LTCC Enumerative Combinatorics

Notes 4 Alex Fink Fall 2015

---

4 Species and exponential generating functions

We now consider labelled structures: when these are of "size" n, there is an underlying set of "labels" of size n. We usually take this set of labels to be $[n]$. As such, a particular kind of labelled structure $\mathscr{A}$ is specified by giving the set of structures labelled by $[n]$, for each $n \in \mathbb{N}$.

But it is better for the notion to capture more structure, namely to capture what it means for there to be a *set* of labels, and the possibility of *relabelling* structures by changing this set, i.e. by replacing this set with another with which it is in bijection. In 1980 André Joyal introduced the notion of a *(combinatorial) species* for this. A species $\mathscr{A}$ is a functor from the category of finite sets with bijections (sometimes called Core(FinSet)) to itself. That is, it assigns to each finite set S a finite set $\mathscr{A}(S)$, the set of S-labelled structures, and to each bijection $f : S \xrightarrow{\sim} T$ of finite sets a corresponding bijection $\mathscr{A}(f) : \mathscr{A}(S) \xrightarrow{\sim} \mathscr{A}(T)$. These bijections must respect identities and composition in Core(FinSet): so $\mathscr{A}(\mathrm{id}_S) = \mathrm{id}_{\mathscr{A}(S)}$ and $\mathscr{A}(f \circ g) = \mathscr{A}(f) \circ \mathscr{A}(g)$.

We continue to allow ourselves the shorthand $\mathscr{A}_n = \mathscr{A}([n])$, and write $a_n = |\mathscr{A}_n|$. The right generating function to use for species is the exponential generating function,

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

Here are some first examples of species:

- The "atomic" species $\mathscr{Z}$, with $\mathscr{Z}_1 = \{\circ\}$ and $\mathscr{Z}_n$ empty for $n \neq 1$, so that $Z(x) = x$.

- The species Set of sets. This is the species that imposes no extra structure: the only set "labelled" by a set S is just S itself. So $|\mathrm{Set}_n| = 1$ for all n, and the generating function is e^x.

- The species Tot of total orders. Since a total order on a finite set can be specified by simply listing the elements from least to greatest, we counted these in Corollary 2.2, getting $|\mathrm{Tot}_n| = n!$. So the generating function is $1/(1-x)$.

- The species Perm of permutations. A permutation of a set S is a bijection $\sigma : S \to S$. When $S = [n]$ these can also be encoded as lists $\sigma(1), \cdots, \sigma(n)$ of all the elements of S, so again $|\mathrm{Perm}_n| = n!$ and the generating function is $1/(1-x)$.

The descriptions of these species are somewhat informal, in that they omit the bijections $\mathscr{A}(F)$. These are supposed to be obvious from the naïve idea of relabelling structures, and we will continue to pass over them in our definitions of operations. But they are an important part of the data.

For instance, although the species of total orders and permutations are equinumerous, they are not isomorphic species because relabelling acts differently (i.e. there is no invertible natural transformation between them). Relabelling a permutation acts on both the domain and codomain copies of S, so amounts to conjugation: if $f : S \xrightarrow{\sim} T$ then $\mathrm{Perm}(f)$ maps $\sigma : S \to S$ to $f \circ \sigma \circ f^{-1}$. On the other hand, relabelling a partial order acts "on only one side". If $<$ is a partial order on S and $f : S \xrightarrow{\sim} T$ then $f(a)\ \mathrm{Tot}(f)(<)\ f(b)$ iff $a < b$; the copy of $[n]$, as it were, that "labels" the position in the ordered list above is unaffected. To see that these are indeed different, note that the action of a nonidentity permutation $f : S \to S$ may fix some permutations but can never fix a partial order.

On the other hand, the species Seq of sequences containing each element of a set once is isomorphic to Tot.

4.1 Operations on exponential generating functions

Let $\mathscr{A}$ and $\mathscr{B}$ be species. The analogues of our sum and product rules for combinatorial classes from the last section are these. (We now use $\cdot$ for multiplication for conformance with the literature.)

Let $\mathscr{A} + \mathscr{B}$ be the species which associates to any set S the disjoint union $\mathscr{A}(S) \dot{\cup} \mathscr{B}(S)$. That is, an $\mathscr{A} + \mathscr{B}$ structure is either an $\mathscr{A}$ structure or a $\mathscr{B}$ structure.

Proposition 4.1 *The generating function of $\mathscr{A} + \mathscr{B}$ is $A(x) + B(x)$.*

Let $\mathscr{A} \cdot \mathscr{B}$ be the species given by

$$(\mathscr{A} \cdot \mathscr{B})(S) = \bigcup_{T \subseteq S} \mathscr{A}(T) \times \mathscr{B}(S \setminus T).$$

That is, to put an $\mathscr{A} \cdot \mathscr{B}$ structure on S, one takes a set partition of S into two parts, and puts an $\mathscr{A}$ structure on one part and a $\mathscr{B}$ structure on the other part.

Proposition 4.2 *The generating function of $\mathscr{A} \cdot \mathscr{B}$ is $A(x)B(x)$.*

Proof If we write c_n for $(\mathscr{A}\cdot\mathscr{B})_n$, then

$$c_n = \sum_k \binom{n}{k} a_k b_{n-k}.$$

So

$$\frac{c_n}{n!} = \sum_k \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!}$$

which is the coefficient of x^n in $A(x)B(x)$.

Suppose now the species $\mathscr{A}$ has $\mathscr{A}(\emptyset) = \emptyset$. The analogue of the free monoid construction on the formula level is $\mathrm{Seq}(\mathscr{A}) := \sum_{k\geq 0} \mathscr{A}^k$, where $\mathscr{A}^k$ again abbreviates $\underbrace{\mathscr{A}\cdot\dots\cdot\mathscr{A}}_{k}$, with $\mathscr{A}^0 = \mathbf{1}$, the species assigning a singleton $\{\varepsilon\}$ to the empty set and empty sets otherwise. Putting a $\mathrm{Seq}(\mathscr{A})$ structure on S entails taking an ordered set partition of S into k parts (for whichever $k \geq 0$), where the parts need not be empty, and then putting a $\mathscr{A}$ structure on each part.

We might also want to do an *unordered* set partition of S, where the parts still need not be empty, and then put some structure $\mathscr{A}$ on each part. We'll call the corresponding species $\mathrm{Set}(\mathscr{A})$.

Proposition 4.3 *The generating function of* $\mathrm{Seq}(\mathscr{A})$ *is* $1/(1-A(x))$.

Proposition 4.4 *The generating function of* $\mathrm{Set}(\mathscr{A})$ *is* $e^{A(x)}$.

The former of these follows from Proposition 4.2, and the latter by introducing factors of $1/k!$.

Let us analyse some familiar species into these operations.

(1) A total order on S is a totally ordered sequence of elements of S; that is, we put an ordered set partition on S and then put the structure on the parts that insists that they be of size 1. So $\mathrm{Tot} \cong \mathrm{Seq}(\mathscr{X})$, with generating function $1/(1-x)$, as just above.

    Of course, $\mathrm{Seq} \cong \mathrm{Seq}(\mathscr{X})$ shares this generating function. Likewise, the species Set is isomorphic to $\mathrm{Set}(\mathscr{X})$.

(2) The species of set partitions is almost Set(Set): set partitions are set partitions with no extra structure on the parts. But we must disallow empty parts; so in fact set partitions are $\mathrm{Set}(\mathrm{Set}_{>0})$, where the species $\mathrm{Set}_{>0}$ is like Set except in assigning the empty set to the empty set. That is, $\mathrm{Set} = \mathbf{1} + \mathrm{Set}_{>0}$.

So the exponential generating function for $\mathrm{Set}_{>0}$ is $e^x - 1$, and that for set partitions is

$$\sum_{n\geq 0} B_n \frac{x^n}{n!} = e^{e^x-1},$$

where the coefficients are the Bell numbers.

In the same vein, the species of ordered set partitions is $\mathrm{Seq}(\mathrm{Set}_{>0})$, so its egf is

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

(3) As in Section 2.5, a permutation decomposes into a union of disjoint cycles. So if Cycle is the species of permutations consisting of a single cycle, then $\mathrm{Perm} \cong \mathrm{Set}(\mathrm{Cycle})$. Writing $C(x)$ for the exponential generating function of cycles, we have

$$\frac{1}{1-x} = e^{C(x)},$$

i.e.

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

Therefore the number of cycles on k elements is $k!/k = (k-1)!$. We could also get this number directly, by bijective considerations: if σ is a cycle on $[k]$, then the values taken on by $\sigma^i(k)$ as i ranges from 1 to $k-1$ are the elements of $[k-1]$ in some order, and there are $(k-1)!$ of these.

The last example shows how to enumerate the connected components of a known structure: if $\mathscr{B} = \mathrm{Set}(\mathscr{A})$, then $A(x) = \log B(x)$.

(4) There are $2^{\binom{n}{2}}$ graphs on n vertices, since a graph is simply a subset of the $\binom{n}{2}$ edges of the complete graph. Therefore the exponential generating function for *connected* graphs on n vertices is

$$\log \sum_{n\geq 0} 2^{\binom{n}{2}} \frac{z^n}{n!},$$

for which I know of no particularly nice closed form.

Moreover, it is easy to refine the enumeration of $\mathscr{B}$-structures by their number of components.

Corollary 4.5 *Let $\mathscr{B} = \mathrm{Set}(\mathscr{A})$, and let $b_{n,k}$ be the number of $\mathscr{B}$-structures on $[n]$ with k components, i.e. the nth coefficient of $\mathscr{A}^k$. Then*

$$\sum_{n,k} b_{n,k} \frac{x^n}{n!} y^k = e^{yA(x)} = B(x)^y.$$

The proofs can be seen as a first application of the theory of *multisort species*, species on tuples of sets $(S_1, \dots, S_\ell)$ which enter into the structure differently. These have a generating function theory in an ℓ-variable formal power series ring, the exponents in each monomial encoding the cardinalities of the underlying sets. We focus on Corollary 4.5. Take each $\mathscr{A}$-structure to be labelled with elements of S_1 in the usual way and also bear a singleton label from S_2. The reason the y^k variable appears without a $k!$ at denominator is that we don't wish to distinguish $\mathscr{B}$-structures in which the different elements of S_2 are allotted to different components.

This applies in example (3) above. Recall that the signless Stirling number of the first kind, $|s(n,k)| = (-1)^{n-k} s(n,k)$, is the number of permutations of n with k cycles. By Corollary 4.5, these have a bivariate generating function

$$\begin{aligned}
\sum_{k,n} |s(n,k)| \frac{x^n}{n!} y^k &= \left(\frac{1}{1-x}\right)^y = (1-x)^{-y} \\
&= \sum_n \binom{-y}{n} (-x)^n \\
&= \sum_n y(y+1)\cdots(y+n-1) \frac{x^n}{n!} \\
&= \sum_n (-1)^n (-y)_n \frac{x^n}{n!}.
\end{aligned}$$

Extracting coefficients of $x^n/n!$ yields the generating function of Proposition 2.13 with the signs altered to make the count signless.

By manipulating the cycle types available, we can count other sets of permutations.

(5) A permutation σ is an *involution*, i.e. satisfies $\sigma = \sigma^{-1}$, if and only if all cycles in σ are of lengths 1 or 2. The egf for these cycles is simply $x + x^2/2$, so the egf for involutions is $e^{x+x^2/2}$.

(6) A *derangement* is a permutation with no fixed points, i.e. no cycles of length 1. The egf for cycles that are not fixed points is $-\log(1-x) - x$, so the generating function for derangements is

$$\sum_n d_n \frac{x^n}{n!} = e^{-\log(1-x)-x} = \frac{e^{-x}}{1-x} = e^{-x} + xe^{-x} + x^2 e^{-x} + \cdots,$$

using d_n for the number of derangements on $[n]$. Expanding and extracting coefficients of x^n gives

$$\frac{d_n}{n!} = \frac{(-1)^n}{n!} + \frac{(-1)^{n-1}}{(n-1)!} + \cdots + \frac{(-1)^0}{0!}$$

i.e.

$$d_n = n!\left(\frac{1}{0!} - \frac{1}{1!} + \cdots \pm \frac{1}{n!}\right) \sim \frac{n!}{e}.$$

In fact d_n is the nearest integer to $n!/e$ for $n \geq 1$, the difference being bounded by the next term $1/(n+1)$.

We can also easily derive a recurrence for the d_n from the above expansion: since only the last term has no counterpart in d_{n-1}, we get

$$d_n = nd_{n-1} + (-1)^n.$$

This is simpler than, though easily obtained from, the recurrence

$$d_n = (n-1)(d_{n-1} + d_{n-2})$$

which arises from the usual technique of deleting n from the structure and relating the result to a smaller structure.

Propositions 4.3 and 4.4, together with the commentary at the end of example (1), presage our next operation on exponential generating functions. Let $\mathscr{A}$ and $\mathscr{B}$ be species with $\mathscr{B}_0 = \emptyset$. Define their *composition* to be the species which assigns to S the set

$$(\mathscr{A} \circ \mathscr{B})(S) = \sum \left(\mathscr{A}(P) \times \prod_{S_i \in P} \mathscr{B}(S_i) \right)$$

where the sum ranges over set partitions $P = \{S_1, \ldots, S_r\}$ of S. That is, an $\mathscr{A} \circ \mathscr{B}$ structure on S is an $\mathscr{A}$-structure whose labels are $\mathscr{B}$-structures, with S comprising the totality of the labels of the $\mathscr{B}$-structures.

For example, if $\mathscr{A}$ is the species Set, the definition boils down to $\mathrm{Set} \circ \mathscr{B} = \mathrm{Set}(\mathscr{B})$. Similarly $\mathrm{Seq} \circ \mathscr{B} = \mathrm{Seq}(\mathscr{B})$. As such, we have seen several examples of composition of species above.

Proposition 4.6 *The generating function of $\mathscr{A} \circ \mathscr{B}$ is $A(B(x))$.*

Proof Letting $\mathrm{Set}_n(\mathscr{B})$ denote the subspecies of $\mathrm{Set}(B)$ where the "outer" set partition has n parts, we have

$$A(B(x)) = \sum_n a_n \frac{B(x)^n}{n!} = \sum_n a_n \,\mathrm{Set}_n(\mathscr{B})$$

where the nth term counts $\mathscr{A} \circ \mathscr{B}$-structures with the $\mathscr{A}$-structure having size n.

The next corollary generalises Corollary 4.5.

Corollary 4.7 *The exponential generating function of $(\mathscr{A} \circ \mathscr{B})$-structures in the indeterminate x, weighted by y^k where k is the size of the $\mathscr{A}$-structure, is $A(yB(x))$.*

Since our earlier operations $\mathrm{Seq}(\mathscr{B})$ and $\mathrm{Set}(\mathscr{B})$ are now unmasked as compositions, there are many examples of compositions above. As a further illustration we give one more.

(7) A *preorder* is a reflexive and transitive relation, and a *partial order* is an antisymmetric preorder.

Given any preorder $\prec$ on a set X, the relation $\sim$ such that $x \sim y$ if and only if $x \prec y$ and $y \prec x$ is an equivalence relation. Moreover, $\prec$ induces a partial order on the quotient $X/{\sim}$. Conversely, any partial order on the (nonempty) sets of a set partition of X can be extended naturally to a preorder $\prec$ on X, by taking $x \prec y$ iff the part of the partition containing x is less than or equal to that containing y. This shows that, if Preord is the species of preorders and PO the species of partial orders, then

$$\mathrm{Preord} = \mathrm{PO} \circ \mathrm{Set}_{>0}.$$

A formula for either of the generating functions involved is still, to my knowledge, an open question.

Exercise The n-cube is the undirected graph whose vertices are binary words of length n, with edges between pairs of words differing in just one position. Let $W(n,m)$ be the set of walks on the n-cube of length m beginning and ending at $000\cdots0$. Describe the species $\mathscr{W}$ with $\mathscr{W}([m]) = \bigcup_n W(n,m)$ as (isomorphic to) a composition. Using Corollary 4.7, give a bivariate generating function for $|W(n,m)|$, and a formula for this number.

Given a species $\mathscr{A}$, let $\mathscr{A}'$ be the species such that $\mathscr{A}'(S) = \mathscr{A}(S \dot\cup \{\circ\})$: that is, an $\mathscr{A}'$ structure on S is a $\mathscr{A}$-structure on the set obtained by adding a new distinguished element $\circ$ to S. Sometimes the extra element is thought of as a "hole" in the structure, and $\mathscr{A}'$ as arising from "puncturing" $\mathscr{A}$.

Particularly often useful is the species $\mathscr{X} \cdot \mathscr{A}'$. Putting this structure on a set S corresponds to partitioning S into a singleton $\{i\}$ (which trivially gets a $\mathscr{X}$-structure) and the remainder, which gets a new distinguished element added and an $\mathscr{A}$-structure imposed. The new element may be identified with i, so the total effect is to put an $\mathscr{A}$-structure on S while also distinguishing one of its elements. This is sometimes spoken of as "rooting" the structure[1]: for instance, if Tree is the species of trees, then $\mathscr{X} \cdot \mathrm{Tree}'$ is the species of rooted trees. We will be counting trees momentarily below.

[1] but presumably not in Australia!

Proposition 4.8 *The generating function of $\mathscr{A}'$ is the derivative $A'(x)$.*

So the generating function of $\mathscr{X} \cdot \mathscr{A}'$ is $xA'(x)$.

(8) Let Cycle be the species of cycles from example (3). A Cycle$'$-structure on a set S is a cycle on S and an extra element $\circ$. But the cycle can be cut at $\circ$ and this element discarded, leaving a ordered sequence on S. Thus Cycle$' \cong$ Seq. This agrees with the equation of generating functions

$$\frac{\mathrm{d}}{\mathrm{d}x} - \log(1-x) = \frac{1}{1-x}.$$

(9) The species Seq itself satisfies the recurrence Seq$' \cong$ Seq $\times$ Seq, by the bijection mapping a sequence of the elements of $S \dot{\cup} \{\circ\}$ to (the subsequence left of $\circ$, the subsequence right of $\circ$).

Writing $S(x)$ for the egf of Seq, we extract the differential equation

$$S'(x) = S(x)^2.$$

This is separable, and we get

$$1 = \frac{S'(x)}{S(x)^2} = -\left(\frac{1}{S(x)}\right)'$$

so $1/S(x) = -x + C$ and

$$S(x) = \frac{1}{C-x},$$

in which the constant C must be 1 to match $|\mathrm{Seq}_0| = 1$.

(10) Let us count the orderings $(w_1, \ldots, w_n)$ of the elements of $[n]$ such that

$$w_1 < w_2 > w_3 < w_4 > \cdots < w_{n-1} > w_n.$$

When $n > 0$ this is manifestly only possible for n odd. In conformance we say there are no such sequences when $n = 0$. These are generally called odd *alternating permutations* (failing to heed the distinction between the species of permutations and sequences).

We wish to use our trusty recurrence-producing technique of deleting n from the structure and analysing the result in smaller structures. This is possible, but there is a technical obstacle. The alternating permutations do not obviously constitute a species, as the symmetric group S_n does not act on them, i.e. set automorphisms of the underlying set $[n]$ destroy the structure. To say unsatisfyingly little, this can be circumvented by using functors from the category of totally ordered sets with bijections, in which context our sum and product and differentiation rules can be parallelly developed.

In any case, denoting this not-a-species by $\mathscr{E}$, we have $\mathscr{E}' \cong 1 + \mathscr{E} \times \mathscr{E}$, since if $\circ$ is taken to be greater than all elements of $[n]$, the bijection mapping an alternating permutation on $[n] \dot\cup \{\circ\}$ to (the subsequence left of $\circ$, the subsequence right of $\circ$) still holds, except that it fails to produce the alternating permutation of length 1. So its exponential generating function satisfies

$$E'(x) = 1 + E(x)^2,$$

whose solution is $E(x) = \tan x$.

The reader can check that the even alternating permutations, defined analogously, have exponential generating function $\sec x$.

4.1.1 Trees

The last operation we will discuss here is the compositional inverse of a power series. We pause to build up a setting in which we will use it, the species of trees.

A *tree* is a connected graph with no cycles. It is straightforward to show that a tree on n vertices contains $n-1$ edges, and that there is a unique path between any two vertices in a tree. Denote the species of trees by Tree. This species has a simple but unexpected formula for its labelled counting problem:

Theorem 4.9 (Cayley, Sylvester) *The number of labelled trees on n vertices is* n^{n-2}.

We will also make heavy use of the species of rooted trees, $\mathrm{RTree} := \mathscr{X} \cdot \mathrm{Tree}'$; from the theorem it will follow that there are n^{n-1} of these on n vertices. Our first proof of Cayley's theorem 4.9 above is due to Joyal, and features in Aigner and Ziegler's *Proofs from the Book*.

Proof As noted above, the species Seq and Perm, of linear orders and permutations, are quite different but are equicardinal: the numbers of each on a set of size n are the same, namely $n!$.

Hence the numbers of structures on any set are also equal for their compositions with the species of rooted trees, $\mathrm{Seq} \circ \mathrm{RTree}$ and $\mathrm{Perm} \circ \mathrm{RTree}$.

Consider an object in $(\mathrm{Seq} \circ \mathrm{RTree})(S)$. This consists of a linear ordering $(T_1, \ldots, T_r)$ of rooted trees. I claim that this is equivalent to a tree with two distinguished vertices: Joyal calls such objects *vertebrates*, with the distinguished vertices the *head* and *tail*. Let x_i be the root of T_i, and augment the collection of trees by the edges $\{x_i, x_{i+1}\}$ for $i = 1, \ldots, r-1$. The resulting graph is a single tree, and becomes a vertebrate by deeming x_1 its head and x_r its tail. Conversely, given a vertebrate with head and tail x and y, there is a unique path from x to y in the tree, the *backbone*, which becomes the linear order, and the remainder of the tree consists of rooted trees attached to the vertices of the path.

Now consider an object in $(\mathrm{Perm} \circ \mathrm{RTree})(S)$. Identify the root of each tree with the corresponding point of the set on which the permutation acts. The resulting structure defines a function f from the point set to itself, where

- if v is a root, then $f(v)$ is the image of v under the permutation;
- if v is not a root, then $f(v)$ is the vertex after v on the unique path from v to the root of the tree to which it belongs.

Conversely, given a function $f : S \to S$, the restruction of f to the set Y of periodic points of f (those points in the image of $f^{\circ n}$ for all n) is a permutation; the pairs $\{v, f(v)\}$ for which v is not a periodic point make up the edges of a family of trees, attached to Y at the point for which the iterated images of v under f first enter Y, which we declare to be their roots.

So the number of trees with two distinguished points is equal to the number of functions from the vertex set to itself. Thus, if there are $F(n)$ labelled trees, we see that

$$n^2 F(n) = n^n,$$

from which Cayley's theorem follows.

The structure RTree is very amenable to recursive description. If we remove the root from a rooted tree, the result consists of an unordered collection of trees, each of which has a natural root (at the neighbour of the root of the original tree). Conversely, given a collection of rooted trees, add a new root, joined to the roots of all the trees in the collection, to obtain a single rooted tree. So we have

$$\mathrm{RTree} \cong \mathscr{X} \cdot (\mathrm{Set} \circ \mathrm{RTree}).$$

Hence the exponential generating function $R(x)$ for rooted trees satisfies

$$R(x) = xe^{R(x)}.$$

This is, formally, a recurrence relation for the coefficients of $R(x)$, and the coefficients of $R(x)$ can be computationally evaluated as such. But by rearranging to

$$x = R(x)e^{-R(x)}$$

we see that $R(x)$ is the inverse under composition of formal power series of the series xe^{-x}. This inverse can be found systematically with the technique of *Lagrange inversion*.

We denote by $[x^a]f(x)$ the coefficient of x^a in a formal power series $f(x)$. This notation is of principal use when $f(x)$ is built up out of other series and operations, so that we have no ready-made notation f_n for its coefficients. (We let this notation $[x^a]$ bind very loosely in the order of operations sense, so that $[x^a]f^n = [x^a](f^n)$ and so on.)

Proposition 4.10 (Lagrange inversion) *Let f be a formal power series over a field of characteristic zero, with $f(0) = 0$ and $f'(0) \neq 0$. Then there is a unique formal power series g such that $g(f(x)) = x$, given by*

$$n[x^n]g(x)^k = k[x^{n-k}]\left(\frac{x}{f(x)}\right)^n.$$

An alternate form of the statement is that

$$[y^n]g(y) = \frac{1}{n!}\left[\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}\left(\frac{x}{f(x)}\right)^n\right]_{x=0}.$$

The general proof of either of these statements takes us far afield, so I will pass over them here.

Note that also $f(g(y)) = y$ for this inverse g, and in fact the formal power series of the proposition, with zero constant term and non-zero linear term, form a group, which has become known as the *Nottingham group*.

Let's use this to count trees. Since $R(x)$ is the compositional inverse of xe^{-x}, we get

$$\begin{aligned} n[x^n]R(x) &= 1 \cdot [x^{n-1}]\left(\frac{x}{xe^{-x}}\right) \\ &= [x^{n-1}]e^{nx} \\ &= \frac{n^{n-1}}{(n-1)!} \end{aligned}$$

and $[x^n]R(x)$ is $1/n!$ times the number of labelled rooted trees on n vertices. So these trees number n^{n-1}, and their unrooted analogues n^{n-2}, proving Cayley's theorem.

4.2 Counting orbits

Let X be a set, and G a group. An *action* of G on X is a group homomorphism ϕ from G to the symmetric group S_X of permutations of X. We suppress the name of the action itself and, given $g \in G$ and $x \in X$, write $g \cdot x$, or simply gx, for $\phi(g)(x) \in X$. That is, our actions are *left actions* (as opposed to *right actions*, where g sends x to xg). Explicitly, we have for all $g, h \in G$ and $x \in X$,

- $1x = x$, where 1 denotes the identity of G;
- $(gh)x = g(hx)$.

The *orbits* of the action are the equivalence classes of the relation $\sim$ on X with $x \sim y$ if $y = gx$ for some $g \in G$. The set of orbits of G on X is denoted X/G; the orbit of x is written Gx.

The perspective on algebra which reigned until the middle of the nineteenth century would have defined a *group* simply as the image of a group action on a set (although not in that language!) Now we call such an image, i.e. a subgroup of S_X, a *permutation group*.

If G is a group acting on a set X, then we can construct actions of G on various auxiliary sets built from X, for example, the set $X \times X$ of ordered pairs of elements of X, the set of subsets of X, the set of functions from X to another set or from another set to X. As one example, G acts on $X \times X$ by the rule

$$g(x, y) = (gx, gy)$$

for $x, y \in X, g \in G$; that is, the element g acts coordinate-wise on ordered pairs, mapping (x, y) to (gx, gy).

The foundational enumerative fact in this context is the Orbit-Stabiliser Theorem. The *stabiliser* G_x of x is the subgroup of all elements of G which fix x.

Theorem 4.11 (Orbit-Stabiliser Theorem) *Let G be a group acting on the finite set X, and $x \in X$. The orbit Gx is in bijection with the set of cosets G/G_x. Thus*

$$|Gx| = |G|/|G_x|.$$

Proof There is in fact a canonical such bijection, the one $hG_x \mapsto hx$ between the image and coimage of the map $G \to X$, $g \mapsto gx$ provided by the first isomorphism theorem. The result on cardinality follows from Lagrange's theorem.

The next proposition is often called *Burnside's lemma*, though Burnside attributed it not to himself but to Frobenius; it was apparently so well known even then that he missed giving attribution to Cauchy, who stated it twelve years earlier. Pólya's name has also been given to the lemma. Given $g \in G$, the symbol X^g denotes the subset of X *fixed* by the action of g, i.e. of $x \in X$ such that $gx = x$.

Proposition 4.12 *Let G be a group acting on the finite set X. Then the number of orbits of G on X is given by the formula*

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

Proof We count in two different ways the pairs (x, g), with $x \in X$, $g \in G$, and $gx = x$. Let there be N such pairs. On the one hand, clearly

$$N = \sum_{g \in G} |X^g|.$$

On the other hand, by the orbit-stabiliser theorem, if Gx is an orbit of size n then its stabiliser has size $|G|/n$, so the number of pairs (y, g) with $gy = y$ for which y lies in Gx is $n \cdot |G|/n = |G|$. So each orbit contributes $|G|$ to the sum, and so $N = |G|k$, where k is the number of orbits. Equating the two values gives the result.

A first example of standard kind is to count the ways to colour the sides of a cube with one of n colours, where two colourings count as the same if they're in the same orbit under rotations of the cube, i.e. if one colouring can be turned in $\mathbb{R}^3$ to coincide with the other. To use Burnside's lemma, we have to examine the 24 rotations of the cube and find the number of colourings fixed by each. (The group of these rotations is the symmetric group S_4, which acts as the group of all permutations on the pairs of opposite vertices of the cube.)

- The identity fixes all n^6 colourings.
- There are three axes of rotation through the mid-points of opposite faces. A rotation through a half-turn about such an axis fixes n^4 colourings: we can choose arbitrarily the colour for the top face, the bottom face, the front and back faces, and the left and right faces (assuming that the axis is vertical). A rotation about a quarter turn fixes n^3 colourings, since the four faces other than top and bottom must all have the same colour. There are three half-turns and six quarter-turns.
- A half-turn about the axis joining the midpoints of opposite edges fixes n^3 colourings. There are six such rotations.
- A third-turn about the axis joining opposite vertices fixes n^2 colourings. There are eight such rotations.

By Burnside's lemma, the number of orbits of colourings is

$$\frac{1}{24}(n^6 + 3n^4 + 12n^3 + 8n^2). \tag{1}$$

A second application is another proof of the generating function for signless Stirling numbers of the first kind $|s(n,k)|$ in example (3) of Section 4.1. Since the identity to be proved is between two polynomials in x, we may assume that x is a positive integer.

Consider the set of functions from $\{1,...,n\}$ to a set X of cardinality x. There are x^n such functions. Let the symmetric group S_n act on these functions by

$$\sigma(f)(i) = f(\sigma^{-1}(i))$$

for $\sigma \in S_n$. The orbits are simply the selections of n things from X, where repetitions are allowed and order is not important. So the number of orbits is

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

We can also count the orbits using Burnside's Lemma. Let g be a permutation in S_n having k cycles. How many functions are fixed by g? Clearly a function f is fixed if and only if it is constant on each cycle of g; its values on the cycles can be chosen arbitrarily. So there are x^k fixed functions. Since the number of permutations with k cycles is $|s(n,k)|$, Burnside's Lemma shows that the number of orbits is

$$\frac{1}{n!}\sum_k |s(n,k)|x^k.$$

Equating the two expressions and multiplying by $n!$ gives the result.

A naïve attempt to count the orbits of a group G on a finite set X might conclude that there are $|X|/|G|$ of them. Of course, that count need not even be an integer; it is only correct when the action is *free*, i.e. only the identity element of G fixes any element of X. There is however a way to save the formula $|X/G| = |X|/|G|$, which we will only briefly sketch here. This is done by changing the meaning of $|\cdot|$ from ordinary set cardinality to a weighted version, weighting each element by the inverse size of its automorphism group,

$$\text{“}|Y|\text{”} := \sum_{y\in Y} \frac{1}{|\operatorname{Aut}(y)|}.$$

For instance, if X is a set of objects labelled by a finite set, reckoned as lacking automorphisms, then any unlabelled object Gx in X/G, written as the orbit of some $x \in X$, inherits an automorphism for each element of the stabiliser of x, and then Burnside's lemma reads "$|X/G|$" = "$|X|$"$/|G|$. This new function "$|\cdot|$" is formally defined in the setting of groupoids and known as *groupoid cardinality*. Some of its many applications are as the "mass" in the Smith-Minkowski-Siegel mass formulas for lattices, and in probability distributions that arise in natural situations such as the Cohen-Lenstra heuristics for class groups.

4.2.1 Counting orbits with weights

As we have taken unlabelled structures to be S_n-orbits of labelled structures, Burnside's lemma is a significant tool in counting them. Under the name *Pólya theory*, the extension of tools like Burnside's lemma to structures with a parameter is a standard topic in enumerative combinatorics. The theory springs from examples like ours with the cube in the previous section. How would we extend this example if we wished to count the colourings *weightedly*? For instance, we might wish to subclassify the colourings by their number of white faces. To adjust the Burnside analyses for these purposes, we would need to track not just how many orbits each group element has but how large each orbit is, so that it counts with the necessary weight when coloured white.

The development is most natural in the context of *symmetric functions*, which I should have liked to introduce properly if I'd expected to have the time to. Instead I'll give a lightning introduction here.

Let R be a divisible ring: $\mathbb{Q}$ or $\mathbb{C}$ or similar are fine choices. The polynomial ring $R[x_1,\dots,x_n]$ bears an action of the symmetric group S_n permuting the variables. The subring $\Lambda^n := R[x_1,\dots,x_n]^{S_n}$ of polynomials fixed by all permutations of the variables, which we call *symmetric functions* in n variables, is itself a polynomial ring: one presentation thereof is $\Lambda^n = R[p_1,\dots,p_n]$ where

$$p_k = x_1^k + \cdots + x_n^k.$$

For notational convenience, we write $p_{k_1\dots k_s} = p_{k_1}\cdots p_{k_s}$. Such products in which no subscript exceeds n form an R-module basis for Λ^n.

Note that the polynomials p_k with $k > n$ are also in Λ^n; they merely fail to be algebraically independent of the generators $p_1,\dots,p_n$. Now given naturals $m < n$, there is an inclusion of rings $\iota : \Lambda^m \hookrightarrow \Lambda^n$ which sends $p_k \in \Lambda^m$ to $p_k \in \Lambda^n$ for all $k \le m$. By what we have just noted, this is a funny inclusion in that the image of p_k for $k > m$ is not p_k, but it is well-defined nonetheless.

We can also define *the* ring of symmetric functions Λ to be the direct limit of the inclusions ι, that is the union of all the Λ^n identified under our inclusions. More informally, an element of Λ is a symmetric polynomial "in infinitely many variables". This Λ is a polynomial ring in countably many generators, $\Lambda = R[p_1, p_2, \dots]$. The rings Λ^n also bear a family of surjections, $\pi : \Lambda^n \twoheadrightarrow \Lambda^m$

for $n > m$ given by $\pi(x_i) = x_i$ for $i \leq m$ and $\pi(x_i) = 0$ for $i > m$, such that

$$\Lambda^m \overset{\iota}{\hookrightarrow} \Lambda^n \overset{\pi}{\twoheadrightarrow} \Lambda^m$$

is the identity. This means that given any R-algebra A and elements $a_1, a_2, \ldots \in A$, all but finitely many of which are zero, there is an evaluation map $\Lambda \to A$ which substitutes a_i for x_i.

Our setup is the following. Suppose S is a set with an action of a group G, and X is a set of "colours". By quotienting out the subgroup of elements which act trivially, we may and will assume that G is a subgroup of the symmetric group on S. As above, there is an induced G-action on the set X^S of set maps $f : S \to X$, that is, colourings of the elements of S. Momentarily fix a bijection $\pi : X \xrightarrow{\sim} [n]$. Then to any such f we can associate a monomial $x^f \in R[x_1, \ldots, x_n]$ recording which colours appear, given as

$$x^f = \prod_{s \in S} x_{\pi(f(s))}.$$

In fact x^f depends only on the G-orbit of f within X^S, so we may as well write it x^{Gf}. The *pattern enumerator* is the sum of this monomial for all possible orbits,

$$F_G = \sum_{Gf \in X^S/G} x^{Gf}.$$

Note that F_G is a symmetric function in $\Lambda^{|S|}$, because the colours in X play symmetric roles. Therefore, F_G is independent of π.

On the other hand, for a subgroup G of the symmetric group on S, we define a *cycle indicator*. If g is a permutation of S, momentarily let $\mathrm{Cyc}(g)$ be the set partition of S given by the cycles into which g decomposes. The cycle indicator is then

$$Z_G = \frac{1}{|G|} \sum_{g \in G} \prod_{C \in \mathrm{Cyc}(g)} p_{|C|},$$

which is a symmetric function in $\Lambda^{|S|}$.

Proposition 4.13 *Let S be a finite set and G a subgroup of its symmetric group. Then $Z_G = F_G$.*

Proof To apply Burnside's lemma, we have to count the functions of given weight fixed by a permutation $g \in G$. As we have seen, a function is fixed by g if and only if it is constant on the cycles of g. Now, functions fixed by g associated the same colour to every point of any cycle. For a particular i-cycle of g, the generating function of the assignments of colours to that cycle is exactly $p_i = x_1^i + \cdots + x_n^i$. So the generating function for all functions fixed by g is $\prod_{C \in \mathrm{Cyc}(g)} p_{|C|}$. Averaging over all $g \in G$ gives the result.

For example, the cycle indicator for the cube above is

$$\frac{1}{24}(p_{111111} + 3p_{2211} + 6p_{411} + 6p_{222} + 8p_{33}).$$

We may evaluate it at $(x_1, x_2, x_3, x_4, \ldots) = (x, 1, 1, 0, 0, 0, \ldots)$, to get a generating function for the number of three-coloured cubes according to the number of white faces: we get

$$\frac{1}{24}\Big((x+2)^6 + 3(x^2+2)^2(x+2)^2 + 6(x^4+2)(x+2)^2 + 6(x^2+2)^3 + 8(x^3+2)^2\Big)$$
$$= 10 + 12x + 16x^2 + 10x^3 + 6x^4 + 2x^5 + x^6.$$

Alternatively, setting n variables to 1 and the remainder to 0 will recover the polynomial of equation 1.

We give an application to tie this to the foregoing material. To compute the cycle indicator of the whole symmetric group, we wish to take a weighted average over all permutations, where the weight of a permutation is the product of p_i for each cycle of size i. So we redo example (3) with the generating functions appropriately weighted, and using the variable y for the formal power series since I've used x in the symmetric functions. Inserting this weight, the generating function for the species of cycles in the ring $\Lambda[[y]]$ becomes

$$\sum_{n\geq 1} \frac{p_n}{n} y^n.$$

This can be formally rewritten as the symmetric power series

$$\sum_{i\geq 1} -\log(1 - x_i y);$$

the expression is legitimate as long as we eventually make a specialisation with all but finitely many x_i specialised to zero. So the egf for permutations with these weights on their cycles, in y, is

$$e^{\sum_{i\geq 1} -\log(1-x_i y)} = \prod_{i\geq 1} \frac{1}{1 - x_i y}.$$

This is the *ordinary* generating function for the cycle indicators of S_n; it is ordinary because the coefficients $1/n!$ in the exponential generating function machinery are absorbed by the definition of the cycle indicator.

This expression is therefore also the ordinary generating function for pattern enumerators for coloured (unlabelled) sets. For example, to count n-coloured sets, we set n of the x_i to 1 and the others to zero, yielding

$$\left(\frac{1}{1-y}\right)^n = \sum_k \binom{-n}{k} (-y)^k = \sum_k \binom{n+k-1}{k} y^k$$

by the binomial theorem. But an n-coloured unlabelled set of size k is a way to put k indistinguishable balls into n distinguishably coloured boxes, i.e. a weak composition of k with n terms. So this replicates our count of Section 2.

Finally, readers interested in representation theory may wish to note an interpretation of the cycle indicator. If G is a subgroup of S_n, consider the induced representation $1_G^{S_n}$ (in characteristic zero), that is, the representation with a basis labelled by the right cosets of G in S_n, on which S_n acts by permutation matrices. Under a standard identification of Λ^n with the vector space generated by characters of S_n-representations, the cycle indicator Z_G is identified with the character of $1_G^{S_n}$. For the details, see sections 7.18 and 7.24 of Stanley's *Enumerative Combinatorics, volume 2*.