18.310 lecture notes September 2, 2013

Generating Functions

Lecturer: Michel Goemans

We are going to discuss enumeration problems, and how to solve them using a powerful tool: generating functions. What is an *enumeration problem*? That's trying to determine the number of objects of size n satisfying a certain definition. For instance, *what is the number of permutations of* $\{1, 2, \ldots, n\}$*?* (answer: $n!$), or *what is the number of binary sequences of length* n*?* (answer: 2^n). Ok, now let us introduce some tools to answer more difficult enumerative questions.

1 What is a generating function?

A generating function is just a different way of writing a sequence of numbers. Here we will be dealing mainly with sequences of numbers (a_n) which represent the number of objects of size n for an enumeration problem. The interest of this notation is that certain natural operations on generating functions lead to powerful methods for dealing with recurrences on a_n.

Definition 1. Let $(a_n)_{n\geq 0}$ be a sequence of numbers. The *generating function* associated to this sequence is the series

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

Also if we consider a class $\mathcal{A}$ of objects to be enumerated, we call *generating function* of this class the generating function

$$A(x) = \sum_{n\geq 0} a_n x^n,$$

where a_n is the number of objects of size n in the class.

Note that the variable x in generating functions doesn't stand for anything but serves as a placeholder for keeping track of the coefficients of x^n.

Example 1. The generating function associated to the class of binary sequences (where the *size* of a sequence is its length) is $A(x) = \sum_{n\geq 0} 2^n x^n$ since there are $a_n = 2^n$ binary sequences of size n.

Example 2. Let p be a positive integer. The generating function associated to the sequence $a_n = \binom{k}{n}$ for $n \leq k$ and $a_n = 0$ for $n > k$ is actually a polynomial:

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

Here the second equality uses the binomial theorem. Thus $A(x) = (1+x)^k$ is the generating function of the subsets of $\{1, 2, \ldots, k\}$ (where the *size* of a subset is its number of elements).

We see on this second example that the generating function has a very simple form. In fact, more simple than the sequence itself. This is the first magic of generating functions: in many natural instances the generating function turns out to be very simple.

Let us now modify this example in connection with probabilities. Suppose we have a coin having probability p of landing on heads and a probability $q = 1-p$ of landing on tails. We toss it k times and denote by a_n the probability of getting exactly n heads. What is the generating function of the sequence (a_n)? Well, it can be seen that $a_n = \binom{k}{n} q^{k-n} p^n$ thus the generating function is

$$A(x) = \sum_{n\geq 0} \binom{k}{n} q^{k-n} p^n x^n = (q+px)^k,$$

using the binomial theorem again.

Now, observe that the generating function is

$$(q+px)(q+px)(q+px)\cdots(q+px),$$

which is just multiplying k times the generating function $(q+px)$ corresponding to a single toss of the coin[1]. This is the second magic of generating functions: the generating function for complicated things can be obtained from the generating function for simple things. We will explain this in details, but first we consider an example.

2 Operations on classes and generating functions

We start with an easy observation. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are disjoint classes of objects, and $\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$ is their union (the symbol $\uplus$ denotes disjoint union). For instance $\mathcal{A}$ could be the set of permutations and $\mathcal{B}$ could be the set of binary sequences. Can we express the generating function of $C(x)$ of $\mathcal{C}$ in terms of the generating function $A(x) = \sum_{n\geq 0} a_n x^n$ of $\mathcal{A}$ and $B(x) = \sum_{n\geq 0} b_n x^n$ of $\mathcal{B}$? Well yes it is simply

$$C(x) = \sum_{n\geq 0}(a_n + b_n)x^n = \sum_{n\geq 0} a_n x^n + \sum_{n\geq 0} b_n x^n = A(x) + B(x)$$

since the number c_n of objects of size n in $\mathcal{C}$ is $a_n + b_n$.

We have just seen addition of generating functions, and we will now look at multiplication of generating functions. Consider the following problem. We have a die with six faces (numbered 1 to 6) and a die with eight faces (numbered 1 to 8). We roll the dice and we consider the sum of the dice. We want to know the number of ways c_n of getting each number n.

We claim that that the generating function $C(x) = \sum_{n\geq 0} c_n x^n$ is given by

$$C(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6) \times (x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8).$$

[1]This is not a coincidence: if we were to expand out the product into a sum, it would be a sum of 2^k terms, each of which takes either a q or a px from each of the k terms in the product. Hence each terms can be seen as a particular sequence of tails (represented by q) and head tosses (represented by px). In this calculation, the x's were a device for keeping track of the number of heads.

Indeed the first part accounts for the possible outcomes of the first die and the second part accounts for the possible outcome of the second die. For instance getting the sum 5 by getting 2 from the first die and 3 from the second die is accounted by the multiplication of the monomial x^2 from the first parenthesis with monomial x^3 from the second parenthesis x^3, etc. Multiplying this out, we get

$$C(x) = 1+2\,x+3\,x^2+4\,x^3+5\,x^4+6\,x^5+7\,x^6+7\,x^7+7\,x^8+6\,x^9+5\,x^{10}+4\,x^{11}+3\,x^{12}+2\,x^{13}+x^{14}.$$

In the above problem we see that multiplying generating function is meaningful. Let us now try to generalize the above reasoning. Given two sets, $\mathcal{A}$ and $\mathcal{B}$ the *Cartesian product* $\mathcal{A}\times\mathcal{B}$ is defined as the set of pairs (a, b) with $a \in \mathcal{A}$ and $b \in \mathcal{B}$. So if $\mathcal{A}$ and $\mathcal{B}$ are finite the cardinality of these sets are related by $|\mathcal{A}\times\mathcal{B}| = |\mathcal{A}| \times |\mathcal{B}|$. We also suppose that the *size* of a pair (a, b) is the size of a plus the size of b.

For instance, in the example above the class $\mathcal{A}$ represents the possible numbers of the first die, so that $\mathcal{A} = \{1, 2, 3, 4, 5, 6\}$ and the class $\mathcal{B}$ represents the possible number of the second die, so that $\mathcal{B} = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Now $\mathcal{C} = \mathcal{A}\times\mathcal{B}$ represents the possible numbers of the two dice. The *size* of a number on the first die is just that number, so the generating function for $\mathcal{A}$ is $A(x) = x+x^2+x^3+x^4+x^5+x^6$ while the generating function for $\mathcal{B}$ is $B(x) = x+x^2+x^3+x^4+x^5+x^6+x^7+x^8$. Now the *size* of a pair of number $(a, b) \in \mathcal{C}$ is the sum of the numbers of the two dice. So we want to determine c_n which is the number of pairs (a, b). We have claimed above that $C(x) = A(x) \times B(x)$. We now prove a generalization of the above relation between generating functions.

Theorem 1. *Let $\mathcal{A}$ and $\mathcal{B}$ be classes of objects and let $A(x)$ and $B(x)$ be their generating functions. Then the class $\mathcal{C} = \mathcal{A}\times\mathcal{B}$ has generating function $C(x) = A(x)B(x)$.*

Proof. Let c_n be the number of objects of size n in the Cartesian product $\mathcal{C} = \mathcal{A}\times\mathcal{B}$. These objects $c = (a, b)$ are obtained by picking an object $a \in \mathcal{A}$ of size $k \le n$ (a_k choices) and an object $b \in \mathcal{B}$ of size $n-k$ (b_{n-k} choices). Thus

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

Now let us consider product of generating functions

$$A(x)B(x) = \left(\sum_{k\ge 0} a_k x^k\right) \times \left(\sum_{k\ge 0} b_k x^k\right).$$

In order to get a monomial x^n in this product, one must multiply a monomial $a_k x^k$ for $k \le n$ from the first sum with a monomial $b_{n-k}x^{n-k}$ from the second sum. Thus one has

$$A(x)B(x) = \sum_{n\ge 0}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) x^n.$$

This completes the proof. □

We denote by $\mathcal{A}^k = \mathcal{A}\times\mathcal{A}\times\cdots\times\mathcal{A}$ the set of k-tuples of elements in $\mathcal{A}$. Using Theorem 1 we see that this class has generating function $A(x)^k = A(x) \times A(x) \times \cdots \times A(x)$ (where $A(x)$ is the generating function of $\mathcal{A}$). For instance, the generating function for the sum of numbers obtained by rolling 4 dice with 6 faces is

$$C(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6)^4.$$

Lastly we define

$$\mathrm{Seq}(\mathcal{A}) = \cup_{k \geq 0} \mathcal{A}^k,$$

as the set of finite sequences of elements in $\mathcal{A}$. For instance if $\mathcal{A} = \{0, 1\}$ then

$$\mathcal{A}^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$$

and $\mathrm{Seq}(\mathcal{A})$ is the set of binary sequences. Because of Theorem 1 we see that the generating function of the class $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$ is

$$C(x) = \sum_{k \geq 0} A(x)^k$$

where $A(x)$ is the generating function of $\mathcal{A}$. Observe also that $\displaystyle\sum_{k \geq 0} A(x)^k = \frac{1}{1 - A(x)}$ since

$$(1 - A(x)) \times \sum_{k \geq 0} A(x)^k = (1 - A(x)) \times (1 + A(x) + A(x)^2 + A(x)^3 + \ldots) = 1.$$

For instance for the binary sequences, $\mathcal{A} = \{0, 1\}$ has generating function $A(x) = 2x$ ($\mathcal{A}$ contains 2 binary sequences of length 1 and nothing else) so the class of binary sequences $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$ has generating function

$$C(x) = \sum_{k \geq 0} A(x)^k = \sum_{k \geq 0} (2x)^k = \frac{1}{1 - 2x}.$$

We will know use these results to treat various problems.

3 Number of ways of giving change

Let us look at the following simple question. Suppose we have 6 pennies, 1 nickel, and 2 dimes. For what prices can we give exact change? and in how many different ways? Let g_n be the number of ways we can give the exact changes for n cents ($g_n = 0$ if we cannot make the change), and let $G(x) = \sum_{n \geq 0} g_n x^n$ be the generating function for this problem.

We claim that

$$G(x) = (1 + x + x^2 + x^3 + x^4 + x^5 + x^6) \times (1 + x^5) \times (1 + x^{10} + x^{20})$$

Indeed here a way of giving change is determined by a triple (a, b, c) where a is the number of pennies, b is the number of nickels, c is the number of dimes. Moreover the "size" is the total number of cents it represents. So by Theorem 1 the generating function $G(x)$ is $A(x)B(x)C(x)$ where $A(x) = (1 + x + x^2 + x^3 + x^4 + x^5 + x^6)$ is the generating function of the change which can be made in pennies, $B(x) = (1 + x^5)$ is the generating function of the change which can be made in nickels, and $B(x) = (1 + x^{10} + x^{20})$ is the generating function of the change which can be made in dimes.

What happens when we have an arbitrary number of dimes, nickels and pennies? Well in this case the generating function for the change which can be made in pennies becomes $A(x) = \sum_{k\geq 0} x^k = \frac{1}{1-x}$, generating function for the change in nickels becomes $B(x) = \sum_{k\geq 0} x^{5k} = \frac{1}{1-x^5}$, and the generating function for the change in dimes becomes $C(x) = \sum_{k\geq 0} x^{10k} = \frac{1}{1-x^{10}}$. So the generating function for the number of ways of giving the change if one has infinitely many pennies, nickels, and dimes is

$$\frac{1}{(1-x)(1-x^5)(1-x^{10})}.$$

4 Dots and Dashes

Now, let's think about another problem. Suppose that you are sending information using a sequence of two symbols, say dots and dashes, and suppose that sending a dash takes 2 units of times, while sending a dot take 1 unit of time. Here the *size* of a message will be defined as the number of units times it takes. So we ask the question: how many different messages can you send in n time units? Let's call this number f_n. We'll figure out for the first few f_n. We have

n	1	2	3	4	5
f_n	1	2	3	5	8
messages	·	··	···	····	·····
		–	–·	–··	–···
			·–	·–·	·–··
				··–	··–·
				––	···–
					––·
					–·–
					·––

You may already recognize the pattern: these are the Fibonacci numbers. But let's see what we can learn about Fibonacci numbers by using generating functions. The recursion for the Fibonacci numbers is

$$f_n = f_{n-1} + f_{n-2}$$

It's not difficult to see why this works. The first symbol must either be a dot or a dash. If the first symbol is a dash, removing it leaves a sequence two units short, and if the first symbol is a dot, removing it leaves a sequence one unit shorter. Adding up these two possibilities gives us the above recursion relation.

Now, how does this connect to generating functions? Let us define

$$F(x) = \sum_{j=0}^{\infty} f_j x^j.$$

What is f_0? It has to be 1, in order to have $f_2 = f_1 + f_0$. This makes sense intuitively: there is one message, the empty message, using zero units of time. What does this recurrence say about $F(x)$? Let's look at the following equations

$$\begin{array}{rcccccccccccccl}
F(x) & = & 1 & + & x & + & 2x^2 & + & 3x^3 & + & 5x^4 & + & 8x^5 & \ldots \\
xF(x) & = & & & x & + & x^2 & + & 2x^3 & + & 3x^4 & + & 5x^5 & \ldots \\
x^2F(x) & = & & & & & x^2 & + & x^3 & + & 2x^4 & + & 3x^5 & \ldots
\end{array}$$

We can see that by multiplying by x and x^2 we have shifted the terms, so that instead of $f_k x^k$ we get $f_{k-1}x^k$ and $f_{k-2}x^k$. We thus get that the equation $f_k = f_{k-1} + f_{k-2}$ nearly corresponds to the equation $F(x) = xF(x) + x^2F(x)$. This isn't quite right. All the terms x^k for $k > 0$ do cancel, but the constant term doesn't. To make the constant term correct, we need to add 1 to the right side, obtaining the correct equation

$$F(x) = xF(x) + x^2F(x) + 1.$$

Now, this can be rewritten as

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

At this point the perceptive reader will have observed that there was a much faster way to obtain the above result. Indeed one can observe that $F(x)$ is the generating function of the set of sequences $\mathrm{Seq}(\mathcal{A})$ of the class $\mathcal{A} = \{dot, dash\}$. Moreover the class $\mathcal{A}$ has generating function $A(x) = x + x^2$ since it has one element of size 1 and one element of size 2. Therefore by the discussion in the previous section on gets $F(x) = \dfrac{1}{1 - A(x)} = \dfrac{1}{1 - x - x^2}$.

Now we want to use the expression of $F(x)$ in order to obtain some information on the coefficients f_n. When we have a polynomial in the denominator of a fraction like this, we can factor the polynomial and express it as the sum of two simpler fractions. That is, we first factor the denominator

$$1 - x - x^2 = (1 - \phi_+ x)(1 - \phi_- x)$$

where $\phi_+ = \frac{1+\sqrt{5}}{2}$ and $\phi_- = \frac{1-\sqrt{5}}{2}$. (Note that ϕ_+ and ϕ_- are not the roots of $1 - x - x^2$, but the inverses of the roots.)

We now use the *method of partial fractions* to rewrite this as

$$F(x) = \frac{a}{1 - \phi_+ x} + \frac{b}{1 - \phi_- x}$$

for some a and b. Elementary algebra gives

$$a = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right),$$

$$b = -\frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right).$$

Now, we need to remember the Taylor series for $1/(1 - \alpha x)$. This is

$$\frac{1}{1 - \alpha x} = 1 + \alpha x + \alpha^2 x^2 + \alpha^3 x^3 + \dots$$

Even if you don't remember Taylor expansions, you should recognize this as the formula for summing a geometric series.

We thus have that, expanding each of the fractions in the expression for $F(x)$ above in a Taylor series,

$$\begin{aligned} F(x) \quad = \quad & a\left(1+\left(\frac{1+\sqrt{5}}{2}\right)x+\left(\frac{1+\sqrt{5}}{2}\right)^2 x^2+\left(\frac{1+\sqrt{5}}{2}\right)^3 x^3+\ldots\right) \\ & +b\left(1+\left(\frac{1-\sqrt{5}}{2}\right)x+\left(\frac{1-\sqrt{5}}{2}\right)^2 x^2+\left(\frac{1-\sqrt{5}}{2}\right)^3 x^3+\ldots\right) \end{aligned}$$

Substituting in the values we know for a and b, we get

$$\begin{aligned} F(x) \quad = \quad & \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)+\left(\frac{1+\sqrt{5}}{2}\right)^2 x+\left(\frac{1+\sqrt{5}}{2}\right)^3 x^2+\left(\frac{1+\sqrt{5}}{2}\right)^4 x^3+\ldots\right) \\ & -\frac{1}{\sqrt{5}}\left(\left(\frac{1-\sqrt{5}}{2}\right)+\left(\frac{1-\sqrt{5}}{2}\right)^2 x+\left(\frac{1-\sqrt{5}}{2}\right)^3 x^2+\left(\frac{1-\sqrt{5}}{2}\right)^4 x^3+\ldots\right) \end{aligned}$$

Now, this gives us a nice expression for f_n, the nth Fibonacci number. We equate the coefficients of x^n on the left- and right-hand sides of this equation. Since the nth Fibonacci number f_n is the coefficient on x^n in $F(x)$, we get

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

Since $\left|\frac{1-\sqrt{5}}{2}\right| < 1$, we can see that the second term goes to 0 as n gets large, and f_n grows as

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

for some C. This means that we have found the asymptotic growth rate for the number of messages that can be encoded by our dots and dashes, and that the number of bits sent per time unit is

$$\log_2 \frac{1+\sqrt{5}}{2}.$$

5 Generalized Recurrence Equations

We've seen an example of a linear recurrence equation in the last section. I'm going to now give a general method for solving linear recurrence equations (also called linear difference equations). If you've taken 18.03, you'll notice that this method looks a lot like the method for solving linear differential equations.

Suppose we have a recurrence equation

$$f_n = \alpha f_{n-1} + \beta f_{n-2} + \gamma f_{n-3}$$

I'm only writing this equation down with three terms, but the generalization to k terms is obvious, and works exactly like you'd expect. How do we solve this? What we do is to write down the generating function

$$F(x) = \sum_{j=0}^{\infty} f_j x^j.$$

Then, using the same reasoning as before, we get an equation for $F(x)$ of the following form:

$$F(x) = \alpha x F(x) + \beta x^2 F(x) + \gamma x^3 F(x) + p(x)$$

where $p(x)$ is a low-degree polynomial that makes this equation work for the first few elements of the sequence, where the recurrence equation doesn't necessarily work (because we don't have an f_{-1} term). For the Fibonacci number example above, we have $p(x) = 1$. Note that if we don't have a $p(x)$ term, we get the solution $F(x) = 0$ which, while its coefficients (all 0's) satisfy the linear recurrence equation, doesn't tell us anything useful. The maximum degree on $p(x)$, if the recurrence equation has three terms, is quadratic (and if the recurrence equation has k terms, is $k - 1$). You can see this by noticing that for the x^3 component and later, the recurrence is guaranteed to work. I'll let you check this fact.

As before, we next obtain

$$F(x) = \frac{p(x)}{1 - \alpha x - \beta x^2 - \gamma x^3}$$

Let's suppose we can factor the denominator as follows:

$$1 - \alpha x - \beta x^2 - \gamma x^3 = (1 - r_1 x)(1 - r_2 x)(1 - r_3 x).$$

I'll leave the question of what happens if you have a double or triple root for a homework problem. We then use the method of partial fractions (which you may remember from Calculus) to get

$$F(x) = \frac{a}{1 - r_1 x} + \frac{b}{1 - r_2 x} + \frac{c}{1 - r_3 x}$$

where a, b, and c are constants.

We can then see, by taking a series expansion for this generating function, that a generic term of our sequence will be

$$f_n = a r_1^n + b r_2^n + c r_3^n.$$

How did we get the roots r_1, r_2 and r_3? They are the zeroes of the polynomial

$$y^3 - \alpha y^2 - \beta y - \gamma = 0.$$

We can see this by taking $y = \frac{1}{x}$, so

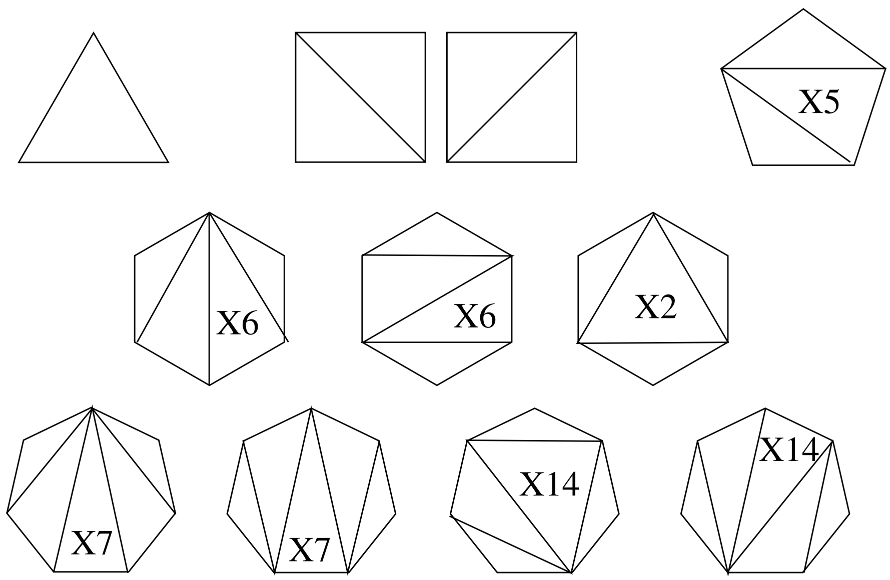
$$\begin{aligned} 1 - \alpha x - \beta x^2 - \gamma x^3 &= x^3(y^3 - \alpha y^2 - \beta y - \gamma) \\ &= x^3(y - r_1)(y - r_2)(y - r_3) \\ &= (1 - r_1 x)(1 - r_2 x)(1 - r_3 x). \end{aligned}$$

6 Chord diagrams

Let's count something harder now. Let's count how many ways there are of putting chords into a k-gon to divide it into triangles. We'll call this number C_{k-2}. The sequence starts as follows:

k	3	4	5	6	7
$j = k - 2$	1	2	3	4	5
C_j	1	2	5	14	42

as you can see from the following figure.



Here, I've illustrated one of each essentially different way of dividing a k-gon into triangles, along with the number of times it must be counted (because of symmetry) for $k \leq 7$.

How can we find a recurrence for this number? Well, for a k-gon, let's look at the triangle through the edge $(k, 1)$, one of the specific sides of the polygon. There must be a third point in this triangle. Call it j. Clearly, we must have $2 \leq j \leq k - 1$. If we remove this triangle, we now have two smaller polygons, a j-sided one, and a $(k - j + 1)$-sided one. We can now divide these polygons up into triangles independently. We thus get that the number of ways of triangulating a k-gon, given that we have a triangle with vertices $1, j, k$, is $C_{j-2}C_{k-j-1}$.

One thing we notice is that for this to be true for $j = 2$ or $j = k - 1$, we have to set $C_0 = 1$. This takes care of the case where j is 2 or $k - 1$, and one of the two smaller polygons is just an edge.

Now, we can get a recurrence. Summing over all the j between 2 and $k - 1$ gives

$$C_{k-2} = \sum_{j=2}^{k-1} C_{j-2}C_{k-j-1}$$

This formula can use some rethinking of the limits. Let's let $k' = k - 2$ and $j' = j - 2$. We get

$$C_{k'} = \sum_{j'=0}^{k'-1} C_{j'}C_{k'-j'-1}$$

which is a nicer looking recurrence relation.

The next question is how we evaluate it using generating functions. Let's look at the generating function for counting these triangulations. That is,

$$G(x) = \sum_{i=0}^{\infty} C_i x^i = 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 + \ldots$$

What happens when we square $G(x)$. We get

$$\begin{aligned} G(x)^2 &= 1 + (1+1)x + (1\cdot 2 + 1\cdot 1 + 2\cdot 1)x^2 + (1\cdot 5 + 1\cdot 2 + 2\cdot 1 + 5\cdot 1)x^3 + \ldots \\ &= \sum_{k=0}^{\infty} x^k \sum_{j=0}^{k} C_j C_{k-j} \end{aligned}$$

You can see that the x^k expression on the right is the right-hand-side of the recurrence relation we found about for C_{k+1} (with $k' = k+1$), so we get

$$G(x)^2 = \sum_{k=0}^{\infty} C_{k+1} x^k.$$

Multiplying by x gives a sum with the x^j coefficient equal to $C_j x^j$. We now have an expression relating $xG(x^2)$ and $G(x)$. We need to make sure we get the smallest terms right. We can check the constant term is the only one that is wrong, and we can fix that by adding 1 to the right hand side, to get the equation

$$G(x) = 1 + xG(x)^2.$$

At this point the perceptive reader will have observed that there was a faster way to obtain the above equation. Indeed one can observe that a chord diagram is either empty (it has no chord), or has one chord dividing 2 chord diagrams. Thus the non empty chord diagrams are made of a Cartesian product of a chord (generating function x), a left chord diagram (generating function $G(x)$), and a right chord diagram (generating function $G(x)$). Using Theorem 1 we get that the non empty chord diagrams have generating function $x \times G(x) \times G(x)$ (while empty chord diagram have generating function 1). This gives

$$G(x) = 1 + xG(x)^2.$$

as above.

We now solve this equation. This is a quadratic equation in $G(x)$, so we can use the quadratic formula to solve for $G(x)$, obtaining

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

We now have a choice. Which of the two roots of this equation should we use. We can figure this out by looking at the first term. We should have $G(0) = 1$. Depending on which root we choose, when we plug in 0 we either get $G(0) = 2/0$, or $G(0) = 0/0$. Clearly, the first option gives the wrong answer. Using l'Hopital's rule, we can figure out that in the second case, we indeed have $G(0) = 1$, so we get

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

This was the fun part; we have an algebraic expression for $G(x)$. We now need to expand it in a power series to find the x^k term. In this case, unfortunately, this happens to be somewhat tedious. We will go through the steps carefully.

The first step is expanding $(1-y)^{1/2}$ in a power series. We use the binomial formula

$$(1-y)^{1/2} = 1 - \binom{1/2}{1} y + \binom{1/2}{2} y^2 - \binom{1/2}{3} y^3 + \binom{1/2}{4} y^4 - \dots.$$

This might look odd if you haven't seen it before, but one can define $\binom{\alpha}{k}$ even when α is not a positive integer: the formula is

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

Simplifying,

$$(1-y)^{1/2} = 1 - \tfrac{1}{2}y + \tfrac{1}{2}(-\tfrac{1}{2})\frac{y^2}{2!} - \tfrac{1}{2}(-\tfrac{1}{2})(-\tfrac{3}{2})\frac{y^3}{3!} + \tfrac{1}{2}(-\tfrac{1}{2})(-\tfrac{3}{2})(-\tfrac{5}{2})\frac{y^4}{4!} + \cdots$$

Now, we need to substitute $y = 4x$, and plug the resulting expression into the formula we got for $G(x)$. We obtain

$$G(x) = \frac{1-\sqrt{1-4x}}{2x} = \frac{1}{2x}\sum_{k=1}^{\infty} \frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2k-7)\cdot(2k-5)\cdot(2k-3)}{2^k}\frac{(4x)^k}{k!}$$

where we have the product of all odd numbers between 1 and $2k-3$ in the numerator, and $k!$ in the denominator. All the $-$ signs cancel out, as they should: since we're counting things, we have to get a positive integer.

How do we simplify this expression? Recall $G(x) = \sum C_k x^k$, so equating coefficients, we get

$$C_k = \frac{1}{2}\cdot\frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2k-5)\cdot(2k-3)\cdot(2k-1)}{2^{k+1}}\cdot\frac{(4)^{k+1}}{(k+1)!}$$

where we have had to replace k by $k+1$ in the above formula to make up for the $\frac{1}{x}$ in front of it. We can cancel out the powers of 2 in the numerator and denominator to get

$$C_k = \frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2k-5)\cdot(2k-3)\cdot(2k-1)}{(k+1)!}2^k.$$

Now, let's multiply the top and bottom of the above expression by $k!$, and write $k!2^k = 2\cdot 4\cdot 6\cdots(2k)$. We get

$$C_k = \frac{1\cdot 2\cdot 3\cdots(2k-1)\cdot(2k)}{(k+1)(k!)^2}.$$

Thus, we have

$$C_k = \frac{1}{k+1}\frac{(2k)!}{(k!)^2} = \frac{1}{k+1}\binom{2k}{k},$$

which is the definition of the k'th Catalan number.

The Catalan numbers turn up in quite a few places (as we've already seen). Prof. Richard Stanley has a section on his webpage (which is also in his book) giving 66 combinatorial interpretations of the Catalan numbers.

Exercise. Give a bijective proof that the number of chord diagrams is given by the Catalan numbers.

7 Diagonals in Pascal's triangle

In this section we use generating function of more than one variable in order to solve a neat problem. Recall that Pascal's triangle is formed by binomial coefficients:

								1								
							1		1							
						1		2		1						
					1		3		3		1					
				1		4		6		4		1				
			1		5		**10**		10		5		1			
		1		6		15		20		**15**		6		1		
	1		7		21		35		35		21		**7**		1	
1		8		28		56		70		56		28		8		**1**

We will compute the sums of the diagonal elements in Pascal's triangle. For example, the sum of the boldface elements above is

$$34 = 1 + 10 + 15 + 7 + 1 = \binom{4}{0} + \binom{5}{2} + \binom{6}{4} + \binom{7}{6} + \binom{8}{8}.$$

How do we do this with generating functions? Let's use a generating function of two variables. Consider the two variable function

$$g(x, y) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \binom{a+b}{a} x^a y^b.$$

Note that the coefficient of $x^a y^b$ of g is precisely the value in row a, column b of Pascal's triangle (both indexed starting from 0). We already saw the sum of the terms in the n'th row of Pascal's triangle

$$(x + y)^n = \sum_{a=0}^{n} \binom{n}{a} x^a y^{n-a}$$

Now, let's sum over all rows.

$$G(x, y) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \binom{a+b}{a} x^a y^b = \sum_{n=0}^{\infty} (x + y)^n = \frac{1}{1 - x - y}$$

where the last equality comes from the sum of a geometric series.

What are we looking for? We're looking for the sum

$$\sum_{j=0}^{m} \binom{m+j}{2j}$$

for all m. The generating function for that would be

$$H(z) := \sum_{m=0}^{\infty} \sum_{j=0}^{m} \binom{m+j}{2j} z^m.$$

We would like to relate H to G somehow (since we have an expression for G). So we'd like to somehow turn $\binom{a+b}{a}$ into $\binom{m+j}{2j}$. Can we do this? Looking at this more closely, it involves setting $a = 2j$ and $b = m - j$. This would involve turning $x^a y^b$ into $x^{2j}y^{m-j}$. So we'd like to get $x^{2j}y^{m-j} = z^m$. Note we can do this if we let $x = z^{1/2}$ and $y = z$. Dealing with square roots is not so nice, so lets square everything and let $z = x^2$, $y = x^2$. Our hope will be that $H(x^2)$ is related to $G(x, x^2)$.

So consider $G(x, x^2)$:

$$G(x, x^2) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\binom{a+b}{a}x^{a+2b}.$$

Now let's compute the coefficient of x^{2m} in $G(x, x^2)$; in other words, the sum of all terms where $a + 2b = 2m$, or $a = 2(m - b)$. We obtain

$$\begin{aligned}\text{coeff. of } x^{2m} \text{ in } G(x, x^2) &= \sum_{b=0}^{m}\binom{2(m-b)+b}{2(m-b)}\\ &= \sum_{j=0}^{m}\binom{m+j}{2j} \qquad \text{with the substution } j = m - b.\end{aligned}$$

Success! This is exactly h_m (or in other words, the coefficient of x^{2m} in $H(x^2)$). Note that we don't have $H(x^2) = G(x, x^2)$; rather, we have

$$G(x, x^2) = H(x^2) + Q(x),$$

where $Q(x)$ consists of only odd powers of x.

At any rate, we can now determine h_n, since we know

$$G(x, x^2) = \frac{1}{1 - x - x^2} = \sum_{r=0}^{\infty} f_r x^r,$$

where f_m is the m'th Fibonacci number. So $h_m = f_{2m}$.

It turns out that once you know that you know the answer, it's easy to prove it by induction. Generating functions have the advantage that we didn't have to guess the answer first for the technique to work.

8 Exponential generating functions

The generating functions we have seen so far are technically known as *ordinary* generating functions. There are other kinds of generating functions. A particularly important kind are *exponential* generating functions.

Definition 2. If $a_0, a_1, a_2, \ldots$ is an infinite sequence of integers, the *exponential generating function* (EGF) of $(a_n)_{n\in\mathbb{N}}$ is the function

$$\hat{A}(x) = \sum_{n=0}^{\infty}\frac{a_n}{n!}x^n.$$

(The notation here of putting a hat on exponential generating functions is not standard, but we will just use it as a reminder that we're not working with the normal generating function.)

Example. The EGF of the sequence $1, 1, 1, \dots$ is

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

The choice of whether to use ordinary or exponential generating functions depends on the situation. Here, we'll see a setting where exponential generating functions are particularly nice.

Let p_n denote the number of permutations of a set of size n. We will define $p_0 = 1$ (this just turns out to be the most convenient definition for us). You should recall that $p_n = n!$. So we can compute the EGF:

$$\hat{P}(x) = \sum_{n=0}^{\infty} \frac{n!}{n!} x^n = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Pretty simple!

A *cycle* is a special kind of permutation. Let π be a permutation of the set $\{1, 2, \dots, n\}$. Consider the sequence

$$r_1 = 1, \quad r_2 = \pi(r_1), \quad r_3 = \pi(r_2), \quad \dots r_n = \pi(r_{n-1}).$$

Then π is a cycle if and only if $(r_1, r_2, \dots, r_n)$ is a reordering of $(1, 2, \dots, n)$.

How many cycles on a set of size n are there? It's not too hard to see that this number (call it c_n) is just $(n-1)!$: we have $n-1$ choices for $r_2 = \pi(1)$ (we can't pick 1), then $n-2$ choices for $\pi(r_2)$ (we can't pick 1 or r_2), and so on. We also define $c_0 = 0$; again this is a matter of convenience for us.

So the generating function for the number of cycles is

$$\hat{C}(x) = \sum_{n=1}^{\infty} \frac{(n-1)!}{n!} x^n = \sum_{n=1}^{\infty} \frac{x^n}{n} = \log\left(\frac{1}{1-x}\right).$$

How did we get that last step? One way is to integrate both sides of the identity

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n.$$

Now consider $\hat{P}(x)$ and $\hat{C}(x)$; you might notice that $\hat{P}(x) = e^{\hat{C}(x)}$. Is this a coincidence? In fact, no! It is a very special case of a much more general (and very useful) fact. The key reason for the relation is that any permutation can be descibed in terms of a disjoint collection of cycles:

Example. Consider the set $S = \{1, 2, 3, 4, 5, 6\}$, and the permutation π on S defined by

$$\pi(1) = 2, \pi(2) = 5, \pi(3) = 6, \pi(4) = 4, \pi(5) = 1, \pi(6) = 3.$$

This can be described by the cycle $1 \to 2 \to 5 \to 1 \cdots$ on the set $\{1, 2, 5\}$, the cycle $3 \to 6 \to 3 \cdots$ on the set $\{3, 6\}$, and the cycle $4 \to 4 \to 4 \cdots$ on the set $\{4\}$.

Rather than dealing with the abstraction of the general setting, we will prove the connection between $\hat{P}(x)$ and $\hat{C}(x)$ in a way that is clearly quite general, and then see some other examples of the same connection.

Theorem 2. *Let $\hat{P}$ be the EGF for the number of permutations $(p_n)_{n\in\mathbb{N}}$, and $\hat{C}$ be the EGF for the number of cycles $(c_n)_{n\in\mathbb{N}}$. Then $\hat{P}(x) = e^{\hat{C}(x)}$.*

Proof. As noted already, any permutation of a set S of size n can be split up into a collection of disjoint cycles. So in order to enumerate all the permutations of S, we can consider all possible *partitions* of S into nonempty pieces, and then choose any cycle we like on each piece of the partition. To understand the number of ways of doing this, let's first count the number of permutations with *exactly* k cycles.

Lemma 1. *The number of permutations of a set of size n with exactly k cycles has generating function*

$$\frac{1}{k!}\left(\hat{C}(x)\right)^k.$$

Proof. We'll first do the case $k = 2$, which is easier to understand.

We'll begin by counting a slightly different thing. Let's count the number of ways of partitioning a set S of size n into 2 cycles, one colored red and the other blue (so we'll count as different two permutations which are exactly the same, but whose cycles are colored differently). Let q_n be this number, and $\hat{Q}(x)$ be the EGF for $(q_n)_{n\in\mathbb{N}}$. What is q_n? Well, we must decide which elements of S will be in the red cycle (the rest will be in the blue), and then we get to pick any cycle on the red part and any cycle on the blue part. Hence

$$q_n = \sum_{T\subseteq S} c_{|T|}\cdot c_{n-|T|}.$$

(You might ask why we are allowing $T = \emptyset$ or $T = S$, since we really want exactly 2 cycles. But recall that $c_0 = 0$, and so it makes no difference if we include these terms in the sum or not.) We can rewrite this by splitting the sum up by the size of T: there are $\binom{n}{t}$ choices of T with $|T| = t$, and so

$$q_n = \sum_{t=0}^{n}\binom{n}{t}c_t c_{n-t}.$$

So now let's write down $\hat{Q}(x)$: Now consider $\hat{Q}(x)$:

$$\begin{aligned}
\hat{Q}(x) &= \sum_{n=0}^{\infty}\frac{q_n}{n!}x^n\\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\left(\sum_{t=0}^{n}\frac{n!}{t!(n-t)!}c_t c_{n-t}\right)x^n\\
&= \sum_{t=0}^{\infty}\sum_{n=t}^{\infty}\frac{c_t x^t}{t!}\cdot\frac{c_{n-t}x^{n-t}}{t!}\\
&= \sum_{t=0}^{\infty}\frac{c_t x^t}{t!}\cdot\sum_{s=0}^{\infty}\frac{c_s x^s}{s!} \qquad \text{substituting } s = n-t\\
&= \hat{C}(x)\cdot\hat{C}(x).
\end{aligned}$$

Now to finish the lemma for $k = 2$, we note that we've overcounted every permutation with 2 cycles twice, since there are 2 ways of coloring the cycles. So the EGF for what we're interested in is $\frac{1}{2!}\hat{C}(x)^2$, as required.

Now let's do the general case. Again, suppose we have k colors; let's count the number of ways of partitioning a set S of size n into k cycles, one of each color (so we'll count as different two permutations which are exactly the same, but whose cycles are colored differently). Let h_n be this number, and $\hat{H}(x)$ be the EGF for $(h_n)_{n\in\mathbb{N}}$. We wish to show that $h_n/n!$ is the coefficient of x^n in $\hat{C}(x)^k$. So let's consider this coefficient of

$$\hat{C}(x) \cdot \hat{C}(x) \cdots \hat{C}(x).$$

It is a sum of various contributions, where each contribution consists of choosing nonnegative integers $n_1, n_2, \ldots, n_k$ that sum to n, and multiplying the x^{n_1} term from the first term in this product with the x^{n_2} term of the second, the x^{n_3} term of the third, etc. So we obtain

$$\begin{aligned}\text{coeff. of } x^n \text{ in } H(x) &= \sum_{n_1,\ldots,n_k,\sum n_i = n} \frac{c_{n_1}}{n_1!} \cdot \frac{c_{n_2}}{n_2!} \cdots \frac{c_{n_k}}{n_k!} \\ &= \frac{1}{n!} \sum_{n_1,\ldots,n_k,\sum n_i = n} \binom{n}{n_1 \quad n_2 \quad \cdots \quad n_k} c_{n_1} c_{n_2} \cdots c_{n_k}.\end{aligned}$$

Here, $\binom{n}{n_1\ n_2\ \cdots\ n_k} := \frac{n!}{n_1! n_2! \cdots n_k!}$ is the multinomial coefficient, and it counts the number of ways of partitioning a set of size n into a piece of size n_1, a piece of size n_2, etc. (If you haven't seen this before, verify this! It's a generalization of the usual binomial coefficients.) So this sum is precisely describing all the ways of placing k disjoint colored cycles, and hence is exactly h_n, as required.

Again, to finish the lemma, we note that we've overcounted every permutation with k cycles precisely $k!$ times, since there are $k!$ ways of coloring the cycles. So the EGF for what we're interested in is $\frac{1}{k!}\hat{C}(x)^k$, as required. □

The rest of the proof is easy now. We just need to consider all possible partitions with any number of parts between 1 and n, so we should sum up the EGF's for each:

$$\hat{P}(x) = \sum_{k=1}^{n} \frac{1}{k!}(\hat{C}(x))^k = e^{\hat{C}(x)};$$

the last step follows by considering the Taylor series of $e^{\hat{C}(x)}$. □

The thing to notice about this proof is that we never used anything specific about the values c_n—all that we used was that any permutation can be uniquely described by a partition into cycles. The theorem in fact holds much more generally. Let's see another example.

A *derangement* on a set S is a permutation π of S such that $\pi(i) \neq i$ for all $i \in S$; no element is kept fixed by the permutation. Let d_n denote the number of derangements on a set of size n, defining $d_0 = 1$.

It should be clear that a derangement can be partitioned into cycles *of length at least 2*. What is the generating function for cycles of length at least 2? Let c'_n be the number of cycles of length at least 2 on a set of size n. Then clearly $c'_0 = c'_1 = 0$, and $c'_n = c_n$ for $n \geq 2$. Since $c_0 = 0$ but $c_1 = 1$, we deduce that the EGF for (c'_n) is $C'(x) = \hat{C}(x) - x = -\log(1-x) - x$.

But now we can deduce that the EGF for $(d_n)_{n\in\mathbb{N}}$ is simply

$$\hat{D}(x) = \exp(C'(x)) = \frac{e^{-x}}{1-x}.$$

From this, a formula for d_n can be determined; we leave this as an exercise.

This relation is certainly not restricted to permutations. For example, any graph on n nodes can be described as a collection of *connected* graphs. If $\hat{G}(x)$ is the generating function for the number of graphs, and $\hat{H}(x)$ the generating function for connected graphs, then $\hat{G}(x) = e^{\hat{H}(x)}$.