Labelled Enumeration

2

The symbolic method discussed in Chap. 1 may not always be suitable in addressing enumeration problems in combinatorial classes where some natural way of distinguishing objects by its labels appears, for example in the class of permutations. In this chapter, the notion of labelled classes is introduced and the power and flexibility of the symbolic method will again be demonstrated in applications to count classes of permutations, set partitions, labelled trees, words and other combinatorial objects.

2.1 Exponential Generating Functions

The **exponential generating function** of a sequence $a = (a_1, a_2, \cdots)$ of complex numbers is

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

We note that, for exponential generating functions, we have

$$a_n = n![z^n]A(z).$$

Exponential generating functions turn out to be a more convenient type of generating functions for labelled classes. The main reason for this is their behaviour with respect to the product. If $A(z)$, $B(z)$ are exponential generating functions of sequences a_n and b_n respectively, then their product

$$A(z)B(z) = \sum_{n\geqslant 0} \left(\sum_{k=0}^{n} \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} \right) z^n$$

is the exponential generating function of the sequence

S. Ball, O. Serra, *A Course in Combinatorics and Graphs*, Compact Textbooks in Mathematics, https://doi.org/10.1007/978-3-031-55384-4_2

$$c_n = n![z^n]A(z)B(z) = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}.$$

The above expression is sometimes called the binomial convolution of the sequences a_n and b_n.

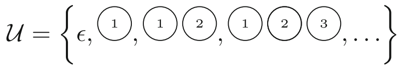
2.2 Labelled Classes

A combinatorial class $\mathcal{A}$ is *labelled* if the objects in $\mathcal{A}$ are labelled graphs, the size of an object being the number of vertices in the graph, and the labels of an object $\alpha \in \mathcal{A}$ of size n are distinct labels in $\{1, \ldots, n\}$ attached to the vertices of the graph.

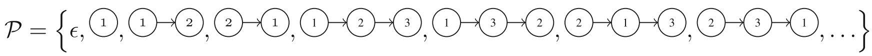
For convenience we also consider the null class ϵ which has a single object of size zero and no labels. We also denote by $\mathcal{N}$ the class with a sole object of size one with label 1.

We first consider some important examples of labelled classes.

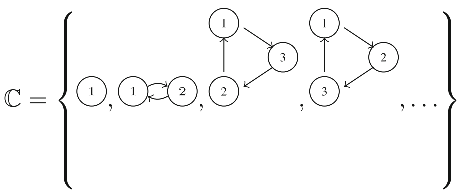
The Class $\mathcal{U}$ of Urns The objects of $\mathcal{U}$ are edgeless labelled graphs. There is a unique way of labelling the n vertices, so there is a unique object of each size.



The Class $\mathcal{P}$ of Permutations The objects of $\mathcal{P}$ are labelled directed paths. There are $n!$ different labelings of a directed path with n vertices.



The Class $\mathcal{C}$ of Cyclic Permutations The objects of $\mathcal{C}$ are labelled directed cycles. There are $(n-1)!$ different labelings of a directed cycle with $n \geqslant 1$ vertices.



In an analogue of the previous section, the **exponential generating function** of a combinatorial class $\mathcal{A}$ is

$$A(z) = \sum_{\alpha} \frac{z^{|\alpha|}}{|\alpha|!} = \sum_{n \geq 0} A_n \frac{z^n}{n!},$$

where A_n is the number of objects in $\mathcal{A}$ of size n.

As we shall see, the exponential generating function is a more convenient generating function for labelled classes.

For the three examples discussed above, the exponential generating functions are

$$U(z) = \sum_{n \geqslant 0} \frac{z^n}{n!} = e^z,$$

$$P(z) = \sum_{n \geqslant 1} z^n = \frac{1}{1-z},$$

and

$$C(z) = \sum_{n \geqslant 1} \frac{z^n}{n} = -\log(1-z).$$

2.3 Labelled Constructions

As for the unlabelled case, the power of the symbolic method relies on the possibility of describing a class in a formal symbolic way by means of elementary operations. The basic ones are described below.

The **sum** $\mathcal{A} + \mathcal{B}$ of two labelled classes $\mathcal{A}$ and $\mathcal{B}$ is simply its disjoint union. Every object in $\mathcal{A} + \mathcal{B}$ inherits its size and labels from its class.

The labelled product is the most interesting operation. We introduce some convenient terminology as follows. Given any sequence of n pairwise distinct natural numbers $a = (a_1, a_2, \ldots, a_n)$, its **reduction** $\rho(a)$ is the sequence $\sigma(1), \ldots, \sigma(n)$, where $\sigma \in \mathrm{Sym}(n)$ is a permutation of $\{1, \ldots, n\}$, with the property that $\sigma(i) < \sigma(j)$ if and only if $a_i < a_j$. In other words, the reduction is an order preserving map onto $\{1, \ldots, n\}$. For example, $\rho(4, 8, 3, 6, 2) = (3, 5, 2, 4, 1)$.

Let $\mathcal{A}, \mathcal{B}$ be labelled classes. In order to define the labelled product we must define a way to label the pairs $(\alpha, \beta) \in \mathcal{A} \times \mathcal{B}$ with labels in $\{1, 2, \ldots, |\alpha| + |\beta|\}$. The idea is to use all possible labels on (α, β) whose reductions on α and β coincide with the original ones. Therefore, we define

$$\alpha * \beta = \{(\alpha', \beta') \in \mathcal{A} \times \mathcal{B} \mid \rho(\alpha') = \alpha,\ \rho(\beta') = \beta\}.$$

For example, if α is labelled $(1, 2)$ and β is labelled $(1, 3, 2)$ then $\alpha * \beta$ consists of the objects (α', β') with labels

$$(1, 2, 3, 5, 4), (1, 3, 2, 5, 4), (1, 4, 2, 5, 3), (1, 5, 2, 4, 3), (2, 3, 1, 5, 4),$$
$$(2, 4, 1, 5, 3), (2, 5, 1, 4, 3), (3, 4, 1, 5, 2), (3, 5, 1, 4, 2), (4, 5, 1, 3, 2).$$

Note that

$$|\alpha * \beta| = \binom{|\alpha| + |\beta|}{|\alpha|}.$$

The **labelled product** of two labelled classes $\mathcal{A}$ and $\mathcal{B}$ is defined as

$$\mathcal{A} * \mathcal{B} = \bigcup_{\substack{\alpha\in\mathcal{A}\\ \beta\in\mathcal{B}}} (\alpha * \beta).$$

If $\mathcal{C} = \mathcal{A} * \mathcal{B}$ then

$$\begin{aligned} C(z) &= \sum_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} \sum_{(\alpha,\beta)\in\alpha*\beta} \frac{z^{|\alpha|+|\beta|}}{(|\alpha|+|\beta|)!} \\ &= \sum_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} \frac{(|\alpha|+|\beta|)!}{|\alpha|!|\beta|!} \frac{z^{|\alpha|+|\beta|}}{(|\alpha|+|\beta|)!} \\ &= \left(\sum_{\alpha\in\mathcal{A}} \frac{z^{|\alpha|}}{|\alpha|!}\right)\left(\sum_{\beta\in\mathcal{B}} \frac{z^{|\beta|}}{|\beta|!}\right) \\ &= A(z)B(z), \end{aligned}$$

which explains the use of exponential generating functions instead of ordinary ones. The labelled product is the natural product operation of labelled combinatorial classes.

The **sequence** $\mathrm{Seq}(\mathcal{A})$ of a labelled class $\mathcal{A}$ is defined as

$$\mathrm{Seq}(\mathcal{A}) = \{\epsilon\} + \mathcal{A} + (\mathcal{A} * \mathcal{A}) + (\mathcal{A} * \mathcal{A} * \mathcal{A}) + \cdots = \bigcup_{k\geqslant 0} \mathrm{Seq}_k(\mathcal{A}),$$

where

$$\mathrm{Seq}_k(\mathcal{A}) = \underbrace{\mathcal{A} * \cdots * \mathcal{A} * \mathcal{A}}_{k \text{ times}}.$$

By the expression of generating functions of labelled products, if $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$ then

$$C(z) = 1 + A(z) + A^2(z) + \cdots = \frac{1}{1 - A(z)}.$$

The class of k**-sets** $\mathrm{Set}_k(\mathcal{A})$ of a labelled class $\mathcal{A}$ is

$$\mathrm{Seq}_k(\mathcal{A})/\sim$$

where two objects are identified by the equivalence relation $\sim$ if they only differ on the ordering of its components. The class of sets of $\mathcal{A}$ is

$$\mathrm{Set}(\mathcal{A}) = \bigcup_{k \geq 0} \mathrm{Set}_k(\mathcal{A}).$$

We observe that, in the labelled product, the number of n-tuples of objects in each every equivalence class of $\mathrm{Seq}_n(\mathcal{A}))/\sim$ is $n!$. Accordingly, if $\mathcal{D} = \mathrm{Set}(\mathcal{A})$ then

$$D(z) = 1 + A(z) + \frac{1}{2!}A^2(z) + \frac{1}{3!}A^3(z) + \cdots = e^{A(z)}.$$

This is known as the exponential formula in classical enumerative combinatorics, to which the symbolic method gives a natural and simple derivation.

We summarise the above relations between operations in labelled classes and the corresponding exponential generating functions in the following theorem.

Theorem 2.1 *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be labelled classes and denote by $A(z)$, $B(z)$, $C(z)$ their exponential generating functions.*

1. *If $\mathcal{A} = \mathcal{B} + \mathcal{C}$, then $A(z) = B(z) + C(z)$.*
2. *If $\mathcal{A} = \mathcal{B} * \mathcal{C}$, then $A(z) = B(z)C(z)$.*
3. *If $\mathcal{A} = \mathrm{Seq}(\mathcal{B})$, then $A(z) = \frac{1}{1-B(z)}$.*
4. *If $\mathcal{A} = \mathrm{Set}(\mathcal{B})$, then $A(z) = e^{B(z)}$.*

2.4 Permutations

Permutations The class of permutations was defined as the class of labelled directed paths. An alternative symbolic description provides a wealth of enumeration possibilities. Recall that a permutation can be expressed as a product of cycles in a unique way.

Let $\mathcal{C}$ denote the class of cyclic permutations. The number C_n of cyclic permutations of $\{1, \ldots, n\}$ of size n in $\mathcal{C}$ is $(n-1)!$. Therefore, the exponential generating function of $\mathcal{C}$ is

$$C(z) = \sum_{n \geqslant 1} (n-1)! \frac{z^n}{n!} = \log \frac{1}{1-z}.$$

A permutation is a set of disjoint cycles. The class $\mathcal{P}$ of permutations has the formal specification

$$\mathcal{P} = \text{Set}(\mathcal{C}).$$

It follows that the exponential generating function is

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

as we have already seen. However, this specification allows for the flexibility of the symbolic method. The following examples illustrate this fact.

Derangements. A **derangement** is a permutation with no fixed points (that is, with no cycles of length one). The formal specification of the class $\mathcal{D}$ of derangements is

$$\mathcal{D} = \text{Set}\left(\sum_{n \geqslant 2} C_n\right),$$

and thus the generating function is

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

Thus, one obtains directly that

$$D_n = n![z^n]e^{-z}\left(\frac{1}{1-z}\right) = n! \sum_{i=0}^{n} (-1)^i \frac{1}{i!} \approx \frac{n!}{e}.$$

Involutions. An **involution** is a permutation σ with the property that σ^2 is the identity. The cycle decomposition of an involution has only cycles of lengths one or two. The class of involutions is

$$\mathcal{I} = \text{Set}(C_1 + C_2),$$

and its exponential generating function is

$$I(z) = \exp(z + \frac{z^2}{2}).$$

Therefore, the number of involutions of size n is

$$i_n = n![z^n]I(z) = n![z^n]e^z e^{z^2/2} = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!}{k!(n-2k)!2^k}.$$

More generally, the class $\mathcal{I}_r$ of permutations satisfying $\sigma^r = 1$ is

$$\mathcal{I}_r = \text{Set}(\sum_{j|r} C_j)$$

and its exponential generating function is

$$I_r(z) = \exp\left(\sum_{j|r} \frac{z^j}{j}\right).$$

Number of Cycles. The class $\mathcal{P}^{(k)}$ of permutations with k disjoint cycles is

$$\mathcal{P}^{(k)} = \text{Set}_k(\mathcal{C}).$$

It has exponential generating function

$$P^{(k)}(z) = \frac{1}{k!}\left(\log \frac{1}{1-z}\right)^k.$$

The number of such permutations of n is the (signless) *Stirling number of first kind*, or the Stirling cycle number,

$$\left[n \atop k \right] = \frac{n!}{k!}[z^n]\left(\log \frac{1}{1-z}\right)^k = \frac{n!}{k!} \sum_{i_1+i_2+\cdots+i_k=n} \frac{1}{i_1 i_2 \cdots i_k}.$$

Some simple values of the Stirling cycle numbers are

$$\left[n \atop 1 \right] = (n-1)!, \left[n \atop n-1 \right] = \binom{n}{2}.$$

Of course,

$$\sum_{k=1}^{n} \left[n \atop k \right] = n!.$$

Number of Cycles and Cycle Lengths. We may specify the set $\mathcal{P}_{A.B}$ of permutations which have cycles with length in $A \subset \mathbb{N}$ and a number of cycles which is an integer in $B \subset \mathbb{N}$. The formal specification is

$$\mathcal{P}_{A.B} = \prod_{i \in B} (\sum_{j \in A} \mathcal{C}_j),$$

from which

$$P_{A,B}(z) = \beta(\alpha(z)),$$

where

$$\alpha(z) = \sum_{a \in A} \frac{z^a}{a!}, \quad \beta(z) = \sum_{b \in B} \frac{z^b}{b!}.$$

2.5 Set Partitions

A partition is a collection of nonempty sets. We have already seen an approach to enumerate partitions by describing them as an unlabelled combinatorial class. However, its description as a labelled combinatorial class is more natural: a partition is a set of disjoint subsets. The labelled class of partitions can be described as

$$\mathcal{P} = \text{Set}(\mathcal{U}_1),$$

where $\mathcal{U}_1$ is the class of urns (excluding the empty object). Therefore, the exponential generating function of the class of partitions is

$$P(z) = \exp(\exp(z) - 1).$$

The total number of partitions of $\{1, \ldots, n\}$ is the Bell number B_n. The exponential generating function provides an expression for this number

$$B_n = n![z^n]P(z) = \frac{n!}{e}[z^n] \sum_{k \geqslant 0} \frac{e^{kz}}{k!} = \frac{n!}{e}[z^n] \sum_{k \geqslant 0} \sum_{m \geqslant 0} \frac{k^m z^m}{m!k!} = \frac{1}{e} \sum_{k \geqslant 0} \frac{k^n}{k!},$$

from which one can obtain asymptotic expressions.

The class $\mathcal{P}^{(k)}$ of partitions into k parts is

$$\mathcal{P}^{(k)} = \text{Set}_k(\mathcal{U}_1).$$

Hence,

$$P^{(k)}(z) = \frac{1}{k!}(e^z - 1)^k.$$

This gives an alternative derivation of the Stirling numbers of second kind,

$$\begin{aligned}\left\{ {n \atop k} \right\} &= \frac{n!}{k!}[z^n](e^z - 1)^k \\ &= \frac{n!}{k!}[z^n]\sum_{m=0}^{k}\binom{k}{m}(-1)^{k-m}e^{mz} \\ &= \frac{n!}{k!}[z^n]\sum_{m=0}^{k}\binom{k}{m}(-1)^{k-m}\sum_{\ell\geqslant 0}\frac{m^\ell}{\ell!}z^\ell \\ &= \frac{1}{k!}\sum_{m=0}^{k}\binom{k}{m}(-1)^{k-m}m^n.\end{aligned}$$

Additional specializations can be obtained as in the case of permutations. For instance, the exponential generating function for the class of partitions with no singletons is

$$\exp(e^z - 1 - z).$$

2.6 Words

Words on an alphabet can also be treated from the perspective of labelled combinatorial classes. A word of length n on an alphabet $A = \{a_1, \ldots, a_r\}$ can be seen as a map

$$f : \{1, \ldots, n\} \to A$$

and it can be specified by the sequence

$$(f^{-1}(a_1), \ldots, f^{-1}(a_r)),$$

a sequence of subsets (including the empty set). Therefore, the class $\mathcal{W}_A$ of words on A can be specified as

$$\mathcal{W}_A = (\mathcal{U})^r,$$

where $\mathcal{U}$ is the class of urns, now including the empty set. This gives

$$W_A(z) = e^{rz} \text{ and } w_{A,n} = r^n,$$

as expected.

If the number of occurrences of the letter i is restricted to a set $A_i \subset \mathbb{N}$ then the symbolic specification is

$$\mathcal{W}_{A;A_1,\ldots,A_r} = \mathcal{U}_{A_1} * \cdots * \mathcal{U}_{A_r}$$

where

$$\mathcal{U}_{A_i} = \sum_{a \in A_i} (\mathbb{N})^a / \sim,$$

and so

$$U_{A_i} = \sum_{a \in A_i} \frac{z^a}{a!}.$$

For example, the class of words on an alphabet of r letters in which each letter appears at least twice has exponential generating function

$$W_A^{\geqslant 2} = (e^z - 1 - z)^r$$

2.7 Labelled Trees

Let $\mathcal{T}$ be the class of rooted labelled trees. There is a distinguished vertex (the root) and the nodes of the trees are labelled. The size of the tree is its number of nodes.

A tree in $\mathcal{T}$ consists of a node and a set of trees. Therefore,

$$\mathcal{T} = \mathcal{N} * \mathrm{Set}(\mathcal{T}).$$

The generating function satisfies the equation

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

The Lagrange inversion formula provides the classical Cayley formula for the number of labelled rooted trees with n vertices.

$$T_n = n^{n-1}$$

The above is the well-known Cayley formula for the number of labelled trees (dividing by the n possible roots of a labelled tree).

2.8 Notes and References

For a more comprehensive look at labelled enumeration, see Flajolet and Sedgewick (2009, Chapter 2). The general symbolic approach to enumeration problems can be traced back to Joyal (1981) who derived the exponential formula for the set construction. There are many identities involving the Stirling numbers, for partitions and for permutations, which have not been explored here. For the Stirling permutation numbers there is a closed formula which is more involved than the one obtained here for the Stirling partition numbers. Cayley's formula for the number of spanning trees has many beautiful proofs, see for example Matousek and Nesetril (2008). The one given here is probably the simplest one.

2.9 Exercises

Exercise 2.1 Compute the number $f_n(r)$ of permutations which have no cycles of length r ($f_n(1)$ is the number of derangements). Prove that $\lim_{n\to\infty} f_n(r)/n! = e^{-1/r}$.

Exercise 2.2 Compute the exponential generating function of the permutations which decompose into even cycles. Analogously for the ones decomposing into odd cycles.

Exercise 2.3 Compute the exponential generating function of the permutations which decompose into an even number of cycles. Analogously for the ones decomposing into an odd number of cycles.

Exercise 2.4 Show that the number of permutations of $\{1, \dots, 2n\}$ whose cycle decomposition contains only even cycles is

$$(2n-1)^2(2n-3)^2 \cdots 3^2.$$

Exercise 2.5 Let $\mathcal{W}^{(k,r)}$ denote the class of words over the alphabet $\{a, b\} \cup \{0, 1\}$ in which every letter appears at most k times and each number appears at least r times. The size of a word is its length.

i. Compute $W_n^{(0,r)}$ for $n > 2(r-1)$, $r \geqslant 1$.
ii. Give an expression of the exponential generating function of $\mathcal{W}^{(k,0)}$.
iii. Compute $W_n^{(2,2)}$ for $n > 6$.

Exercise 2.6 Compute the exponential generating function of set partitions with an odd number of blocks.

Exercise 2.7 Compute the exponential generating function of rooted labelled trees such that the root has exactly k descendants. Find the number of such trees with n nodes.

Exercise 2.8 Let $\mathcal{T}_2$ denote the class of rooted binary labelled trees, that is, every node has zero or two childs, the size being the number of nodes. Let $\mathcal{F}_2$ denote the class of forests in which each connected component is an object in $\mathcal{T}_2$.

i. Find the exponential generating function of $\mathcal{T}_2$.
ii. Give an expression of $T_{2,n}$ in terms of the Catalan numbers. What is the value of $T_{2,n}$ for $n = 1, 2, 3, 4, 5$?
iii. Use the Bürman–Lagrange formula to obtain an expression of $F_{2,n}$, and compute the first four values.

Exercise 2.9 A star is a tree where all but at most one vertex is a leaf. Let $\mathcal{S}$ be the class of rooted labelled star forests (a forest is a set of trees). The size of a star forest is the number of vertices it has.

i. Give a symbolic description of $\mathcal{S}$, the exponential generating function of the class and derive the number of rooted labelled star forests with n vertices. Compute the first few values of these numbers.
ii. Use the above to count the number of maps $f : \{1, \dots, n\} \to \{1, \dots, n\}$ which are idempotent: $f(f(x)) = f(x)$ for all x.
iii. Let $\mathcal{I}^{(3)}$ be the class of maps $f : \{1, \dots, n\} \to \{1, \dots, n\}$ such that $f^3 = f$, the size of a map being n. Give a symbolic description, the exponential generating function of the class and an expression for the number of such maps on $\{1, \dots, n\}$. Compute the first few values of these numbers.

Exercise 2.10 Let $\mathcal{P}_{k,2}$ be the class of partitions of a set into k parts, each of them has cardinality at least two.

1. Find the exponential generating function of $\mathcal{P}_{k,2}$.
2. Show that the number $s_{n,k}$ of doubly surjective maps $f : \{1, \dots, n\} \to \{1, \dots, k\}$ (every pre-image has cardinality at least two) is

$$\sum_{i,j,l:i+j+l=k} \binom{n}{j} \frac{k!}{i!l!} (-1)^{i+j} k^{n-j}.$$

3. Give a formula for the number w_n of words of length n on the alphabet $\{a_1, \dots, a_k\}$ such that each symbol appears at least twice.