11. Introduction to Exponential Generating Functions.

We have seen several applications of generating functions – more specifically, of ordinary generating functions. Exponential generating functions are of another kind and are useful for solving problems to which ordinary generating functions are not applicable.

Ordinary generating functions arise when we have a (finite or countably infinite) set of objects S and a weight function $\omega : S \to \mathbb{N}^r$. Then the ordinary generating function $\Phi_{\mathcal{G}}^{\omega}(\mathbf{x})$ is defined and we can proceed with calculations. Exponential generating functions arise in a somewhat more complicated situation. The basic idea is that they are used to enumerate "combinatorial structures on finite sets". In this section I will try to give you some idea of what this means without getting bogged down in an axiomatic development. In the process, we will be able to derive enough of the theory to solve some interesting problems. A more formal treatment of the subject is postponed until Chapter 12, which includes proper foundations of the theory as well as discussion of some subtle issues which we can't even speak about until the language is developed. But all that is for later! Right now, let's concentrate on the general ideas, and save the niggling for when we've already got the big picture.

So – a "combinatorial structure on a finite set" – just what does that mean? Graphs are good examples: a graph $G = (V, E)$ consists of a finite set V together with some additional structure, in this case a set E of two–element subsets of V. Endofunctions are also good examples: an *endofunction* is a finite set V together with some additional structure, in this case a function $\phi : V \to V$ from V to itself. Generally, a "combinatorial structure on a finite set" means a finite set together with some additional information defined in terms of that set.

Of course, we are interested in counting things. So we will not consider just one (combinatorial) structure (on a finite set), but an entire family of related structures, called a *class* of structures. For example, we can consider the class $\mathcal{G}$ of all graphs. This consists of all finite sets V and all graph structures $G = (V, E)$ on these finite sets. It's a pretty big thing! (In fact, it is way too big even to be a set – but that's another story.) The point about the class $\mathcal{G}$ is that to each finite set X it associates the finite set $\mathcal{G}_X$ of all graphs which have X as their set of vertices. Notice that if $X \neq Y$ are two different finite sets then $\mathcal{G}_X \cap \mathcal{G}_Y = \varnothing$, since if $G \in \mathcal{G}_X \cap \mathcal{G}_Y$ then $X = V(G) = Y$. Also notice that if $\#X = n$ then $\#\mathcal{G}_X = 2^{n(n-1)/2}$, so that $\#\mathcal{G}_X$ depends only on $\#X$. These properties are the essential ones that we abstract to define a class of structures.

Definition 11.1 (Classes of Structures)**.** A *class* $\mathcal{A}$ of structures associates to every finite set X another finite set $\mathcal{A}_X$, in such a way that the following two conditions are satisfied:
(i) if $X \neq Y$ are distinct finite sets then $\mathcal{A}_X \cap \mathcal{A}_Y = \varnothing$;
(ii) if X and Y are finite sets with $\#X = \#Y$, then $\#\mathcal{A}_X = \#\mathcal{A}_Y$.

(In Section 12 we will enrich this definition and speak about "natural" classes, but this will suffice for now.) The interpretation of the class $\mathcal{A}$ is that $\mathcal{A}_X$ is the finite set of combinatorial structures in the class $\mathcal{A}$ which are defined in terms of the finite set X of "vertices".

Now we see the kind of enumeration problem that exponential generating functions are designed to solve: given a class $\mathcal{A}$ of structures, determine $\#\mathcal{A}_X$ for all finite sets X. That is, determine how many $\mathcal{A}$–structures are defined on each finite set. Of course, by condition (ii) of Definition 11.1, this amounts to determining $\#\mathcal{A}_{N_n}$ for all $n \in \mathbb{N}$. The notation $\mathcal{A}_{N_n}$ is a bit awkward, so let's use $\mathcal{A}_n := \mathcal{A}_{N_n}$ to mean the same thing.

We could put all the numbers $\#\mathcal{A}_n$ for $n \in \mathbb{N}$ together into a generating function:

$$\sum_{n=0}^{\infty} (\#\mathcal{A}_n) x^n$$

but it turns out that this is not the way to do it. The reason why this is no good is that the combinatorial operations we will use to analyze and manipulate classes of structures are not reflected by algebraic operations on these power series. Instead, the proper way to translate the combinatorics into algebra in this situation is as follows.

Definition 11.2 (Exponential Generating Functions)**.** Let $\mathcal{A}$ be a class of structures. The *exponential generating function of* $\mathcal{A}$ is

$$A(x) := \sum_{n=0}^{\infty} (\#\mathcal{A}_n) \frac{x^n}{n!}.$$

(This definition will be embellished a little later to include more indeterminates, but this is the essential form.)

Let's illustrate this with a few cheap examples for which we already know the answer.

Example 11.3. First, consider the class $\mathcal{S}$ of permutations: to each finite set X it associates the finite set $\mathcal{S}_X$ of all bijections $\sigma : X \to X$ from X to X. Condition (i) is easy, and condition (ii) follows from Example 2.2, so that $\mathcal{S}$ satisfies Definition 11.1. Way back in Theorem 2.1 we saw that $\#\mathcal{S}_n = n!$ for all $n \in \mathbb{N}$, so that the exponential generating function for the class of permutations is

$$S(x) := \sum_{n=0}^{\infty} (\#\mathcal{S}_n)\frac{x^n}{n!} = \sum_{n=0}^{\infty} n!\frac{x^n}{n!} = \frac{1}{1-x}.$$

Example 11.4. Second, consider the class $\mathcal{C}$ of *cyclic permutations*. Recall from Exercise 2.2 that $\#\mathcal{C}_0 = 0$ and $\#\mathcal{C}_n = (n-1)!$ for all $n \geq 1$. Condition (i) is easy, and you should think about how to verify condition (ii). Therefore, $\mathcal{C}$ is a class and its exponential generating function is

$$C(x) := \sum_{n=0}^{\infty} (\#\mathcal{C}_n)\frac{x^n}{n!} = \sum_{n=1}^{\infty} (n-1)!\frac{x^n}{n!} = \log\left(\frac{1}{1-x}\right)$$

by Example 7.9(a).

Example 11.5. Third, consider the class $\mathcal{E}$ of *finite sets*: to each finite set X this associates the set $\mathcal{E}_X := \{X\}$ containing only X. This seems like much ado about nothing, but it will be very useful in a little while. That is, $\mathcal{E}$ is the class of finite sets with no additional structure. Conditions (i) and (ii) are clear in this case. The exponential generating function for the class $\mathcal{E}$ is

$$E(x) := \sum_{n=0}^{\infty} (\#\mathcal{E}_n)\frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x)$$

by Example 7.9(b).

Notice that we have the relation

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

among these power series. Using the names of the exponential generating functions, that is $S(x) = E(C(x))$. This suggests that some combinatorial relation exists among the classes $\mathcal{S}$, $\mathcal{C}$, and $\mathcal{E}$ – a relation which it would be sensible to denote by something like $\mathcal{S} \equiv \mathcal{E}[\mathcal{C}]$. In fact this is the case – a permutation is equivalent to a finite set of pairwise disjoint cyclic permutations. Our first task is to develop enough of the theory to make sense of an expression like $\mathcal{S} \equiv \mathcal{E}[\mathcal{C}]$.

The usefulness of this theory stems from the ability to identify combinatorial relations among classes – as above – and then to translate these into functional equations for the corresponding exponential generating functions. After that, one can apply algebraic techniques such as the Lagrange Implicit Function Theorem to extract the coefficients of these generating functions, thereby solving the relevant enumeration problems. To realize this program we need to discuss several operations on classes of structures. Then we will have developed enough technique to analyze some nontrivial problems and derive some interesting results. (On a first pass through these constructions it might help to read the first several carefully and then skim quickly through the rest. After seeing how they are applied in a few problems one can return and reread them all carefully.)

Definition 11.6 (Equivalence of Classes)**.** Two classes $\mathcal{A}$ and $\mathcal{B}$ are *(numerically) equivalent* if $\#\mathcal{A}_X = \#\mathcal{B}_X$ for every finite set X. Of course, this is the case if and only if their exponential generating functions are equal: $A(x) = B(x)$. We use the notation $\mathcal{A} \equiv \mathcal{B}$ to denote that $\mathcal{A}$ and $\mathcal{B}$ are equivalent.

(This concept of equivalence will be superseded in Section 12 by the much more interesting concept of "natural equivalence", but this will do for now.)

Definition 11.7 (Local Finiteness and Sums of Classes)**.** Let $(\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \ldots)$ be a (finite or infinite) sequence of classes. We say that this sequence is *locally finite* provided that for every finite set X, at most finitely many of the sets $(\mathcal{A}_X^{(i)} : i \geq 1)$ are not empty. If this is the case then the set

$$\mathcal{B}_X := \bigcup_{i=1}^{\infty} \left(\{i\} \times \mathcal{A}_X^{(i)} \right)$$

is a finite union of finite sets, so that $\mathcal{B}_X$ is a finite set. A typical element of $\mathcal{B}_X$ is an ordered pair (i, α) with $i \geq 1$ and $\alpha \in \mathcal{A}_X^{(i)}$. The presence of the first coordinate ensures that the sets $\{i\} \times \mathcal{A}_X^{(i)}$ are pairwise disjoint. Since the union defining $\mathcal{B}_X$ is a disjoint union, we have

$$\#\mathcal{B}_X = \sum_{i=1}^{\infty} \#\mathcal{A}_X^{(i)}.$$

Conditions (i) and (ii) of Definition 12.1 can now be verified easily for $\mathcal{B}$. In summary, for a locally finite sequence of classes $(\mathcal{A}^{(i)} : i \geq 1)$ the class $\mathcal{B}$ is well–defined, and is called the *sum* of the sequence. The exponential generating function of $\mathcal{B}$ is

$$B(x) = \sum_{i=1}^{\infty} A(x).$$

The sum of classes $\mathcal{A}^{(i)}$ for $i \geq 1$ is usually denoted by

$$\bigoplus_{i=1}^{\infty} \mathcal{A}^{(i)}.$$

We could use another set of indices for the classes $\mathcal{A}^{(i)}$ as well, rather than $\{1, 2, \ldots\}$. For example, with only two classes we would just write $\mathcal{A}^{(1)} \oplus \mathcal{A}^{(2)}$, and so on.

As an example of this definition, consider the class $\mathcal{A} \oplus \mathcal{A}$ (in which $\mathcal{A}$ is any class). For a finite set X, a typical element of $(\mathcal{A} \oplus \mathcal{A})_X$ is of the form (i, α) with $i \in \{1, 2\}$ and $\alpha \in \mathcal{A}_X$. So $\mathcal{A} \oplus \mathcal{A}$ is the class of $\mathcal{A}$–structures each of which has been given one of two "colours": $i = 1$ means "red" while $i = 2$ means "blue". The exponential generating function of $\mathcal{A} \oplus \mathcal{A}$ is $2A(x)$ as it should be.

Definition 11.8 (Subclasses and Difference of Classes)**.** Two classes $\mathcal{A}$ and $\mathcal{B}$ are such that $\mathcal{A}$ *is a subclass of* $\mathcal{B}$ provided that for every finite set X, $\mathcal{A}_X \subseteq \mathcal{B}_X$. In this case we can define a class $\mathcal{B} \smallsetminus \mathcal{A}$, called $\mathcal{B}$ *minus* $\mathcal{A}$, by saying that

$$(\mathcal{B} \smallsetminus \mathcal{A})_X := \mathcal{B}_X \smallsetminus \mathcal{A}_X$$

for every finite set X. The exponential generating function of $\mathcal{B} \smallsetminus \mathcal{A}$ is $B(x) - A(x)$.

Now, properly speaking, $\mathcal{A}$ is not a subclass of $\mathcal{A} \oplus \mathcal{B}$. However, for any finite set X there is an injective function

$$\begin{array}{rcl} \mathcal{A}_X & \to & (\mathcal{A} \oplus \mathcal{B})_X \\ \alpha & \mapsto & (1, \alpha) \end{array}$$

so $\mathcal{A}$ is equivalent to a subclass of $\mathcal{A} \oplus \mathcal{B}$. From the point of view of exponential generating functions this is good enough. There is some subtlety to this way of comparing classes of structures, to which we return in Section 12.

Definition 11.9 (Superposition of Classes)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be any two classes. The *superposition* of $\mathcal{A}$ and $\mathcal{B}$ is the class $\mathcal{A}\&\mathcal{B}$ defined as follows: for any finite set X,

$$(\mathcal{A}\&\mathcal{B})_X := \mathcal{A}_X \times \mathcal{B}_X.$$

That is, an $(\mathcal{A}\&\mathcal{B})$–structure on X is an ordered pair (α, β) in which α is an $\mathcal{A}$–structure on X and β is a $\mathcal{B}$–structure on X. Certainly $\#(\mathcal{A}\&\mathcal{B})_X = (\#\mathcal{A}_X)(\#\mathcal{B}_X)$, which implies condition (ii), and condition (i) is also easily verified. There is no elementary formula for the exponential generating function of $\mathcal{A}\&\mathcal{B}$ in terms of $A(x)$ and $B(x)$. Nonetheless, superposition is sometimes useful and will be important in Section 12.

Definition 11.10 (Products of Classes)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be any two classes. The *product* of $\mathcal{A}$ and $\mathcal{B}$ is the class $\mathcal{A} * \mathcal{B}$ defined as follows: for any finite set X,

$$(\mathcal{A} * \mathcal{B})_X := \bigcup_{S \subseteq X} \left(\mathcal{A}_S \times \mathcal{B}_{(X \smallsetminus S)}\right).$$

That is, an $(\mathcal{A} * \mathcal{B})$–structure on the set X is an ordered pair (α, β) in which α is an $\mathcal{A}$–structure on some subset $S \subseteq X$, and β is a $\mathcal{B}$–structure on the complementary subset $X \smallsetminus S$ of X. Notice that condition (i) of the definition of classes $\mathcal{A}$ and $\mathcal{B}$ implies that the union defining $(\mathcal{A} * \mathcal{B})_X$ is a disjoint union. Thus, using condition (ii) as well, we calculate that the cardinality of this set is, for an n–element set X,

$$\#(\mathcal{A} * \mathcal{B})_X = \sum_{k=0}^{n} \binom{n}{k} (\#\mathcal{A}_k)(\#\mathcal{B}_{n-k}).$$

Since this depends only on $\#X = n$, $\mathcal{A} * \mathcal{B}$ satisfies condition (ii) of the definition of a class. Verification of condition (i) for $\mathcal{A} * \mathcal{B}$ is left as a good exercise. From multiplying the above equation by $x^n/n!$ and summing over all $n \in \mathbb{N}$, one easily

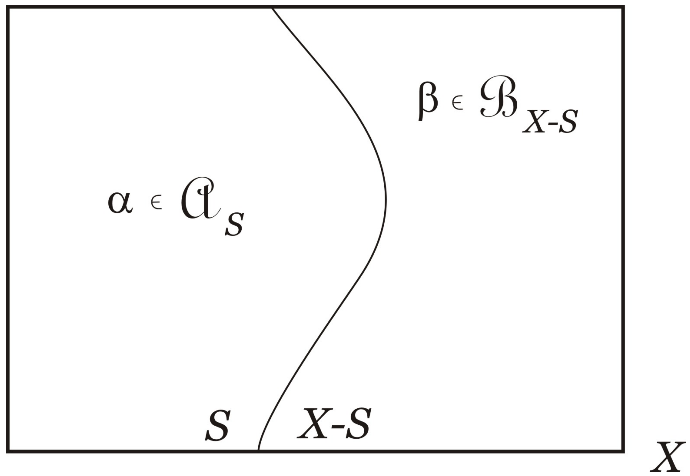

FIGURE 11.1. a structure from the class $\mathcal{A} * \mathcal{B}$.

calculates that the exponential generating function of $\mathcal{A} * \mathcal{B}$ is $A(x)B(x)$. Figure 11.1 illustrates the generic form of a structure from the class $\mathcal{A} * \mathcal{B}$.

Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be classes of structures. The classes $(\mathcal{A} * \mathcal{B}) * \mathcal{C}$ and $\mathcal{A} * (\mathcal{B} * \mathcal{C})$ are not equal, but they are equivalent. That is, for any finite set X we have $\#((\mathcal{A} * \mathcal{B}) * \mathcal{C})_X = \#(\mathcal{A} * (\mathcal{B} * \mathcal{C}))_X$ because of the following bijection:

$$\begin{array}{rcl} (\mathcal{A} * \mathcal{B}) * \mathcal{C} & \equiv & \mathcal{A} * (\mathcal{B} * \mathcal{C}) \\ ((\mathcal{A} * \mathcal{B}) * \mathcal{C})_X & \rightleftharpoons & (\mathcal{A} * (\mathcal{B} * \mathcal{C}))_X \\ ((\alpha, \beta), \gamma) & \leftrightarrow & (\alpha, (\beta, \gamma)) \end{array}$$

This extends similarly to any finite number of factors, so that we can speak about iterated products such as $\mathcal{A} * \mathcal{B} * \mathcal{C} * \cdots * \mathcal{D}$ unambiguously, at least modulo the equivalence relation. (The concept of natural equivalence is used in the next section to strengthen this sense in which we can say that the product $*$ is associative.)

Definition 11.11 (Powers of Classes)**.** Let $\mathcal{A}$ be a class of structures. By iterating the product construction we may define the *powers of* $\mathcal{A}$ to be the products of $\mathcal{A}$ with itself any finite number of times. That is, $\mathcal{A}^1 := \mathcal{A}$ and for all $k \geq 1$, $\mathcal{A}^{k+1} := \mathcal{A}^k * \mathcal{A}$. For all $k \geq 1$, the exponential generating function of $\mathcal{A}^k$ is $A(x)^k$, as can be seen by induction on k. We would like to define $\mathcal{A}^0$ as well – this should have exponential generating function $A(x)^0 = 1$ and be such that $\mathcal{A}^0 * \mathcal{A} \equiv \mathcal{A}$. The class $\mathcal{E}_0$ defined by putting

$$(\mathcal{E}_0)_X := \begin{cases} \{\varnothing\} & \text{if } X = \varnothing, \\ \varnothing & \text{if } X \neq \varnothing. \end{cases}$$

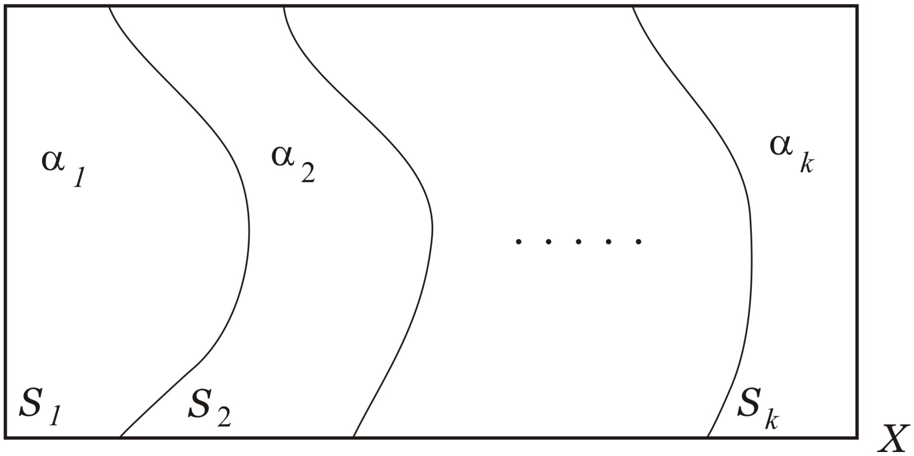

FIGURE 11.2. a structure from the class $\mathcal{A}^k$.

for each finite set X does have exponential generating function $E_0(x) = 1$. This class $\mathcal{E}_0$ is known as the class of *empty*, or *null* structure. By the product formula the classes $\mathcal{E}_0 * \mathcal{A}$ and $\mathcal{A}$ have the same generating function, so they are equivalent. For any natural numbers $j, k \in \mathbb{N}$ there is an equivalence $\mathcal{A}^j * \mathcal{A}^k \equiv \mathcal{A}^{j+k}$. An $\mathcal{A}^k$–structure on a set X is an ordered sequence of k $\mathcal{A}$–structures $(\alpha_1, \ldots, \alpha_k)$ which are pairwise disjoint and cover the set X. Figure 11.2 illustrates the generic form of a structure from the class $\mathcal{A}^k$.

Definition 11.12 (Finite Strings and Connected Classes)**.** Let $\mathcal{A}$ be a class of structures, and consider the sequence $(\mathcal{A}^j : \ j \in \mathbb{N})$ of powers of the class $\mathcal{A}$. We can form the sum of this sequence if (and only if) it is locally finite. This, however, need not be the case. For instance, if $\mathcal{A}_\varnothing \neq \varnothing$ then let $\alpha \in \mathcal{A}_\varnothing$. In this case, for each $k \in \mathbb{N}$, the sequence $(\alpha, \ldots, \alpha)$ of length k is an element of $\mathcal{A}^k_\varnothing$. This shows that if $\mathcal{A}_\varnothing \neq \varnothing$ then the sequence of powers of $\mathcal{A}$ is not locally finite. The converse is also true, and is left as an important exercise: if $\mathcal{A}_\varnothing = \varnothing$ then the sequence of powers of $\mathcal{A}$ is locally finite. We say that the class $\mathcal{A}$ is *connected* when $\mathcal{A}_\varnothing = \varnothing$. (This rather odd–sounding choice of terminology is explained after Example 11.19.) If $\mathcal{A}$ is a connected class then the powers of $\mathcal{A}$ form a locally finite sequence, and we define the *class of finite strings of $\mathcal{A}$–structures* to be the class

$$\mathcal{A}^* := \bigoplus_{k=0}^{\infty} \mathcal{A}^k.$$

By what has gone before, the exponential generating function of $\mathcal{A}^*$ is

$$A^*(x) := \sum_{k=0}^{\infty} A(x)^k = \frac{1}{1 - A(x)}.$$

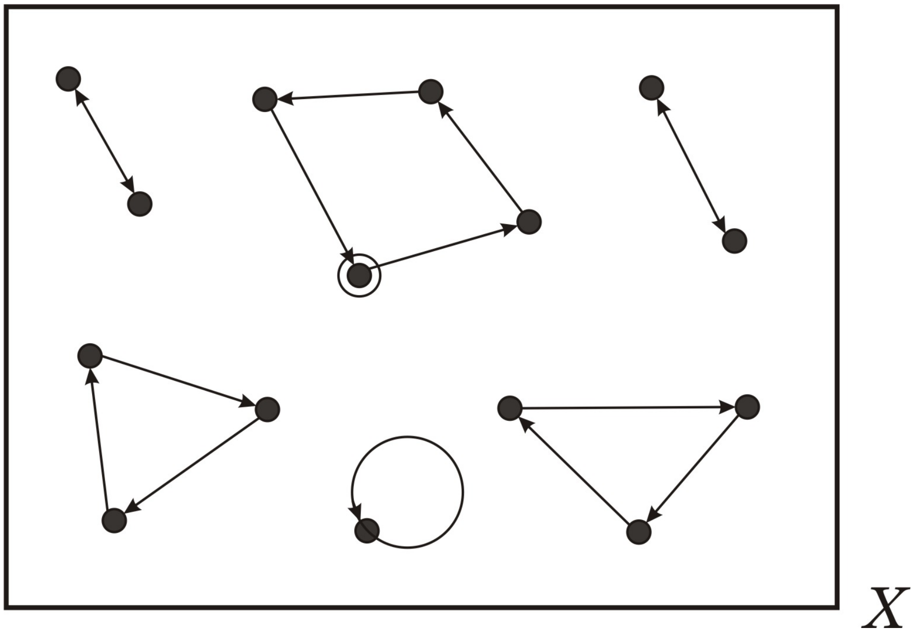

FIGURE 11.3. a rooted permutation.

Notice that the condition that $\mathcal{A}$ is connected corresponds to the condition that $[x^0]A(x) = 0$, which is exactly what is required for the composition of $A(x)$ into $1/(1-x)$ to be well–defined in the ring $\mathbb{Q}[[x]]$ of formal power series.

Definition 11.13 (Rooted Structures)**.** Let $\mathcal{A}$ be a class of structures. The class of *rooted* $\mathcal{A}$*–structures* is denoted by $\mathcal{A}^\bullet$ and defined as follows: for every finite set X,

$$\mathcal{A}^\bullet_X := \mathcal{A}_X \times X.$$

That is, a rooted $\mathcal{A}$–structure on the set X is an ordered pair (α, v) with $\alpha \in \mathcal{A}_X$ and $v \in X$. Conditions (i) and (ii) are easily verified for $\mathcal{A}^\bullet$. The exponential generating function of $\mathcal{A}^\bullet$ is

$$A^\bullet(x) := \sum_{n=0}^{\infty} (\#\mathcal{A}^\bullet_n)\frac{x^n}{n!} = \sum_{n=0}^{\infty} n(\#\mathcal{A}_n)\frac{x^n}{n!} = x\frac{d}{dx}A(x).$$

Notice that $\mathcal{A}^\bullet$ is always a connected class. We picture a structure in $(\alpha, v) \in \mathcal{A}^\bullet_X$ as an $\mathcal{A}$–structure α on the set X with one "special" or *root* vertex $v \in X$ circled. For example, Figure 11.3 illustrates a structure from the class $\mathcal{S}^\bullet$ – that is, a "rooted permutation".

Example 11.14 (The Classes of k–Sets)**.** For each natural number $k \in \mathbb{N}$, define a class $\mathcal{E}_k$ as follows. For every finite set X,

$$(\mathcal{E}_k)_X := \begin{cases} \{X\} & \text{if } \#X = k, \\ \varnothing & \text{if } \#X \neq k. \end{cases}$$

Notice that in the case $k = 0$ this agrees with the previous definition of $\mathcal{E}_0$. Conditions (i) and (ii) are easily verified, as is the fact that the exponential generating function of $\mathcal{E}_k$ is $E_k(x) = x^k/k!$. This $\mathcal{E}_k$ is called the class of *k–sets*. The intuitive content of this definition is that, given $k \in \mathbb{N}$ and a finite set X, there is exactly one way for X to be a k–element set if $\#X = k$ (it is what it is), and there is no way for X to be a k–element set if $\#X \neq k$. The sequence $(\mathcal{E}_k : \ k \in \mathbb{N})$ is locally finite. Comparing exponential generating functions we see that the sum $\bigoplus_{k=0}^{\infty} \mathcal{E}_k$ is equivalent to the class $\mathcal{E}$ of Example 11.5.

The class $\mathcal{E}_1$ of 1–sets, or *singletons*, is used so frequently that it deserves special attention. Since $\mathcal{E}_1(x) = x$ we also use the notation $\mathcal{X} := \mathcal{E}_1$ for this class.

Definition 11.15 (k–Sets of Structures)**.** Let $\mathcal{A}$ be a class of structures. We define the class $\mathcal{E}_k[\mathcal{A}]$ as follows. For any finite set X, the finite set $(\mathcal{E}_k[\mathcal{A}])_X$ is the image of the set $\mathcal{A}_X^k$ under the following function:

$$\begin{array}{ccc} \mathcal{A}_X^k & \longrightarrow & (\mathcal{E}_k[\mathcal{A}])_X \\ (\alpha_1, \ldots, \alpha_k) & \mapsto & \{\alpha_1, \ldots, \alpha_k\} \end{array}$$

That is, a structure in $(\mathcal{E}_k[\mathcal{A}])_X$ is an unordered k–element set of pairwise disjoint $\mathcal{A}$–structures which cover X. Each element of $(\mathcal{E}_k[\mathcal{A}])_X$ is the image of $k!$ different elements of $\mathcal{A}_X^k$ under this function, from which it follows that $\#(\mathcal{E}_k[\mathcal{A}])_X = (\#\mathcal{A}_X^k)/k!$. Conditions (i) and (ii) are easily verified, as is the fact that the exponential generating function of $\mathcal{E}_k[\mathcal{A}]$ is $A(x)^k/k!$. Figure 11.4 illustrates the generic form of a structure from the class $\mathcal{E}_k[\mathcal{A}]$.

Theorem 11.16 (The Exponential Formula)**.** *Let $\mathcal{A}$ be a class of structures. If $\mathcal{A}$ is connected then $(\mathcal{E}_k[\mathcal{A}] : \ k \in \mathbb{N})$ is a locally finite sequence of classes. The class*

$$\mathcal{E}[\mathcal{A}] := \bigoplus_{k=0}^{\infty} \mathcal{E}_k[\mathcal{A}]$$

has exponential generating function $\exp(A(x))$.

Proof. Since $\mathcal{A}$ is connected, $\mathcal{A}_\varnothing = \varnothing$. That is, every $\mathcal{A}$–structure α uses at least one vertex. Therefore, if X is a finite set and $\{\alpha_1, \ldots, \alpha_k\} \in (\mathcal{E}_k[\mathcal{A}])_X$, then $k \leq \#X$. The reason for this is that the α_i (for $1 \leq i \leq k$) have pairwise disjoint vertex–sets which cover X, and each of these vertex–sets has at least one element. Therefore, if $k > \#X$ then $(\mathcal{E}_k[\mathcal{A}])_X = \varnothing$. This shows that the sequence $(\mathcal{E}_k[\mathcal{A}] : \ k \in \mathbb{N})$ is locally finite. The formula for the exponential generating function of $\mathcal{E}[\mathcal{A}]$ follows from Definitions 11.7 and 11.15. $\square$

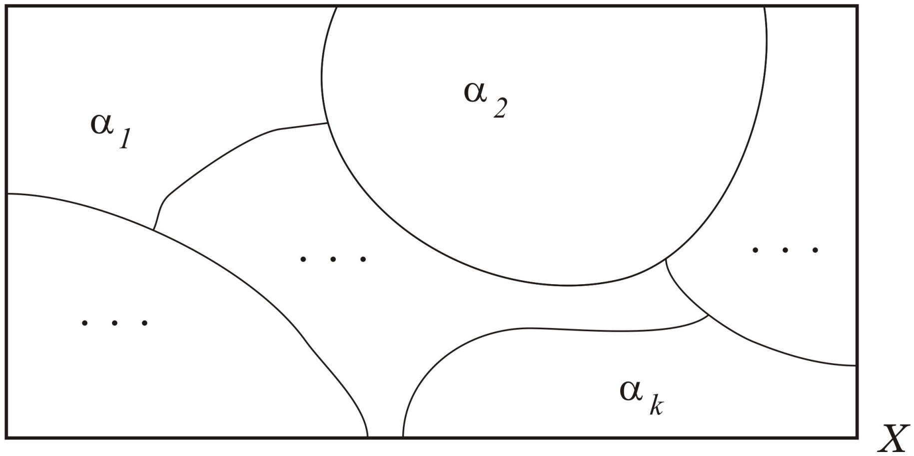

FIGURE 11.4. a structure from the class $\mathcal{E}_k[\mathcal{A}]$.

We have finally developed enough technology to explain the cryptic formula $\mathcal{S} \equiv \mathcal{E}[\mathcal{C}]$ in the paragraph after Example 11.5. We will develop some more theory later in this section and in the next, but we can already do quite a bit with what we have.

Example 11.17. Let $\mathcal{J}$ be the class of (simple, undirected) graphs in which every connected component is a cycle. (These are sometimes called "two–factors".) See Figure 11.5 for an example. Let $\mathcal{H}$ be the class of graphs which are cycles. (Check conditions (i) and (ii) for these classes.) Since each graph in $\mathcal{J}$ can be uniquely decomposed as a disjoint union of an unordered set of cycles, $\mathcal{J} \equiv \mathcal{E}[\mathcal{H}]$. We can obtain the exponential generating function $H(x)$ directly, as follows. Since a (simple) graph cycle must have at least three vertices, we have $\#\mathcal{H}_0 = \#\mathcal{H}_1 = \#\mathcal{H}_2 = 0$. For each $n \geq 3$, a graph cycle may be directed consistently in one of two ways, each of which yields a cyclic permutation. This leads to the equations $\#\mathcal{H}_n = (\#\mathcal{C}_n)/2 = (n-1)!/2$ for all $n \geq 3$. Therefore

$$\begin{aligned} H(x) &= \sum_{n=3}^{\infty} \frac{(n-1)!}{2} \frac{x^n}{n!} = \frac{1}{2} \sum_{n=3}^{\infty} \frac{x^n}{n} \\ &= \frac{1}{2} \log\left(\frac{1}{1-x}\right) - \frac{x}{2} - \frac{x^2}{4}. \end{aligned}$$

By the Exponential Formula we have $J(x) = \exp(H(x))$, and we conclude that

$$J(x) = \frac{\exp(-x/2 - x^2/4)}{\sqrt{1-x}}.$$

Cool! Getting an answer to the enumeration problem $\#\mathcal{J}_n = n![x^n]J(x)$ remains a challenge, however. (It is not really difficult, but the answer is slightly unpleasant.)

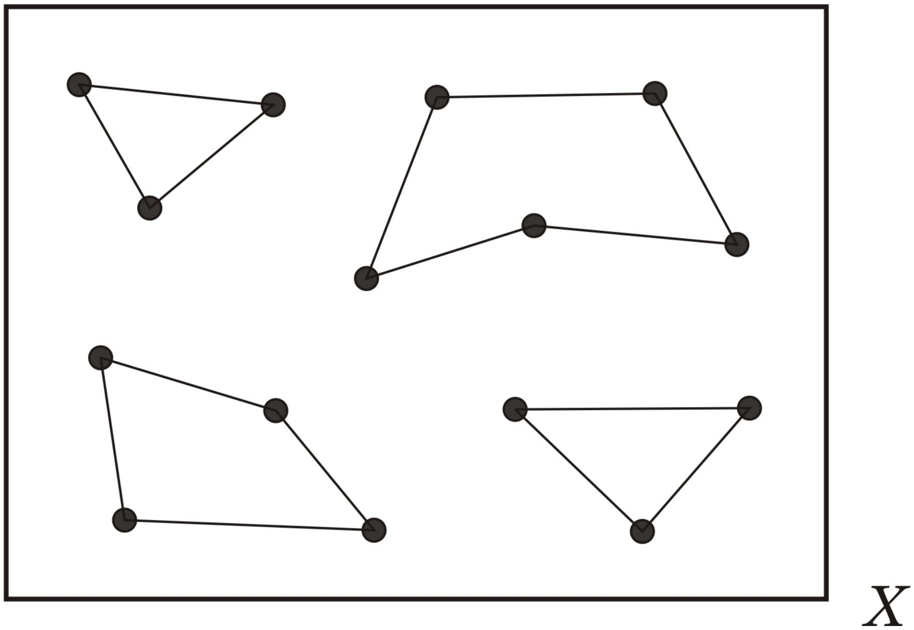

FIGURE 11.5. illustration of Example 11.17.

Example 11.18 (Labelled Trees)**.** Let $\mathcal{T}$ be the class of graphs which are trees. (Check conditions (i) and (ii). Verifying (ii) is not trivial at this point – how would you do it? The ideas of Chapter 12 provide the method.) Using Definition 11.13, $\mathcal{T}^\bullet$ is the class of rooted trees. We may delete the root vertex from a rooted tree and, for each connected component of the remaining graph, root that component at its unique vertex that was adjacent to the deleted vertex. We obtain an unordered set of pairwise disjoint rooted trees, none of which uses the deleted vertex. See Figure 11.6 for an example. Conversely, from a designated vertex v and a set of rooted trees which are pairwise disjoint and do not use v, we may join v to the root of each tree by an edge and root the resulting tree at v. That is, we have bijections for each finite set X:

$$\begin{array}{rcl} \mathcal{T}^\bullet_X & \rightleftharpoons & (\mathcal{X} * \mathcal{E}[\mathcal{T}^\bullet])_X \\ (T, v) & \leftrightarrow & (v, \{(S_1, w_1), \ldots, (S_k, w_k)\}) \end{array}$$

To be more precise, in passing from the LHS to the RHS we let $\{S_1, \ldots, S_k\}$ be the connected components of $T \smallsetminus \{v\}$, and for each $1 \leq i \leq k$ we let w_i be the unique vertex of S_i which is adjacent to v in T. Conversely, in passing from the RHS to the LHS we already have v, and we let

$$V(T) := \{v\} \cup V(S_1) \cup \cdots \cup V(S_k)$$

and

$$E(T) := E(S_1) \cup \cdots \cup E(S_k) \cup \{\{v, w_i\} : 1 \leq i \leq k\}.$$

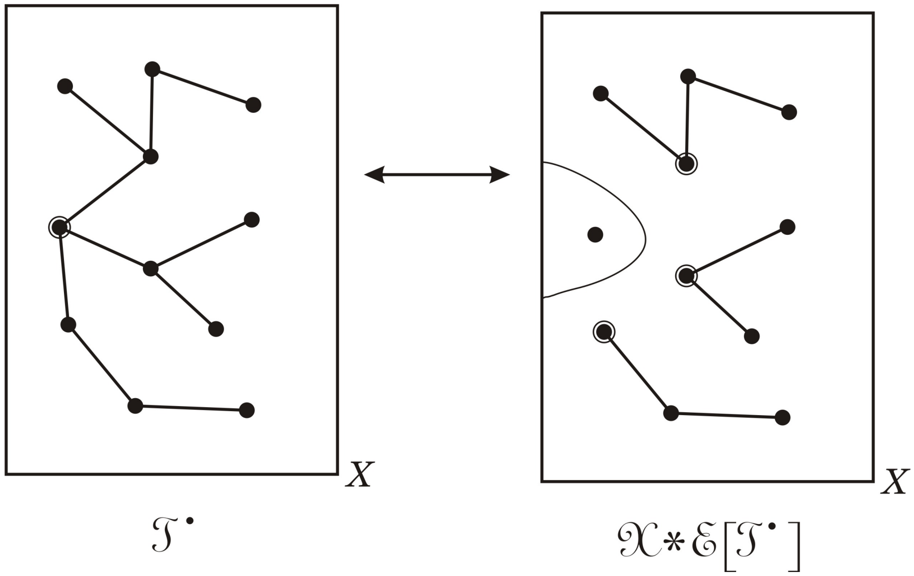

FIGURE 11.6. illustration of Example 11.18.

These bijections establish the following equivalence of classes:

$$\mathcal{T}^\bullet \equiv \mathcal{X} * \mathcal{E}[\mathcal{T}^\bullet].$$

By applying the forgoing theory, we obtain a functional equation for the exponential generating function:

$$T^\bullet(x) = x\exp(T^\bullet(x)).$$

The number of rooted trees on any n–element set is $\#\mathcal{T}^\bullet_n = n![x^n]T^\bullet(x)$. We can obtain this coefficient by applying the Lagrange Implicit Function Theorem, in this case with $R(x) = T^\bullet(x)$, $F(u) = u$, and $G(u) = \exp(u)$. Thus, we calculate that

$$\begin{aligned}\#\mathcal{T}^\bullet_n &= n![x^n]T^\bullet(x) = (n-1)![u^{n-1}]\exp(u)^n \\ &= (n-1)![u^{n-1}]\exp(nu) = n^{n-1}.\end{aligned}$$

Since each tree with n vertices can be rooted at any of its n vertices, we conclude that the number of (unrooted) trees on a set of n vertices is $\#\mathcal{T}_n = n^{n-2}$ for each $n \geq 1$.

Example 11.19 (Stirling numbers of the second kind)**.** Recall from Example 3.5 that for $n \in \mathbb{N}$ and $0 \leq k \leq n$, $S(n,k)$ denotes the number of set partitions of N_n with exactly k parts. We define the class $\mathcal{P}tn$ of set partitions by saying that for every finite set X, $\mathcal{P}tn_X$ is the set of all set partitions of X. (Check that this satisfies conditions (i) and (ii) of Definition 11.1.) To keep track of the number of parts in a set partition, we will use a **bivariate** generating function

$$Ptn(x,y) := \sum_{n=0}^{\infty} \left(\sum_{\pi \in \mathcal{P}tn_n} y^{\#\pi} \right) \frac{x^n}{n!}.$$

(Notice that this is "exponential in x" and "ordinary in y".) A set partition of X is a finite set of pairwise disjoint nonempty finite sets which partition X. That is, letting

$$\mathcal{E}_{\geq 1} := \bigoplus_{k=1}^{\infty} \mathcal{E}_k$$

denote the class of nonempty sets, we have an equivalence

$$\mathcal{P}tn \equiv \mathcal{E}[\mathcal{E}_{\geq 1}]$$

of classes. Since the exponential generating function of $\mathcal{E}$ is $\exp(x)$, by considering how the indeterminate y enters the formula we find that

$$Ptn(x,y) = \exp(y \exp(x) - y).$$

For any $n, k \in \mathbb{N}$, the fact that

$$S(n,k) = n![x^n y^k] \exp(y \exp(x) - y)$$

can be used to give another proof of Excercise 3.11.

Examples 11.17 and 11.19 illustrate the reason why a class $\mathcal{A}$ for which $\mathcal{A}_\varnothing = \varnothing$ is said to be "connected". In Example 11.17, the connected components of the graphs from the class $\mathcal{J}$ were graphs from the connected class $\mathcal{H}$. In Example 11.19, the parts ("connected components") of the set partitions from the class $\mathcal{P}tn$ were from the connected class $\mathcal{E}_{\geq 1}$. The classes $\mathcal{J}$ and $\mathcal{P}tn$ are not connected. The structures from these classes are built up as disjoint unions of pieces, each of which is from a connected class. In short, structures from a connected class often serve as the connected components out of which structures from other classes are constructed.

Definition 11.20 (Composition of Classes)**.** The exponential formula is the prototypical special case of composition of classes. Let $\mathcal{A}$ and $\mathcal{B}$ be classes of structures, with $\mathcal{A}$ connected (*i.e.* $\mathcal{A}_\varnothing = \varnothing$). We define the *composition of* $\mathcal{A}$ *into* $\mathcal{B}$ to be the class $\mathcal{B}[\mathcal{A}]$ defined as follows. Fix a finite set X. A $\mathcal{B}[\mathcal{A}]$–structure on X consists of a pair (ξ, β) such that
- ξ is an $\mathcal{E}[\mathcal{A}]$–structure on X, and
- β is a $\mathcal{B}$–structure on ξ.

Remember, ξ is a finite set (of $\mathcal{A}$–structures), so this makes sense. In set–theoretic notation, the definition is

$$\mathcal{B}[\mathcal{A}]_X := \bigcup_{\xi \in \mathcal{E}[\mathcal{A}]_X} (\{\xi\} \times \mathcal{B}_\xi).$$

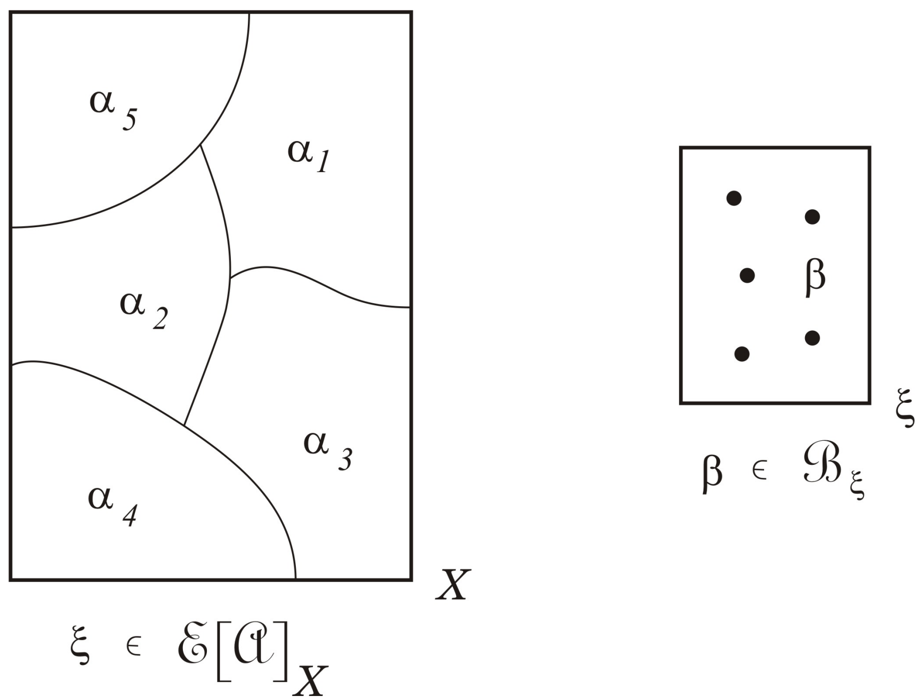

FIGURE 11.7. a structure from the class $\mathcal{B}[\mathcal{A}]$.

Verification of condition (i) for $\mathcal{B}[\mathcal{A}]$ is left as an exercise. If ξ is an element of $(\mathcal{E}_k[\mathcal{A}])_X$ then ξ is a k–element set, so there are $\#\mathcal{B}_k$ different $\mathcal{B}$–structures on ξ. Therefore

$$\#\mathcal{B}[\mathcal{A}]_X = \sum_{k=0}^{\infty}(\#\mathcal{E}_k[\mathcal{A}]_X)(\#\mathcal{B}_k).$$

This verifies condition (ii) for $\mathcal{B}[\mathcal{A}]$.

(The notations $\mathcal{E}_k[\mathcal{A}]$ and $\mathcal{B}[\mathcal{A}]$ are slightly inconsistent when $\mathcal{B} = \mathcal{E}_k$, but the two constructions are "naturally equivalent" in the sense of Section 12.) Figure 11.7 attempts to illustrate the generic form of a structure from the class $\mathcal{B}[\mathcal{A}]$.

Theorem 11.21 (The Compositional Formula)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be classes with $\mathcal{A}$ connected. Then the exponential generating function of $\mathcal{B}[\mathcal{A}]$ is $B(A(x))$.*

Proof. We calculate that

$$\begin{aligned}
\sum_{n=0}^{\infty}(\#\mathcal{B}[\mathcal{A}]_n)\frac{x^n}{n!} &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}(\#\mathcal{E}_k[\mathcal{A}]_n)(\#\mathcal{B}_k)\frac{x^n}{n!} \\
&= \sum_{k=0}^{\infty}(\#\mathcal{B}_k)\sum_{n=0}^{\infty}(\#\mathcal{E}_k[\mathcal{A}]_n)\frac{x^n}{n!} \\
&= \sum_{k=0}^{\infty}(\#\mathcal{B}_k)\frac{A(x)^k}{k!} = B(A(x)).
\end{aligned}$$

□

Example 11.22. Let $\mathcal{Q}$ be the class of graphs which are connected, have exactly one cycle, have maximum degree at most three, and are such that each vertex of degree three is on the unique cycle. We will determine $\#\mathcal{Q}_n$ for all $n \in \mathbb{N}$. Fix a finite set X and consider a graph $\gamma \in \mathcal{Q}_X$. Directing each cut–edge of γ towards the unique cycle C of γ shows how we may regard γ as a collection of nonempty directed paths which have been fit together "inside" the graph cycle C. See Figure 11.8 for an example. Conversely, given an unordered set of pairwise disjoint nonempty directed paths partitioning X and a graph cycle on this set of paths, a graph in $\mathcal{Q}_X$ is constructed by connecting the terminal vertices of these paths according to the graph cycle. A nonempty directed path is equivalent to a structure from the class $\mathcal{X} * \mathcal{X}^*$, so that this analysis gives an equivalence of classes

$$\mathcal{Q} \equiv \mathcal{H}[\mathcal{X} * \mathcal{X}^*]$$

with the class $\mathcal{H}$ as in Example 11.17. Therefore

$$\begin{aligned}
Q(x) &= H\left(\frac{x}{1-x}\right) \\
&= \frac{1}{2}\left[\log\left(\frac{1}{1-x/(1-x)}\right) - \frac{x}{1-x} - \frac{x^2}{2(1-x)^2}\right] \\
&= \frac{1}{2}\left[\log\left(\frac{1}{1-2x}\right) - \log\left(\frac{1}{1-x}\right) - \frac{x}{1-x} - \frac{x^2}{2(1-x)^2}\right] \\
&= \frac{1}{2}\left[\sum_{n=1}^{\infty}\frac{2^n x^n}{n} - \sum_{n=1}^{\infty}\frac{x^n}{n} - \sum_{j=0}^{\infty}x^{j+1} - \frac{1}{2}\sum_{j=0}^{\infty}\binom{j+1}{j}x^{j+2}\right] \\
&= \frac{1}{2}\left[(2-1-1)x + \sum_{n=2}^{\infty}\left(\frac{2^n-1}{n} - 1 - \frac{n-1}{2}\right)x^n\right] \\
&= \sum_{n=2}^{\infty}\left(\frac{(2^{n+1}-n^2-n-2)(n-1)!}{2}\right)\frac{x^n}{n!}.
\end{aligned}$$

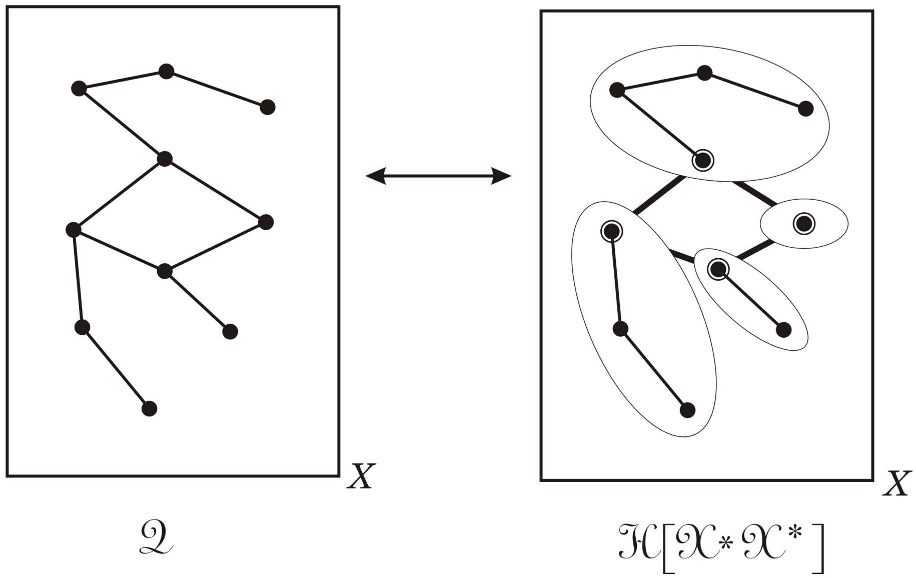

FIGURE 11.8. illustration for Example 11.22.

Hence, the number of graphs in $\mathcal{Q}$ with vertex set $\{1, 2, \ldots, n\}$ is $\#\mathcal{Q}_0 = \#\mathcal{Q}_1 = 0$, and

$$\#\mathcal{Q}_n = (n-1)!\left(2^n - 1 - \binom{n+1}{2}\right),$$

for all $n \geq 2$. In particular, $\#\mathcal{Q}_2 = 0$ as well, as it should be.

Example 11.23 (Endofunctions and Doubly–Rooted Trees)**.** In this example we obtain the formula $\#\mathcal{T}_n = n^{n-2}$ (for all $n \geq 1$) of Example 11.18 by a more direct combinatorial argument.

Let $\mathfrak{F}$ denote the class of endofunctions. It is clear that $\#\mathfrak{F}_n = n^n$ for all $n \in \mathbb{N}$. Let $\mathcal{R} = \mathcal{T}^\bullet$ be the class of rooted trees, and fix a finite set X. Directing each edge of a rooted tree $\gamma \in \mathcal{R}_X$ towards the root of γ, and putting a directed loop at the root of γ, determines the functional directed graph of an endofunction on X. This determines an injective function $\mathcal{R}_X \to \mathfrak{F}_X$ for every finite set X. Thus, $\mathcal{R}$ may be regarded as a subclass of $\mathfrak{F}$.

The functional directed graph of an arbitrary endofunction is the disjoint union of an unordered set of (weakly) connected components. Letting $\mathfrak{C}$ denote the class of endofunctions for which the functional directed graph is connected, we have $\mathfrak{F} \equiv \mathcal{E}[\mathfrak{C}]$.

The functional directed graph of an endofunction in $\mathfrak{C}$ may be uniquely decomposed as the disjoint union of an unordered set of endofunctions in $\mathcal{R}$, with the loops at the roots replaced by a cyclic permuation of the roots. This gives an equivalence

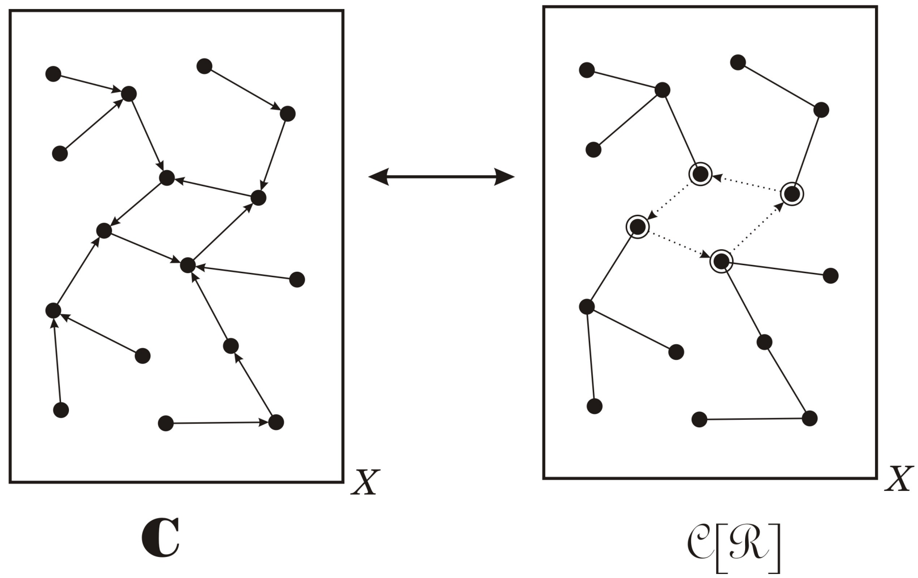

FIGURE 11.9. the equivalence $\mathfrak{C} \equiv \mathcal{C}[\mathcal{R}]$.

$\mathfrak{C} \equiv \mathcal{C}[\mathcal{R}]$, in which $\mathcal{C}$ is the class of cyclic permutations. See Figure 11.9 for an example.

Thus we have $\mathfrak{F} \equiv \mathcal{E}[\mathcal{C}[\mathcal{R}]]$, and so

$$F(x) = \exp\left(\log\left(\frac{1}{1-R(x)}\right)\right) = \frac{1}{1-R(x)}.$$

Since $\#\mathfrak{F}_n = n^n$ for all $n \in \mathbb{N}$, we conclude that

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

Now consider the class $\mathcal{R}^\bullet = (\mathcal{T}^\bullet)^\bullet$ of *doubly–rooted* trees. The rootings are done sequentially, so the root vertices are ordered: one first, the other second. We denote the first root with a circle and the second root with a triangle. Also, the two root vertices may coincide. If γ is a doubly–rooted tree on the set X then γ contains a unique directed path ℓ from the second root to the first root. Every edge of γ not on ℓ may be directed towards ℓ; also put a directed loop at each vertex on ℓ. This decomposes γ uniquely as the disjoint union of a nonempty unordered set of rooted trees in $\mathcal{R}$, and a total order on the set of their roots. See Figure 11.10 for an example.

This gives a natural equivalence

$$\mathcal{R}^\bullet \equiv (\mathcal{X} * \mathcal{X}^*)[\mathcal{R}] \equiv \bigoplus_{k=1}^{\infty} \mathcal{R}^k,$$

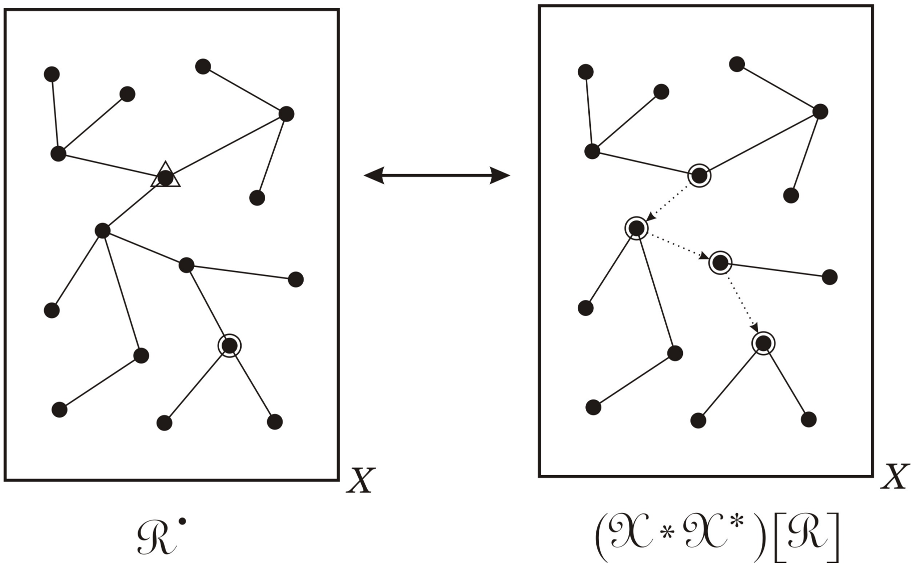

FIGURE 11.10. the equivalence $\mathcal{R}^{\bullet} \equiv (\mathcal{X} * \mathcal{X}^*)[\mathcal{R}]$.

and hence

$$T^{\bullet\bullet}(x) = R^{\bullet}(x) = \frac{R(x)}{1 - R(x)} = F(x) - 1.$$

Comparing this with the above formulas, we get

$$\left(x\frac{d}{dx}\right)^2 T(x) = \sum_{n=1}^{\infty} n^2 (\#\mathcal{T}_n)\frac{x^n}{n!} = \sum_{n=1}^{\infty} n^n \frac{x^n}{n!}.$$

Therefore, for all $n \geq 1$ we have $\#\mathcal{T}_n = n^{n-2}$.

Example 11.24 (Expected Number of Leaves in a Tree)**.** Let $\mathcal{T}$ be the class of trees, let X be a finite set, and let $T \in \mathcal{T}_X$ be a tree with vertex–set X. Denote by $\ell(T)$ the number of leaves (vertices of degree at most one) in T. Assume that $\#X = n$. Among all the n^{n-2} trees in $\mathcal{T}_X$, what is the average value of $\ell(T)$? That is, how many leaves should we expect to see on a tree with n vertices? First of all, since the designation of a root vertex $v \in X$ does not change the number of leaves of T, we may just as well compute the average of $\ell(T)$ as (T, v) varies over all n^{n-1} rooted trees in $\mathcal{T}^{\bullet}_n$. This allows us to use the recursive structure $\mathcal{R} \equiv \mathcal{X} * \mathcal{E}[\mathcal{R}]$ of the class $\mathcal{R} = \mathcal{T}^{\bullet}$ of rooted trees.

Consider the bivariate generating function

$$R(x, y) := \sum_{n=0}^{\infty} \left(\sum_{(T,v) \in \mathcal{R}_n} y^{\ell(T)} \right) \frac{x^n}{n!}.$$

For any $n \geq 1$,

$$L_n := n![x^n]\,\frac{\partial}{\partial y}R(x,y)\bigg|_{y=1} = \sum_{(T,v)\in\mathcal{R}_n} \ell(T)$$

is the total number of leaves among all the rooted trees in $\mathcal{R}_n$. Thus, the answer to our question is L_n/n^{n-1}.

Notice that for $(T,v) \in \mathcal{R}_n$, the leaves of T are:

- those vertices which have no children in (T,v), that is, the terminal vertices of (T,v), and
- if the root vertex has at most one child, then the root is also a leaf.

This special case for the root vertex is kind of annoying, so let's ignore it for the moment. For a rooted tree $(T,v) \in \mathcal{R}_n$, let $\tau(T,v)$ be the number of vertices which have no children in T, and let

$$B(x,y) := \sum_{n=0}^{\infty}\left(\sum_{(T,v)\in\mathcal{R}_n} y^{\tau(T,v)}\right)\frac{x^n}{n!}.$$

To obtain an expression for $B(x,y)$, we use the recursive structure $\mathcal{R} \equiv \mathcal{X} * \mathcal{E}[\mathcal{R}]$. In this equivalence, if (T,v) corresponds to $(v, \{(S_1,w_1),\ldots,(S_k,w_k)\})$, then

$$\tau(T,v) = \begin{cases} 1 & \text{if } k=0, \\ \tau(S_1,w_1)+\cdots+\tau(S_k,w_k) & \text{if } k\geq 1. \end{cases}$$

Keeping track of $\tau(T,v)$ through the equivalence $\mathcal{R} \equiv \mathcal{X} * \mathcal{E}[\mathcal{R}]$ yields the functional equation

$$\begin{aligned} B(x,y) &= x\left[y + \sum_{k=1}^{\infty}\frac{B(x,y)^k}{k!}\right] \\ &= x(y + \exp(B(x,y)) - 1) \end{aligned}$$

for the generating function $B(x,y)$.

Now we can handle the special case of the root vertex, by noticing that if (T,v) corresponds to $(v, \{(S_1,w_1),\ldots,(S_k,w_k)\})$ as above, then

$$\ell(T) = \begin{cases} 1 & \text{if } k=0, \\ 1+\tau(S_1,v_1) & \text{if } k=1, \\ \tau(S_1,v_1)+\cdots+\tau(S_k,v_k) & \text{if } k\geq 2. \end{cases}$$

From this, we see that

$$\begin{aligned} R(x,y) &= x\left[y + yB(x,y) + \sum_{k=2}^{\infty}\frac{B(x,y)}{k!}\right] \\ &= x(y + yB(x,y) + \exp(B(x,y)) - 1 - B(x,y)). \end{aligned}$$

The functional equation for $B(x,y)$ and the equation for $R(x,y)$ in terms of $B(x,y)$ are of the form to which the Lagrange Implicit Function Theorem applies, in this case with $\mathbb{K} = \mathbb{Q}(y)$, $G(u) = \exp(u) + y - 1$, and $F(u) = \exp(u) + (y-1)(1+u)$. Notice that $B(x,y) = xG(B(x,y))$ and $R(x,y) = xF(B(x,y))$, and that $F'(u) = (d/du)F(u) = \exp(u) + y - 1 = G(u)$. Thus we calculate that

$$\begin{aligned}
L_n &= n![x^n]\,\frac{\partial}{\partial y}R(x,y)\bigg|_{y=1} = n!\,\frac{\partial}{\partial y}[x^n]xF(B(x,y))\bigg|_{y=1} \\
&= n!\,\frac{\partial}{\partial y}[x^{n-1}]F(B(x,y))\bigg|_{y=1} = n(n-2)!\,\frac{\partial}{\partial y}[u^{n-2}]F'(u)G(u)^{n-1}\bigg|_{y=1} \\
&= n(n-2)![u^{n-2}]\,\frac{\partial}{\partial y}(\exp(u)+y-1)^n\bigg|_{y=1} \\
&= n(n-2)![u^{n-2}]n\exp((n-1)u) \\
&= \frac{n^2(n-2)!(n-1)^{n-2}}{(n-2)!} = n^2(n-1)^{n-2}.
\end{aligned}$$

Hence, finally, we see that the average number of leaves among all trees on the set $\{1,2,\ldots,n\}$ is

$$\frac{n^2(n-1)^{n-2}}{n^{n-1}} = (n-1)\left(1-\frac{1}{n}\right)^{n-3} \sim \frac{n}{e},$$

asymptotically as $n \to \infty$. Informally speaking, in a large random tree one expects that something close to the fraction $1/e \approx 0.36787944\ldots$ of the vertices are leaves.

---

11. Exercises.

---

1. Let $(\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \ldots)$ be a locally finite sequence of classes. Show that $\mathcal{B} := \bigoplus_{i=1}^{\infty} \mathcal{A}^{(i)}$ satisfies condition (i) of Definition 11.1.

---

2. Let $\mathcal{A}$ and $\mathcal{B}$ be classes of structures. Show that $\mathcal{A} * \mathcal{B}$ satisfies condition (i) of Definition 11.1.

---

3. Show that if $\mathcal{A}$ is a connected class (*i.e.* $\mathcal{A}_{\varnothing} = \varnothing$) then the powers of $\mathcal{A}$ form a locally finite sequence.

---

4. Let $\mathcal{A}$ and $\mathcal{B}$ be classes, with $\mathcal{A}$ connected. Show that $\mathcal{B}[\mathcal{A}]$ satisfies condition (i) of Definition 11.1.

---

5. Recall that a derangement is a permutation with no fixed points. Let $\mathcal{D}$ be the class of derangments.

(a) Derive the exponential generating function

$$D(x) = \frac{\exp(-x)}{1-x}.$$

(b) Use part (a) to give another solution for Example 2.6.

---

6. For a permutation $\sigma \in \mathcal{S}_n$, let $c(\sigma)$ be the number of cycles of σ. What is the average value of $c(\sigma)$ among all $n!$ permutations in $\mathcal{S}_n$?

---

7. Use the formula of Example 11.19 to give another solution for Exercise 3.11.

---

8. Fix a positive integer k. For a permutation σ, let $c(\sigma, k)$ be the number of cycles in σ of length exactly k.

(a) Obtain an algebraic formula for the bivariate exponential generating function

$$S(x,y) = \sum_{n=0}^{\infty} \left(\sum_{\sigma \in \mathcal{S}_n} y^{c(\sigma,k)} \right) \frac{x^n}{n!}.$$

(b) Show that the average number of cycles of length k among all $n!$ permutations in $\mathcal{S}_n$ is

$$\begin{cases} 1/k & \text{if } k \leq n, \\ 0 & \text{if} k > n. \end{cases}$$

---

9. Let $\mathcal{Y}$ be the class of (nonrooted) labelled trees in which each vertex has degree either 1 or 3. Show that for all $k \geq 0$, $\#\mathcal{Y}_{2k+1} = 0$ and

$$\#\mathcal{Y}_{2k+2} = \frac{(2k)!}{2^k} \binom{2k+2}{k}.$$

10(a) Let $\mathcal{A}$ be the class of rooted labelled trees such that each vertex has at most two children. Show that the exponential generating function for $\mathcal{A}$ is

$$A(x) = \frac{1}{x} - 1 - \frac{1}{x}\sqrt{1-2x-x^2}.$$

(b) Let $\mathcal{B}$ be the class of rooted labelled trees which are in $\mathcal{A}$ and are such that the root vertex has at most one child. Show that the exponential generating function for $\mathcal{B}$ is

$$B(x) = 1 - \sqrt{1 - 2x - x^2}.$$

(c) Let $\mathcal{U}$ be the class of endofunctions $f : X \to X$ such that for every $v \in X$, $\#f^{-1}(v) \leq 2$. Show that the exponential generating function for $\mathcal{U}$ is

$$U(x) = \frac{1}{\sqrt{1 - 2x - x^2}}.$$

(d) Use part (c) to obtain a formula for the number of endofunctions $f : N_n \to N_n$ in the class $\mathcal{U}$, for every natural number $n \in \mathbb{N}$.

11(a) Derive the following formula

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

(b) Let $\mathcal{Q}$ be the class of endofunctions in which each cycle has odd length. Explain the following formulas which implicitly determine the exponential generating function $Q(x)$ of $\mathcal{Q}$:

$$\begin{cases} Q(x) &= \sqrt{\frac{1+R(x)}{1-R(x)}}, \\ R(x) &= x \exp(R(x)). \end{cases}$$

---

12. A *triangle–tree* is a connected graph in which every edge is in exactly one cycle, and this cycle has length three (see Figure 11.11). Show that the number of triangle–trees with vertex–set N_n is 0 when n is even, and is

$$\frac{(2k)!(2k+1)^{k-1}}{k!2^k}$$

when $n = 2k + 1$ is odd.
(Hint: Describe the recursive structure of the class of rooted triangle–trees.)

---

13. Let $\mathfrak{F}$ be the class of endofunctions, and for $\phi \in \mathfrak{F}_X$, let $p(\phi)$ denote the number of fixed points of ϕ: that is, the number of $v \in X$ such that $\phi(v) = v$.

(a) Obtain functional equations which determine the exponential generating function

$$F(x, y) = \sum_{n=0}^{\infty} \left(\sum_{\phi \in \mathfrak{F}_n} y^{p(\phi)} \right) \frac{x^n}{n!}.$$

(b) Among all endofunctions $\phi \in \mathfrak{F}_n$, what is the average value of $p(\phi)$?

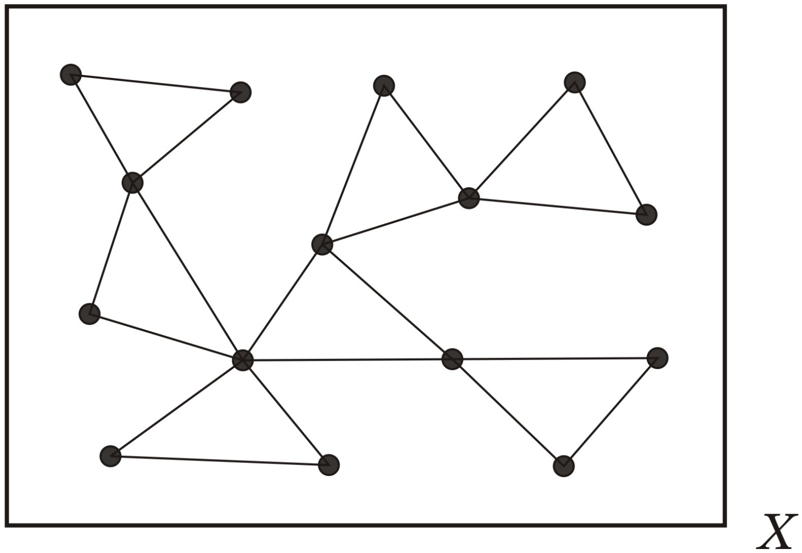

FIGURE 11.11. a triangle–tree.

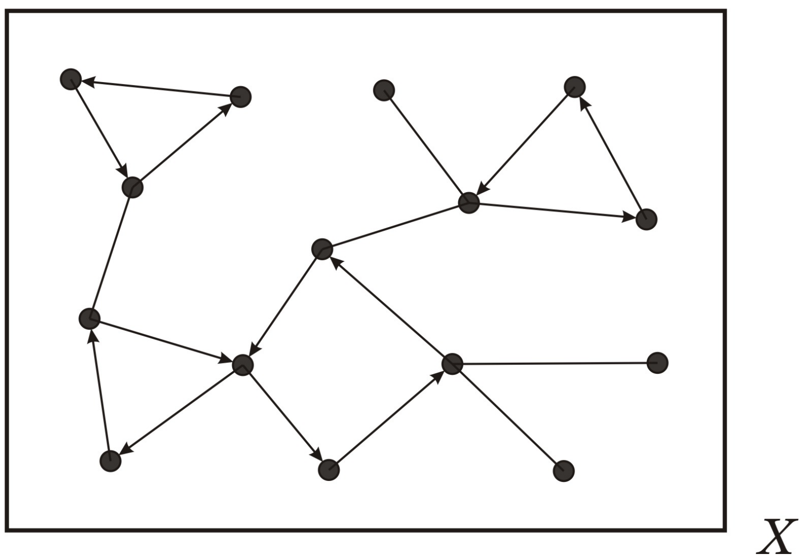

FIGURE 11.12. an oriented cactus.

---

14. A *cactus* is a connected graph such that each edge is in at most one cycle. Equivalently, it is a connected graph in which each block (2–connected component) is either an edge or a cycle. An *oriented cactus* is a cactus in which each cycle has been directed in one of its two strongly connected orientations. (See Figure 11.12.)

(a) Show that for all $n \geq 1$, the number of oriented cacti on the set $\{1, \ldots, n\}$ is

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

(b) Derive a functional equation that implicitly determines the exponential generating function for the class of rooted non–oriented cacti.

(c) For each $n \in \mathbb{N}$, what is the number of (non–rooted, non–oriented) cacti on the set N_n?

15(a) Show that the number of rooted trees on the set N_n which have exactly k terminal vertices is

$$(n-k)! \binom{n}{k} S(n-1, n-k).$$

(b)* Find a combinatorial (bijective) proof of this result.

16.* Revisit Exercise 8 of Chapter 8. Find a combinatorial proof of Exercise 8.8(b).

17.* For $n, k \in \mathbb{N}$, let $q(n,k)$ be the number of connected graphs with k edges and vertex–set $\{1, 2, \ldots, n\}$; also let $Q_n(t) := \sum_{k=0}^{n(n-1)/2} q(n,k) t^k$.

(a) Explain an efficient algorithm for computing $Q_n(t)$.
(Hint: the generating function $\sum_{n=0}^{\infty} Q_n(t) x^n / n!$ is related to an easily determined series.)

(b) If you know Maple or another computer algebra application, write some code and crank out $Q_8(t)$. (Or do it by pencil and paper! ;-)

---

11. Endnotes.

Here are three books that treat exponential generating functions in detail:

- I.P. Goulden and D.M. Jackson, "Combinatorial Enumeration," John Wiley & Sons, New York, 1983.

- R.P. Stanley, "Enumerative Combinatorics, vol. II," Cambridge U.P., Cambridge, 1999.

• H.S. Wilf, "Generatingfunctionology," Academic Press, New York, 1994.

That the number of labelled trees with n vertices is n^{n-2} is attributed to Cayley, in 1889. His solution is a bit sketchy, however. A bijective proof was given in 1918 by Prüfer. See page 51 of

• N.L. Biggs, E.K. Lloyd, and R.J. Wilson, "Graph theory: 1736–1936," Clarendon Press, Oxford, 1976.

There are other kinds of generating functions besides the ordinary generating functions and exponential generating functions we have discussed. For example, in number theory it is frequently useful to encode a sequence $a_1, a_2, \ldots$ of integers by means of its *Lambert series*:

$$\sum_{k=1}^{\infty} a_k \frac{x^k}{1 - x^k}.$$

It is not difficult to verify that for every $n \geq 1$, the coefficient of x^n in this series is

$$\sum_{d|n} a_d,$$

the sum being over all positive divisors of n. For an introduction to several of the various forms of generating functions, see

• P. Doubilet, G.-C. Rota, and R.P. Stanley, *On the foundations of combinatorial theory. VI. The idea of generating function*, in "Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability Vol. II: Probability theory," Univ. California Press, Berkeley, 1972.

• R.P. Stanley, *Generating functions*, in "Studies in Combinatorics" (G.-C. Rota, ed.), Math. Assoc. America, Washington, 1978.

12. Foundations of Exponential Generating Functions.

In Section 11 we introduced the theory of exponential generating functions and applied it to solve several enumeration problems. There were some shortcomings of that discussion, though, which we address in this section. The main point of dissatisfaction is that the concept of equivalence used in Section 11, while adequate for numerical purposes, is much too coarse to discern the more interesting properties of classes of structures. The underlying problem is that Definition 11.1 does not really provide an adequate foundation for the theory even though, as we saw, much of it can be developed satisfactorily at that level of detail. Thus, we begin with a more sophisticated definition of a "natural" class of structures.

Definition 12.1 (Natural Classes of Structures). A *natural class of structures* $\mathcal{A}$ associates to each finite set X another finite set $\mathcal{A}_X$, so that the following conditions hold.
[***V**] There is an algorithm $\mathsf{V}_{\mathcal{A}}$ which takes as input an $\mathcal{A}$–structure $\alpha \in \mathcal{A}_X$ and returns as output the set X on which α is defined.
[***C1**] For finite sets X and Y and a bijection $f : X \rightarrow Y$, there is an *induced bijection* $f_* : \mathcal{A}_X \rightarrow \mathcal{A}_Y$.
[***C2**] For the identity bijection $\mathrm{id}_X : X \rightarrow X$, the induced bijection is the identity bijection

$$(\mathrm{id}_X)_* = \mathrm{id}_{\mathcal{A}_X} : \mathcal{A}_X \rightarrow \mathcal{A}_X.$$

[***C3**] For bijections $f : X \rightarrow Y$ and $g : Y \rightarrow Z$, the induced bijections are compatible with composition of functions:

$$(g \circ f)_* = g_* \circ f_* : \mathcal{A}_X \rightarrow \mathcal{A}_Z.$$

If we want to emphasize the class $\mathcal{A}$ which is being used to produce the induced bijection f_* then we will write $f_{\mathcal{A}}$ instead.

This looks like a lot to require, but we will see that **everything** in Section 11 satisfies these much more demanding conditions.

First of all, I want to explain the intuitive content of the axioms. The axiom [*V] says that if one is handed an arbitrary structure α from the class $\mathcal{A}$ then there is a computation $\mathsf{V}_{\mathcal{A}}$ which can be performed with this input in order to determine the set (of vertices) on which α is defined. The axiom [*C1] says that given a bijection $f : X \rightarrow Y$, if we take the set of $\mathcal{A}$–structures defined on X and "relabel the vertices according to f" then we obtain a bijection f_* from $\mathcal{A}_X$ to $\mathcal{A}_Y$. The key idea is that f_* is the result of changing the names of the vertices only, leaving the additional structure (from the class $\mathcal{A}$) unchanged. Axiom [*C2] is then an obvious requirement – if we do not change the names of the vertices at all then each $\mathcal{A}$–structure on X must be mapped to itself. Axiom [*C3] is likewise a necessary requirement for this interpretation of the induced bijections, which I leave to you to ponder.

Example 12.2 (Graphs)**.** As an example, consider the class $\mathcal{G}$ of graphs. Let's verify the axioms to show that $\mathcal{G}$ is a natural class. Given a graph $\gamma = (V, E)$, we can determine its vertex–set $\mathsf{V}_{\mathcal{G}}(\gamma) := V$, so that axiom [*V] holds. For the remaining axioms, let $f : X \to Y$ be a bijection, and consider $\gamma = (X, E) \in \mathcal{G}_X$. We define

$$f_*(\gamma) := (Y, \{\{f(v), f(w)\} : \ \{v, w\} \in E\}).$$

This defines a function $f_* : \mathcal{G}_X \to \mathcal{G}_Y$. It is now a routine if somewhat tedious matter to check that this definition satisfies axioms [*C1], [*C2], and [*C3].

Example 12.3 (Endofunctions)**.** As another example, consider the class $\mathfrak{F}$ of endofunctions. Let's verify the axioms to show that $\mathfrak{F}$ is a natural class. Given an endofunction ϕ, we can determine its vertex–set $\mathsf{V}_{\mathfrak{F}}(\phi) := \mathrm{dom}(\phi)$, since the domain of a function is implicit in its definition. Thus, axiom [*V] holds. For the remaining axioms, let $f : X \to Y$ be a bijection, and consider $\phi \in \mathfrak{F}_X$. We define

$$f_*(\phi) := f \circ \phi \circ f^{-1} : Y \to Y$$

This defines a function $f_* : \mathfrak{F}_X \to \mathfrak{F}_Y$. It is now a routine if somewhat tedious matter to check that this definition satisfies axioms [*C1], [*C2], and [*C3].

The next issue is to show that these axioms imply the conditions of Definition 11.1.

Proposition 12.4. *Let $\mathcal{A}$ be a natural class of structures. Then $\mathcal{A}$ satisfies the conditions of Definition 11.1.*

Proof. Let X and Y be finite sets, and assume that $\alpha \in \mathcal{A}_X \cap \mathcal{A}_Y$. From axiom [*V] we have $\mathsf{V}_{\mathcal{A}}(\alpha) = X$ and $\mathsf{V}_{\mathcal{A}}(\alpha) = Y$, so that $X = Y$. This implies condition (i). For the second condition, let X and Y be finite sets such that $\#X = \#Y$. By Proposition 1.1 there is a bijection $f : X \to Y$. By axiom [*C1] there is thus a bijection $f_* : \mathcal{A}_X \to \mathcal{A}_Y$. By Proposition 1.1 again, it follows that $\#\mathcal{A}_X = \#\mathcal{A}_Y$, verifying condition (ii). □

The main problem with Section 11 is that the concept of equivalence of classes used there was much too weak to be really interesting. The correct idea of equivalence takes some getting used to but enables us to see some beautiful subtleties which were invisible before. In order to define it we first need to properly describe a few things that we've already done.

Definition 12.5 (Natural Transformations)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be natural classes. A *natural transformation τ from $\mathcal{A}$ to $\mathcal{B}$* is denoted by $\tau : \mathcal{A} \Rightarrow \mathcal{B}$ and is defined as follows. For each finite set X there is a function $\tau_X : \mathcal{A}_X \to \mathcal{B}_X$, and these satisfy

the following axiom:
[*T] Let $f : X \to Y$ be a bijection between finite sets. Then the diagram

$$\begin{array}{ccc} \mathcal{A}_X & \xrightarrow{f_{\mathcal{A}}} & \mathcal{A}_Y \\ \tau_X \downarrow & & \downarrow \tau_Y \\ \mathcal{B}_X & \xrightarrow{f_{\mathcal{B}}} & \mathcal{B}_Y \end{array}$$

commutes. In less visual terms, this means that $\tau_Y \circ f_{\mathcal{A}} = f_{\mathcal{B}} \circ \tau_X$ as functions from $\mathcal{A}_X$ to $\mathcal{B}_Y$.

The intuitive content of this definition is as follows. Each τ_X is a "procedure for changing an $\mathcal{A}$–structure on X into a $\mathcal{B}$–structure on X". Commutativity of the diagram says that τ_X and τ_Y really are the same procedure: only the names of the elements of the underlying vertex–set have been changed according to the bijection $f : X \to Y$. This does not change the effect of the transformation τ. That is, the transformation τ does not depend on the names of the elements of the set underlying the structures on which it acts.

Example 12.6. Let $\mathcal{R} := \mathcal{T}^\bullet$ be the class of rooted trees, and let $\mathfrak{F}$ be the class of endofunctions. In Example 11.22 we considered $\mathcal{R}$ to be a subclass of $\mathfrak{F}$. More precisely, we were considering a natural transformation $\tau : \mathcal{R} \Rightarrow \mathfrak{F}$ defined as follows. For a finite set X, the function $\tau_X : \mathcal{R}_X \to \mathfrak{F}_X$ takes as input a rooted tree $(T, v) \in \mathcal{R}_X$ and returns as output the following endofunction $\tau_X(T, v) := \phi \in \mathfrak{F}_X$. We let $\phi(v) := v$, and for every other $w \in X$ we let $\phi(w)$ be the unique parent of w in the rooted tree (T, v). Care must be taken to verify the axiom [*T], but again it follows directly from the definitions. All the functions τ_X are injective in this example, so we can consider $\mathcal{R}$ as a subsclass of $\mathfrak{F}$ *via* the natural transformation τ.

Example 12.7. We define a natural transformation η from the class $\mathfrak{F}$ of endofunctions to the class $\mathcal{G}$ of graphs, as follows. For any finite set X, the function $\eta_X : \mathfrak{F}_X \to \mathcal{G}_X$ is defined as follows. For any endofunction $\phi \in \mathfrak{F}_X$, the graph $\eta_X(\phi)$ is defined as follows: $\eta_X(\phi) := (X, E)$ in which

$$E := \{\{v, w\} \subseteq X : \ v \neq w \text{ and either } w = \phi(v) \text{ or } v = \phi(w)\}.$$

It is a good exercise to check that axiom [*T] holds in this example. Notice that in general the functions η_X are neither surjective nor injective.

Definition 12.8 (Natural Equivalence)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be natural classes. A *natural equivalence* between $\mathcal{A}$ and $\mathcal{B}$ is a pair of natural transformations $\tau : \mathcal{A} \Rightarrow \mathcal{B}$ and $\rho : \mathcal{B} \Rightarrow \mathcal{A}$ such that for every finite set X, the functions $\tau_X : \mathcal{A}_X \to \mathcal{B}_X$ and $\rho_X : \mathcal{B}_X \to \mathcal{A}_X$ are mutually inverse bijections. If there exists a natural equivalence between the classes $\mathcal{A}$ and $\mathcal{B}$ then we say that these classes are *naturally equivalent*, and denote this relation by $\mathcal{A} \equiv \mathcal{B}$.

This supersedes the notation used in Section 11 but does not contradict it – in every case in which the symbol $\equiv$ was used in Section 11 the classes are in fact **naturally** equivalent.

Next, we revisit the constructions of Section 11 for natural classes.

Definition 12.9 (Sums of Classes)**.** Let $(\mathcal{A}^{(j)} :\ j \geq 1)$ be a locally finite sequence of natural classes. Then the sum $\mathcal{B} := \bigoplus_{j=1}^{\infty} \mathcal{A}^{(j)}$ is a natural class. We check the axioms of Definition 12.1 for $\mathcal{B}$. For an arbitrary $\mathcal{B}$–structure $\beta = (i, \alpha)$ we let $\mathsf{V}_{\mathcal{B}}(\beta) := \mathsf{V}_{\mathcal{A}^{(i)}}(\alpha)$. For a finite set X, if $\beta \in \mathcal{B}_X$ then $\alpha \in \mathcal{A}_X^{(i)}$, so that $\mathsf{V}_{\mathcal{B}}(\beta) = X$ as desired, by axiom [*V] for $\mathcal{A}^{(i)}$. This verifies [*V] for $\mathcal{B}$. Consider a bijection $f : X \to Y$ between finite sets. We define the induced bijection $f_{\mathcal{B}} : \mathcal{B}_X \to \mathcal{B}_Y$ as follows: for each $\beta = (i, \alpha) \in \mathcal{B}_X$, let $f_{\mathcal{B}}(\beta) := (i, f_{\mathcal{A}^{(i)}}(\alpha))$. Since each $f_{\mathcal{A}^{(i)}}$ is a bijection from $\mathcal{A}_X^{(i)}$ to $\mathcal{A}_Y^{(i)}$, it follows that $f_{\mathcal{B}}$ is a bijection from $\mathcal{B}_X$ to $\mathcal{B}_Y$. This establishes [*C1] for $\mathcal{B}$. Axioms [*C2] and [*C3] follow from the construction of the induced bijections $f_{\mathcal{B}}$, as can be checked.

Definition 12.10 (Subclasses and Difference of Classes)**.** For natural classes $\mathcal{A}$ and $\mathcal{B}$, in order to say that $\mathcal{A}$ *is a subclass of* $\mathcal{B}$ we not only need $\mathcal{A}_X \subseteq \mathcal{B}_X$ for every finite set X, but also we require that $\mathsf{V}_{\mathcal{A}}(\alpha) = \mathsf{V}_{\mathcal{B}}(\alpha) = X$ for all $\alpha \in \mathcal{A}_X$, and that $f_{\mathcal{A}}(\alpha) = f_{\mathcal{B}}(\alpha) \in \mathcal{A}_Y$ for all $\alpha \in \mathcal{A}_X$ and $f : X \to Y$. In this case, the difference $\mathcal{B} \smallsetminus \mathcal{A}$ is again a natural class, with operations $\mathsf{V}_{\mathcal{B}\smallsetminus\mathcal{A}}$ and $f_{\mathcal{B}\smallsetminus\mathcal{A}}$ induced from the corresponding operations on $\mathcal{B}$. Thus, $\mathcal{B} \smallsetminus \mathcal{A}$ is also a subclass of $\mathcal{B}$.

Definition 12.11 (Superposition of Classes)**.** For natural classes $\mathcal{A}$ and $\mathcal{B}$, the superposition $\mathcal{A}\&\mathcal{B}$ is also a natural class. For an $(\mathcal{A}\&\mathcal{B})$–structure (α, β) we may define $\mathsf{V}_{\mathcal{A}\&\mathcal{B}}(\alpha, \beta) := \mathsf{V}_{\mathcal{A}}(\alpha)$ to satisfy axiom [*V]. If $(\alpha, \beta) \in (\mathcal{A}\&\mathcal{B})_X$ and $f : X \to Y$ is a bijection, then

$$f_{\mathcal{A}\&\mathcal{B}}(\alpha, \beta) := (f_{\mathcal{A}}(\alpha), f_{\mathcal{B}}(\beta))$$

satisfies axioms [*C1], [*C2], and [*C3].

Definition 12.12 (Products and Powers of Classes)**.** For natural classes $\mathcal{A}$ and $\mathcal{B}$, the product $\mathcal{A} * \mathcal{B}$ is also a natural class. For an $(\mathcal{A} * \mathcal{B})$–structure (α, β) we may define $\mathsf{V}_{\mathcal{A}*\mathcal{B}}(\alpha, \beta) := \mathsf{V}_{\mathcal{A}}(\alpha) \cup \mathsf{V}_{\mathcal{B}}(\beta)$ to satisfy axiom [*V]. If $(\alpha, \beta) \in (\mathcal{A} * \mathcal{B})_X$ and $f : X \to Y$ is a bijection, then

$$f_{\mathcal{A}*\mathcal{B}}(\alpha, \beta) := (f_{\mathcal{A}}(\alpha), f_{\mathcal{B}}(\beta))$$

satisfies axioms [*C1], [*C2], and [*C3]. Iterating this construction shows that for any natural class $\mathcal{A}$, each of the powers $\mathcal{A}^k$ is also a natural class.

Definition 12.13 (Finite Strings and Connected Classes)**.** If $\mathcal{A}$ is a connected natural class then the sequence $(\mathcal{A}^k :\ k \in \mathbb{N})$ of powers of $\mathcal{A}$ is locally finite, and it follows from Definitions 12.9 and 12.12 that $\mathcal{A}^*$ is a natural class.

Definition 12.14 (Rooted Structures)**.** For a natural class $\mathcal{A}$, the class $\mathcal{A}^\bullet$ is also natural. For an $\mathcal{A}^\bullet$–structure (α, v) we may define $\mathsf{V}_{\mathcal{A}^\bullet}(\alpha, v) := \mathsf{V}_{\mathcal{A}}(\alpha)$ to satisfy axiom [*V]. If $(\alpha, v) \in \mathcal{A}^\bullet_X$ and $f : X \to Y$ is a bijection, then

$$f_{\mathcal{A}^\bullet}(\alpha, v) := (f_{\mathcal{A}}(\alpha), f(v))$$

satisfies axioms [*C1], [*C2], and [*C3].

Definition 12.15 (Composition of Classes)**.** If $\mathcal{A}$ and $\mathcal{B}$ are natural classes, with $\mathcal{A}$ connected, then $\mathcal{B}[\mathcal{A}]$ is a natural class. First consider the special case of the class $\mathcal{E}[\mathcal{A}]$. For an $\mathcal{E}[\mathcal{A}]$–structure ξ we may define

$$\mathsf{V}_{\mathcal{E}[\mathcal{A}]}(\xi) := \bigcup_{\alpha \in \xi} \mathsf{V}_{\mathcal{A}}(\alpha)$$

to satisfy axiom [*V]. If $f : X \to Y$ is a bijection, then the induced bijection $f_{\mathcal{E}[\mathcal{A}]} : \mathcal{E}[\mathcal{A}]_X \to \mathcal{E}[\mathcal{A}]_Y$ may be defined as follows. For each $\xi \in \mathcal{E}[\mathcal{A}]_X$ let

$$f_{\mathcal{E}[\mathcal{A}]}(\xi) := \{(f|_{\mathsf{V}_{\mathcal{A}}(\alpha)})_{\mathcal{A}}(\alpha) : \ \alpha \in \xi\}.$$

That is, for each $\mathcal{A}$–structure α in ξ, let $f|_{\mathsf{V}_{\mathcal{A}}(\alpha)}$ be the restriction of f to the subset $\mathsf{V}_{\mathcal{A}}(\alpha) \subseteq X$. This gives a bijection from $\mathsf{V}_{\mathcal{A}}(\alpha)$ to some subset of Y, and we take the image of α under the corresponding induced bijection $(f|_{\mathsf{V}_{\mathcal{A}}(\alpha)})_{\mathcal{A}}$ of $\mathcal{A}$–structures. One can check the axioms [*C1], [*C2], and [*C3] for this construction. Moreover, there is another bijection to be considered, since both $\xi \in \mathcal{E}[\mathcal{A}]_X$ and $f_{\mathcal{E}[\mathcal{A}]}(\xi)$ are finite sets. Namely, there is a unique bijection $\widehat{f} : \xi \to f_{\mathcal{E}[\mathcal{A}]}(\xi)$ with the property that for all $\alpha \in \xi$,

$$\mathsf{V}_{\mathcal{A}}(\widehat{f}(\alpha)) = \{f(v) : \ v \in \mathsf{V}_{\mathcal{A}}(\alpha)\}.$$

This merely says that $\widehat{f}(\alpha)$ is the part of $f_{\mathcal{E}[\mathcal{A}]}(\xi)$ which corresponds to α under the renaming of vertices $f : X \to Y$. Now we can check the axioms in the general case $\mathcal{B}[\mathcal{A}]$ of composition of classes. Given a $\mathcal{B}[\mathcal{A}]$–structure (ξ, β) we may define $\mathsf{V}_{\mathcal{B}[\mathcal{A}]}(\xi, \beta) := \mathsf{V}_{\mathcal{E}[\mathcal{A}]}(\xi)$ to satisfy axiom [*V]. If $(\xi, \beta) \in \mathcal{B}[\mathcal{A}]_X$ and $f : X \to Y$ is a bijection, then we let

$$f_{\mathcal{B}[\mathcal{A}]}(\xi, \beta) := (f_{\mathcal{E}[\mathcal{A}]}(\xi), \widehat{f}_{\mathcal{B}}(\beta)).$$

We have used the bijection $\widehat{f} : \xi \to f_{\mathcal{E}[\mathcal{A}]}(\xi)$ to induce the bijection $\widehat{f}_{\mathcal{B}} : \mathcal{B}_\xi \to \mathcal{B}_{f_{\mathcal{E}[\mathcal{A}]}(\xi)}$ which is then applied to the $\mathcal{B}$–structure $\beta \in \mathcal{B}_\xi$. One can check the axioms [*C1], [*C2], and [*C3] for this construction, but it is rather involved.

In summary, we have seen that all the constructions of Section 11 can be carried through more precisely for natural classes, and yield natural classes in return. This level of detail is not always appropriate when solving particular problems, but it is important to establish the foundations of the theory. It is a worthwhile exercise to review Section 11 and understand why each of the equivalences discussed there is in fact a natural equivalence.

There are classes which are equivalent but are not naturally equivalent – so from now on we speak of *numerical equivalence* when referring to the weaker relation. Notice that the class $\mathcal{S}$ of permutations and the class $\mathcal{L} := \mathcal{X}^*$ of total orders are numerically equivalent, since they both have exponential generating function $S(x) = (1-x)^{-1} = L(x)$. The obvious question arises as to whether or not these classes are naturally equivalent. In fact they are not naturally equivalent, and much more is true: **there are no natural transformations from $\mathcal{S}$ to $\mathcal{L}$.** This is a very strong statement! There seem to be many possibilities for attempting to define a natural transformation from $\mathcal{S}$ to $\mathcal{L}$. How can we be sure that none of them will work? The key idea is that with a natural transformation $\tau : \mathcal{A} \Rightarrow \mathcal{B}$ information can only be lost (or at best preserved – never added) when passing from $\alpha \in \mathcal{A}_X$ to $\tau_X(\alpha) \in \mathcal{B}_X$. As a consequence of this, $\tau_X(\alpha)$ has at least as many "symmetries" as α has. We'll next make these cryptic comments precise, and then apply this strategy to the classes $\mathcal{S}$ and $\mathcal{L}$.

Definition 12.16 (Automorphisms)**.** Let $\mathcal{A}$ be a natural class, X a finite set, and $\alpha \in \mathcal{A}_X$. An *automorphism* of α is a permutation $\sigma \in \mathcal{S}_X$ such that $\sigma_*(\alpha) = \alpha$. Here we are regarding $\sigma : X \rightarrow X$ as a bijection, and considering the induced bijection $\sigma_* : \mathcal{A}_X \rightarrow \mathcal{A}_X$ guaranteed by axiom [*C1]. Let $\mathrm{aut}(\alpha)$ denote the set of all automorphisms of α.

Proposition 12.17. *Let $\mathcal{A}$ be a natural class, X a finite set, and $\alpha \in \mathcal{A}_X$. Then* $\mathrm{aut}(\alpha)$ *is a subgroup of* $\mathcal{S}_X$*.*

Proof. Certainly $\mathrm{aut}(\alpha)$ is a subset of $\mathcal{S}_X$, and hence is finite. By axiom [*C2] we have $(\mathrm{id}_X)_* = \mathrm{id}_{\mathcal{A}_X} : \mathcal{A}_X \rightarrow \mathcal{A}_X$, so that $(\mathrm{id}_X)_*(\alpha) = \alpha$ and therefore $\mathrm{id}_X \in \mathrm{aut}(\alpha)$. If $\pi, \sigma \in \mathrm{aut}(\alpha)$ then – using axiom [*C3] – we have

$$(\pi \circ \sigma)_*(\alpha) = \pi_*(\sigma_*(\alpha)) = \pi_*(\alpha) = \alpha,$$

so that $\pi \circ \sigma \in \mathrm{aut}(\alpha)$ as well. This shows that $\mathrm{aut}(\alpha)$ is a finite set of permutations which contains the identity permutation and is closed under functional composition. From this it follows that $\mathrm{aut}(\alpha)$ is a group, necessarily a subgroup of $\mathcal{S}_X$. □

Definition 12.18. Let $\mathfrak{P}$ denote the class of *permutation groups*: for any finite set X, $\mathfrak{P}_X$ is the set of all subgroups Γ of $\mathcal{S}_X$. Given a permutation group Γ, we can determine its identity element ι and define $\mathsf{V}_{\mathfrak{P}}(\Gamma) := \mathrm{dom}(\iota)$, the domain of the bijection ι. This verifies axiom [*V]. We leave as an exercise the verification of the axioms [*C1] through [*C3], noting only that for a bijection $f : X \rightarrow Y$, the induced bijection is defined by

$$f_*(\Gamma) := \{f \circ \pi \circ f^{-1} : \ \pi \in \Gamma\}.$$

Proposition 12.19. *For any natural class $\mathcal{A}$, the construction* $\mathrm{aut} : \mathcal{A} \Rightarrow \mathfrak{P}$ *is a natural transformation.*

Proof. We check the axiom [*T] for aut. Let $f : X \to Y$ be a bijection between finite sets, and let $\alpha \in \mathcal{A}_X$. We must chack that $f_{\mathfrak{P}}(\mathrm{aut}(\alpha)) = \mathrm{aut}(f_{\mathcal{A}}(\alpha))$. Consider an arbitrary permutation $\sigma \in \mathcal{S}_Y$. Then $\sigma \in \mathrm{aut}(f_{\mathcal{A}}(\alpha))$ if and only if $\sigma_{\mathcal{A}}(f_{\mathcal{A}}(\alpha)) = f_{\mathcal{A}}(\alpha)$. On the other hand, $\sigma \in f_{\mathfrak{P}}(\mathrm{aut}(\alpha))$ if and only if $\sigma = f \circ \pi \circ f^{-1}$ for some $\pi \in \mathrm{aut}(\alpha)$. This condition implies that

$$\sigma_{\mathcal{A}}(f_{\mathcal{A}}(\alpha)) = (f_{\mathcal{A}} \circ \pi_{\mathcal{A}} \circ f_{\mathcal{A}}^{-1} \circ f_{\mathcal{A}})(\alpha) = f_{\mathcal{A}}(\alpha),$$

using axioms [*C2] and [*C3]. This shows that $f_{\mathfrak{P}}(\mathrm{aut}(\alpha))$ is a subset of $\mathrm{aut}(f_{\mathcal{A}}(\alpha))$. Conversely, for any $\sigma \in \mathrm{aut}(f_{\mathcal{A}}(\alpha))$, let $\pi := f^{-1} \circ \sigma \circ f$. Now

$$\pi(\alpha) = f_{\mathcal{A}}^{-1}(\sigma_{\mathcal{A}}(f_{\mathcal{A}}(\alpha))) = f_{\mathcal{A}}^{-1}(f_{\mathcal{A}}(\alpha)) = \alpha,$$

so that $\pi \in \mathrm{aut}(\alpha)$. Also, since

$$f \circ \pi \circ f^{-1} = f \circ f^{-1} \circ \sigma \circ f \circ f^{-1} = \sigma,$$

this shows that $\mathrm{aut}(f_{\mathcal{A}}(\alpha))$ is a subset of $f_{\mathfrak{P}}(\mathrm{aut}(\alpha))$ as well. This completes the proof. ☐

Proposition 12.20. *Let $\mathcal{A}$ and $\mathcal{B}$ be natural classes, and let $\tau : \mathcal{A} \Rightarrow \mathcal{B}$ be a natural transformation. For any finite set X and $\alpha \in \mathcal{A}_X$,* $\mathrm{aut}(\alpha)$ *is a subgroup of* $\mathrm{aut}(\tau(\alpha))$.

Proof. Consider any $\sigma \in \mathrm{aut}(\alpha)$, so that $\sigma_{\mathcal{A}}(\alpha) = \alpha$. By axiom [*T], we have $\sigma_{\mathcal{B}}(\tau(\alpha)) = \tau(\sigma_{\mathcal{A}}(\alpha)) = \tau(\alpha)$, so that $\sigma \in \mathrm{aut}(\tau(\alpha))$ as well. ☐

Example 12.21. Now we can finally show that there is no natural transformation from $\mathcal{S}$ to $\mathcal{L}$. Suppose that there were such a transformation $\tau : \mathcal{S} \Rightarrow \mathcal{L}$. Fix $n \in \mathbb{N}$ and let $\iota \in \mathcal{S}_n$ be the identity permutation on N_n. Then $\ell := \tau(\iota)$ is a total order on N_n and, by Proposition 12.20, $\mathrm{aut}(\iota)$ is a subgroup of $\mathrm{aut}(\ell)$. However, as is easily seen, $\mathrm{aut}(\iota) = \mathcal{S}_n$ and $\mathrm{aut}(\ell) = \{\iota\}$. For all $n \geq 2$ it is impossible for $\mathcal{S}_n$ to be a subgroup of $\{\iota\}$, and therefore the hypothetical τ does not exist.

We close our discussion of exponential generating functions with an example which could have been included at the end of Section 11. However, the combinatorics is a little complicated and the algebra is **very** complicated, so we have postponed it until now. We limit ourselves to merely sketching the main steps, leaving the verification of details and the substantial algebraic manipulations as good exercises.

Definition 12.22 (Nested Set Systems)**.** A *nested set system on the set X* is a pair (X, Δ) in which X is a finite set and Δ is a set of subsets of X such that
• If $A, B \in \Delta$ then either $A \subseteq B$ or $B \subseteq A$ or $A \cap B = \varnothing$.
(An example is illustrated in Figure 12.1.) Let $\mathcal{N}_X$ denote the set of all nested set systems on the set X. Axioms [*V], [*C1], [*C2], [*C3] are easily verified, so that this defines the natural class $\mathcal{N}$ of nested set systems.

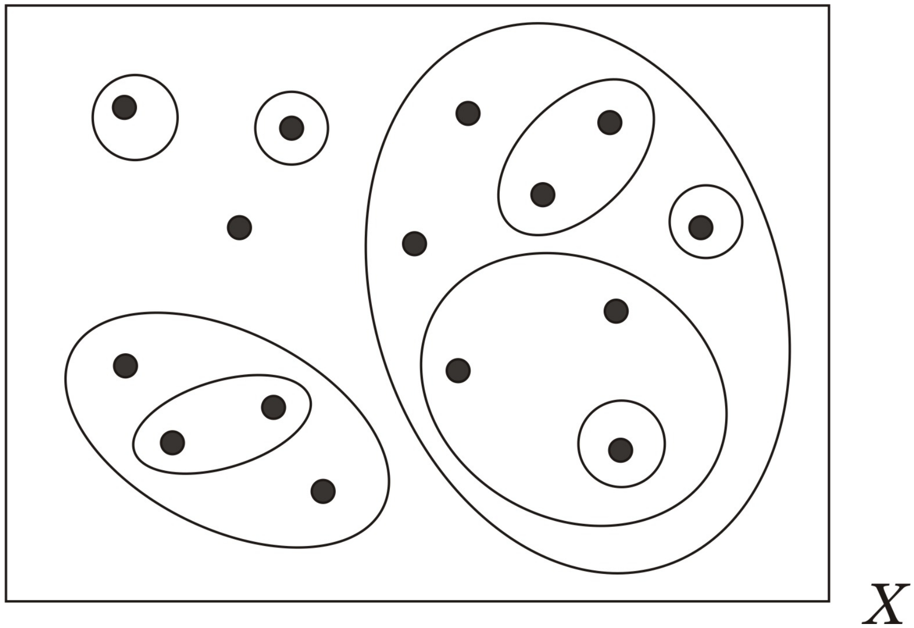

FIGURE 12.1. a nested set system.

To analyze the class $\mathcal{N}$, first notice that $(\varnothing, \varnothing)$ and $(\varnothing, \{\varnothing\})$ are two different $\mathcal{N}$–structures on the empty set. Also, if (X, Δ) is a nested set sytem and $v \in X$ is such that $\{v\} \notin \Delta$, then $(X, \Delta \cup \{\{v\}\})$ is also a nested set system. Let's say that a nested set sytem (X, Δ) is *proper* if $A \in \Delta$ implies $\#A \geq 2$, and denote by $\mathcal{M}$ the subclass of $\mathcal{N}$ consisting of the proper set systems. The *proper part* of $(X, \Delta) \in \mathcal{N}_X$ is (X, Δ°) in which $\Delta^\circ := \{A \in \Delta : \ \#A \geq 2\}$. Let $\mathcal{P}$ be the natural class such that for any finite set X,

$$\mathcal{P}_X := \{(X, A) : \ A \subseteq X\}.$$

For each finite set X, define a function

$$\begin{array}{rcl} \tau_X : \mathcal{N}_X & \to & \mathcal{N}_\varnothing \times \mathcal{P}_X \times \mathcal{M}_X \\ (X, \Delta) & \mapsto & ((\varnothing, \{A \in \Delta : \ A = \varnothing\}), (X, \{v \in X : \ \{v\} \in \Delta\}), (X, \Delta^\circ)). \end{array}$$

We leave it as an exercise to show that this construction defines a natural transformation $\tau : \mathcal{N} \Rightarrow \mathcal{N}_\varnothing * (\mathcal{P} \& \mathcal{M})$, and that moreover this transformation is one part of a natural equivalence $\mathcal{N} \equiv \mathcal{N}_\varnothing * (\mathcal{P} \& \mathcal{M})$.

To further analyze the class $\mathcal{M}$, let's say that a nested set system (X, Δ) is a *cell* if it is proper and $X \in \Delta$. Let $\mathcal{Q}$ be the class of cells. If (X, Δ) is a cell then $(X, \Delta \smallsetminus \{X\})$ is a proper nested set system such that:

- $\#X \geq 2$ (since if $\#X \leq 1$ then (X, Δ) would not have been proper), and
- X is not in $\Delta \smallsetminus \{X\}$.

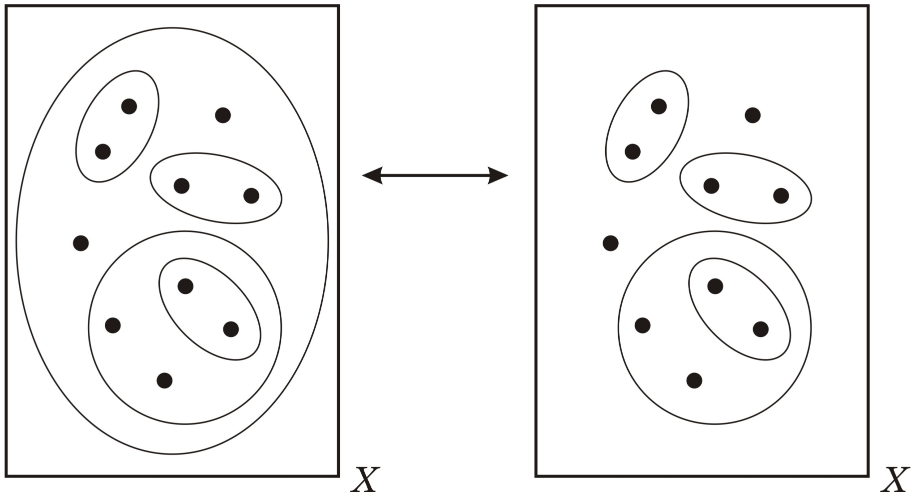

FIGURE 12.2. the equivalence $\mathcal{Q} \equiv \mathcal{M} \smallsetminus (\mathcal{E}_0 \oplus \mathcal{E}_1 \oplus \mathcal{Q})$.

A little thought shows that this gives rise to a natural equivalence

$$\mathcal{Q} \equiv \mathcal{M} \smallsetminus (\mathcal{E}_0 \oplus \mathcal{E}_1 \oplus \mathcal{Q}).$$

See Figure 12.2 for an example of this equivalence.

Finally, an arbitrary proper nested set system can be expressed uniquely as the disjoint union of a collection of cells and singleton vertices, so that

$$\mathcal{M} \equiv \mathcal{E} * \mathcal{E}[\mathcal{Q}].$$

These recursive relations among the classes $\mathcal{N}$, $\mathcal{M}$, and $\mathcal{Q}$ lead to functional equations relating the exponential generating function

$$N(x,y) := \sum_{n=0}^{\infty} \left(\sum_{(N_n,\Delta)\in\mathcal{N}_n} y^{\#\Delta} \right) \frac{x^n}{n!}$$

to the analogous exponential generating functions $M(x,y)$ and $Q(x,y)$ for the subclasses $\mathcal{M}$ and $\mathcal{Q}$, respectively. The remainder of the solution to this enumeration problem (determining $\#\mathcal{N}_n$ for all $n \in \mathbb{N}$) is relegated to a series of exercises.

12. Exercises.

1. In the following, $\mathcal{A}$ denotes a connected class, $\mathcal{C}$ is the class of cycles, $\mathfrak{F}$ is the class of endofunctions, $\mathcal{L} := \mathcal{X}^*$ is the class of total orders, $\mathcal{R} := \mathcal{T}^\bullet$ is the class of rooted trees, and $\mathcal{S}$ is the class of permutations. Prove the following natural equivalences.

(a) $\mathcal{C}^\bullet \equiv \mathcal{X} * \mathcal{L}$.

(b) $\mathcal{L}^\bullet \equiv \mathcal{L} * \mathcal{X} * \mathcal{L}$.

(c) $\mathcal{E}[\mathcal{A}]^\bullet \equiv \mathcal{E}[\mathcal{A}] * \mathcal{A}^\bullet$.

(d) $\mathcal{C}[\mathcal{R}]^\bullet \equiv \mathcal{L}^\bullet[\mathcal{R}]$.

(e) What expression for $\mathfrak{F}^\bullet$ follows from (c) and (d)?

(f) $\mathcal{S}\&\mathcal{L} \equiv \mathcal{L}\&\mathcal{L}$.

2. The classes $\mathcal{T}^{\bullet\bullet\bullet}$ of triply–rooted trees and $\mathfrak{F}^\bullet$ of rooted endofunctions are numerically equivalent, since $\#\mathcal{T}^{\bullet\bullet\bullet} = n^{n+1} = \#\mathcal{F}^\bullet_n$ for all $n \in \mathbb{N}$. Are these classes naturally equivalent?

3.* Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{D}$ be natural classes such that $\mathcal{A} \oplus \mathcal{D} \equiv \mathcal{B} \oplus \mathcal{D}$. Prove that $\mathcal{A} \equiv \mathcal{B}$. (This is trivial for numerical equivalence! For natural equivalence it is quite subtle.)

Exercises 4, 5, and 6 could have been put in Chapter 11, but as they are rather more difficult than the eaarlier ones I've chosen to postpone them until now.

4. A *labelled plane tree* (LPT) is a tree with vertex–set $\{1, 2, \ldots, n\}$ (for some $n \in \mathbb{N}$) which is embedded in the plane as a planar graph. Two embeddings are considered the same if and only if they are "ambient isotopic"; this means that the whole plane may be stretched and squished like a rubber sheet to bring one embedding onto the other. Folding or tearing is not allowed. For example, of the LPTs pictured in Figure 12.3, the center one is equivalent to the one on the left, but not to the one on the right. Let $h(n)$ be the number of (equivalence classes of) LPTs with vertex–set $\{1, 2, .., n\}$, for each $n \in \mathbb{N}$. The first few values are $h(0) = 0$, $h(1) = 1$, $h(2) = 1$, $h(3) = 9$, and $h(4) = 20$. Determine $h(n)$ for all $n \in \mathbb{N}$.

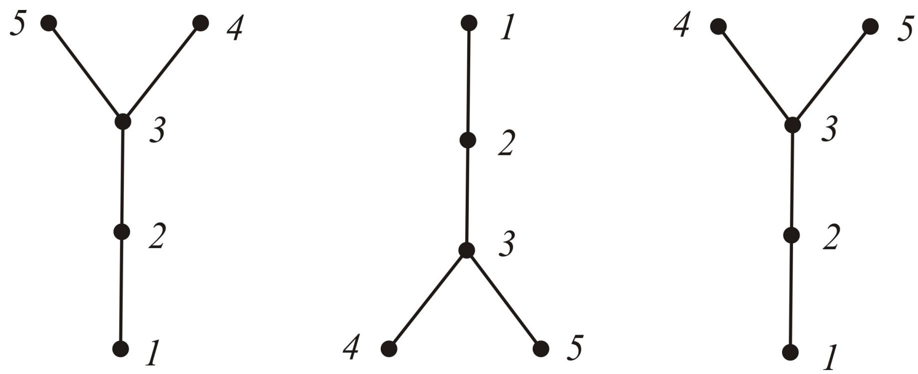

FIGURE 12.3. labelled plane trees.

5. For a finite graph $G = (V, E)$, the *Wiener index* of G is defined to be

$$W(G) := \sum_{v \neq w} \mathrm{dist}_G(v, w),$$

in which the sum is over all unordered pairs of distinct vertices of G, and $\mathrm{dist}_G(v, w)$ denotes the distance between v and w in G. Give a formula for the average value of $W(T)$ as T ranges over the set of all trees with vertex-set $\{1, 2, \ldots, n\}$, for each $n \in \mathbb{N}$.

---

6. For a finite graph $G = (V, E)$, the *productivity* of G is defined to be

$$\pi(G) := \prod_{v \in V} \deg_G(v),$$

in which the product is over all vertices of G, and $\deg_G(v)$ denotes the number of edges incident with v in G. (This is the number of endofunctions $\phi : V \to V$ such that $\{v, \phi(v)\} \in E$ for all $v \in V$.) Give a formula for the average value of $\pi(T)$ as T ranges over the set of all trees with vertex–set $\{1, 2, \ldots, n\}$, for each $n \in \mathbb{N}$.

The remaining exercises outline a solution to the enumeration of nested set–systems.

---

7. Show that the class $\mathcal{N}$ of nested set systems is a natural class.

---

8. Show that the functions $\tau_X : \mathcal{N}_X \to \mathcal{N}_\varnothing \times \mathcal{P}_X \times \mathcal{M}_X$ constructed above define a natural transformation which is one part of a natural equivalence $\mathcal{N} \equiv \mathcal{N}_\varnothing * (\mathcal{P} \& \mathcal{M})$.

---

9. Give detailed justifications of the natural equivalences

$$\mathcal{Q} \equiv \mathcal{M} \smallsetminus (\mathcal{E}_0 \oplus \mathcal{E}_1 \oplus \mathcal{Q})$$

and

$$\mathcal{M} \equiv \mathcal{E} * \mathcal{E}[\mathcal{Q}]$$

discussed above.

---

10. Derive functional equations relating $N(x,y)$, $M(x,y)$ and $Q(x,y)$ from the natural equivalences among $\mathcal{N}$, $\mathcal{M}$, and $\mathcal{Q}$.

---

11. Use the functional equations of Exercise 10 to show that

$$N(x,y) = \frac{(1+y)^2}{y} R\left(\frac{y}{1+y}\exp\left(\frac{x-y+xy}{1+y}\right)\right),$$

in which $R(t) = t\exp(R(t))$. (Hint: the change of variables $z := y/(1+y)$ is very useful.)

---

12. A *normed ring* is a commutative ring R with a *norm function* $|\cdot| : R \to [0,\infty)$ that satisfies the following axioms:

- $|0| = 0$ and $|1| = 1$,
- for all $a, b \in R$, $|ab| = |a||b|$, and
- for all $a, b \in R$, $|a+b| \leq |a| + |b|$.

For example, $\mathbb{C}$ with the usual modulus function is a normed ring.

When R is a normed ring, we may relax the definition of convergence of a sequence of formal power series $f_k(x)$ as $k \to \infty$, as follows. We require that for all $n \in \mathbb{N}$ there exists a constant $A_n \in R$ such that for every real $\varepsilon > 0$ there exists a $K = K(n,\varepsilon)$ such that for all $k \geq K$,

$$|A_n - [x^n]f_k(x)| < \varepsilon.$$

If this holds then the formal power series $F(x) = \sum_{n=0}^{\infty} A_n x^n$ is the limit of the sequence $(f_k(x) :\ k \geq 1)$.

(a) Show that if $(f_k(x) :\ k \geq 1)$ converges in the sense of Section 7, then it converges in the above sense.

(b) Give an example of a sequence of formal power series in $\mathbb{R}[[x]]$ which converges in the above sense but not in the sense of Section 7.

(c) Extend this definition of convergence to sequences of formal Laurent series over a normed ring.

13(a) By setting $y = 1$ in Exercise 11, obtain an expression for $\sum_{n=0}^{\infty}(\#\mathcal{N}_n)x^n/n!$ as the limit of a sequence of formal power series in $\mathbb{R}[[x]]$ which converges in the sense of Exercise 12.

13(b) Deduce that for all $n \in \mathbb{N}$:

$$\#\mathcal{N}_n = 4 \sum_{k=1}^{\infty} \frac{k^{n+k-1}}{k! 2^k e^{k/2}}.$$

---

12. Endnotes.

Our approach to the foundations of exponential generating functions follows the ground–breaking paper of Joyal:

- A. Joyal, *Une théorie combinatoire des séries formelles,* Adv. in Math. **42** (1981), 1–82.

Further developments of this theory are explained in

- F. Bergeron, G. Labelle, and P. Leroux, "Combinatorial Species and Tree–Like Structures," Cambridge U.P., Cambridge, 1998.

Nested set–systems are naturally equivalent to certain other combinatorial structures. Their enumeration was first given in

- J. P. Hayes, *Enumeration of fanout–free Boolean functions*, J. ACM, **23** (1976), 700-709.

See also

- K. L. Kodandapani and S. C. Seth, *On combinational networks with restricted fan–out*, IEEE Trans. Computers, **27** (1978), 309-318.

- L. R. Foulds and R. W. Robinson, *Determining the asymptotic number of phylogenetic trees*, pp. 110-126 of "Combinatorial Mathematics VII (Newcastle, August 1979)", ed. R. W. Robinson, G. W. Southern and W. D. Wallis. Lect. Notes Math., **829**. Springer, 1980.

These last three references were found with the help of Neil Sloane's *On–Line Encyclopedia of Integer Sequences*:

`http://www.research.att.com/~njas/sequences/index.html`

13. A Combinatorial Proof of the Lagrange Implicit Function Theorem.

In this section we give a proof of LIFT which uses slightly weaker hypotheses than that in Section 8. It also gives some combinatorial insight into the "meaning" of the formula which is not apparent from the previous algebraic proof.

Theorem 13.1 (LIFT)**.** *Let $\mathbb{K}$ be a commutative ring which contains the rational numbers $\mathbb{Q}$. Let $F(u)$ and $G(u)$ be formal power series in $\mathbb{K}[[u]]$ with $[u^0]G(u) \neq 0$.*
(a) *There is a unique (nonzero) formal power series $R(x)$ in $\mathbb{K}[[x]]$ such that*

$$R(x) = x\, G(R(x)).$$

(b) *The constant term of $R(x)$ is 0, and for all $n \geq 1$,*

$$[x^n]F(R(x)) = \frac{1}{n}[u^{n-1}]F'(u)G^n(u).$$

(Notice that we do not require $[u^0]G(u)$ to be invertible in $\mathbb{K}$.)

Proof. We prove this combinatorially by interpreting these formal power series as exponential generating functions for "generic" classes of structures. More precisely, we will interpret both sides combinatorially and define bijections which imply that

$$n![x^n]F(R(x)) = (n-1)![u^{n-1}]F'(u)G^n(u)$$

for all $n \geq 1$. To do this, let $f_0, f_1, f_2, \ldots$ and $g_0, g_1, g_2, \ldots$ be infinitely many indeterminates which commute pairwise and are algebraically independent over $\mathbb{K}$. Form the power series

$$F(u) = \sum_{n=0}^{\infty} f_n \frac{u^n}{n!} \quad \text{and} \quad G(u) = \sum_{n=0}^{\infty} g_n \frac{u^n}{n!}.$$

Thinking of these as exponential generating functions for classes $\mathcal{F}$ and $\mathcal{G}$, we see that f_n represents all $\mathcal{F}$–type structures on an n–element set; analogously for g_n as well. The fact that the f_i–s and g_j–s are indeterminates means that these series do not satisfy any special identities – any algebraic formula which can be proved valid for them must also be true if the indeterminates are specialized to have particular values; for example, if $f_n = \#\mathcal{A}_n$ for a particular class $\mathcal{A}$ of structures. It is in this sense that $F(u)$ and $G(u)$ are *generic exponential generating functions.*

In what sense, though, can we talk about a *generic class of structures*, which is what we want $\mathcal{F}$ (and $\mathcal{G}$) to be? In fact, a generic class $\mathcal{F}$ is just the class $\mathcal{E}$ of finite sets – the only difference is in what goes into the exponential generating function. We just agree to mark an n–element set with the indeterminate f_n, for each $n \in \mathbb{N}$.

If we specialize $f_n = \#\mathcal{A}_n$ for some class $\mathcal{A}$, then indeed there are f_n choices for putting an $\mathcal{A}$–structure on an n–element set, and the generic $F(u)$ specializes to the particular $A(u)$ in this case.

Now consider the natural equivalence $\mathcal{R} \equiv \mathcal{X} * \mathcal{G}[\mathcal{R}]$; this (implicitly) defines a class $\mathcal{R}$ for which the exponential generating function satisfies $R(x) = xG(R(x))$. An $\mathcal{R}$–structure on the set X consists of a rooted tree T with vertex–set X and, for each vertex $w \in X$, a $\mathcal{G}$–structure on the set of children of w in T. But the generic class $\mathcal{G}$ is just $\mathcal{E}$, so $\mathcal{R}$ is the class of all rooted labelled trees (RLTs). The only novelty is in how the indeterminates g_j enter the formula for the exponential generating function of $\mathcal{R}$.

Let $c(T, v, w)$ denote the number of children of the vertex w in the RLT (T, v). Recalling that a $\mathcal{G}$–structure on a j–element set is marked by g_j, we see that the equation $R(x) = xG(R(x))$ has the unique solution

$$R(x) = \sum_{n=0}^{\infty} \left(\sum_{(T,v)\in\mathcal{R}_n} \mathbf{g}^{(T,v)} \right) \frac{x^n}{n!},$$

in which

$$\mathbf{g}^{(T,v)} := \prod_{w\in \mathsf{V}_{\mathcal{R}}(T,v)} g_{c(T,v,w)}.$$

Notice that $[x^1]R(x) = g_0 \neq 0$ so that $R(x)$ is nonzero, and from $R(x) = xG(R(x))$ it follows that $[x^0]R(x) = 0$. This proves statement (a) in LIFT.

To prove statement (b), we analyze that formula combinatorially. The power series $F(R(x))$ may be interpreted using composition of classes. This must be the exponential generating function for $\mathcal{F}[\mathcal{R}]$, the class of forests of rooted labelled trees. In this generating function, f_r indicates a forest with exactly r connected components. Thus we have arrived at a combinatorial interpretation of the LHS of LIFT: $n![x^n]F(R(x))$ is the sum over all forests of RLTs with vertex–set $\{1, 2, .., n\}$, in which each forest φ contributes the monomial

$$M(\varphi) := f_{\#\varphi} \prod_{(T,v)\in\varphi} \mathbf{g}^{(T,v)}.$$

(Here we think of φ as a set of RLTs.) That is,

$$F(R(x)) = \sum_{n=0}^{\infty} \left(\sum_{\varphi\in\mathcal{F}[\mathcal{R}]_n} M(\varphi) \right) \frac{x^n}{n!}.$$

The next step is to find a similar interpretation for the RHS of LIFT. But this is easy! $F'(u)G^n(u)$ is u^{-1} times the exponential generating function of the class $\mathcal{F}^\bullet * \mathcal{G}^n$. A structure from this class on a finite set X is (naturally equivalent to) an ordered $(n+2)$–tuple $\sigma = (A, v, B_1, \ldots, B_n)$ in which $A, B_1, \ldots, B_n$ are pairwise disjoint subsets of X which have X as their union, and v is a designated element of A. The contribution of σ to the exponential generating function is the monomial

$$m(\sigma) := f_{\#A} \prod_{i=1}^{n} g_{\#B_i}.$$

Therefore,

$$\begin{aligned}(n-1)![u^{n-1}]F'(u)G^n(u) &= n^{-1}n![u^n]uF'(u)G^n(u) \\ &= \frac{1}{n} \sum_{\sigma \in (\mathcal{F}^\bullet * \mathcal{G}^n)_n} m(\sigma)\end{aligned}$$

is $1/n$ times the sum of $m(\sigma)$ over all σ from the class $\mathcal{F}^\bullet * \mathcal{G}^n$ on the set $\{1, 2, \ldots, n\}$. That factor of $1/n$ is kind of annoying, but we can move it to the LHS by considering the class $\mathcal{F}[\mathcal{R}]^\bullet$ instead.

In summary, our strategy for proving LIFT is to show that for $n \geq 1$,

$$n!n[x^n]F(R(x)) = n![u^n]uF'(u)G^n(u)$$

by constructing a bijection between two sets. On the LHS is the set $\mathcal{F}[\mathcal{R}]^\bullet_n$ of pairs (φ, w) in which φ is a forest of RLTs with vertex–set $\{1, 2, .., n\}$ and w is a designated vertex of φ. On the RHS is the set $(\mathcal{F}^\bullet * \mathcal{G}^n)_n$ with elements σ as described above in the case $X = \{1, 2, \ldots, n\}$. Moreover, in this bijection, if (φ, w) corresponds to σ, then we require that $M(\varphi) = m(\sigma)$. In tabular form:

$$\begin{aligned}\mathcal{F}[\mathcal{R}]^\bullet_n &\rightleftharpoons (\mathcal{F}^\bullet * \mathcal{G}^n)_n \\ (\varphi, w) &\leftrightarrow \sigma = (A, v, B_1, \ldots, B_n) \\ M(\varphi) &= m(\sigma)\end{aligned}$$

Such a bijection will suffice to prove LIFT.

Before defining the bijection we're looking for, I want to remark that its construction is **not natural** in the sense of Section 12 – it will use the numerical order of the labels $\{1, 2, \ldots, n\}$ of the vertices in the underlying set. But this doesn't bother me too much. Notice that we are not trying to show that the classes $\mathcal{F}[\mathcal{R}]^\bullet$ and $\mathcal{F}^\bullet * \mathcal{G}^n$ are naturally equivalent – they are not even numerically equivalent! We're just relating one coefficient of $F(R(x))$ to one coefficient of $F'(u)G^n(u)$.

First think about defining a function from $\mathcal{F}[\mathcal{R}]^\bullet_n$ to $(\mathcal{F}^\bullet * \mathcal{G}^n)_n$. Let φ be a forest of RLTs with vertex–set $\{1, 2, \ldots, n\}$, and let w be a vertex of φ. The corresponding $\sigma = (A, v, B_1, \ldots, B_n)$ will be defined below (eventually). Notice that since we require that $m(\sigma) = M(\varphi)$, this gives us a big hint as to how to construct σ from (φ, w). Since we must have $\#A = \#\varphi$, let A be the set of root vertices of the components of φ, and let v be the root vertex of the component of φ which contains w. To define the sets B_i is a little bit tricky, and this is where the clever combinatorics in this proof comes into play.

We'll start by defining a function BFS from $\mathcal{F}[\mathcal{R}]_n$ to $(\mathcal{F} * \mathcal{G}^n)_n$ using the following algorithm, which makes use of a list L and a (first–in first–out) queue Q.

```
FUNCTION:  BFS from F[R]_n to (F * G^n)_n;
INPUT:  φ;
initially L and Q are empty, and i := 1;
let A be the set of root vertices of the components of φ;
put the vertices in A on the list L in ascending numerical order;
repeat while L is not empty:
  copy the first vertex of L onto Q;
  delete the first vertex from L;
  repeat while Q is not empty:
    let C_i be the set of children of the first vertex of Q;
    put the vertices of C_i onto Q in ascending numerical order;
    increment i ← i + 1;
    delete the first vertex from Q;
  end repeat;
end repeat;
OUTPUT:  (A, C_1, ..., C_n).
```

What this is doing is breadth–first search on each component of φ, taking the components in ascending order of root labels, and recording the set of children of each vertex. (I suggest that you run the algorithm by hand on an arbitrary example with 15 vertices and three components.) Observe that since each vertex in $\{1, 2, \ldots, n\}$ is deleted from the queue exactly once, a sequence of n sets $(C_1, \ldots, C_n)$ is produced. Also notice that if $(B_1, \ldots, B_n)$ is a listing of $(C_1, \ldots, C_n)$ in any order and $\sigma = (A, v, B_1, \ldots, B_n)$, then $m(\sigma) = M(\varphi)$ as required.

From the output $(A, C_1, \ldots, C_n)$ we can recover the forest φ by the following inverse algorithm, again using a list L and a queue Q.

```
FUNCTION:  FOREST from a subset of (F * G^n)_n to F[R]_n;
INPUT:  (A, C_1, ..., C_n);
initially φ has vertices N_n and no edges;
initially L and Q are empty, and i := 1;
put the vertices in A on the list L in ascending numerical order;
repeat while L is not empty:
  copy the first vertex of L onto Q;
  delete the first vertex from L;
  mark the first vertex of Q as a root vertex of φ;
  repeat while Q is not empty:
    join the first vertex of Q to each vertex of C_i by an edge of φ;
```

```
    put the vertices of C_i onto Q in ascending numerical order;
    increment i ← i + 1;
    delete the first vertex from Q;
  end repeat;
end repeat;
OUTPUT:  φ.
```

The algorithm *FOREST* is not well–defined on all of $(\mathcal{F} * \mathcal{G}^n)_n$. It is amusing and instructive to find an example input with $n = 10$, say, which causes the algorithm to malfunction. Nonetheless, it is not difficult to verify that for any $\varphi \in \mathcal{F}[\mathcal{R}]_n$, we have

$$FOREST(BFS(\varphi)) = \varphi.$$

In order to construct the bijection for proving LIFT, it will help to understand exactly which $(n+1)$–tuples in $(\mathcal{F} * \mathcal{G}^n)_n$ are in the image of the function BFS.

For this, let's start with the construction of the sets $(C_1, \ldots, C_k)$ for just the first tree T_1 of φ (the one with the smallest root label). Here, k is the number of vertices of T_1. This corresponds to the first pass through the outer `repeat` loop in the algorithm defining BFS. For each $1 \leq i \leq k$, let $c_i := \#C_i - 1$. Notice that after the i–th vertex has been deleted from the queue Q, the number of vertices remaining on Q is $1 + c_1 + \cdots + c_i$; this is true initially (the case $i = 0$) and the elements of C_i are appended to Q just before the i–th vertex is deleted from Q. So the sequence $(c_1, \ldots, c_k)$ satisfies the following conditions:

- each entry is an integer $c_i \geq -1$;
- if $1 \leq i < k$ then the partial sum $c_1 + \cdots + c_i$ is nonnegative; and
- $c_1 + \cdots + c_k = -1$.

Such a sequence is called a *simple Raney sequence.* We shall also say that a sequence $(C_1, \ldots, C_k)$ of sets is a *simple Raney sequence* when $(\#C_1 - 1, \ldots, \#C_k - 1)$ is. As a result, we see that if $(A, C_1, \ldots, C_n) = BFS(\varphi)$ for some $\varphi \in \mathcal{F}[\mathcal{R}]_n$, then $(C_1, \ldots, C_n)$ is the concatenation of $\#A$ simple Raney sequences.

Lemma 13.2. *Let $(b_1, \ldots, b_k)$ be a sequence of integers such that each $b_i \geq -1$ and $b_1 + \cdots + b_k = -1$. Then there is exactly one cyclic shift of the b_i which is a simple Raney sequence.*

Proof. For $0 \leq i \leq k$, let $s_i := b_1 + \cdots + b_i$, so that $s_0 := 0$ and $s_k := -1$. Let s^* denote the minimum of $\{s_0, \ldots, s_k\}$, and let j be the first index at which $s_j = s^*$. The cyclic shift $c_i := b_{i+j}$ (subscripts modulo k for $1 \leq i \leq k$) of the b_i is easily seen to be a simple Raney sequence. (See Figure 13.1.) Any other cyclic shift of the b_i is seen to have some partial sum which is at most -2, and hence fails to be a simple Raney sequence. □

Definition 13.3. A finite sequence $(c_1, \ldots, c_n)$ of integers is an *r–fold Raney sequence* provided that the following conditions hold:

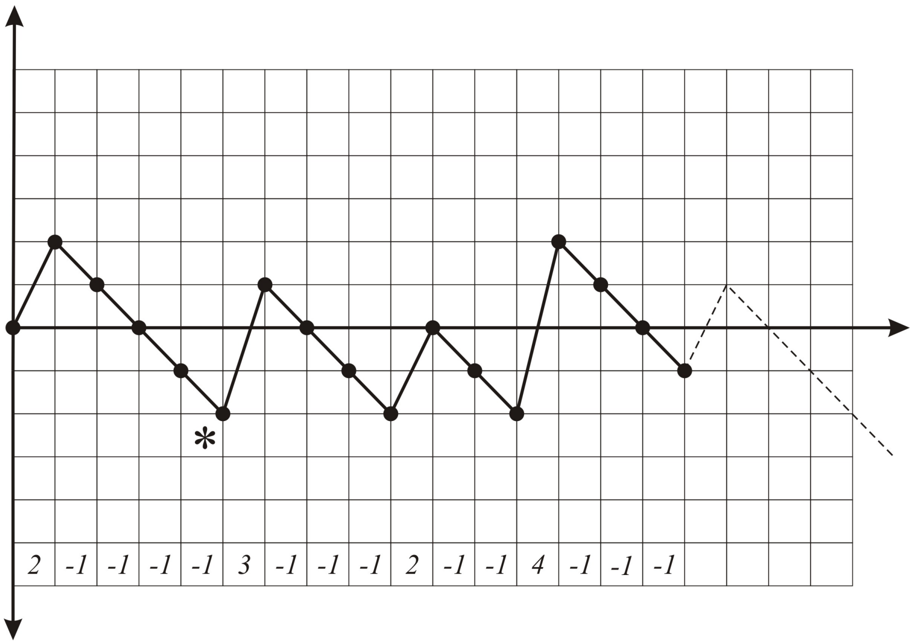

FIGURE 13.1. the unique cyclic shift of Lemma 13.2.

• each entry is an integer $c_i \geq -1$;
• if $1 \leq i < n$ then $c_1 + \cdots + c_i > -r$; and
• $c_1 + \cdots + c_k = -r$.
We shall also say that a sequence $(C_1, \ldots, C_n)$ of sets is an *r–fold Raney sequence* when $(\#C_1 - 1, \ldots, \#C_k - 1)$ is.

It is easy to see that the concatenation of r simple Raney sequences is an r–fold Raney sequence. In part, the following lemma asserts that the converse is also true. (The proof is left as an exercise.)

Lemma 13.4. *Let $\theta = (c_1, \ldots, c_n)$ be an r–fold Raney sequence, for some $r \geq 1$.*
(a) *Then θ has a unique expression as the concatenation $\theta = \rho_1 \cdots \rho_r$ of r simple Raney sequences, called the* blocks *of θ.*
(b) *A cyclic shift of θ is an r–fold Raney sequence if and only if it is obtained by a cyclic shift of the blocks of θ; that is, $\rho_{j+1} \cdots \rho_r \rho_1 \cdots \rho_j$ for some $0 \leq j \leq r-1$.*
(c) *Let $\beta = (b_1, \ldots, b_n)$ be a sequence of integers $b_i \geq -1$ such that $b_1 + \cdots + b_n = -r$. Then there are exactly r cyclic shifts of β which are r–fold Raney sequences.*

Lemma 13.4 gives us enough leverage to finish the proof of LIFT. The idea is to define $(B_1, \ldots, B_n)$ to be a cyclic shift of $(C_1, \ldots, C_n)$ which, along with v, encodes the choice of the root vertex w of (φ, w).

To see this, consider any rooted forest $(\varphi, w) \in \mathcal{F}[\mathcal{R}]^{\bullet}_n$. Let v be the root of the tree containing w, and construct $(A, C_1, \ldots, C_n)$ by applying the algorithm BFS to

φ. We have seen that $(C_1, \ldots, C_n)$ is an r–fold Raney sequence. Let $\rho_1\rho_2\cdots\rho_r$ be the factorization of this r–fold Raney sequence into its blocks, corresponding to the components of φ. It remains to decide which of the sets C_i is to become the first set B_1 – then the cyclic shift taking the C_i–s to the B_i–s is determined, and we have constructed $\sigma = (A, v, B_1, \ldots, B_n)$.

Now, if v is the s–th vertex of A in ascending numerical order $(1 \leq s \leq r)$ then consider the block ρ_s. If w is the p–th vertex (in ascending numerical order) of the s–th component T_s (in ascending order of root labels) of φ, then B_1 is chosen to be the p–th set in the s–th block of $(C_1, \ldots, C_n)$.

Conversely, given $\sigma = (A, v, B_1, \ldots, B_n)$ we must decide which of the sets B_i to choose for the first set C_1. Given $(B_1, \ldots, B_n)$, thought of as a cyclic list of sets, let $\rho_1\rho_2\cdots\rho_r$ be the cyclic list of simple Raney sequences forming its block decomposition. If v is the s–th vertex of A then choose the indexing of these blocks so that B_1 is in the block ρ_s. Now let C_1 be the first set in ρ_1. This determines $(C_1, \ldots, C_n)$ and by applying the algorithm $FOREST$ we construct φ. Now if B_1 is the p–th set in the block ρ_s, let w be the p–th vertex of the s–th component of φ.

The constructions given above provide a pair of mutually inverse bijections between the sets $\mathcal{F}[\mathcal{R}]^\bullet_n$ and $(\mathcal{F}^\bullet * \mathcal{G}^n)_n$, such that if (φ, w) corresponds to σ then $M(\varphi) = m(\sigma)$. There are a number of details to check – that the hypotheses of the lemmas are satisfied when required, that the algorithms always terminate properly on the given input, that they produce the expected output, etc. – but these are left to the diligent reader. This completes the combinatorial proof of LIFT. $\square$

---

13. Exercises.

---

1. Prove Lemma 13.4.

---

13. Endnotes.

This proof of LIFT is a variation of that given by Raney:

- G.N. Raney, *Functional composition patterns and power series reversion*, Trans. Amer. Math. Soc. **94** (1960), 441–451.

There are multivariate versions of LIFT as well. For a state–of–the–art version and some pointers to prior literature, see

- I.P. Goulden and D.M. Kulkarni, *Multivariable Lagrange inversion, Gessel–Viennot cancellation, and the matrix tree theorem*, J. Combin. Theory Ser. A **80** (1997), 295–308.