Step 3 is also easy. We solve for $C(z)$ by the quadratic formula:

$$C(z) \;=\; \frac{1 \pm \sqrt{1-4z}}{2z}\,.$$

But should we choose the + sign or the − sign? Both choices yield a function that satisfies $C(z) = zC(z)^2 + 1$, but only one of the choices is suitable for our problem. We might choose the + sign on the grounds that positive thinking is best; but we soon discover that this choice gives $C(0) = \infty$, contrary to the facts. (The correct function $C(z)$ is supposed to have $C(0) = C_0 = 1$.) Therefore we conclude that

$$C(z) \;=\; \frac{1 - \sqrt{1-4z}}{2z}\,.$$

Finally, Step 4. What is $[z^n]\,C(z)$? The binomial theorem tells us that

$$\sqrt{1-4z} \;=\; \sum_{k\geqslant 0}\binom{1/2}{k}(-4z)^k \;=\; 1 + \sum_{k\geqslant 1}\frac{1}{2k}\binom{-1/2}{k-1}(-4z)^k\,;$$

hence, using (5.37),

$$\begin{aligned}\frac{1-\sqrt{1-4z}}{2z} \;&=\; \sum_{k\geqslant 1}\frac{1}{k}\binom{-1/2}{k-1}(-4z)^{k-1}\\ &=\; \sum_{n\geqslant 0}\binom{-1/2}{n}\frac{(-4z)^n}{n+1} \;=\; \sum_{n\geqslant 0}\binom{2n}{n}\frac{z^n}{n+1}\,.\end{aligned}$$

The number of ways to parenthesize, C_n, is $\binom{2n}{n}\frac{1}{n+1}$.

So the convoluted recurrence has led us to an oft-recurring convolution.

We anticipated this result in Chapter 5, when we introduced the sequence of *Catalan numbers* $\langle 1, 1, 2, 5, 14, \ldots\rangle = \langle C_n\rangle$. This sequence arises in dozens of problems that seem at first to be unrelated to each other [46], because many situations have a recursive structure that corresponds to the convolution recurrence (7.66).

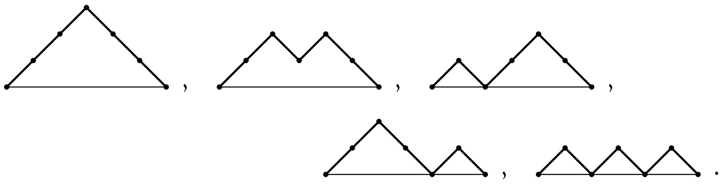
For example, let's consider the following problem: How many sequences $\langle a_1, a_2 \ldots, a_{2n}\rangle$ of +1's and −1's have the property that

$$a_1 + a_2 + \cdots + a_{2n} \;=\; 0$$

and have all their partial sums

$$a_1, \quad a_1 + a_2, \quad \ldots, \quad a_1 + a_2 + \cdots + a_{2n}$$

nonnegative? There must be n occurrences of +1 and n occurrences of −1. We can represent this problem graphically by plotting the sequence of partial sums $s_n = \sum_{k=1}^{n} a_k$ as a function of n: The five solutions for $n = 3$ are



These are "mountain ranges" of width $2n$ that can be drawn with line segments of the forms $\diagup$ and $\diagdown$. It turns out that there are exactly C_n ways to do this, and the sequences can be related to the parenthesis problem in the following way: Put an extra pair of parentheses around the entire formula, so that there are n pairs of parentheses corresponding to the n multiplications. Now replace each '$\cdot$' by $+1$ and each ')' by -1 and erase everything else. For example, the formula $x_0 \cdot \bigl((x_1 \cdot x_2) \cdot (x_3 \cdot x_4)\bigr)$ corresponds to the sequence $\langle +1, +1, -1, +1, +1, -1, -1, -1 \rangle$ by this rule. The five ways to parenthesize $x_0 \cdot x_1 \cdot x_2 \cdot x_3$ correspond to the five mountain ranges for $n = 3$ shown above.

Moreover, a slight reformulation of our sequence-counting problem leads to a surprisingly simple combinatorial solution that avoids the use of generating functions: How many sequences $\langle a_0, a_1, a_2, \ldots, a_{2n} \rangle$ of $+1$'s and -1's have the property that

$$a_0 + a_1 + a_2 + \cdots + a_{2n} \;=\; 1\,,$$

when all the partial sums

$$a_0, \quad a_0 + a_1, \quad a_0 + a_1 + a_2, \quad \ldots, \quad a_0 + a_1 + \cdots + a_{2n}$$

are required to be *positive*? Clearly these are just the sequences of the previous problem, with the additional element $a_0 = +1$ placed in front. But the sequences in the new problem can be enumerated by a simple counting argument, using a remarkable fact discovered by George Raney [302] in 1959: *If $\langle x_1, x_2, \ldots, x_m \rangle$ is any sequence of integers whose sum is $+1$, exactly one of the cyclic shifts*

$$\langle x_1, x_2, \ldots, x_m \rangle, \quad \langle x_2, \ldots, x_m, x_1 \rangle, \quad \ldots, \quad \langle x_m, x_1, \ldots, x_{m-1} \rangle$$

has all of its partial sums positive. For example, consider the sequence $\langle 3, -5, 2, -2, 3, 0 \rangle$. Its cyclic shifts are

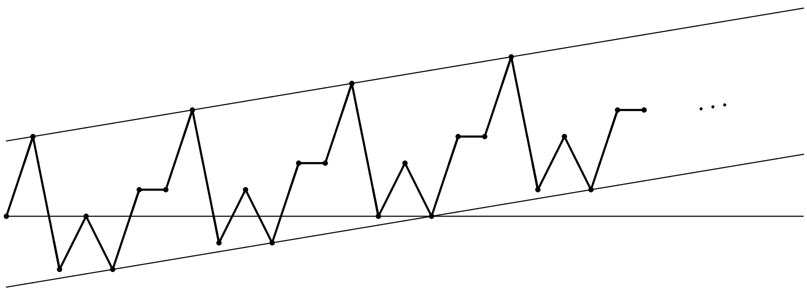
$$\begin{array}{ll} \langle 3,-5,2,-2,3,0\rangle & \langle -2,3,0,3,-5,2\rangle \\ \langle -5,2,-2,3,0,3\rangle & \langle 3,0,3,-5,2,-2\rangle \;\checkmark \\ \langle 2,-2,3,0,3,-5\rangle & \langle 0,3,-5,2,-2,3\rangle \end{array}$$

and only the one that's checked has entirely positive partial sums.

Raney's lemma can be proved by a simple geometric argument. Let's extend the sequence periodically to get an infinite sequence

$$\langle x_1, x_2, \ldots, x_m, x_1, x_2, \ldots, x_m, x_1, x_2, \ldots \rangle ;$$

thus we let $x_{m+k} = x_k$ for all $k \geqslant 0$. If we now plot the partial sums $s_n = x_1 + \cdots + x_n$ as a function of n, the graph of s_n has an "average slope" of $1/m$, because $s_{m+n} = s_n + 1$. For example, the graph corresponding to our example sequence $\langle 3, -5, 2, -2, 3, 0, 3, -5, 2, \ldots \rangle$ begins as follows:



Ah, if stock prices would only continue to rise like this.

The entire graph can be contained between two lines of slope $1/m$, as shown; we have $m = 6$ in the illustration. In general these bounding lines touch the graph just once in each cycle of m points, since lines of slope $1/m$ hit points with integer coordinates only once per m units. The unique lower point of intersection is the only place in the cycle from which all partial sums will be positive, because every other point on the curve has an intersection point within m units to its right.

(Attention, computer scientists: The partial sums in this problem represent the stack size as a function of time, when a product of $n+1$ factors is evaluated, because each "push" operation changes the size by $+1$ and each multiplication changes it by -1.)

With Raney's lemma we can easily enumerate the sequences $\langle a_0, \ldots, a_{2n} \rangle$ of $+1$'s and -1's whose partial sums are entirely positive and whose total sum is $+1$. There are $\binom{2n+1}{n}$ sequences with n occurrences of -1 and $n+1$ occurrences of $+1$, and Raney's lemma tells us that exactly $1/(2n+1)$ of these sequences have all partial sums positive. (List all $N = \binom{2n+1}{n}$ of these sequences and all $2n+1$ of their cyclic shifts, in an $N \times (2n+1)$ array. Each row contains exactly one solution. Each solution appears exactly once in each column. So there are $N/(2n+1)$ distinct solutions in the array, each appearing $(2n+1)$ times.) The total number of sequences with positive partial sums is

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

Example 5: A recurrence with m-fold convolution.

We can generalize the problem just considered by looking at sequences $\langle a_0, \ldots, a_{mn} \rangle$ of $+1$'s and $(1-m)$'s whose partial sums are all positive and