

A crash course on generating functions

by

MARCEL K. GOH

10 JULY 2020

1. Introduction

When counting discrete structures, we often run into sequences that are recursively defined. For example, suppose we have the recurrence

$$R_n = \begin{cases} 1, & \text{if } n = 0; \\ 2R_{n-1} + 3, & \text{if } n \geq 1. \end{cases}$$

We might want to know how the sequence (R_n) grows as n gets large. Now, this particular recurrence is easy enough to solve by the substitution method. But often, it is easier to work with the *generating function* of a sequence. For a sequence (A_n) , this is the formal power series

$$A(z) = \sum_{n \geq 0} A_n z^n.$$

In his textbook, titled *generatingfunctionology*, H. S. Wilf famously wrote: “A generating function is a clothesline on which we hang up a sequence of numbers for display.” That is exactly how we should treat the function $A(z)$ for now; in later sections, we will discuss how this series can be treated as an analytic object.

Let us try to find the generating function for the sequence (R_n) defined above. Right off the bat, we have

$$R(z) = \sum_{n \geq 0} R_n z^n = R_0 + \sum_{n \geq 1} R_n z^n.$$

Plugging in the recursive definition of R_n , we get

$$\begin{aligned} R(z) &= 1 + \sum_{n \geq 1} (2R_{n-1} + 3)z^n \\ &= 1 + 2 \sum_{n \geq 1} R_{n-1} z^n + 3 \sum_{n \geq 1} z^n \\ &= 1 + 2z \sum_{n \geq 0} R_n z^n + 3 \left(\sum_{n \geq 0} z^n - 1 \right) \end{aligned} \tag{1}$$

The first summation on the right-hand side is $R(z)$ and the second one is well-known, so we derive the generating function

$$\begin{aligned} R(z) &= 1 + 2zR(z) + 3 \left(\frac{1}{1-z} - 1 \right) \\ (1-2z)R(z) &= \frac{1+2z}{1-z} \\ R(z) &= \frac{1+2z}{(1-z)(1-2z)}. \end{aligned}$$

Here we can use the partial fraction expansion to get

$$\begin{aligned} R(z) &= \frac{1+2z}{(1-z)(1-2z)} = \frac{4}{1-2z} - \frac{3}{1-z} \\ &= 4 \sum_{n \geq 0} (2z)^n - 3 \sum_{n \geq 0} z^n \\ &= \sum_{n \geq 0} (2^{n+2} - 3)z^n, \end{aligned} \tag{2}$$

implying that $R_n = 2^{n+2} - 3$. This could have been found by easier methods, but this example demonstrates how the generating function of a sequence provides a bridge from the recurrence to a closed form. In later sections, we will see that if we only want asymptotic estimates, then we can actually stop at the generating function equation found in (1).

2. The Symbolic Method and Ordinary Generating Functions

Deriving generating functions is a rather drawn-out process, and it would not be very enjoyable to have to work from scratch every time. Luckily, generating functions are often made up of simpler ones, so we can use the *symbolic method* to arrive directly at a generating function equation. This is the subject of the monolithic textbook *Analytic Combinatorics*, by P. Flajolet and R. Sedgewick. In these notes, we will often skip some details and proofs; anything we gloss over here can certainly be found somewhere in *Analytic Combinatorics*. We will refer to this book from time to time and abbreviate it by ‘AC’.

We define a *combinatorial class* \mathcal{A} to be a countable set with a *size function* $|\cdot| : \mathcal{A} \rightarrow \mathbf{N}_0$ for which the number of elements of any given size is finite. Here are examples of combinatorial classes:

- i) The set of all strings of 0 and 1, with size given by string length.
- ii) The set of all permutations. A permutation of size n is a one-to-one correspondence
- iii) The set of all binary trees, with size given by the number of nodes. from $[1 \dots n]$ to itself.

Ordinary generating functions. The *ordinary generating function*, or OGF, of a combinatorial class \mathcal{A} is defined to be the formal power series

$$A(z) = \sum_{a \in \mathcal{A}} z^{|a|} = \sum_{n \geq 0} A_n z^n,$$

where A_n is the number of elements in \mathcal{A} of size n . It is easy to see that the two summations are equivalent characterisations of $A(z)$, and $A_n = [z^n]A(z)$ is the coefficient of z^n in $A(z)$. We now define the *neutral class* \mathcal{E} containing only one element, of size 0. Thus the generating function of \mathcal{E} is 1. Likewise, we define the *atomic class* \mathcal{Z} , which contains a single element of size 1. The generating function of \mathcal{Z} is z .

When classes \mathcal{A} and \mathcal{B} are disjoint, we will denote their *union* by

$$\mathcal{A} + \mathcal{B} = \mathcal{A} \cup \mathcal{B}.$$

It is easy to see that the generating function of $\mathcal{A} + \mathcal{B}$ is $A(z) + B(z)$. In a similar vein, let the *product* of two combinatorial classes be given by

$$\mathcal{A} \times \mathcal{B} = \{(a, b) : a \in \mathcal{A}, b \in \mathcal{B}\};$$

the generating function of $\mathcal{A} \times \mathcal{B}$ is $A(z) \cdot B(z)$. We will denote the product of n copies of \mathcal{A} by \mathcal{A}^n (this is \mathcal{E} if $n = 0$), which has the generating function $A(z)^n$.

A more complex construction is the *sequence class*. For a class \mathcal{A} , this is

$$\text{SEQ}(\mathcal{A}) = \mathcal{E} + \mathcal{A} + (\mathcal{A} \times \mathcal{A}) + (\mathcal{A} \times \mathcal{A} \times \mathcal{A}) + \dots$$

It follows that the generating function of $\text{SEQ}(\mathcal{A})$ is

$$S(z) = 1 + A(z) + A(z)^2 + A(z)^3 + \dots = \frac{1}{1 - A(z)}.$$

More generally, if Ω is a subset of \mathbf{N}_0 , we define $\text{SEQ}_\Omega(\mathcal{A})$ as the class $\sum_{n \in \Omega} \mathcal{A}^n$, with generating function $\sum_{n \in \Omega} A(z)^n$. A list of further constructions can be found in Fig. I.18 of AC.

Integers and compositions. The set of all positive integers is a combinatorial class. To get its generating function, we note that the positive integers can be associated with nonempty sequences of unlabelled atoms:

$$\{1, 2, 3, \dots\} \cong \{\bullet, \bullet\bullet, \bullet\bullet\bullet, \dots\}$$

This gives the construction $\mathcal{I} = \text{SEQ}_{\geq 1}(\mathcal{Z})$, which immediately tells us that $I(z) = z/(1-z)$ and $I_n = 1$ for all $n \geq 1$.

A *composition* of an integer n is expression of n as a sum of a sequence of positive integers (so the order of the terms matters). For example, there are four compositions of the number 3:

$$1 + 1 + 1 \quad 1 + 2 \quad 2 + 1 \quad 3 \quad (3)$$

Let \mathcal{C} be the set of all integer compositions. We can describe a composition as a sequence of integers, so

$$\mathcal{C} = \text{SEQ}(\mathcal{I}).$$

From this, we find that

$$C(z) = \frac{1}{1 - z/(1-z)} = \frac{1-z}{1-2z}. \quad (4)$$

Reexpressing this as the series

$$C(z) = \frac{1}{1-2z} - \frac{z}{1-2z} = \sum_{n \geq 0} (2z)^n - \sum_{n \geq 0} 2^n z^{n+1} = 1 + \sum_{n \geq 1} 2^{n-1} z^n,$$

we find that $C_0 = 1$ and $C_n = 2^{n-1}$ for $n \geq 1$. Ignoring C_0 , this makes sense because there is a one-to-one correspondence between compositions and a subset of $n-1$ possible bars between n balls: the compositions in (3) correspond to

$$\bullet | \bullet | \bullet \quad \bullet | \bullet \bullet \quad \bullet \bullet | \bullet \quad \bullet \bullet \bullet .$$

Fibonacci numbers. Let \mathcal{I}^* denote the cute combinatorial class $\{1, 2\}$, which clearly has the generating function $I^*(z) = z + z^2$. Let \mathcal{T} be the set of partitions using only the numbers 1 and 2 (for simplicity's sake, we also include the neutral object, i.e. $T_0 = 1$). We have

$$\begin{aligned} 1 &= 1, \\ 2 &= 1 + 1 = 2, \\ 3 &= 1 + 1 + 1 = 1 + 2 = 2 + 1, \\ 4 &= 1 + 1 + 1 + 1 = 1 + 1 + 2 = 1 + 2 + 1 = 2 + 1 + 1 = 2 + 2, \end{aligned}$$

and so on. After computing a few more terms, we notice the pattern

$$(T_n)_{n \geq 0} = (1, 1, 2, 3, 5, 8, 13, \dots).$$

This is almost the infamous Fibonacci sequence; we have $T_n = F_{n+1}$. Hence $T(z)$ must be the generating function of the Fibonacci numbers, divided by z . We can easily derive this function with the symbolic method, since a composition using only 1s and 2s is simply a sequence of elements of \mathcal{I}^* . We conclude that

$$\mathcal{T} = \text{SEQ}(\mathcal{I}^*) \quad \text{and} \quad T(z) = \frac{1}{1 - I^*(z)} = \frac{1}{1 - z - z^2}. \quad (5),$$

so we know that the generating function of the Fibonacci numbers is

$$F(z) = \frac{z}{1 - z - z^2}. \quad (6)$$

Catalan numbers. Let \mathcal{B} be the class of all binary trees. This combinatorial class can be described recursively as follows: “A binary tree is either empty, or else it consists of a root node adjoined to two more binary trees.”. An empty binary tree has size 0, so it is represented by \mathcal{E} and a node has size 1, so we represent it by \mathcal{Z} . We arrive at the symbolic description

$$\mathcal{B} = \mathcal{E} + \mathcal{Z} \times \mathcal{B} \times \mathcal{B},$$

whence we derive a functional equation

$$B(z) = 1 + zB(z)^2$$

that the generating function must satisfy. Treating $B(z)$ as a variable and employing the quadratic formula, we find that

$$B(z) = \frac{1 - \sqrt{1 - 4z}}{2z}. \quad (7)$$

(Since we know that $B(0) = B_0 = 1$, the positive branch of the square root is invalid. On the other hand, l'Hospital's rule can be applied to the negative version to get a limit of 1.) From here, one could use Newton's generalisation of the binomial theorem to get that

$$B_n = \frac{1}{n+1} \binom{2n}{n}, \quad (8)$$

and indeed, $B(z)$ is the generating function of the Catalan numbers. From here, Stirling's approximation can be applied to give asymptotics, but the techniques of later sections will allow us to characterise the asymptotic growth directly from the generating function (7).

Making change. Suppose a cashier has exactly five loonies, three toonies, and four \$5 bills in her till. How many ways can she give \$23 in change? To solve this, we assign each coin/bill a size proportional to its value; so a loonie is described by \mathcal{Z} , a toonie by \mathcal{Z}^2 , and a \$5 bill by \mathcal{Z}^5 . Now we can describe the set of all possibilities with the combinatorial class

$$\mathcal{M} = \text{SEQ}_{\leq 5}(\mathcal{Z}) \times \text{SEQ}_{\leq 3}(\mathcal{Z}^2) \times \text{SEQ}_{\leq 4}(\mathcal{Z}^5).$$

This immediately gives the generating function

$$M(z) = (1 + z + z^2 + z^3 + z^4 + z^5)(1 + z^2 + z^4 + z^6)(1 + z^5 + z^{10} + z^{15} + z^{20}),$$

and the number of ways to make \$23 is exactly $[z^{23}]M(z)$. Note that if the cashier had an unlimited supply of coins and bills, we would have

$$\mathcal{M}^* = \text{SEQ}(\mathcal{Z}) \times \text{SEQ}(\mathcal{Z}^2) \times \text{SEQ}(\mathcal{Z}^5).$$

and the number of ways to make n dollars out of these denominations is

$$[z^n] \left(\frac{1}{1-z} \right) \left(\frac{1}{1-z^2} \right) \left(\frac{1}{1-z^5} \right).$$

The change-making problem was used as a primary example in G. Pólya's 1956 paper on the applications of generating functions.

3. Labelled Structures and Exponential Generating Functions

Ordinary generating functions deal with atoms that are unlabelled, meaning these atoms are indistinguishable from one another. When we want to count labelled objects, we will use an *exponential generating function*, or EGF. For a combinatorial class \mathcal{A} , this is the formal power series

$$A(z) = \sum_{a \in \mathcal{A}} \frac{z^{|a|}}{|a|!} = \sum_{n \geq 0} A_n \frac{z^n}{n!}.$$

In particular, $A_n = [z^n/n!]A(z)$ for any EGF $A(z)$.

The symbolic method applies to EGFs as well. As before, \mathcal{E} denotes the neutral class, with a single object of size 0 and \mathcal{Z} denotes the atomic class, with a single labelled object of size 1. If \mathcal{A} and \mathcal{B} are disjoint combinatorial classes with EGFs $A(z)$ and $B(z)$, then the EGF of their union $\mathcal{A} + \mathcal{B}$ is $A(z) + B(z)$. Instead of an ordinary Cartesian product, we have the notion of a *labelled product*, in which atoms are relabelled

in all consistent ways. For example, a valid relabelling of the labelled triple $(1, 3, 2)$ is $(2, 7, 5)$, because the relative ordering of the atoms is unchanged. The set of all labelled pairs of elements from \mathcal{A} and \mathcal{B} is denoted $\mathcal{A} \star \mathcal{B}$ and its EGF is $A(z) \cdot B(z)$.

The labelled sequence class $\text{SEQ}(\mathcal{A})$, consists of sequences of elements of \mathcal{A} . It has the EGF $1/(1-A(z))$, just as with the OGF, but notice that the EGF counts a lot more things. The easiest way to see this is to consider the generating function of $\text{SEQ}(\mathcal{Z})$, which is $S(z) = 1/(1-z)$. When considered as an OGF, there is only one element of each given size, since $[z^n]S(z) = 1$ for all n . But when taken as an EGF, we find that there are $n!$ objects of size n , because we have $[z^n/n!]S(z) = n!$ for all n . Of course, this makes sense since permuting the atoms in a sequence of length n creates a new labelled object, but does not change the unlabelled one.

There are two more constructions that will be useful to define on labelled classes. First, we let $\text{SET}_k(\mathcal{A})$ denote the class of all sets of size k with elements in \mathcal{A} . This is like a sequence of length k , except that the order is not important, meaning we can divide by a factor of $k!$. The generating function of $\text{SET}_k(\mathcal{A})$ is therefore $A(z)^k/k!$. Now we let

$$\text{SET}(\mathcal{A}) = \mathcal{E} + \mathcal{A} + \text{SET}_2(\mathcal{A}) + \text{SET}_3(\mathcal{A}) + \cdots$$

and derive the generating function for $\text{SET}(\mathcal{A})$:

$$T(z) = \sum_{k \geq 0} \frac{A(z)^k}{k!} = e^{A(z)} \quad (9)$$

The last construction we will handle is the class of k -cycles with elements in \mathcal{A} , denoted $\text{CYC}_k(\mathcal{A})$. There are k ways to represent a cycle as a sequence of k objects, so the generating function of $\text{CYC}_k(\mathcal{A})$ is $A(z)/k$. Now we define the class of unrestricted cycles

$$\text{CYC}(\mathcal{A}) = \text{CYC}_1(\mathcal{A}) + \text{CYC}_2(\mathcal{A}) + \text{CYC}_3(\mathcal{A}) + \cdots$$

and find that the generating function of $\text{CYC}(\mathcal{A})$ is

$$U(z) = \sum_{k \geq 1} \frac{A(z)^k}{k} = \log \frac{1}{1-A(z)}. \quad (10)$$

There are many other interesting EGF constructions; see, for example, Fig. II.18 in AC.

Note that a permutation is simply a set of cycles:

$$\text{SEQ}(\mathcal{Z}) = \text{SET}(\text{CYC}(\mathcal{Z}))$$

Of course, the cycle decomposition of a permutation is well-known, but generating functions give us another way of seeing that this is true, since

$$\frac{1}{1-z} = \exp \left(\log \frac{1}{1-z} \right).$$

Now is as good a time as any to state, without proof, an important classic theorem that will be used in the next example.

Theorem L (*Lagrange inversion theorem*). *If a generating function $g(z)$ satisfies the equation $z = f(g(z))$, with $f(0) = 0$ and $f'(0) \neq 0$, then for an arbitrary function $H(u)$, we have*

$$[z^n]H(g(z)) = \frac{1}{n}[u^{n-1}]H'(u) \left(\frac{u}{f(u)} \right)^n.$$

In particular, for $H(u) = u$ we have

$$[z^n]g(z) = \frac{1}{n}[u^{n-1}] \left(\frac{u}{f(u)} \right)^n.$$

(This particular form of Lagrange's inversion theorem is often called the *Bürmann form*.) ■

Cayley trees. We are now equipped to count rooted non-plane labelled trees, sometimes called Cayley trees. Children of a given node are not ordered and nodes in these trees have unrestricted degree, so we have the specification

$$\mathcal{T} = \mathcal{Z} \star \text{SET}(\mathcal{T}).$$

This immediately gives the functional equation

$$T(z) = ze^{T(z)}$$

and by Lagrange's inversion theorem, with $f(u) = u/e^u$ and $H(u) = u$, we obtain

$$n![z^n]T(z) = n! \left(\frac{1}{n} [u^{n-1}] e^{un} \right) = n! \cdot \frac{1}{n} \cdot \frac{n^{n-1}}{(n-1)!} = n^{n-1}.$$

Note that there are n^{n-2} *unrooted* Cayley trees, since there are n choices for the root. This is called Cayley's formula and the standard combinatorial proof involves establishing a bijection with Prüfer sequences; see Prüfer (1918).

Derangements. A derangement is a permutation in which no element is fixed. This can equally be characterised as a permutation with no cycle of length 1. We saw earlier that permutations is simply a set of cycles, so disallowing cycles of length 1 gives the specification

$$\mathcal{D} = \text{SET}(\text{CYC}_{\geq 1}(\mathcal{Z})).$$

We remove singleton cycles from the generating function by simply subtracting the generating function of $\text{CYC}_1(\mathcal{Z})$ (which is simply z , because a singleton cycle is just an atom) from $\text{CYC}(\mathcal{Z})$, so

$$D(z) = \exp \left(\log \frac{1}{1-z} - z \right) = \frac{e^{-z}}{1-z}. \quad (11)$$

We will revisit this generating function in the context of a word problem once we have the tools to analyse its asymptotics.

4. Counting Parameters with Bivariate Generating Functions

Some situations may require us to do more than just count the number of objects of a certain size. In addition, we might like to know how many of a certain substructure is embedded in an object of a combinatorial class. We can use a *bivariate generating function* or BGF to accomplish this. (BGFs are part of a larger class of *multivariate generating functions*, but we will keep things simple in this crash course. See Chapter III of AC for many more constructions.)

Consider a two-dimensional array of numbers $(a_{n,k})$, where n counts the number of objects of a certain size and k counts some other combinatorial parameter. For example, we might have

- i) The number of n -bit strings of 0s and 1s that have exactly k 0s.
- ii) The number of n -letter permutations that have exactly k cycles.
- iii) The number of binary trees of size n that have k leaves.

Symbolically, suppose we have a combinatorial class \mathcal{A} with an extra scalar parameter $\chi : \mathcal{A} \rightarrow \mathbf{N}_0$. The ordinary and exponential BGFs are defined, respectively, to be the power series

$$\sum_{a \in \mathcal{A}} z^{|a|} u^{\chi(a)} \quad \text{and} \quad \sum_{a \in \mathcal{A}} \frac{z^{|a|}}{n!} u^{\chi(a)},$$

which can be rewritten as

$$\sum_{n,k} a_{n,k} z^n u^k \quad \text{and} \quad \sum_{n,k} a_{n,k} \frac{z^n}{n!} u^k,$$

where $a_{n,k}$ is the number of objects of size a with $\chi(a) = k$. Note that substituting $u = 1$ into a BGF returns the ordinary counting sequence for the class \mathcal{A} .

Bitstrings. To warm up, we will consider a very simple example. Let \mathcal{B} be the class of all binary strings, where for $b \in \mathcal{B}$, its size n is given by its length and $\chi(b)$ is the number of zeroes in the string. There are two distinguishable atoms, \mathcal{Z}_0 and \mathcal{Z}_1 , but we introduce a u only when a zero is present, so we have the specification

$$\mathcal{B} = \text{SEQ}(u\mathcal{Z}_0 + \mathcal{Z}_1)$$

with generating function

$$B(z, u) = \frac{1}{1 - (u+1)z}. \quad (12)$$

When we substitute $u = 1$, we get

$$B(z, 1) = \frac{1}{1 - 2z},$$

which is the generating function of 2^n , the number of binary strings with n bits. The number of n -bit binary strings with k zeroes can now be extracted. It is equal to

$$[u^k][z^n]B(z, u) = [u^k](u+1)^n = \binom{n}{k},$$

by an application of the binomial theorem. Hence the probability that an n -bit binary string has k zeroes is

$$\frac{[u^k][z^n]B(z, u)}{[z^n]B(z, 1)} = \frac{\binom{n}{k}}{2^n}. \quad (13)$$

Probabilities and moments. The above example has an elementary result, but it illustrates how BGFs allow us to extract probabilities. For all a in a class \mathcal{A} , we have

$$\mathbf{P}\{\chi(a) = k \mid |a| = n\} = \frac{[u^k][z^n]A(z, u)}{[z^n]A(z, 1)}. \quad (14)$$

We can also calculate higher moments with the BGF. Notice that if we take the partial derivative with respect to u , the χ values of each term pop down into the coefficient of each term:

$$\frac{\partial}{\partial u} A(z, u) = \frac{\partial}{\partial u} \sum_{a \in \mathcal{A}} z^{|a|} u^{\chi(a)} = \sum_{a \in \mathcal{A}} \chi(a) z^{|a|} u^{\chi(a)-1}$$

Of course, the exponents of the variable u have changed, but letting $A_u(z, u)$ denote the partial derivative with respect to u and evaluating at $u = 1$, we get the cumulated cost

$$A_u(z, u)|_{u=1} = \sum_{n \geq 0} \left(\sum_{a \in \mathcal{A}_n} \chi(a) \right) z^n, \quad (15)$$

whence we can calculate the mean over \mathcal{A}_n , the class of all $a \in \mathcal{A}$ of size n , to be

$$\mathbf{E}_{\mathcal{A}_n}\{\chi\} = \frac{[z^n]A_u(z, u)|_{u=1}}{[z^n]A(z, 1)}. \quad (16)$$

These formulas are valid even when the generating functions are exponential, because the $n!$ factors cancel out. More generally, by taking repeated derivatives, we can calculate factorial moments of χ . Let $A_u^r(z, u)$ denote the r th partial derivative of A with respect to u . We have

$$\mathbf{E}_{\mathcal{A}_n}\{\chi(\chi-1)\cdots(\chi-r+1)\} = \frac{[z^n]A_u^r(z, u)|_{u=1}}{[z^n]A(z, 1)}. \quad (17)$$

For example, by linearity of expectation, the second moment is easily seen to be

$$\begin{aligned} \mathbf{E}_{\mathcal{A}_n}\{\chi^2\} &= \mathbf{E}_{\mathcal{A}_n}\{\chi + \chi(\chi-1)\} \\ &= \mathbf{E}_{\mathcal{A}_n}\{\chi\} + \mathbf{E}_{\mathcal{A}_n}\{\chi(\chi-1)\} \\ &= \frac{[z^n]A_u(z, u)|_{u=1}}{[z^n]A(z, 1)} + \frac{[z^n]A_u^2(z, u)|_{u=1}}{[z^n]A(z, 1)}. \end{aligned} \quad (18)$$

This can be easily used to calculate the variance and standard deviation of χ , from the usual formula

$$\mathbf{V}_{\mathcal{A}_n}\{\chi\} = \sigma(\chi)^2 = \mathbf{E}_{\mathcal{A}_n}\{\chi^2\} - \mathbf{E}_{\mathcal{A}_n}\{\chi\}^2. \quad (19)$$

Cycles in a permutation. Recall that a permutation is a set of cycles. If we mark every cycle with a u , we get the specification

$$\mathcal{P} = \text{SET}(u \text{CYC}(\mathcal{Z})),$$

which gives the bivariate exponential generating function

$$P(z, u) = \exp\left(u \log \frac{1}{1-z}\right) = \exp\left(\log \frac{1}{(1-z)^u}\right) = \left(\frac{1}{1-z}\right)^u \quad (20)$$

for the number of n -letter permutations with exactly k cycles. We can calculate the partial derivatives

$$P_u(z, u) = \left(\frac{1}{1-z}\right)^u \log \frac{1}{1-z}$$

Here the random variable χ counts the number of cycles, and we have

$$\mathbf{E}_{\mathcal{P}_n}\{\chi\} = \frac{[z^n]P_u(z, u)|_{u=1}}{[z^n]P(z, 1)} = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = H_n.$$

The harmonic numbers H_n can be expanded as $\log n + \gamma + O(1/n)$, so we find that there are roughly $\log n$ cycles in a random permutation on n letters. A slightly harder analysis of Taylor coefficients retrieves the second moment. From this, it can be shown that $\mathbf{V}_{\mathcal{P}_n}\{\chi\} \sim \log n$ as well.

External nodes in a binary tree. Hidden inside the specification for binary trees, we can see that we are actually only counting internal nodes, since we have given \mathcal{E} , the external node, the generating function of 1 (meaning it has size 0). If we mark it with a u , we arrive at

$$\mathcal{B} = u\mathcal{E} + \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$$

and the bivariate ordinary generating function satisfies

$$B(z, u) = u + zB(z)^2.$$

Using the quadratic formula once again, we find that the generating function is

$$B(z, u) = \frac{1 - \sqrt{1 - 4uz}}{2z},$$

which checks out, since setting $u = 1$ gives exactly (7). We have the partial derivatives

$$P_u(z, u) = \frac{1}{\sqrt{1-4uz}} \quad \text{and} \quad P_u^2(z, u) = \frac{2z}{(1-4uz)^{3/2}}.$$

With the substitution $u = 1$, these are both relatively well-known generating functions. The first is the generating function for the central binomial coefficients $\binom{2n}{n}$ and the second is $z \frac{\partial}{\partial z} P_u(z, 1)$, so it has coefficients $n \binom{2n}{n}$ (the powers of n drop down into the coefficient when differentiating, and multiplying by z restores the correct power). We calculated in (8) that the total number of binary trees with n internal nodes is $\binom{2n}{n}/(n+1)$, so applying our method gives us

$$\mathbf{E}_{\mathcal{B}_n}\{\chi\} = \frac{\binom{2n}{n}}{\frac{1}{n+1}\binom{2n}{n}} = n+1$$

and

$$\mathbf{E}_{\mathcal{B}_n}\{\chi^2\} = \frac{n \binom{2n}{n}}{\frac{1}{n+1}\binom{2n}{n}} + n+1 = n(n+1) + n+1 = (n+1)^2.$$

So the average number of external nodes in a tree with n internal nodes is $n+1$ and since

$$\mathbf{V}_{\mathcal{B}_n}\{\chi\} = \mathbf{E}_{\mathcal{B}_n}\{\chi^2\} - \mathbf{E}_{\mathcal{B}_n}\{\chi\}^2 = (n+1)^2 - (n+1)^2 = 0,$$

we discover that, in fact, every tree with n internal nodes has exactly $n+1$ external nodes. (We have performed what is probably the most convoluted proof of this simple fact, but it is a good exercise in calculating moments using bivariate generating functions.)

References

Philippe Flajolet and Robert Sedgewick, *Analytic Combinatorics* (New York: Cambridge University Press, 2009).

George Pólya, “On picture-writing,” *American Mathematical Monthly* **63** (1956), 689–697.

Heinz Prüfer, “Neuer Beweis eines Satzes über Permutationen,” *Archiv der Mathematik und Physik* **27** (1918), 142–144.

Herbert S. Wilf, *generatingfunctionology* (Boston: Academic Press, 1994).