A \), the next position \( p' \) also has this form, with augmented set \( A' \) given by \( \{i - 1: i \in A\} \) (with arithmetic modulo \( k \)).

Since there are at least \( k - 1 \) piles in such a position, the new pile is the largest in \( p' \). For \( 1 \leq i \leq k - 1 \), the pile indexed by \( p_i \) becomes the new pile \( p_{i-1}' \), with value \( p_i - 1 \), so \( i + 1 \) \( \in \) \( A' \) if and only if \( i \in A \). Finally, there are \( k \) piles in \( p \) if and only if \( 0 \in A \). Thus \( p_{k-1}' = k - 1 \) \( \in \) \( A' \) if and only if \( 0 \in A \), which is as desired since \( 0 \equiv k \) (mod \( k \)).

#### 3.4.48. In Bulgarian solitaire, the partition of \( \binom{k + 1}{2} \) given by \( \lambda_1 = k - 1 \), \( \lambda_i = k - i + 1 \) for \( 2 \leq i \leq k \), and \( \lambda_{k+1} = 1 \) takes \( k(k-1) \) steps to reach the fixed point \( \mu \) given by \( \mu_i = k - i + 1 \) for \( 1 \leq i \leq k \). Equate partitions with their Ferrers diagrams. The fixed point has \( k \) full slants (diagonals). The initial partition \( \lambda \) has \( k - 1 \) full slants, one gap in slant \( k \) at the top, and one dot in slant \( k + 1 \) at the bottom.

As long as the new part taking one from each old part can serve as the largest part in the new partition, the effect of one step in Bulgarian solitaire is to rotate each slant down to the left (see Example 3.4.17). After \( k \) steps, the gap in slant \( k \) rotates back to the top, but the dot in slant \( (k + 1) \) falls one step short and sits in the row before its original row. Hence the partition after \( k \) steps is \( \lambda_1, \ldots, \lambda_{k-1}, \lambda_k + 1 \). The next \( k \) steps have the same effect, producing \( \lambda_1, \ldots, \lambda_{k-2}, \lambda_{k-1} + 1, \lambda_k \).

After \( k - 2 \) sets of \( k \) steps, the dot in slant \( (k + 1) \) has moved from row \( k + 1 \) to row \( 3 \) when the gap in slant \( k \) is at the top. The partition is now \( \lambda_1, \lambda_2, \lambda_3 + 1, \lambda_4, \ldots, \lambda_k \) (this includes only the first three terms when \( k = 3 \)). At this point, \( k - 1 \) additional steps move the gap in slant \( k \) to the bottom and the dot in slant \( k + 1 \) to the top, producing \( \lambda_1 + 2, \lambda_2, \ldots, \lambda_{k-1} \). There are only \( k - 1 \) parts, and the largest part is \( k + 1 \). Hence on the next step the slants do not cycle, but the next step is the staircase fixed point \( k, k - 1, \ldots, 1 \). The number of steps taken is \( (k - 2)k + k - 1 + 1 \), which equals \( (k - 1) \).

#### 3.4.49. Bounds on \( d(\lambda) \), the number of moves of Bulgarian solitaire needed to reach a partition that lies on a cycle when starting from \( \lambda \).

Let \( F(\lambda) \) denote the Ferrers diagram of a partition \( \lambda \).

- If \( F(\lambda) \subseteq F(\mu) \), then \( d(\lambda) \leq d(\mu) \). We use induction on \( d(\lambda) \). If \( d(\lambda) = 0 \), then \( \lambda \) already lies on a cycle in Bulgarian solitaire, and the statement is trivially true. Let \( \lambda' \) and \( \mu' \) be the images of \( \lambda \) and \( \mu \) after one step,
Since \( d(\lambda') = d(\lambda) - 1 \) and \( d(\mu') = d(\mu) - 1 \), it suffices by the induction hypothesis to prove \( F(\lambda') \subseteq F(\mu') \).

A necessary and sufficient condition for \( F(\lambda) \subseteq F(\mu) \) is that (with 0s added so that both have the same number of parts), the parts of the two partitions can be listed in some order \( x_1, \ldots, x_t \) and \( y_1, \ldots, y_t \) so that \( x_i \leq y_i \) for \( 1 \leq i \leq t \). By simultaneously permuting the two lists, we may assume that \( x_1, \ldots, x_t \) is in nondecreasing order (as \( \lambda_1, \ldots, \lambda_t \)). If \( y_i < y_{i+1} \), then \( x_{i+1} \leq x_i < y_{i+1} \), so \( x_{i+1} \leq y_i \) and \( x_i < y_{i+1} \). Hence \( y_i \) and \( y_{i+1} \) can be interchanged without destroying the ordering property. After placing \( y_1, \ldots, y_t \) in nonincreasing order, we have \( F(\lambda) \subseteq F(\mu) \).

By this observation, we do not need to worry about the order of the parts in \( \lambda' \) and \( \mu' \). Listing all the nonzero parts, we have \( \lambda = (\lambda_1, \ldots, \lambda_s) \) and \( \mu = (\mu_1, \ldots, \mu_t) \) with \( s \leq t \). The condition \( F(\lambda) \subseteq F(\mu) \) also yields \( \lambda_i \leq \mu_i \) for \( 1 \leq i \leq s \). The parts in \( \lambda' \) (not necessarily in order) are \( s \) and \( \lambda_i - 1 \) for \( 1 \leq i \leq s' \), where \( s' \) is the number of parts in \( \lambda \) that are at least 2. The parts in \( \mu' \) (not necessarily in order) are \( t \) and \( \mu_i - 1 \) for \( 1 \leq i \leq t' \), where \( t' \geq s' \). In the given order, the \( j \)th part in \( \lambda' \) is at most the \( j \)th part in \( \mu' \). Hence \( F(\lambda') \subseteq F(\mu') \), by the preceding paragraph.

b) Given that \( d(\lambda) \leq k(k - 1) \) when \( \lambda \) is a partition of \( \binom{k+1}{2} \), also \( d(\lambda) \leq k(k - 1) \) when \( \lambda \) is a partition of \( n \) with \( \binom{k}{2} < n < \binom{k+1}{2} \). Simply augment \( \lambda \) to a partition \( \mu \) of \( \binom{k+1}{2} \) by adding dots on the boundary of the Ferrers diagram. Now \( d(\lambda) \leq d(\mu) \leq k(k - 1) \). (Comment: There are various proofs that \( d(\lambda) \leq k(k - 1) \) when \( \lambda \) is a partition of \( \binom{k+1}{2} \), but they are lengthy.