Math 249 Notes

Haydn Gwyn, Naveen Durvasula

Spring 2020

1 Organizing numbers

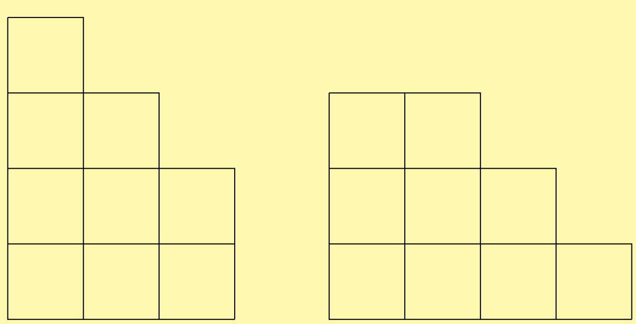
Miscellaneous notation:

- $(n)_k$: the kth falling factorial
- $\binom{n}{k}$: n choose k. We define this for negative cases by letting $\binom{n}{k} = \frac{(n)_k}{k!}$.
- $\left\langle {n \atop k} \right\rangle$: # of k-multisets from $[n]$.
- $S(n, k)$: # of partitions of a set $[n]$ into k blocks
- $p(n, k)$: # partitions of n into k positive parts. The parts $(\lambda_1 \geq \cdots \geq \lambda_k > 0)$ are written as a weakly decreasing sequence with $|\lambda| = \sum_i \lambda_i = n$.

Definition 1.1 (The Twelvefold Way)
We begin by thinking about how to count maps $[k] \mapsto [n]$.

	$D \to D$	$I \to D$	$D \to I$	$I \to I$
Any map	n^k	$\left\langle {n \atop k} \right\rangle$	$\sum_{m \leq n} S(k, m)$	$p_{\leq m}(k) := \sum_{m \leq n} p(k, m)$
Injection	$(n)_k$	$\binom{n}{k}$	1 if $k \leq n$ else 0	1 if $k \leq n$ else 0
Surjection	$k!S(k, n)$	$\left\langle {n \atop n-k} \right\rangle$	$S(k, n)$	$p(k, n)$

Definition 1.2 (Ferrers/Young diagrams)
The *Ferrers diagram* of a partition is a left- and bottom-justified arrangment of squares in which the number of squares in each row represents the parts of the partition.



The Ferrers diagram for the partition $(3, 3, 2, 1)$ and its transpose $(4, 3, 2)$

2 Ordinary generating functions

Suppose we wanted to count k element subsets A of an n element set. This is given by $\binom{n}{k}$. We introduce a new method of counting the same quantity. We do this by weighting A by the quantity $x^{|A|}$. The contribution for any individual element is given by $(1+x)$. as if the element is not in the set, 1 is contributed, else x is contributed. Thus, we consider the polynomial $(1+x)^n$. This is known as the *ordinary generating function* of the function $x \mapsto \binom{n}{x}$.

Example 2.1 (Multisets)
We have n fixed elements and we want to pick a multiset A. Again, we want to weight the set by $x^|A|$. As we have a multiset, we can choose how many times to include each element, encapsulating the number k of times that x_i is included with x_i^k. Since k can take on any non-negative integer value, we have the generating function

$$\prod_{i=1}^{n}(1+x_i+x_i^2+\cdots) = \prod_{i=1}^{n}\frac{1}{1-x_i} = \frac{1}{(1-x_1)(1-x_2)\cdots(1-x_n)}$$

where each term in the expansion corresponds to precisely one multiset. Collapsing the x_i to a single variable x, we find that the coefficients of the expansion of

$$\frac{1}{(1-x)^n} = (1-x)^{-n} = \sum_k \binom{-n}{k}(-x)^k = \sum_k (-1)^k \frac{(-n)_k}{k} x^k = \sum_k \frac{n(n+1)(n+2)\cdots(n+k-1)}{k!} x^k$$

represent the number of k-multisets with elements coming from $[n]$. We thus have that the number of k-multisets with elements coming from $[n]$ is

$$\left\langle {n \atop k} \right\rangle = \binom{n+k-1}{k}$$

The existence of the nice formula found in Example 2.1 suggests that the problem may yield to a combinatorial approach. Indeed, we can identify a k-multiset from $[n]$ with *weak compositions* of the number k. That is, ordered n-tuples $(a_1, a_2, \cdots, a_n)$ satisfying

$$a_1 + a_2 + \cdots + a_n = k$$

where $a_i \geq 0$. The partial sums of this sequence are between 0 and k inclusive, but they may not be distinct. If, however, we replace a_i with $a_i + 1$ for each i, the partial sums will be distinct and range from 1 to $n+k$ inclusive. There are clearly $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$ ways to choose these partial sums, which is the value we're looking for.

As it turns out, the kth rising factorial comes about when counting the number of *distributions* from a k element set into an n element set. These distributions are much like placing books onto a bookshelf, where we have k elements that we place into n indistinguishable shelves that maintain an ordering. Initially, there are n different shelves we can place the books into. However, once we place the first book, there are two spots that we may place the books as we have divided one shelf in two. Thus, the number of ways to place k books into the n shelves is $(n)^k$. If we make each of the k elements indistinguishable, we must divide by $k!$ which again gives us the multiset formula $\frac{(n)^k}{k!}$.

2.1 Formal power series

We don't want to have to worry about things like radii of convergence when dealing with polynomials. Series like $\sum n!x^n$ have combinatorial value as generating functions. We define a formal power series as simply a sum $\sum_n f(n)x^n$. Thus, we interpret these polynomials as simply a list of numbers. We may multiply formal power series as only finitely many terms may generate any given term in the product. We may divide formal power series in some cases. For example, if we have a power series $g := 1 + a_1x + a_2x^2 + \cdots$ that begins with a 1, we may take its multiplicative inverse as

$$\frac{1}{1+g} = 1 = g + g^2 - g^3 + \cdots$$

We'd like for all of our analytical results to hold for formal power series, and indeed they do, but we include a note on justifying this. The commonplace approach to dealing with this issue is to simply prove all of the results of single variable calculus with formal power series; this is not so hard, merely time-consuming, and so we accept the fact that the things we have proven analytically in fact hold formally.

2.2 Partitions and Stirling numbers

We won't get nice numbers for partitions $p(n, k)$, however, we can get nice numbers for its generating functions. Thus, we may easily compute $p(n, k)$. We now try to count all possible partitions λ by $x^{|\lambda|}t^{\ell(\lambda)}$ where $\ell(\lambda)$ is its length and $|\lambda|$ is the sum. When choosing how many parts have 1, we get the geometric series $1 + xt + \cdots$ yielding the sum $\frac{1}{1-tx}$. When choosing how many have 2, we get the geometric series $1 + x^2t + (x^2t)^2 + \cdots$ yielding the sum $\frac{1}{1-tx^2}$. Thus, we end up with the generating function

$$\prod_{i=1}^{\infty} \frac{1}{1-tx^i} = \sum_{n,k} p(n,k)x^nt^k$$

If we instead only care about the number of partitions independent of the length of the partition, we can change our weighting scheme to forget the length of the partition by collapsing t to 1. This gives the generating function

$$\prod_{i=1}^{\infty} \frac{1}{1-x^i} = \sum p(n)x^n = 1 + x + 2x^2 + 3x^3 + 5x^4 + \cdots$$

We can be even more flexible, and try to compute the number of odd partitions using the generating function

$$\prod_{\{i|i\equiv 1 \bmod 2\}} \frac{1}{1-x^i} = \sum p_{\text{odd}}(n)x^n = 1 + x + x^2 + 2x^3 + \cdots$$

We can be clever in yet another way, considering the number of partitions of n into distinct parts. The generating function for $p_d(n)$ is given by

$$\prod_{i=1}^{\infty}(1 + x^i) = 1 + x + x^2 + 2x^3 + \cdots$$

As it turns out, these two generating functions are equal. Thus, although not obvious combinatorially, the number of ways to partition a number into odd parts is equal to the number of ways to partition a number into distinct parts! We can show this nicely algebraically as

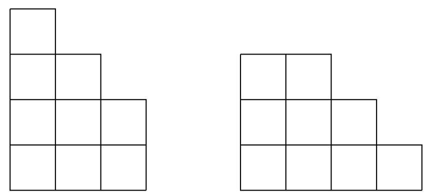
$$\prod_{\{i|i\equiv 1 \bmod 2\}} \frac{1}{1-x^i} = \frac{\prod_{\{i|i\equiv 0 \bmod 2\}} 1 - x^i}{\prod_i 1 - x^i} = \frac{\prod_i(1 - x^{2i})}{\prod_i(1 - x^i)} = \prod_i(1 + x^i)$$

which is precisely the generating function for the number of distinct partitions. Now we turn our attention to set of identities due to Rogers and Ramanujan which count the number of partitions with distinct, non consecutive parts.

Theorem 2.2 (Rogers-Ramanujan Identities)
We have

$$\prod_{i\equiv 1,4 \bmod 5}\frac{1}{1-x^i}=\sum_{m=0}^{\infty}\frac{x^{m^2}}{\prod_{i=1}^{m}(1-x_i)}$$
$$\prod_{i\equiv 2,3 \bmod 5}\frac{1}{1-x^i}=\sum_{m=0}^{\infty}\frac{x^{m^2+m}}{\prod_{i=1}^{m}(1-x_i)}$$

Recall that we may depict partitions using a *Ferrers diagram:*



Let Λ_n denote the set of partitions on n elements. For any $\lambda \in \Lambda_n$, we define the transpose operation ${}^* : \Lambda_n \mapsto \Lambda_n$. For any partition λ, we define λ^* as the partition given by geometrically "transposing" the Young diagram. Algebraically, this is given by letting the mth part of the transpose be equal to the number of parts of λ with length less than or equal to m. While algebraically, it is not necessarily intuitive that taking the transpose twice yields the original object, geometrically this is obvious. We can write down a generating function for the number of partitions with parts all at most m rather simply:

$$\sum p_{\le m}(n)x^n=\prod_{i=1}^{m}\frac{1}{1-x^i}$$

Seeing as the transpose operation bijects partitions with parts of size at most m and partitions with at most m parts, we can observe immediately that this generating function also encapsulates the number of partitions with at most m parts. It follows that the generating function for partitions of length precisely m is

$$x^m\prod_{i=1}^{m}\frac{1}{1-x^i}$$

(this follows by incrementing each part of a partition with at most m parts). We are led finally to the identity

$$\prod_{i=1}^{\infty}\frac{1}{1-x^it}=\sum_{m=0}^{\infty}\frac{t^mx^m}{\prod_{i=1}^{m}(1-x^i)}$$

On the left hand side, we have clearly the generating function

$$\sum_{\lambda}x^{|\lambda|}t^{\ell(\lambda)}$$

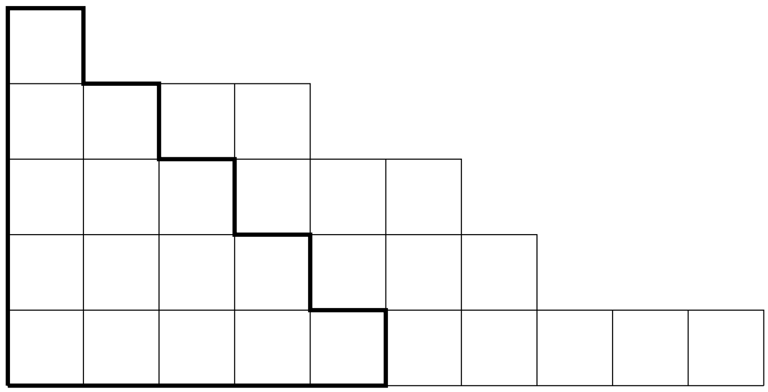
and on the right hand side, we've merely partitioned (so to speak) by the length of our partitions. Let's try to repeat this process for partitions with distinct parts. Geometrically, we can see that for partitions with distinct parts, we must have a strictly decreasing "staircase" structure which we call δ_m, and pasted to its right is another *weakly* decreasing partition of at most m parts μ. That is to say, to any λ with distinct parts with $\ell(\lambda)=m$ we can associate a general partition μ with $|\mu|=|\lambda|-\binom{m+1}{2}$. It follows that the generating function for partitions of length m with distinct parts can be written in the form

$$x^{\binom{m+1}{2}}\prod_{i=1}^{m}\frac{1}{1-x^i}$$

Putting this together, we get the partition identity

$$\prod_{i\equiv 1 \bmod 2} \frac{1}{1-x^i} = \prod_{i=1}^{\infty} 1+x^i = \sum_{m=0}^{\infty} x^{\binom{m+1}{2}} \prod_{i=1}^{m} \frac{1}{1-x^i}$$

We can view this identity as a simpler form of the Rogers-Ramanujan identities - a "zeroth Rogers-Ramanujan identity."



A partition with 5 distinct parts and a labelling of μ_5.

Returning to the Rogers-Ramanujan identities, we consider what the set of length n partitions with parts congruent to 1 or 4 mod 5 is. We can see that the m^2 term has combinatorial significance as a Young diagram with m^2 blocks is a staircase which skips every other step. Thus, the first identity is given by the set of distinct and non-consecutive partitions. Similarly, the second identity has can be decomposed into a staircase with distinct, non-consecutive parts > 1. We've now covered binomial coefficients, multichoose coefficients, and partition numbers (to a certain extent). The remaining set of numbers that appears in the twelvefold way table is the Stirling numbers of the second kind. We define

$$S(n,k) := \#\{\text{partitions of } [n] \text{ into } k \text{ disjoint nonempty blocks}\}$$

We can quickly establish a recurrence relation on the Stirling numbers by considering the block containing the number n. If n is in a block on its own, we can eliminate the block to obtain a partition of $[n-1]$ into $k-1$ parts. Otherwise, n is not in a block on its own, so we simply remove n from its block to yield of partition of $[n-1]$ into k parts. That being said, there are k partitions of $[n]$ that yield the same k-partition of $[n-1]$ when n is removed (as there are k blocks from which n could have been removed). So we end up with the recurrence

$$S(n,k) = kS(n-1,k) + S(n-1,k-1)$$

Now, note that one can find the total number of set partitions of $[n]$ simply by summing $S(n,k)$ over k. In particular, we have

$$B_n = \sum_k S(n,k)$$

where we use B_n because these numbers are called the *Bell numbers*. We show the following generating function for the Stirling numbers when we fix a given k.

$$\sum_n S(n,k)x^n = \frac{x^k}{\prod_{j=1}^{k}(1-jx)}$$

Rather than generating a recurrence, we aim to show this more combinatorially. Consider the following coding scheme for a partition: upon adding a new element i, we either append $*$ to the code if we create a new block, or we append the number j corresponding to the block that we added the new element. Once a new block is created, we number the block with the element that first instantiated the block. Clearly, the set of codes and the set of partitions is in bijection. We consider restraints on the set of codes.

- The number of $*$ elements is equal to k
- The first element must be a $*$, and each subsequent element may be either a $*$ or a number less than or equal to the number of $*$s observed thus far.

Thus, a code appears as

$$*\{1\}^* * \{1,2\}^* * \{1,2,3\}^* \cdots$$

This appears as k different Cartesian product terms, one after each $*$. We consider each Cartesian product term. Each $*$ contributes a weight of x. There is therefore $j^r x^{r+1}$ associated with $\{1 \cdots j\}^r$, yielding $\frac{x}{1-jx}$ weight for the set $\{1 \cdots j\}^*$. Putting the Cartesian products together, we get the desired result $\prod_{j=1}^k \frac{x}{1-jx}$.

In homage to undergraduate combinatorics, we observe the following:

$$f_k(x) := \sum_n S(n,k)x^n = \sum_n kS(n-1,k)x^n + S(n-1,k-1)x^n = kxf_k(x) + xf_{k-1}(x)$$

and it follows that

$$f_k(x) = \frac{x}{1-kx} f_{k-1}(x)$$

from which the form of the generating function follows inductively.

Now, this ordinary generating function has a certain novelty, but there turns out to be a nicer approach. Namely, we have

$$\sum_n S(n,k)\frac{x^n}{n!} = \frac{(e^x-1)^k}{k!}$$

What falls quite easily out of this generating function is the generating function for the Bell numbers, which we obtain by summing the above *exponential generating function* over all k:

$$\sum_n B_n x^n = e^{e^x-1}$$

We will develop the machinery for proving that these generating functions are correct later. For now, we turn our attention to yet another "generating function":

$$\sum_k S(n,k)(x)_k$$

as it turns out, this generating function for the Stirling numbers is equal to x^n. To see this, we rewrite the left hand side as

$$\sum_k S(n,k)k!\binom{x}{k}$$

Now we can note that this counts the number of mappings from $[n]$ to $[x]$. This is because we can choose k elements to be in the image of our mapping for any k (giving $\binom{x}{k}$), then we choose an unordered partition of $[n]$ (giving $S(n,k)$), and then we order the blocks of our partition. It is clear the x^n is the number of mappings from $[n]$ to $[x]$ as well, which proves the identity. The numbers $S(n,k)$ are referred to as the Stirling numbers of the 2nd kind. We may also define *Stirling numbers of the* 1*st kind* $s(n,k)$ as

$$\sum_k s(n,k)x^k := (x)_n$$

There is a matrix interpretation of the relationship of the two kinds of Stirling numbers. The matrix with nk entries given by $S(n,k)$ is inverse to the matrix with nk entries given by $s(n,k)$. This follows from combining the two identities above to obtain

$$\sum_k S(n,k)\sum_j s(k,j)x^j = x^n$$

and taking coefficients of x^n on both sides gives the dot product of a row of one matrix with a column of the other. Now what is the significance of these Stirling numbers of the first kind? We can observe rather obviously that

$$\sum(-1)^k s(n,k)x^k = -x(-x-1)\cdots(-x-n+1)$$

and hence

$$\sum(-1)^{n-k} s(n,k)x^k = x(x+1)\cdots(x+n-1)$$

and it is clear that all coefficients in these series are positive. Hence, we have written down the generating function for $c(n,k) = |s(n,k)| = (-1)^{n-k}s(n,k)$. We call these number the *signless Stirling numbers of the first kind*. We now show that in fact $c(n,k)$ counts the number of permutations in S_n that have k cycles. We use a similar coding scheme as before. To construct a given permutation, we can add elements in one by one, either inserting the element into its own cycle or inserting it into an existing cycle. If we insert it into its own cycle, we have increased the number of cycles, and hence need to increase the exponent on x by one. Otherwise, we have k choices for where to insert the next element, where k is the number of elements already inserted. We thus multiply by $(x+k)$ when we insert the $k+1$st element, which leads us to the product

$$\sum c(n,k)x^k = (x+1)(x+2)\cdots(x+n-1)$$

2.2.1 Stirling numbers as an algebraic tool

We can in fact use Stirling numbers as a means to compute certain sums. Consider the sum

$$\sum_{i=0}^{N} i^k$$

for some fixed k. We can note by taking finite differences that this will be a degree-$(k+1)$ polynomial in N, but it is not so easy to compute this explicitly. This is what we set out to do. We exploit the fact that $(x)_k$ is easier to sum than x^k. In particular, we may observe

$$\sum_{i=0}^{N}(i)_k = k!\sum_{i=0}^{N}\binom{i}{k}$$

The right summand, by the hockey stick identity (or by seeing the sum as a partition of the $(k+1)$-subsets of $[N+1]$ by their largest element) is equal to $\binom{N+1}{k+1}$. Hence, we have

$$\sum_{i=0}^{N}(i)_k = k!\binom{N+1}{k+1} = \frac{(N+1)_{k+1}}{k+1}$$

Now we have nice expressions for sums of falling factorials, which comprise a basis of the set of polynomials. We now aim to resolve our original sum:

$$\sum_{i=0}^{N} i^k = \sum_{i=0}^{N}\sum_{m} S(k,m)(i)_m = \sum_{m} S(k,m)\sum_{i=0}^{N}(x)_m = \sum_{m} S(k,m)\frac{(N+1)_{m+1}}{m+1}$$

3 Combinatorial Species

We look at combinatorial structures we can place on a set of n elements that remain invariant under a "relabeling" or isomorphism of the set to itself. We give a few examples before giving the general definition

- Partitions of a set A into k blocks (for a specified k) – there are $S(n,k)$ of these, and relabeling the elements does not change the number of these
- Weighted partitions on a set A, counted by $t^{\#\text{blocks}}$ – the generating function $\sum S(n,k)t^k$ describes this

- Linear orderings (bijections $f : [n] \xrightarrow{\simeq} A$) of a set A – there are $n!$ of these
- Bijections from A to A: there are $n!$ of these
  - Despite the fact that there is an obvious bijection between the set of linear orderings of A and the set of bijections from A to itself, these two sets in fact represent distinct combinatorial species, as any bijection between the two is necessarily label-dependent.
- Maps from A to $[k]$ (not necessarily bijections). The number of these mappings is k^n.
- Trees on the vertex set A. The number of these is n^{n-2}. We could also have *rooted* spanning trees, of which there are n^{n-1}.

It turns out that the most natural generating function to associate with these combinatorial objects is the exponential generating function. This is because we want to eventually divide by $n!$, as the map from the total set of objects to the isomorphism classes (by assumption) is $n!$ to 1. Allow us to revisit the examples from above and offer exponential generating functions for them (a couple of which we have already seen). Recall that

$$\sum_{n=0}^{\infty} S(n,k)\frac{x^n}{n!} = (e^x - 1)^k/k! \qquad \sum_{n,k} S(n,k)t^k \frac{x^n}{n!} = e^{(t(e^x-1)}$$

In fact, the generating function on the right encapsulates all generating functions of the form on the left. For linear orderings, we have

$$\sum n!\frac{x^n}{n!} = \sum x^n = \frac{1}{1-x}$$

which is also the generating function for permutations (as expected). Maps $A \to [k]$ are given by the generating function

$$\sum k^n \frac{x^n}{n!} = e^{kx} = (e^x)^k$$

For rooted spanning trees, we have $T_n = n^{n-1}$, and

$$T(x) = \sum T_n \frac{x^n}{n!}$$

satisfies

$$T(x) = xe^{T(x)}$$

from which it follows quite interestingly that $T(x)$ is the inverse of the function xe^{-x}. The coefficients of the latter function are easy to calculate, and the theory of calculating the coefficients of an inverse series given the coefficients of a series is well-studied; we will prove the Lagrange inversion formula using trees (combinatorically, *not* analytically) later.

We give an example of something that is not a species, but still has a nice exponential generating function. These are *zig-zag permuatations* of $[n]$. As we shall see, the ordering of these numbers affects the number of these permutations. The zig zag permutations are permutations a such that

$$a_1 < a_2 > a_3 < a_4 \cdots \{< \text{ if } n \text{ is even else } >\}a_n$$

Let z_n be the number of these permutations (these are called *Euler numbers*). Miraculously, it turns out that

$$\sum_{n=0}^{\infty} z_n \frac{x^n}{n!} = \sec(x) + \tan(x)$$

This is not a species, as it requires some structure (namely a total ordering) on the label set A. That is, if the set A were a set of fruits, we could not identity a permutation of them as zig-zag. Alternatively, we can observe that permuting the labels in a zig-zag permutation may not necessarily yield a zig-zag permutation, which again is sufficient for these objects not to be combinatorial species. We now define formally what a combinatorial species is.

Definition 3.1 (Combinatorial Species)
For any finite set of labels A, a *species* associates with it a finite set $F(A)$. The generating function coefficient $f_n := |F(A)|$ for any A with $|A| = n$. The requirement we have is that any bijective map $A \xrightarrow{\simeq} B$ induces a bijective map $F(A) \xrightarrow{\simeq} F(B)$. We also assert that if $g : A \to B$ and $h : B \to C$ are bijections, then $h \circ g$ is an eqivalent bijection from A to C. We call F, in this spirit, a *functor* from the category of finite sets with morphisms as bijections to itself.

We can now define the *equivalence* of combinatorial species in terms of the isomorphism of functors in the category of finite sets. Combinatorially, isomorphic species are exactly equivalent. Previously, we discussed the relationship between sums and products of generating functions and disjoint unions and Cartesian products of sets respectively. We elaborate this more formally using the language of species. We define the sum of two species F and G

$$(F + G)(A) = F(A) \coprod G(A)$$

Note that category theoretically, this corresponds to taking the coproduct. Note that $f_n + g_n = (f + g)_n$ for "added" species F and G. We define the generating function

$$(F + G)(x) := \sum |(F + G)[n]| \frac{x^n}{n!} = F(x) + G(x)$$

Taking the product of species is not as simple, as taking the product term by term of the formal power series does not correspond to taking the product of the two functions. The purpose of the factorials in the denominators of the terms of exponential generating functions will become a bit clearer now, as we see that

$$\left(\sum_k f_k \frac{x^k}{k!} \right) \left(\sum_k g_k \frac{x^k}{k!} \right) = \sum_k \sum_{j=0}^{k} \binom{k}{j} f_j g_{k-j} \frac{x^n}{n!}$$

We can now formally define the product species

$$(F \cdot G)(A) := \coprod_{A = A_1 \coprod A_2} F(A_1) \times F(A_2)$$

Useful in studying species are the family of *indicator species*. We define

- $\mathbf{1}(x) = 1$ – one structure on $A = \emptyset$ and no others
- $X(x) = x$ – one structure on $A \iff |A| = 1$
- $E(x) = e^x$ – one structure on any A
- $(E - 1)(x) = e^x - 1$ – one structure on $A \neq \emptyset$

We can now apply the above to study linear orderings. We can observe (in terms of species):

$$L = \mathbf{1} + X \cdot L$$

and it follows immediately that the generating function for L is $\frac{1}{1-x}$, which tells us that there are $n!$ linear orderings on an n-set. Let us consider another example. Let $F_k(A)$ denote the set of partitions of the set A into k ordered non-empty blocks (note the distinction from regular partitions, as we have given an ordering to the blocks). Note that this is equivalent to the set of surjections from A onto $[k]$. Recalling that $E - 1$ is the indicator species with precisely one structure on every nonempty set, we may write

$$F_k(x) = (e^x - 1)^k$$

where $F_k(x)$ is the generating function of $F_k(A)$. If we then remove the requirement that the blocks be ordered, we find that the generating function for the Stirling numbers of the second kind is

$$\sum_n S(n, k) \frac{x^n}{n!} = \frac{(e^x - 1)^k}{k!}$$

From this, we can find an exact formula for the Stirling numbers of the second kind

$$\begin{aligned}
S(n,k) &= \frac{n!}{k!}[x^n](e^x-1)^k \\
&= \frac{n!}{k!}[x^n]\sum_{i=0}^{k}\binom{k}{i}(-1)^{k-i}e^{ix} \\
&= \frac{n!}{k!}\sum_{i=0}^{k}\binom{k}{i}(-1)^{k-i}\frac{i^n}{n!} \\
&= \sum_{i=0}^{k}\frac{(-1)^{k-i}i^n}{i!(k-i)!}
\end{aligned}$$

Analogously, we could have recovered the same formula by using the inclusion-exclusion principle. We consider another example. Let $B(A)$ denote the set of partitions of the set A. We can consider B as a sum of species

$$B = \sum_{k=0}^{\infty} S_k$$

This yields as the generating function

$$B(x) = e^{e^x-1}$$

If we let $B_t(A) = B(A)$, but also counting a partition π with $t^{|\pi|}$. We have that the generating function

$$B_t(x) = \sum_{n,k} S(n,k)t^k\frac{x^n}{n!} = e^{t(e^x-1)}$$

We've now seen how product species can be used to resolve previously quite annoying problems. What about species *composition*? We consider this from the vantage point of generating functions, and we claim that $F \circ G$ only "converges" (i.e. has finite coefficients) if F is a polynomial or G has no constant term. The former restriction is clearly less convenient for us, so we turn our attention to formal power series with $G(0) = 0$ (that is, G has no structures on the empty set).

Definition 3.2 (Composite Species)
Formally, let F and G be species, and $G(\emptyset) = \emptyset$. We define $(F \circ G)(A)$ to be the species given by first partitioning A into a set of unordered (nonempty) blocks, and imposing a G structure on each of the blocks. We then place an F structure on the set of blocks. Clearly, this object commutes functorially under bijections from the labels to themselves. Thus, we have

$$(F \circ G)(A) = \coprod_{\pi \text{ partition of } A} F(\pi) \times \prod_{B\in\pi} G(B)$$

We aim to verify that the generating function

$$(F \circ G)(x) = F(G(x))$$

First, let $G_k(A)$ be the structure given by partitioning A into k unordered blocks and placing a G structure on each. We have

$$G_k(A) = \{\pi \vdash A, \ell(\pi) = k\}$$

We use the notation $\pi \vdash A$ to indicate that π partitions A. Observing that the set of *ordered* partitions of A is equivalent to $G(A)^k$, we determine that the exponential generating function for G_k is $G(x)^k/k!$. As $(F \circ G)$ is simply the disjoint union over these structures, we have that

$$(F \circ G)(x) = \sum f_k\frac{G(x)^k}{k!} = F(x) \circ G(x)$$

We can now reinterpret the work we previously did with counting partitions (using product species). In particular, we see that $B(x) = e^{e^x-1}$, the generating function for the Bell numbers, is reminiscent of the generating function for $E \circ (E-1)$, and naturally we may consider a partition as a set of nonempty subsets of A. Another example of a composite partition includes the set of *double partitions* $Q(A)$, in which we first partition the set, then partition the partitions. Thus, the set of double partitions is given by the species

$$Q(A) = E \circ (E \circ (E-1) - 1)$$

yielding the generating function

$$Q(x) = e^{e^{e^x-1}-1}$$

The circular orderings $C(A)$ also form a composite partition. These are given by the equivalence class linear orderings when the starting element in the ordering is forgotten. Thus, we have that $|C[n]| = (n-1)!$. A permutation on the set A can be decomposed into cycles, or circular orderings. Thus, we have that

$$P = E \circ C$$

so its generating function

$$P(x) = e^{C(x)}$$

As we have already found that

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

we can derive that

$$C(x) = \log \frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{x^n}{n} = -\log(1-x)$$

Now rather than simply considering permutations of A, we consider weighted permutations. In particular, we weight by the number of cycles of a permutation and have

$$P_t(A) = \sum_{w \in P(A)} t^{\#\text{cycles}(w)} = \sum_k c(n,k) t^k$$

Where $c(n,k)$ is the number of permutations of A with k cycles for $|A| = n$. We can see rather simply (referring to the above) that

$$P_t(x) = e^{tC(x)} = e^{-t\log(1-x)} = (1-x)^{-t}$$

Recalling that

$$P_t(x) = \sum_{n,k} c(n,k) t^k x^n / n!$$

we can see using the binomial theorem that

$$c(n,k) = t(t+1)\cdots(t+n-1)$$

We can use a similar technique to count derangements more efficiently. In particular, derangements comprise a species (note that $D(\sigma(A)) = D(A)$ for a permutation σ). We can note that

$$D(A) = E \circ (C - X)$$

where we have used that any derangement has a cycle decomposition with no one-element cycles. This leads to

$$D(x) = e^{-\log(1-x)-x} = \frac{e^{-x}}{1-x} = e^{-x}(1 + x + x^2 + \cdots)$$

The coefficient of x^n in this series is rather simply seen to be

$$n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \approx \frac{n!}{e}$$

where the approximation is for n large. As it turns out, this is the same formula one would get by using inclusion-exclusion. We now try to count permutations by their specific cycle structure. We construct a mixed generating function that keeps track of the number of cycles of each length. Formally, we count permutations $\omega \in S_n$ by the generating function

$$p_1^{\#1-\text{cycles}} p_2^{\#2-\text{cycles}} p_3^{\#3-\text{cycles}} \cdots p_k^{\#k-\text{cycles}}$$

For example, when $n = 3$, we have the ordinary generating function

$$p_1^3 + 3p_1p_2 + 2p_3$$

for permutations on 3 letters. We are interested in studying the generating function

$$\sum_{n,\omega \in S_n} p_{\tau(\omega)} \frac{x^n}{n!}$$

We can do this quite efficiently by repeating what we did before, but replacing C with $C_{\mathbf{p}}$ where

$$C_{\mathbf{p}}(p_1, \cdots; x) = \sum_{k>0} p_k \frac{x^k}{k!}$$

Thus, the generating function of the composition

$$\begin{aligned}
(E \circ C_{\mathbf{p}})(x) &= e^{C_{\mathbf{p}}(\mathbf{p};x)} \\
&= e^{\sum_{k>0} p_k \frac{x^k}{k}} \\
&= \prod_{k>0} e^{p_k \frac{x^k}{k}} \\
&= \prod_{k>0} \sum_{r=0}^{\infty} \frac{p_k^r x^{rk}}{k^j r!}
\end{aligned}$$

With this function being written down, we would like to determine the number of permutations of a given cycle type λ. This is nothing but the coefficient of $p_\lambda \frac{x^n}{n!}$ in the series above. This is

$$\frac{n!}{\prod_k k^{r_k} r_k!}$$

The product in the denominator of this quantity is often notated as z_λ, and we will encounter it again later. Another way to derive this same formula is through a group theoretic argument. If permutations have the same cycle type, they belong to the same conjugacy class. We may identify cycle types with conjugacy classes and observe that the number of permutations with a given cycle type is the size of the conjugacy class corresponding to it. Seeing as the stabilizer of one of these conjugacy classes is nothing but the centralizer of any of its elements (by definition), we want the centralizer of a permutation in a given conjugacy class $Z(\omega)$ to have z_λ elements, where $\lambda := \lambda(\omega)$. This centralizer can be described by the semi-direct product of all permutations on the cycle decomposition of the permutation and the cyclic group on each cycle (as we can move around the elements within a given cycle, or we can swap any pair of cycles - transpositions generate the symmetric group on the cycles).

We look at another example. Let $T(A)$ denote the species of all rooted trees on a vertex set A. Its generating function is given by

$$T(x) = x + 2\frac{x^2}{2!} + 9\frac{x^3}{3!} + \cdots$$

We eventually see that letting $T(x) = \sum t_n \frac{x^n}{n!}$, we get that $t_n = n^{n-1}$. We may construct a rooted tree by first defining the root, then partitioning the remaining elements and constructing a rooted tree on each partition. This gives us the species identity

$$T = X \cdot (E \circ T)$$

thus giving the generating function identity $T(x) = xe^{T(x)}$ that we claimed was easy to derive earlier. With this identity established, we aim to compute an explicit formula for the number of rooted trees on n vertices. Recall that we may rewrite this as

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

Thus, $T(x)$ is the inverse function of xe^{-x}. We could use the Lagrange inversion formula to determine the coefficients of the expansion of the inverse of xe^{-x}, but we opt instead to offer a combinatorial tree-based argument for the Lagrange inversion formula. If we are to do this, we must offer a species-based approach to compute t_n.

In terms of intuition, it helps that we already know (think) that $t_n = n^{n-1}$. Consider the species $M(A) = \{f : A \mapsto A\}$ consisting of all maps from A to itself. Not requiring injectivity and surjectivity, we have that $|M([n])| = n^n$, which is very close to what we "think" t_n is. If we can show that $xT'(x)+1 = M(x)$, we can show that indeed $t_n = n^{n-1}$. We can depict any element of $M(A)$ as a directed graph on the elements of A. For any element of A, we can follow outgoing edges (there is one outgoing edge from each vertex) until we run into a cycle. Clearly, every vertex leads into a unique cycle, and hence we may partition elements of A by the cycle to which they lead. Further, considering the elements of A that lead to a given element $r \in A$ that is in a cycle, we find that said elements comprise a tree rooted at r. Hence, any map $A \to A$ can be represented as a collection of rooted trees whose roots are grouped into cycles. Notably, a group of cycles is nothing more than a permutation of the roots of the trees (it is the cycle decomposition), and hence we may write

$$M = P \circ T$$

Thus, we have that $M(x) = \frac{1}{1-T(x)}$, as the exponential generating function for permutations is $P(x) = \frac{1}{1-x}$. Let's try to arrive at what we want to show by first differentiating T. We have that $T = xe^T$. Thus,

$$\begin{aligned}
T' &= e^T + xT'e^T = e^T + TT' \\
xT' &= xe^T + xTT' = T + xTT' \\
x(1-T)T' &= T \\
xT' &= \frac{T}{1-T} = \frac{1}{1-T} - 1 = M(x) - 1
\end{aligned}$$

as desired. We turn our attention now to the number of children of each vertex in a rooted tree.

Theorem 3.3 (Cayley's Formula)
If T is a tree on A, we define $c_T(a)$ to be the number of children of $a \in A$. Then we have

$$\sum_{T \in T([n])} \prod_{i \in [n]} x_i^{c_T(i)} = \left(\sum_{i \in [n]} x_i \right)^{n-1}$$

Proof. Notice on both sides we have polynomials of degree $n-1$. Therefore, when expanded, every term on either side is missing at least one variable. It is therefore sufficient to show that both sides are equivalent when an arbitrary x_i is omitted. As both sides are symmetric in i, we need only show this for a single case – we choose $x_n = 0$. If $x_n = 0$ on the right, it is easy to see that we get

$$(x_1 + \cdots + x_{n-1})^{n-1} = (x_1 + \cdots + x_{n-1})(x_1 + \cdots + x_{n-1})^{n-2} = (x_1 + \cdots + x_{n-1}) \sum_{T \in T([n-1])} \prod_{i \in [n-1]} x_i^{c_T(i)}$$

where the latter equality follows from inductive hypothesis. For the left hand side, the only nonzero terms after setting $x_n = 0$ are those that correspond to trees in which n is a leaf. These trees can be counted simply by considering the parent of n, which could be any node $1, \cdots, n-1$. It becomes quite evident then that

$$\sum_{T \in T([n])} \prod_{i \in [n]} x_i^{c_T(i)} \Big|_{x_n = 0} = (x_1 + \cdots + x_{n-1}) \sum_{T \in T([n-1])} \prod_{i \in [n]} x_i^{c_T(i)}$$

□

We now try to modify Cayley's formula so that we only keep track of the number of nodes that have k children. Thus, we write the modified generating function

$$\sum_{T \in T([n])} \prod a_k^{\#\text{ vertices with } k \text{ children}}$$

Comparing to the original form of Cayley's formula, we can see that terms of the form $x_1^{\mu_1} \cdots x_n^{\mu_n}$ contribute a term $a_{\mu_1} \cdots a_{\mu_n}$ to this modified generating function. Employing the original form of Cayley's formula and the multinomial theorem, we have

$$\begin{aligned}
\sum_{T \in T([n])} \prod a_k^{\#\text{ vertices with } k \text{ children}} &= \sum_{\mu_1 + \cdots + \mu_n = n-1} \binom{n-1}{\mu_1, \cdots, \mu_n} a_{\mu_1} \cdots a_{\mu_n} \\
&= (n-1)! \sum_{\mu_1 + \cdots + \mu_n = n-1} \frac{a_{\mu_1}}{\mu_1!} \cdots \frac{a_{\mu_n}}{\mu_n!} \\
&= (n-1)! [x^{n-1}] A(x)^n \\
&= \left[\frac{x^{n-1}}{(n-1)!} \right] A(x)^n
\end{aligned}$$

Where $A(x)$ is the exponential generating function with coefficients given by a_k. We call A a *generic species* given by placing one structure of weight a_k on each set of size k. If we let T_A be the species of rooted trees, weighted by $\prod a_k^{\#\text{ vertices with } k \text{ children}}$. We can write the species identity $T_A = X \cdot (A \circ T_A)$. We have that

$$T_A(\mathbf{a}; x) = \sum_n \left(\sum_{T \in T([n])} \prod a_k^{\#\text{ vertices with } k \text{ children}} \right) \frac{x^n}{n!}$$

Because we are working with a generic species, we are able to prove general statements about relations between exponential generating functions of various species. In particular, we find that

$$T_A(x) = X \cdot (A \circ T_A)(x) \implies \frac{T_A(X)}{A \circ T_A(X)} = x \implies \frac{x}{A(x)} \circ T_A(x) = x$$

Thus, we are computing the functional inverse of $\frac{x}{A(x)}$. Noting that the coefficients of $F \circ G$ will converge if G has no constant term, we can consider arbitrary formal power series $xB(x)$ and $x/A(x)$ which are functional inverses. We have that

$$\begin{aligned}
\left[\frac{x^n}{n!}\right] xB(x) &= \left[\frac{x^{n-1}}{(n-1)!}\right] A(x)^n \\
\left[\frac{x^{n+1}}{(n+1)!}\right] xB(x) &= \left[\frac{x^n}{n!}\right] A(x)^{n+1} \\
(n+1)! \, [x^n] \, B(x) &= n! \, [x^n] \, A(x)^{n+1} \\
[x^n] \, B(x) &= \frac{1}{n+1} \, [x^n] \, A(x)^{n+1}
\end{aligned}$$

The latter formula is known as the *Lagrange Inversion Formula*. Let's check this formula on our previous example of counting rooted trees. In this case, $A(x) = e^x$ and $xB(x) = T(x)$. Applying the Lagrange inversion formula gives that

$$[x^n]B(x) = \frac{1}{n+1}[x^n]e^{(n+1)x} = \frac{(n+1)^{n-1}}{n!} = \frac{(n+1)^n}{(n+1)!}$$

We have $[x^n]B(x) = [x^{n+1}]T(x)$, from which our desired formula follows.

4 Catalan Numbers

Definition 4.1 (Catalan Numbers)
The nth Catalan number

$$\begin{aligned} C_n &:= \#\text{balanced () strings of order } n \\ &:= \#\text{Dyck paths of order } n \end{aligned}$$

We can see that $C_3 = 5$ as we have the balanced strings

$$((())) \quad (()()) \quad ()(()) \quad (())() \quad ()()()$$

Corresponding to this parenthetical representation are *Dyck paths* which consist of lattice paths from $(0, n)$ to $(n, 0)$ that do not go above the diagonal. Clearly, these two sets are in bijection as we can construct a Dyck path from a balanced string by going down every time we see a left parenthesis and right whenever we see a right parenthesis.

We first see that the Catalan numbers are related to the number of unlabelled ordered rooted forests on n vertices, or equivalently, the number of unlabelled ordered rooted trees on $n+1$ vertices, which consists of trees that are further endowed with a linear ordering on the children of each vertex. We can associate balanced strings of parentheses with these trees recursively by letting the minimal balanced substrings (substrings whose substrings are not balanced) represent children of the root.

We can note some nice connections between Catalan numbers and various combinatorial objects by identifying bijections between balanced strings of parentheses and said objects. We claim first that the number of binary trees on $2n+1$ vertices with $n+1$ leaves equals the nth Catalan number. In particular, we may read each left parenthesis in a string as "create a left child of the current node" and each right parenthesis as "move up the tree until finding a node without a right child and add a right child." Starting from just a root node, this describes a bijection from strings of parentheses to the aforementioned set of binary trees. Another object counted by the Catalan numbers is triangulations of an $n+2$-gon.

We now aim to write the generating function for the nth Catalan number. We choose to write an ordinary generating function

$$C(x) = \sum_n C_n x^n$$

We can see rather simply that any balanced string of parentheses consists of minimally balanced substrings which take the form

$$(\langle \text{balanced string of parentheses} \rangle)$$

The generating function for these is nothing other than $xC(x)$, and because we can have an arbitrary number of these, we may conclude that

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

Alternatively, we can note that any balanced string of parentheses has a first minimally balanced substring as described above and then has a "remainder" which is another general balanced string of parentheses. This yields

$$C(x) = 1 + x(C(x))^2$$

where the 1 comes from the empty string (which is especially necessary because the remainder could be empty). This latter equation is quite clearly equivalent to the former, and either way, we can derive

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

There is also a species way to define this, and this method gives motivation for why we use an ordinary generating function for the Catalan numbers. Let $T_L(A)$ denote the species of *ordered* labelled rooted trees on A. Interestingly, due to the fact that we have ordered the trees, each specific tree has trivial automorphism group. Thus, there are $n!$ labelled rooted trees for each unlabelled rooted tree. Thus, since $T_L(A)$ will have a nice exponential generating function, C will have a nice ordinary generating function. Since we can construct the set of ordered labeled trees inductively by first choosing a root, then imposing a linear ordering on the children, and subsequently constructing an ordered labelled rooted tree on each child, we get the species identity $T_L = X \cdot (L \circ T_L)$.

We can observe now how an explicit formula for the Catalan numbers falls out of this ordinary generating function by using the binomial theorem. We have

$$\begin{aligned}
(1-4x)^{1/2} &= \sum_n \binom{1/2}{n}(-4)^n x^n \\
&= \sum_n \frac{\frac{1}{2}\cdot\frac{-1}{2}\cdot\frac{-3}{2}\cdots\frac{-(2n-3)}{2}}{n!}(-4)^n x^n \\
&= -\sum_n 2^n \frac{1\cdot 3\cdot 5\cdot\cdots\cdot(2n-3)}{n!}x^n \\
&= -\sum_n 2^n \frac{(2n-2)!}{2^{n-1}(n-1)!n!}x^n \\
&= -\sum_n \frac{(2n-2)!}{2n!(n-1)!}x^n \\
&= -\sum_n 2\frac{\binom{2n-2}{n-1}}{n}x^n
\end{aligned}$$

This leads to the explicit formula

$$C_n = \frac{\binom{2n}{n}}{n+1}$$

We can use this formula to compute a few Catalan numbers:

n	C_n
0	1
1	1
2	2
3	5
4	14
5	42

A more elegant way to arrive at this formula comes from applying the Lagrange Inversion formula. Recall that if $xB(x)$ and $x/A(x)$ are inverse functions, then

$$[x^n]B(x) = \frac{1}{n+1}[x^n]A(x)^{n+1}$$

We can derive from earlier equations that

$$xC(x) - x^2C(x)^2 = x$$

from which we conclude that $xC(x)$ is the inverse of the function $x - x^2$. Hence, we can set $B(x) = C(x)$ and $A(x) = \frac{1}{1-x}$ and apply the Lagrange inversion formula to yield

$$C_n = \frac{1}{n+1}[x^n]\frac{1}{(1-x)^{n+1}}$$

The series on the right side of the equation merely contains multinomial coefficients, and we have previously demonstrated n multichoose k to equal $\binom{n+k-1}{k}$, which yields

$$C_n = \frac{1}{n+1}[x^n]\frac{1}{(1-x)^{n+1}} = \frac{1}{n+1}\binom{n+1+n-1}{n} = \frac{1}{n+1}\binom{2n}{n}$$

These two methods for finding the Catalan numbers explicitly have been fun, but not as fun as playing "pin the tail on the Dyck path." Consider a Dyck path from the coordinate $(0,n)$ to $(n+1,0)$ where the final direction is necessarily going to the right. That is, we have taken a regular Dyck path from $(0,n)$ to $(n,0)$, and attached a "tail" pointing east. We can associate with these Dyck paths bit-string words w that have n 0s, $n+1$ 1s, and end in a 1. We first note that there are $\binom{2n}{n}$ total words corresponding to paths from $(0,n)$ to $(n+1,0)$, and $n+1$ words corresponding the same rotation class. We now show that only one of these corresponds to a Dyck path with a tail. Taking our Dyck path, we can extend it infinitely by duplication on either side. At the end of each horizontal "step" (i.e. a 1 in the binary word), we can place a mark. As n and $n+1$ are relatively prime, each period of marks corresponding to the same rotation lie on the same same line of slope $\frac{n}{n+1}$. Further, each marked point lies on its *own* line of this slope (as two lattice points cannot lie on a line of this slope without being at least $n+1$ away from each other on the x axis). Thus, there is a unique "highest" mark, and considering the line of slope $\frac{n}{n+1}$ that lies tangent to it, we have constructed a Dyck path with a tail, as all steps lie underneath it. Thus, we get our familiar formula for the Catalan numbers

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

We can frame the Lagrange inversion formula in relation to Dyck paths and Catalan numbers. In particular, we can view the Lagrange inversion formula as stating

$$b_n = \frac{1}{n+1}\sum a_{r_1}a_{r_2}\cdots a_{r_{n+1}}$$

where the sum on the right hand side is over non-negative $r_1,\cdots,r_{n+1}$ whose sum is n. We can rather cleverly identify the r_i as representing the lengths of the runs of zeros in the previously described bit strings (i.e. lengths of vertical segments in the associated paths). We can hence consider the prior sum as being a sum over Dyck paths, and for each Dyck path, we sum over all rotations the term $a_{r_1}a_{r_2}\cdots a_{r_{n+1}}$. Given

$$A(x) = a_0 + a_1x + \dots$$

We therefore have that

$$B(x) = a_0 + a_0a_1x + (a_0^2a_2 + a_0a_1^2)x^2 + (a_0^3a_3 + 3a_0^2a_1a_2 + a_0a_1^3)x^3$$

where we get the coefficient terms by counting the tailed Dyck paths from $(0,n)$ to $(n+1,0)$. We divide by $n+1$ automatically here as only one out of $n+1$ rotations is a Dyck path. A common idea once we are able to count something is to establish a *q-analogue* for it. For example, we can consider the q-analogue of function composition

$$\sum f_ng(x)^n \Longrightarrow \sum f_ng(x)g(qx)\cdots g(q^{n-1}x)$$

In the case of the above formulation of the Lagrange inversion formula, we can consider injecting powers of q with exponents equal to the area between a given Dyck path and the "highest" Dyck path - namely, a staircase. We can perform a more specific analysis by viewing the Catalan numbers as a sum over all Dyck paths of a certain length of 1 (i.e. each summand is 1). The modification we make is that we change each summand to be $q^{|\delta\backslash\lambda|}$, which is q raised to the power of the area between the Dyck path and the highest Dyck path. This would yield a q-analogue for the Catalan numbers:

$$C_n(q) = \sum_\lambda q^{|\delta\backslash\lambda|}$$

We could alternatively employ the formulas

$$(k)_q = 1 + q + \cdots + q^{k-1} \qquad (k)_q! = (1)_q\cdots(k)_q$$

and then defining

$$C_n(q)' = \frac{1}{(n+1)_q}\binom{2n}{n}_q$$

either approach yields a viable q-analogue to the Catalan numbers, and the two q-analogues are both in common use.

5 The Cycle Index

We aim now to introduce the cycle index of a combinatorial species. Consider a species F; we've already established our ability to associate a function $F(x)$ to F. We are interested in observing how permuting the elements of a set A changes the structures that F associates to A. In particular, we'd like to see which structures of F are fixed by a given element $w \in \mathrm{Sym}(A)$. Recall the cycle type of a permutation w - a decreasing tuple containing the sizes of the cycles comprising w - as well as the cycle indicator of w - the monomial in $p_1, p_2, \cdots$, where the exponent of p_i is the number of cycles of length i in w.

Definition 5.1
The *cycle index* of a combinatorial species F is given by

$$Z_F(p_1, p_2, \cdots) = \sum_n \frac{1}{n!} \sum_{w \in \mathrm{Sym}(n)} |F([n])^w| p_{\tau(w)}$$

where $|F([n])^w|$ is the number of elements of $F([n])$ fixed by w and $p_{\tau(w)}$ is the cycle indicator of w.

The cycle index Z_F contains strictly more information than the exponential generating function, as setting $p_1 = x$ yields only one permutation: the identity. Thus, we are simply counting the number of structures on n elements, as all elements are fixed.

Theorem 5.2 (The lemma that is *not* Burnside's)
The number of orbits of a group G acting on a set X is given by

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

where $|X^g|$ denotes the set of points $x \in X$ fixed by g.

As a notational convention, we define

$$Z_F[x] := Z_F(x, x^2, x^3, \dots)$$

Thus, we have that

$$Z_F[x] = \sum_n \frac{1}{n!} \sum_{w \in S_n} |F([n])^w| x^n$$

and by *Burnside's* lemma, we see the inner term is precisely equal to the number of orbits of S_n on $F([n])$. In particular, we can observe that

$$\frac{1}{n!}\sum_{w \in S_n} |F([n])^w|$$

is the average number of F-structures on A fixed by an element of the symmetric group, which is hence the number of orbits of $F(A)$ under action by the symmetric group. But orbits under relabelling are really just unlabelled structures, so in fact we have that

$$Z_F[x] = \sum_n x^n \#\{\text{unlabelled } F\text{-structures}\}$$

that is, from the exponential generating function for labelled structures we have uncovered the ordinary generating function for unlabelled structures. Before we go into examples, we first note that just as operations on species corresponded to operations on generating functions, they also correspond to operations on the cycle indices of the species (showing that the product of cycle indices is the cycle index of a product species requires a bit of care, but it boils down to considering the product of the cycle indices term-by-term and decomposing into products of terms from the individual cycle indices). Further, we note that we can reexpress the prior sum as

$$\sum_n \frac{1}{n!} \sum_{a \in F([n])} \sum_{w \in \mathrm{Stab}(a)} p_{\tau(w)}$$

and the innermost sum can be considered as some special quantity relating to each element of $F([n])$.

As an example, we look at linear orderings. In a linear ordering, there are no automorphisms, as relabeling the ordering yields a different ordering. We can derive the cycle index of L in terms of the indicator X and 1 using the same identity

$$L = 1 + X \cdot L$$

We first analyze the indicator species. The cycle index Z_1 of 1 is trivially 1. Similarly, $Z_X = p_1$. Thus, we get that

$$Z_L = 1 + p_1 \cdot Z_L \implies Z_L = L(p_1) = \frac{1}{1 - p_1}$$

Next, we look at ordered rooted trees. Previously, we had the identity

$$T = X \cdot (L \circ T)$$

Although we haven't yet dealt with composition, we will see that this implies that

$$Z_T = p_1 \cdot \left(\frac{1}{1 - p_1} \right)$$

From this, we can derive that

$$Z_T = T(p_1) = p_1 C(p_1) = x C(x)$$

where $C(x)$ is the Catalan number generating function.

As another example, let's determine the cycle index for the trivial species E. Any permutation will fix the single structure on $[n]$, and we have

$$Z_E = \sum_n \frac{1}{n!} \sum_{w \in \mathrm{Sym}(n)} p_{\tau(w)}$$

We can be a bit clever and consider replacing p_i with $p_i x^i$, which would yield

$$Z_E(xp_1, x^2 p_2, \cdots) = \sum_n \sum_{w \in \mathrm{Sym}(n)} p_{\tau(w)} \frac{x^n}{n!}$$

Noting additionally that $P = E \circ C$, where C is the species for cycles, we can recall

$$C(x) = \sum \frac{x^k}{k} \qquad C(p_1, p_2, \cdots; x) = \sum p_k \frac{x^k}{k}$$

from which it follows that

$$\sum_n \sum_{w \in \mathrm{Sym}(n)} p_{\tau(w)} \frac{x^n}{n!} = \exp\left(\sum p_k \frac{x^k}{k} \right)$$

This function is of particular import, and hence we define

$$\Omega = \exp\left(\sum p_k \frac{x^k}{k} \right) = \sum_{n \geq 0} h_n$$

where h_n is the complete symmetric polynomial of degree n (consisting of the sum of all degree n monomials in $p_1, \cdots, p_n$). We check that when we plug in $p_1 = x$, and $p_2, \dots, p_n = 0$, we get the generating function e^x. If we set $p_1 = x$, we get that

$$Z_E(x, 0, 0, \dots) = \exp x = e^x$$

Similarly, we can compute $Z_E(x, x^2, \dots)$ as

$$Z_E(x, x^2, \dots) = \exp \sum_{k>0} \frac{x^k}{k} = \exp \log \frac{1}{1-x} = \frac{1}{1-x}$$

which makes sense, as this is the ordinary generating function, and by setting all $p_i = x$, we have multiplied by $n!$ in each sum term. Next we look at the species of permutations. We note that changing the labels is equivalent to conjugation. Thus, the set of points $P([n])^w$ is given by the set of g that do not change under permutation. That is, $P([n]) = Z(w)$, the centralizer of w. Thus, we get that

$$Z_P = \sum_n \sum_{w \in S_n} |Z(w)| p_{\tau(w)} = \sum_n \sum_{w \in S_n} z_{\tau(w)} p_{\tau(w)}$$

as

$$\frac{n!}{|Z(w)|} = |\text{conjugacy class of } w| = \frac{n!}{z_{\tau(w)}}$$

Furthermore, as

$$\frac{1}{n!} \sum_{w \in S_n} |Z(w)| p_{\tau(w)} = \sum_{|\lambda| = n} p_\lambda$$

we can solve for Z_P as

$$Z_P = \sum_{|\lambda|=n} p_\lambda = \prod_i \frac{1}{1-p_i}$$

We can verify that

$$Z_P(x, 0, 0, \dots) = \frac{1}{1-x}$$

Next, we can see that *unlabeled* permutations are equivalent to partitions, as

$$Z_P(x, x^2, x^3, \dots) = \prod_i \frac{1}{1-x^i}$$

which is the ordinary generating function for partitions.

Now, we can only get so far without understanding the relation between cycle indices and species composition.

Definition 5.3 (Plethystic Substitution)
We define the *Plethystic Substitution* on a function A in variables $a_1, a_2, \cdots$.

$$p_k[A] = A \Big|_{a \mapsto a^k}$$

As an example,

$$p_k \left[\sum x_i \right] = \sum x_i^k$$

We redefine our notation, letting

$$Z[A] := Z \Big|_{p_k \mapsto p_k[A]}$$

We see that this aligns with our previous definition, as

$$Z[x] = Z(x, x^2, x^3, \cdots)$$

as we replace p_k with $p_k[x] = x^k$.

Theorem 5.4
If F is a species and A the ordinary generating function for some weighted set X, then $Z_F[A]$ is the ordinary generating function for unlabelled F structures decorated by A. Formally, this constitutes mapping the elements of the set of F structures to elements in X, and weighting the set by the product of the weights of the corresponding decorations.

Recall that
$$Z_F = \sum_n \frac{1}{n!} \sum_{w \in S_n} |F([n])^w| p_{\tau(w)}$$
If our theorem is true, then this should give us the ordinary generating function for unlabelled F structures with each vertex weighted by x. We can see that this holds as
$$Z_F[x] = \sum_n \left(\frac{1}{n!} \sum_{w \in S_n} |F([n])^w| \right) x^n$$
which by Burnside's lemma reduces to the ordinary generating function unlabeled structures, as this corresponds to the orbit under action of the symmetric group. We let $F_Y(S)$ denote the set of Y-decorated F structures on S. That is,
$$F_Y(S) = F(S) \times Y^S$$
We have
$$Z_{F_Y}(p_1, p_2, \cdots ; a_1, a_2, \cdots) = \sum_n \frac{1}{n!} \sum_{w \in S_n} |F_Y([n])^w| p_{\tau(w)}$$
We say more formally, that some $(f, d) \in F_Y$ is *fixed* iff $f \in F([n])^w$ and d is constant on the cycles of w. To count the number of elements in the latter case, we simply take the p_ks denoting the length of each cycle and replace it with the plethystic evaluation of A, as each cycle of length k contributes y^k in weight. Thus, we have that
$$|F_Y([n])|^w = |F([n])^w| \cdot p_{\tau(w)}[A]$$
And in particular, we derive that
$$Z_{F_Y}[x] = Z_F[Ax]$$
which follows from
$$\sum_n \frac{1}{n!} |F_y([n])| p_{\tau(w)}[x] = \sum_n \frac{1}{n!} \sum_{w \in S_n} |F([n])^w| p_{\tau(w)}[A] p_{\tau(w)}[x] = \sum_n \frac{1}{n!} |F([n])^w| p_{\tau(w)}[Ax]$$

We are now ready to define the composition of cycle indices. We begin by defining
$$p_k * W := W \Big|_{p_\ell \mapsto p_{k\ell}}$$
from which we can define for cycle indices $Z(p_1, \ldots, p_n)$ and $W(p_1, \ldots, p_n)$,
$$Z * W := W \Big|_{p_k \mapsto p_k * w}$$
Thus, we have that
$$(Z * W)[A] = Z[W[A]]$$
As it turns out, the above property characterizes $Z * W$. That is, the plethystic evaluation operator $Z[A]$ of a function Z determines Z itself. This is because if $A = \sum_{i=1}^{\infty} x_i$, then the $p_k[A]$ are linearly independent.

Theorem 5.5
Let F and G be combinatorial species. Then
$$Z_{F \circ G} = Z_F * Z_G$$

Proof. It suffices to show that for any A, we have

$$Z_{F\circ G}[A] = (Z_F * Z_G)[A] = Z_F[Z_G[A]]$$

A is nothing but the weighted ordinary generating function for some set, so the left-hand side of the above equation refers to the set of unlabelled F structures on unlabelled G structures decorated by A. The right-hand side consists of unlabelled F structures decorated by unlabelled G structures that are decorated by A. It is relatively simple to see that these two things are the same. $\square$

Note that this same technique allows us to demonstrate in a new way our previous result

$$Z_{F\cdot G} = Z_F \cdot Z_G$$

We return to the trivial species

$$Z_E = \Omega = \exp \sum \frac{p_k}{k} = \sum_\lambda \frac{p_\lambda}{z_\lambda}$$

We note that the plethystic substitution

$$\Omega[A+B] = \Omega[A]\Omega[B]$$

As an example, consider

$$\Omega[x_1 + x_2 + \dots] = \prod \frac{1}{1-x_i} = \prod \Omega[x_i]$$

As another example, recall the Bell species $B(S)$ given by the set of partitions of S. We have that

$$B = E \circ (E-1)$$

from which we get that

$$Z_B = \Omega * (\Omega - 1)$$

We can use this to far more efficiently evaluate the ordinary generating function for integer partitions, as this is given precisely by *unlabeled* set partitions. We can solve

$$Z_B[x] = \Omega[\Omega[x]-1] = \Omega[x + x^2 + \dots] = \prod_{i=1}^{\infty} \frac{1}{1-x^i}$$

as before.

As yet another example, we consider *unlabelled* rooted trees. Previously, we counted ordered rooted trees, yielding the generating function for the Catalan numbers. Let t_n denote the number of unlabelled rooted trees. Let $T_u(x)$ denote the ordinary generating function for these unlabelled rooted trees. The first few terms are given by

$$T_u(x) = t_1 x + t_2 x^2 + \dots$$

We have the species identity

$$T = X \cdot (E \circ T)$$

Thus, we get that

$$Z_T = p_1 \cdot (\Omega * Z_T)$$

This is quite a powerful tool, but we will use it here to simply compute T_u using $T_u = Z_T[x]$. In particular, $p_1[x] = x$, and

$$\Omega * Z_T = \exp \sum_k p_k * Z_T[x]/k = \exp \sum_k p_k * Z_T[x^k]/k$$

which yields

$$Z_T[x] = x \exp \sum_{k>0} \frac{1}{k} Z_T[x^k]$$

While this is not a nice closed-form solution, it does allow us to compute these values by successive approximation.

We can apply our tools to study isomorphism classes of simple graphs with a given vertex set S. In particular, we will make a mixed generating function weighting graphs on S by q^e, where e is the number of edges of the graph. It's relatively simple to find that

$$F_G(x;q) = 1 + x + (1+q)x^2/2 + (1+q)^3x^3/3! + \cdots = \sum (1+q)^{\binom{n}{2}} \frac{x^n}{n!}$$

We aim then to find the cycle index $Z_G[x;q]$. This will allow us to count graphs up to isomorphism. That is, isomorphism classes of graphs.

6 Cycle indices for graphs

We now focus on computing the cycle indices for (simple) graphs. We denote by $G(S)$ the species consisting of graphs with vertex set S. We want to not only keep track of the number of graphs, but also the number of edges we have in the graph. Thus, we will weight our generating function with $q^{|E|}$. This gives us the mixed exponential generating function

$$F_G(x,q) = \sum (1+q)^{\binom{n}{2}} \frac{x^n}{n!}$$

We can similarly mix ordinary generating coefficients for q into the the cycle index Z_G, letting

$$Z_G(p_1,\ldots,q) = \sum_n \frac{1}{n!} \sum_{w\in S_n} |G([n])^w|_q p_{\tau(w)}$$

where $|G([n])^w|_q$ denotes the ordinary generating function weighted by q. As the number of permutations of a given cycle type is given by $\frac{n!}{z_\tau}$, we can rewrite the above as

$$\sum_n \sum_{|\tau|=n} \frac{1}{z_\tau} |G([n])^{w_\tau}|_q p_\tau$$

In the way of identifying the graphs fixed by a certain permutation, we consider two cycles, of lengths τ_i and τ_j, in a given permutation. Observe that if there is an edge between a vertex of cycle i and a vertex of cycle j in a certain graph, then in order for it to be preserved by the permutation, all edges in the orbit of this single edge must be included in the graph. This yields $\mathrm{lcm}(\tau_i,\tau_j)$ edges. Because there are $\tau_i\tau_j$ total pairs of vertices that we could choose from these two cycles, the number of possible cycles (i.e. orbits of edges under the permutation) is

$$\frac{\tau_i\tau_j}{\mathrm{lcm}(\tau_i\tau_j)} = \gcd(\tau_i,\tau_j)$$

This means that each pair of cycles in a given permutation contributes a factor of

$$(1+q^{\mathrm{lcm}(\tau_i,\tau_j)})^{\gcd(\tau_i,\tau_j)}$$

We could alternatively consider an edge within a single cycle i. The orbit of any particular edge under the action of the permutation is of size τ_i (if τ_i is even, then one orbit has size $\tau_i/2$), and the number of cycles (orbits) is $\lfloor \tau_i/2 \rfloor$. Thus, we get that for odd τ_i, the weighting is given by

$$(1+q^{\tau_i})^{(\tau_i-1)/2}$$

and for even τ_i we have a weighting

$$(1+q^{\tau_i})^{\tau_i/2-1}(1+q^{\tau_i/2})$$

Hence, we find that the cycle index $Z_G(p_1, p_2, \cdots; q)$ for graphs is

$$\sum_n \sum_{|\tau|=n} \frac{1}{z_\tau} p_\tau \cdot \left(\prod_{i<j} (1 + q^{\mathrm{lcm}(\tau_i, \tau_j)})^{\gcd(\tau_i, \tau_j)} \right) \left(\prod_{\tau_i \text{ odd}} (1 + q^{\tau_i})^{\frac{\tau_i - 1}{2}} \right) \left(\prod_{i \text{ even}} \left((1 + q^{\tau_i})^{\frac{\tau_i}{2} - 1} (1 + q^{\frac{\tau_i}{2}}) \right) \right)$$

We try to obtain a similar result for connected graphs. We note that as a graph may be decomposed into connected components, we have the species identity

$$G = E \circ G_{\mathrm{conn}}$$

This gives us immediately that

$$Z_G = \Omega * Z_{G_{\mathrm{conn}}}$$

We can similarly state that

$$Z_G(p; k) = \Omega * Z_{G_{\mathrm{conn}}}(p; q)$$

where the above plethysm is given by

$$p_k * f(p; q) = f(p_k, p_{2k}, \ldots; q^k)$$

Intuitively, it is better to think of the p_k as operators on power series. Now that we have written this identity, we can find the cycle index for connected graphs if we are somehow able to compute the "plethystic inverse" of Ω. We can write this problem more generally as solving for C, given B, in the equation

$$1 + B = \Omega * C$$

That is to say, can we find a Λ such that

$$\Lambda * B = C$$

We note that we can think of p_1 as a plethystic identity, as

$$p_1 * A = A = A * p_1$$

Thus, we aim to find Λ such that

$$\Lambda * (\Omega - 1) = p_1 = (\Omega - 1) * \Lambda$$

It is not immediately obvious that a right and left inverse exists for Ω, but by successive approximations, we can show that these exist. As they both exist, they must be the same. We have

$$\Gamma * \Lambda = 1 + p_1$$

so

$$\exp \sum_k \frac{1}{k} p_k * \Lambda = 1 + p_1$$

Now noting that $p_k * \Lambda = \Lambda(p_k, p_{2k}, \cdots)$, we find

$$\sum_k \frac{1}{k} p_k * \Lambda = p_1 - \_1^2/2 + \cdots = \sum (-1)^{m-1} \frac{p_1^m}{m}$$

To solve this equation for Λ, we attempt to find a candidate solution L for each term of the summation. We can begin by letting $L = p_1$. This gives us that

$$\sum \frac{1}{k} p_k * L = p_1 + \frac{p_2}{2} + \frac{p_3}{3} + \frac{p_4}{4} + \frac{p_5}{5} + \frac{p_6}{6}$$

We need to get rid of these extra terms, so let's first add a $-\frac{p_2}{2}$ term. This subtracts the even terms from the above sequence. By subtracting $-\frac{p_3}{3}$, we get rid of the multiples of 3 (although we've now subtracted 6 twice!). By subtracting $-\frac{p_5}{5}$, we get rid of the multiples of 5. We now have to add back a $\frac{p_6}{6}$. In general, we get that powers of primes go away, but products of distinct primes have to be added back to account for this. This brings us to a number-theoretic concept known as *Möbius inversion*. Formally, the *Möbius function*

$$\mu(n) := \begin{cases} (-1)^\ell & n = p_1 \dots p_\ell \\ 0 & \text{otherwise} \end{cases}$$

It is a well-known result that

$$\sum_{\ell | n} \mu(\ell) = \delta_{n,1}$$

where δ is the Kronecker delta function. To prove this, let $S(n)$ be the sum above over the divisors of n. Without loss of generality, assume $\mu(n) \neq 0$. Let p be a prime divisor of n. Then

$$\sum_{\ell | n} \mu(\ell) = \sum_{\ell | n/p} \mu(\ell) + \sum_{\ell | n/p} -\mu(\ell) = 0$$

where the first sum includes factors of n not divisible by p and the second sum includes factors of n divisible by p.

We can observe now that

$$L = \sum_{\ell=1}^{\infty} \mu(\ell) \frac{f(p_\ell)}{\ell}$$

whence it follows that

$$\sum \frac{1}{k} p_k * L = \sum_{k,\ell} \mu(\ell) \frac{f(p_{k\ell})}{k\ell} = f(p_1)$$

where $f(p_1) = \log(1 + p_1)$

7 Symmetric Functions

Definition 7.1 (Symmetric polynomial)
A *symmetric polynomial* over a (probably commutative) ring R in n variables is a polynomial $P \in R[X_1, X_2, \cdots, X_n]$ invariant under any R-isomorphism fixing $\{X_1, \cdots, X_n\}$.

Some examples of symmetric polynomials include

- The sum of the kth powers of variables

$$x_1^k + x_2^k + \cdots + x_n^k = p_k(x_1, \dots, x_n)$$

- The kth *elementary symmetric polynomial* in $x_1, \cdots, x_n$:

$$e_k = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \prod_{j=1}^{k} a_{i_j}$$

We may also consider symmetric polynomials in infinitely many variables. However, we must be careful in our definition, as if we enforce that the polynomial is invariant under the action of S_∞ on the variables, we must have infinitely many terms in the polynomial, since if x_1 is present, so must $x_2, \dots, x_\infty$. We therefore consider formal series. Let R be the ring of symmetric formal series in $x_1, x_2, \cdots$. What do we mean, in general, by a symmetric formal series? Vaguely, we call a series symmetric if it is invariant under action by S_∞ on $x_1, x_2, \cdots$, but we have two conceptions of S_∞. We have a "thick" S_∞, which consists of all bijections $\{X_1, X_2, \cdots\} \to \{X_1, X_2, \cdots\}$, and we have a "thin" S_∞ which contains all finite permutations (i.e. that fix all but finitely many variables). This thin S_∞ is generated by the set of transpositions. The thin S_∞ is therefore a subgroup of the thick S_∞ given by a direct limit of finite permutation groups. Consider a finite partition λ. We define

$$m_\lambda = x_1^{\lambda_1} \cdots x_\ell^{\lambda_\ell} + \text{ all similar terms}$$

where "all similar terms" refers to all *distinct* monomials in the orbit of the first term under action by S_∞. Observe that regardless of which conception of S_∞ we use, this notion will be the same. For example, we have

$$p_k = x_1^k + x_2^k + \cdots = m_{(k)}$$

and

$$e_k = \sum_{|I|=k} x_I = m_{(1,1,\cdots,1)}$$

where $x_I = \prod_{i \in I} x_i$. A "less trivial" example is

$$m_{(2,1)} = x_1x_2^2 + x_1^2x_2 + x_1x_3^2 + x_1^2x_3 + x_1x_4^2 + x_1^2x_4 + x_2x_3^2 + x_2^2x_3 + \cdots$$

where we include obscenely many terms to demonstrate the care we've taken to assure that the $\cdots$ in fact refers to all terms. It's quite evident at this point that

$$R = \{\text{(potentially infinite) linear combinations } \sum_\lambda c_\lambda m_\lambda\}$$

where c_λ is an element of the ground ring (from now on, we simply use $\mathbb{Z}$). It is important to note that the m_λ do not form a (Hamel) basis, as we can have infinite linear combinations. We consider the subring $\Lambda \subset R$ given by the set of $f \in R$ of bounded degree. It follows that the elements of Λ are finite linear combinations of m_λs. Thus, the set of m_λ form a basis of Λ (as a free module over $\mathbb{Z}$). We refer to the elements of Λ as *symmetric functions*. We can speak of *homogeneous symmetric functions of degree d* which are symmetric functions all of whose terms are of degree d. We then have

$$\Lambda = \bigoplus_{d \geq 0} \Lambda_d$$

where Λ_d is the set of homogeneous symmetric functions of degree d. It follows that Λ is graded. We have that $\{m_\lambda \mid |\lambda| = d\}$ forms a basis for Λ_d. There is a projection homomorphism from Λ (or equivalently $\Lambda_\mathbb{Z}$ to $\mathbb{Z}[X_1, \cdots, X_n]^{S_n}$ obtained by setting $X_{n+1}, X_{n+2}, \cdots$ to 0. We can therefore think of symmetric functions in n variables as reductions of symmetric functions in infinite variables.

Next, we consider a different basis for symmetric functions. We consider the e_k mentioned above

$$e_k := m_{(1^k)}$$

and further define

$$e_\lambda := \prod_{i=1}^{|\lambda|} e_{\lambda_i}$$

Consider any e_λ. We know the m_λ functions to be a basis of Λ, so we can determine the coefficients of the m_μ in the representation of e_λ. That is, we'd like to determine

$$[x^{\mu_1}x^{\mu_2}\cdots x^{\mu_\ell}]e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_m}$$

We can do this by filling in an $m \times \ell$ matrix whose ith row is an indicator vector for the product of x_i's we "take" from e_{λ_i}. In particular, the number of 1s in row i is simply λ_i, and the number of 1s in the ith column is simply μ_i. Any matrix of this form represents a single term (by which we mean a term with *coefficient 1*) in e_λ. Hence, the coefficient of m_μ in e_λ is the number of 0-1 matrices with row-sums given by λ and column sums given by μ. Note that matrices of this form are in bijection with 0-1 matrices with row-sums given by μ and column sums given by λ (under the transpose mapping), so we have

$$[m_\mu]e_\lambda = [m_\lambda]e_\mu$$

As an example, we consider degree 3 symmetric functions.

	m_{111}	m_{21}	m_3
e_3	1	0	0
e_{21}	3	1	0
e_{111}	6	3	1

Notably, this matrix is lower triangular and as 1s along the diagonal. This comes from the fact that $[m_{\lambda^*}]e_\lambda = 1$ where λ and λ^* are transpose. If we could show that this is the case in general, then we would have that the matrix is invertible over the integers (as it would have determinant 1). Before we get around to showing that the e_λ functions form a basis, we make note of a partial ordering on partitions of n. We say $\lambda \leq \mu$ if the ith partial sum of λ is less than or equal to the ith partial sum of μ for all i (we pad partitions with zeros to give λ and μ the same length). As it turns out, this partial ordering is a total ordering for $n < 6$, so we use as our example the partitions of 6.

Below is a diagram of the partial ordering for partitions of length 6.

$$\begin{tikzcd}
& (6) \arrow[d, dash] & \\
& (51) \arrow[d, dash] & \\
& (42) \arrow[dl, dash] \arrow[dr, dash] & \\
(33) \arrow[dr, dash] & & (411) \arrow[dl, dash] \\
& (321) \arrow[dl, dash] \arrow[dr, dash] & \\
(222) \arrow[dr, dash] & & (3111) \arrow[dl, dash] \\
& (2211) \arrow[d, dash] & \\
& (21111) \arrow[d, dash] & \\
& (111111) &
\end{tikzcd}$$

A few of intuitive notes about this partial ordering

- $\lambda \leq \mu \Leftrightarrow \mu^* \leq \lambda^*$. This is not necessarily obvious.
- If μ and λ are such that $\mu \leq \kappa \leq \lambda \implies \kappa \in \{\mu, \lambda\}$ and $\mu \leq \lambda$, then the Young diagram of μ can be obtained from the Young diagram of λ by moving a single square up by one row (recall that Young diagrams are left and bottom-justified). This explains the previous note.
- If we consider the weaker partial order $\leq^*$ where $\mu \leq^* \lambda$ if μ can be obtained from λ by moving a square up a row, then the ordering $\leq$ is the transitive closure of $\leq^*$. We leave the proof of this (i.e. that $\lambda < \mu \implies \exists \nu \,\text{s.t.}\, \lambda < \nu \leq \mu$) to the enthusiastic reader.

Now we prove that the previously considered matrix is in fact triangular. We claim that if $[m_\mu]e_\lambda \neq 0$, then $\mu \leq \lambda^*$. Observe that in the first k entries of row i, there are at most $\min(k, \lambda_i)$ entries equal to 1. The sum over all rows of these first k entries is precisely the kth partial sum of μ, which is hence bounded by

$$\sum_i \min(k, \lambda_i)$$

But this sum is precisely equal to the kth partial sum of λ^*, so it follows that $\mu \leq \lambda^*$ (as k is arbitrary). We have shown now that if $[m_\mu]e_\lambda \neq 0$, then $\mu \leq \lambda^*$, meaning that all entries of the aforementioned matrix that are above the main diagonal (these entries correspond to μ and λ such that $\mu > \lambda^*$) are 0. Further, it is simple to see that entries on the main diagonal of this matrix are equal to one, which gives us that this matrix is lower diagonal with determinant one. It follows that the elementary symmetric polynomials comprise a basis of the algebra of symmetric polynomials.

We have from this that $\{e_\lambda \mid |\lambda| = d\}$ forms a basis of $\Lambda_{(d)}$, and furthermore, we get that $\{e_\lambda \mid \lambda_1 \leq n\}$ (just as $\{m_\lambda \mid \ell(\lambda) \leq n\}$ was) a basis for $\mathbb{Z}[x_1, \dots, x_n]^{S_n}$. Thus, we have that

$$\mathbb{Z}[x_1, \dots, x_n]^{S_n} \simeq \mathbb{Z}[e_1, \dots, e_n]$$

This is also known as the *fundamental theorem of symmetric functions.*

In general,

$$\prod (\alpha_i - \alpha_j)^2$$

is known as the *discriminant.* For example, in the degree 2 polynomial $(x - \alpha_1)(x - \alpha_2)$, we have that the coefficient of x is given by $\alpha_1 + \alpha_2$. Likewise, we have that

$$\begin{aligned}(\alpha_1 - \alpha_2)^2 &= \alpha_1^2 + \alpha_2^2 - 2\alpha_1\alpha_2 \\ &= m_2(\alpha_1, \alpha_2) - 2m_{11}(\alpha_1, \alpha_2) \\ &= e_1^2 - 4e_2\end{aligned}$$

For the monic polynomial we had before, $e_1^2 - 4e_2 = b^2 - 4ac$ as $e_1 = -b$ and $e_2 = c$.

Next, we define the *complete homogenous polynomials* given by the sum of all monomials of degree k, which we may write as

$$h_k = \sum_{|\lambda|=k} m_\lambda$$

We get, for example, that

$$h_1 = x_1 + x_2 + \dots, \qquad h_2 = x_1^2 + x_1x_2 + x_2^2 + x_1x_3 + \dots$$

We can write h_2 in terms of the elementary symmetric functions as

$$h_2 = e_1^2 - e_2$$

We define

$$h_\lambda = \prod_{i=1}^{|\lambda|} h_{\lambda_i}$$

As one would expect, the h_λ form a basis over Λ. We prove this by relating the generating function of the h_n to those of the e_n. The generating function for the elementary symmetric function is given by

$$E(t) = \sum_{n=0}^{\infty} t^n e_n(x) = \prod_i (1 + tx_i) \quad e_0(x) = 1$$

Likewise, we have that the generating function for the h_n

$$H(t) = \sum_{n=0}^{\infty} t^n e_n(x) = \prod_i (1 + tx_i) \quad h_0(x) = 1$$

In particular, we can observe that $H(t)E(-t) = 1$, whence it follows that

$$\sum_{k+\ell=n} (-1)^\ell h_k e_\ell = 0$$

for $n > 0$. This gives us a relation between the complete homogeneous symmetric polynomials and the elementary symmetric polynomials. This relation gives us an inductive method for calculating the h_n polynomials from the e_n polynomials *and vice-versa.* This tells that the subring of the symmetric polynomials generated by the h_λ is the entire ring and that the h_n are algebraically independent. From this, we get that

- The h_λ form a basis of the algebra of symmetric functions.
- $\Lambda = \mathbb{Z}[h_1, \ldots, h_n]$
- $\mathbb{Z}[x_1, \ldots, x_n]^{S_n} \simeq \mathbb{Z}[h_1, \ldots, h_n]$

Having offered two nice bases for the algebra of symmetric functions, we can consider the mapping $\omega : \Lambda \to \Lambda$ given by $\omega(h_k) = e_k$ and $\omega(e_k) = h_k$ (these relations clearly induce the mapping on all functions since the e_k and the h_k each form a basis). The ring homomorphism ω is an involution and is a fundamental tool/object in the study of symmetric functions.

We alternatively show the computation of the coefficient of m_μ in h_λ. We have that

$$[m_\mu]h_\lambda = [x^\mu]h_\lambda$$

Our process is going to be very similar. We again consider a matrix representation. Again, our row sums must be λ and our column sums must be μ, but this time, h_λ is not necessarily a monomial. Thus, we get that the transformation matrix is a non-negative integer matrices with row sums λ and column sums μ. Furthermore, these matrices in fact must be *symmetric* (the same must hold for the es).

Consider the generating function

$$\prod_{i,j} \frac{1}{1 - x_i y_j} = \prod_j \left(\prod_i \frac{1}{1 - y_j x_i} \right)$$

We note that the term inside the parentheses is given precisely by $H(y_j) = \sum y_j^n h_n(x)$. Thus, we may write this product above as

$$\sum_\lambda m_\lambda(y) h_\lambda(x) = \sum_{\lambda,\mu} m_\lambda(y) m_\mu(x)$$

By a similar argument we get that

$$\prod_{i,j} (1 + x_i y_j) = \sum_\lambda m_\lambda(y) e_\lambda(x)$$

Now we hope to introduce an inner product $\langle, \rangle : \Lambda \otimes \Lambda \mapsto \mathbb{Z}$ We want this inner product to be defined such that h_μ is the dual basis to m_λ. Thus, we define

$$\langle m_\lambda, h_\mu \rangle = \delta_{\lambda\mu}$$

It follows from this definition that

$$\langle h_\lambda, h_\mu \rangle = [m_\mu]h_\lambda = [m_\lambda]h_\mu = \langle h_\mu, h_\lambda \rangle$$

We first note a generic fact of bilinear forms (which applies also to free modules over a commutative ring). Suppose we have a non-degenerate pairing (not necessarily symmetric)

$$\langle, \rangle : W \otimes V \mapsto K$$

where K is the ground field. Let $w_i \in W$ and $v_i \in V$ be dual bases under this pairing. We can think of the pairing as a linear functional $W \mapsto V^*$ which takes a vector and returns the functional, that when given an element of V, returns the value of the pairing with w. Since the pairing is non-degenerate, this map is a bijection. We would like to claim that

$$\sum w_i \otimes v_i \in W \otimes V$$

depends only on the pairing.

Going back to our product above, we can see that it belongs to $\Lambda_d(X) \otimes \Lambda_d(Y)$, as it is doubly symmetric in x and y. Thus, in our sum

$$\sum_{\lambda,\mu} m_\lambda(y) m_\mu(x)$$

the product really functions as a tensor product, and as described above, it is uniquely described given the pairing that we defined above. Thus, in fact we get that the product

$$\prod_{i,j} \frac{1}{1 - x_i y_j} = \sum_\lambda u_\lambda(y) v_\lambda(x)$$

for *all* dual bases u_λ and v_λ for $\langle , \rangle$ with degree $|\lambda|$.

We've now made note of four distinct bases for the algebra of symmetric functions. We'd like to make note of a slight distinction between the power sum symmetric functions and the other three bases - the monomial symmetric functions, the elementary symmetric functions, and the complete homogeneous symmetric functions - in terms of structure. This distinction is that the other three bases form $\mathbb{Z}$-bases of $\Lambda_\mathbb{Z}$, while we will in fact see that the power sum symmetric functions comprise a $\mathbb{Q}$-basis of $\Lambda_\mathbb{Q}$. We define

$$p_k := m_{(k)} = \sum_i x_i^k$$

As we have done before, we define

$$p_\lambda = \prod_{i=1}^{|\lambda|} p_{\lambda_i}$$

We want to show that $\{p_\lambda\}$ forms a $\mathbb{Q}$ basis for $\Lambda_\mathbb{Q}$. There are a few ways to show this. We start by writing p_λ as a sum of monomials, writing

$$p_\lambda = \sum c_{\lambda\mu} m_\mu$$

In general, the terms of p_λ can be thought of as picking some "bunches" of λ and adding these bunches together. We have

$$p_\lambda = c_{\lambda\lambda} m_\lambda + \sum_{\lambda<\mu} c_{\lambda\mu} m_\mu$$

where the ordering that we reference is the refinement ordering, in which $\mu < \lambda$ if μ can be obtained from λ by splitting up some block or blocks of λ into two or more parts, rather than the typical ordering on partitions (in fact, however, the typical ordering is a subset of the refinement ordering). In the case of equality, we have that the number of ways we can pick λ is given by the number of ways we can label the x_i so as to give the same partition. That is to say, we don't get a coefficient $c_{\lambda\lambda} = 1$, we instead get

$$c_{\lambda\lambda} = \prod r_i!$$

where r_i denotes the number of partitions there are of size i. We can write, as we have done previously, a matrix for conversion from the monomial symmetric functions to the power sum symmetric functions. Take the example for degree-3 polynomials:

	m_{111}	m_{21}	m_3
p_{111}	6	3	1
p_{21}	0	1	1
p_3	0	0	1

We have, as we did before, a triangular matrix with nonzero entries along the diagonal, but in this case, the determinant is non-unital, and hence the inverse of this matrix contains rational (non-integral) numbers.

Similar to how we used a relationship between e_k and h_k to show that h_λ was a basis, we will find a relationship between the h_k and the p_k to show that the p_λ form a ($\mathbb{Q}$) basis.

$$H(t) = \sum h_n(x)t^n = \prod \frac{1}{1 - tx_i} = \Omega[tx]$$

We can observe now that

$$\begin{aligned} \log H(t) &= \sum_i \log \frac{1}{1 - tx_i} \\ &= \sum_i \sum_{k>0} \frac{t^k x_i^k}{k} \\ &= \sum_{k>0} \frac{t^k p_k(x)}{k} \end{aligned}$$

which allows us to conclude that

$$H(t) = \exp \sum p_k \frac{t^k}{k} = \Omega[tx]$$

Whereas previously, the p_k were plethystic evaluation operators, we now have a similar formula where the p_k are truly polynomials (they are power sum). Setting $t = 1$ and recalling that

$$\prod \frac{1}{1 - x_i} = \exp \sum \frac{p_k}{k} = Z_E = \Omega[x]$$

and also that

$$\prod \frac{1}{1 - x_i} = \sum h_n(x)$$

we get that

$$h_n = \sum_{|\lambda|=n} \frac{p_\lambda}{z_\lambda}$$

(using our earlier computation for the degree n term of Ω). Now recall the previously defined involution ω which swaps the elementary symmetric polynomials and the complete homogeneous polynomials. We have that

$$\omega H(t) = E(t) = \sum e_n t^n = \prod (1 + tx_i) = H(-t)^{-1}$$

and taking the logarithm of both sides yields

$$\omega \log H(t) = -\log H(-t)$$

or equivalently,

$$\omega \sum p_k \frac{t^k}{k} = -\sum p_k \frac{(-t)^k}{k}$$

Thus, it follows that

$$\left[\frac{t^k}{k}\right] \omega p_k = (-1)^{k-1} p_k$$

and hence that

$$\omega p_\lambda = \prod (-1)^{\lambda_i - 1} p_\lambda = (-1)^{n-\ell(\lambda)} p_\lambda$$

Note that $(-1)^{n-\ell(\lambda)}$ is the sign of a permutation of cycle type λ. So we have another identity:

$$e_n = \sum_{|\lambda|=n} (-1)^{n-\ell(\lambda)} \frac{p_\lambda}{z_\lambda}$$

This identity, coupled with the prior identity giving a formula for the complete homogeneous symmetric polynomials, suggests a deep connection between the symmetric functions of fixed degree and the symmetric group.

We tie up a few loose ends on power sums. It now makes sense to think of plethystic evaluation as an operation on Λ. For

$$f \in \Lambda_{\mathbb{Q}} = \mathbb{Q}\left[p_1, p_2, \cdots\right]$$

we have that $f[A] = f\Big|_{p_k \mapsto p_k[A]}$ and $f[X] = f(x_1, x_2, \dots)$ were $X = \sum x_i$. In the way of determining $f[X]$, we consider instead

$$f[-\epsilon X]\Big|_{\epsilon=-1}$$

we can observe that

$$p_k[-\epsilon X]\mid_{\epsilon=-1} = -\epsilon^k p_k(X_1, X_2, \cdots)\Big|_{\epsilon=-1} = (-1)^{k-1} p_k$$

so

$$\omega f = f[-\epsilon X]\Big|_{\epsilon=-1} = (-1)^d f[-X]$$

for f homogeneous of degree d.

We further note that we may write

$$\Omega[XY] = \prod \frac{1}{1 - x_i y_j} = \sum_\lambda u_\lambda(y) v_\lambda(x)$$

for u_λ and v_λ dual bases. We may further write

$$\Omega = \sum_\lambda \frac{p_\lambda}{z_\lambda}$$

as before, thus

$$\Omega[XY] = \sum_\lambda \frac{p_\lambda\left[XY\right]}{z_\lambda} = \sum_\lambda \frac{p_\lambda\left[X\right] p_\lambda\left[Y\right]}{z_\lambda}$$

which resembles our above sum! Thus, we have that the p_λ are (nearly) self dual, we need only scale by a factor of z_λ. More formally, we have that

$$\langle p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} z_\lambda = \delta_{\lambda\mu} z_\mu$$

where the second equality follows since we multiply by $\delta_{\lambda\mu}$ We can in particular observe that

$$\langle h_n, p_\lambda \rangle = \sum_{|\mu|=n} \left\langle \frac{p_\mu}{z_\mu}, p_\lambda \right\rangle = 1 \text{ for all } \lambda$$

and also

$$\langle e_n, p_\lambda \rangle = (-1)^{n-\ell(\lambda)} = \text{sign of } \tau_\lambda$$

7.1 Schur Functions

We introduce this very important class of symmetric polynomials using Jacobi's "bi-alternate" formula.

Definition 7.2

An *antisymmetric polynomial* $f(x_1, x_2, \cdots, x_n)$ is a polynomial such that

$$f(\cdots, x_i, \cdots, x_j, \cdots) = -f(\cdots, x_j, \cdots, x_i, \cdots)$$

That is, switching two variables changes the sign of the polynomial.

We observe immediately that if λ does not have n distinct parts (one of the parts may be 0), then

$$[x^\lambda]f = 0$$

If λ *does* have n distinct parts, then x^λ can occur in f, and if it does, then it induces $n!$ total terms with coefficients induced by the condition that f is antisymmetric. Said induced polynomials comprise a basis for the antisymmetric polynomials of degree n. As an example, we have

$$a_{210} = x_1^2x_2 - x_1x_2^2 + x_1x_3^2 - x_2x_3^2 + x_2^2x_3 - x_1^2x_3$$

In general, we may write

$$a_\mu = \begin{vmatrix} x_1^{\mu_1} & \dots & x_n^{\mu_n} \\ \vdots & \ddots & \vdots \\ x_1^{\mu_n} & \dots & x_n^{\mu_n} \end{vmatrix}$$

As an example, we can observe the smallest partition with n distinct parts, namely $\rho = (n-1, n-2, \cdots, 1, 0)$. The determinant giving a_ρ is called the *Vandermonde determinant* and is in fact equal to

$$\pi_{i<j}(x_i - x_j)$$

This fact is called the *Vandermonde identity* (apparently); the proof is not so hard. As a side note, the Vandermonde determinant is the square root of the discriminant of the polynomial with roots $x_1, \cdots, x_n$.

If we take any partition λ and add ρ to it, we necessarily get a partition with distinct parts (and in fact, we may use this as the general form of a partition with distinct parts). As such, we have that the $a_{\lambda+\rho}$ polynomials comprise a basis for the antisymmetric functions. Further, we can see that $a_{\lambda+\rho}$ is divisible by a_ρ for all λ, meaning that $\frac{a_{\lambda+\rho}}{a_\rho}$ is a symmetric polynomial of degree $|\lambda|$ (the divisibility comes from the fact that setting two variables equal in any antisymmetric function must make the function equal 0). Further, given any symmetric function, we can multiply by a_ρ to get an antisymmetric function. We have now given a bijection between symmetric and antisymmetric polynomials.

Definition 7.3 (Schur functions)
The *Schur functions* are the symmetric functions given by

$$S_\lambda(x_1, \dots, x_n) = \frac{a_{\lambda+\rho}}{a_\rho}$$

As an example, we see that

$$S_{1^n} = \frac{a_{(n,1)}}{a_\rho} = x_1 \cdot x_2 \dots$$

Due to the aforementioned bijection between symmetric and antisymmetric functions, we have trivially that the Schur functions form a basis for the symmetric polynomials. As another example, we show how to get other polynomials in terms of the Schur polynomials. We take as an example $a_\rho \cdot m_2$. We know that the product must be some linear combination of $a_{\rho+(2)}$ and $a_{\rho+(11)}$. We first note that when consider the product terms, adding a power of two to any of the terms except for the first two will yield 0, as will end up with two x_i of equal degree. We can see that the coefficient of $a_{\rho+(2)}$ must be 1, as when we add 2 to the exponent of x_1, we get this exactly. When adding 2 the the exponent of x_2, we see that we nearly get the right term, but it is written in the wrong order where the order $n-1$ term appears first. To fix this, we must swap x_1 and x_2, which yields a negative sign as transpositions have negative sign. Thus, we get that

$$a_{\rho+(2)} - a_{\rho+(11)} = a_\rho \cdot m_2 \implies m_2 = S_{(2)} - S_{(1)}$$

It is true in general that

$$S_{(k)} = h_k \qquad S_{(1^k)} = e_k$$

these facts are not obvious, but they can be proven with techniques similar to what we have just discussed.

Consider now

$$a_{\lambda+\rho}(x_1, x_2, \cdots, x_{n-1}, 0)$$

Considering the determinant that gives $a_{\lambda+\rho}$, we have that if $\lambda_n \neq 0$, then setting $x_n = 0$ makes the determinant equal to 0. If $\lambda_n = 0$, then all but the last entry of the rightmost column of the matrix become 0, and hence the determinant of the matrix reduces to the determinant of the upper left $(n-1) \times (n-1)$ block of the matrix. The determinant of this block is in fact equal to

$$a_{(\lambda_1,\cdots,\lambda_{n-1})+\rho_{n-1}}(x_1, \cdots, x_{n-1}) \cdot x_1 \cdots x_{n-1}$$

because we have removed all terms with x_n (which yields $a_{(\lambda_1,\cdots,\lambda_{n-1})}$ and increased the exponent of all variables by 1 (which gives the $x_1 \cdots x_{n-1}$. We can perform a similar computation to compute $S_\lambda(x_1, x_2, \cdots, x_{n-1}, 0)$. We have that

$$S_\lambda(x_1, x_2, \cdots, x_{n-1}, 0) = \frac{a_{\lambda+\rho_n}(\cdots, 0)}{a_{\rho_n}(\cdots, 0)} = \frac{a_{\lambda+\rho_{n-1}}}{a_{\rho_{n-1}}} = S_\lambda(x_1, \cdots, x_{n-1})$$

with the special exception that the value is 0 if $\lambda_n > 0$.

We can also write

$$S_\lambda(x_1, x_2, \cdots) = \sum_\mu K_{\lambda\mu} m_\mu = S_\lambda[x_1 + \cdots + x_n] = \begin{cases} S_\lambda(x_1, \cdots, x_n) & n \geq \ell(\lambda) \\ 0 & n < \ell(\lambda) \end{cases}$$

where we call the $K_{\lambda\mu}$ the *Kostka Coefficients*. Notably, these coefficients do not depend on n. We may therefore write in Λ, the set of symmetric polynomials in infinite variables as

$$S_\lambda = \sum_\mu K_{\lambda\mu} \cdot m\mu$$

Writing $S_\lambda = \frac{a_{\lambda+\rho}}{a+\rho}$, we have

$$S_\lambda(x_1, \ldots, x_n) = \frac{\sum_{w\in S_n}(-1)^{\ell(w)} x^{w(\lambda+\rho)}}{\prod_{i<j}(x_i - x_j)} = \sum_{w\in S_n} w\left(\frac{x^{\lambda+\rho}}{\prod_{i<j}(x_i - x_j)}\right)$$

Rewriting this using the identity

$$\prod_{i<j}(x_i - x_j) = x^\rho \prod_{i<j}(1 - \frac{x_j}{x_i})$$

, we may rewrite the above as

$$S_\lambda(x_1, \ldots, x_n) = \sum_{w\in S_n} w\left(\frac{x^\lambda}{\prod_{i<j}(1 - \frac{x_j}{x_i})}\right)$$

which is also known as the *Weyl character formula.* The *character* of an irreducible representation is given by the trace of an element of the Borel subgroup under the representation. In general, the characters are symmetric polynomials (as traces are invariant under conjugation), and we can see that in fact, they are basis elements of Λ given by the Schur polynomials. As it turns out, $K_{\lambda\mu}$ refers to the dimension of the μ weight space of an irreducible representation V_λ. Thus, we can show (even non-combinatorially), that the Kostka coefficients are nonnegative.

We continue our study of the Schur functions by introducing the *Bernstein operators* B_m which simply adds a part to a partition λ. Formally, we have that $B_m \cdot S_\lambda = S_{(m;\lambda)}$. Under this convention, we need to determine how the Schur functions are defined for non-partitions. Suppose we have a sequence λ that is not necessarily a partition. If, upon adding ρ, we have all distinct parts, then we may write $\lambda + \rho = w^{-1}(\nu + \rho)$ where $w \in S_n$ and ν *is* a valid partition. We may now write $\nu = w(\lambda + \rho) - \rho$, and we define

$$S_\lambda = (-1)^{\ell(w)} S_\nu$$

Now, we can write

$$S_{(m;\lambda)} = \sum_{w\in S_n} w\left(\frac{x_1^m x_2^{\lambda_1}\cdots x_n^{\lambda_{n-1}}}{\prod_{j>1}(1-x_j/x_i)\cdot\prod_{1<i<j}(1-x_j/x_i)}\right)$$

which we may simplify (by collecting terms with a common denominator) to

$$S_{(m;\lambda)} = \sum_i \frac{x_i^m s_\lambda[X-x_i]}{\prod_{j\neq i}(1-x_j/x_i)} = \frac{x_1^m S_\lambda(x_2,\cdots,x_n)}{\prod_{j\neq 1}(1-x_j/x_1)}$$

We independently note that

$$\Omega[Xz] = \prod_{i=1}^n \frac{1}{1-x_iz} = \sum_i \frac{1}{1-x_iz}\prod_{j\neq i}\frac{1}{1-\frac{x_j}{x_i}}$$

Notice that if we take any function of z^{-1} and multiply by $\frac{1}{1-x_iz}$, we have that the constant term

$$[z^0]\left[f(z^{-1})\cdot\frac{1}{(1-x_iz)}\right] = f(x_i)$$

Going back to our previous expression, notice that we have a very similar expression

$$S_{(m;\lambda)} = \frac{x_1^m S_\lambda(x_2,\cdots,x_n)}{\prod_{j\neq 1}(1-x_j/x_1)} = [z^0]\left[\Omega[Xz]z^{-m}S_\lambda[X-z^{-1}]\right] = [z^m]\Omega[Xz]S_\lambda[X-z^{-1}]$$

as the latter two product terms in the first equality convert $\frac{1}{1-x_iz}$ into $x_i^m S_\lambda[X-x_i]$. We thus define, for some arbitrary $g\in\Lambda$ (as opposed to simply Schur polynomials),

$$B_m g(x) := [z^m]\Omega[Xz]S_\lambda[X-z^{-1}]$$

As a final detail, we note that, under the inner product on Λ, we have (as a proposition), that

$$\langle\Omega[AX], g(X)\rangle = g[A]$$

We can show this as follows. We may write

$$\Omega[AX] = \sum_\lambda h_\lambda[A]m_\lambda(X)$$

we chose h_λ and m_λ, but this equality in fact holds for any pair of dual bases. Thus,

$$\langle\Omega[AX], g(X)\rangle = \langle\sum_\lambda h_\lambda[A]m_\lambda(X), g(X)\rangle$$

from which the result follows. This enables us to define the adjoint operator $\theta^\perp$ to θ by $\langle\theta^\perp f, g\rangle = \langle f, \theta g\rangle$. We claim that

$$\Omega[AX]^\perp f = f[X+A]$$

We have that

$$\langle\Omega[AX]^\perp f, g\rangle = \langle f, \Omega[AX]g\rangle$$

Now we make the important note that the expression $\Omega[BX]$ is just as general as $g(x)$, and hence we need to demonstrate that

$$\langle f[X+A], \Omega[BX]\rangle = \langle f, \Omega[AX]\Omega[BX]\rangle$$

The left-hand side of this is nothing but $f[B+A]$, and the right-hand side simplifies to $\langle f, \Omega[(A+B)X]\rangle = f[A+B]$, and hence the two are equal, proving the prior claim.

Having introduced the Bernstein raising operators for Schur functions, we hope to explore the properties of these operators some more. We had

$$B_m = [z^m]\Omega[zX]\Omega[-z^{-1}X]^\perp$$

and we found that for the Schur fucntion S_λ, we have

$$B_m \cdot S_\lambda = S_{(m;\lambda)}$$

that is, B_m adds a part to the partition λ. We have the convention

$$S_\mu = \sum w\left(\frac{x^\mu}{\prod_{i<j}(1 - x_j/x_i)}\right)$$

for any $\mu \in \mathbb{Z}^n$. Note that if $\mu + \rho$ has any repeated entries, we'll have $S_\mu = 0$. Otherwise, $\mu + \rho$ is some permutation of $\lambda + \rho$, for λ a partition, and hence we have $S_\mu = S_\lambda \cdot \operatorname{sign}(w)$. where w sends λ to μ. Falling out of these Bernstein operators, for example, is that

$$S_{(m)} = B_m \cdot 1 = h_m(x)$$

We hope now to address the multiplication of a Schur polynomial by an elementary polynomial. We have

$$S_\lambda = B_{\lambda_1} \cdot B_{\lambda_2} \cdot \dots \cdot B_{\lambda_\ell} \cdot 1$$

We need to understand how multiplication and perping interact. The result we have is that

$$\Omega[AX]^\perp\Omega[BX] = \Omega[AB]\Omega[BX]\Omega[AX]^\perp$$

so when we commute two operators like this, we find a "constant factor." To explain this, we note that

$$\begin{aligned}\Omega[AX]^\perp\Omega[BX]f(x) &= \Omega[B(x + A)]f(x + A)\\ &= \Omega[AB]\Omega[BX]\Omega[AX]^\perp f\end{aligned}$$

recall that the adjoint operator of $\Omega[AX]$ substitutes x for $x + A$ (this explains the first line), and also Ω sends sums to products. Now note that

$$\Omega[uX] = \sum h_n(X)u^n$$

and, recalling the behavior of Ω when we negate the argument (see earlier in these notes), we find that

$$\Omega[-uX] = \sum(-1)^k e_k(x)u^k = \sum e_k(x)(-u)^k$$

Now we take $\Omega[-uX]$ and the generating function for the Bernstein operators. We have

$$\begin{aligned}\Omega[-uX]\Omega[zX]\Omega[-z^{-1}X] &= \Omega[zX]\Omega[-z^{-1}X]^\perp\Omega[-uX]\Omega[-uz^{-1}]\\ &= \Omega[zX]\Omega[-z^{-1}X]^\perp\Omega[-uX](1 - uz^{-1})\end{aligned}$$

We can now take the coefficient of $(-u)^k z^m$ on both sides. This gives us

$$e_k B_m = B_m e_k + B_{m+1}e_{k-1}$$

so now we have some way of commuting the e_k's and the B_m's. We now turn our attention back to the Schur polynomials. We find that

$$\begin{aligned}e_k S_\lambda &= e_k B_{\lambda_1} \cdot B_{\lambda_2} \cdot \dots \cdot B_{\lambda_\ell} \cdot 1\\ &= \sum B_{\lambda_1+\epsilon_1} B \cdots B_{\lambda_\ell+\epsilon_\ell} e_j\end{aligned}$$

where ϵ_i is either 0 or 1 (representing which of the two summands in the above formula we choose) and

$$j + \sum \epsilon_i = k$$

That is, the e_j rectifies any difference in degree between the terms of the summation and k. We can write e_j as $S_{(1^j)}$, and we find that the $B \cdots Be_j$ is a Schur function. This corresponds to an ℓ-tuple, composed of the parts of λ possibly boosted by 1, but this ℓ-tuple is not necessarily a partition, as if λ has two consecutive equal parts, the latter of which is boosted, we get a followed by $a+1$ in the resulting ℓ-tuple. This is ostensibly an issue, but when we consider that the Weyl character function of a partition with repeated parts is 0, we realize that these terms simply don't matter, seeing as if we have a followed by $a+1$ in μ, then $\mu + \rho$ will have at least one repeated part. We finally arrive at the *Pieri rule*:

$$e_k s_\lambda = \sum_\mu s_\mu$$

where μ is a partition such that μ/λ is a vertical k-strip (i.e. we've added k 1's to the partition λ - including to any 0 components at the end of λ - to obtain μ). With this rule in mind, we can consider $e_k^\perp$. We have

$$\Omega[-uX]^\perp = \sum e_k(x)^\perp (-u)^k$$

and we observe that (playing the same game as before)

$$\begin{aligned}
[\Omega[-uX]^\perp \Omega[zX]\Omega[-z^{-1}X]^\perp &= \Omega[-uz]\Omega[zX]\Omega[-z^{-1}X]^\perp \Omega[-uX]^\perp \\
&= (1-uz)\Omega[-z^{-1}X]^\perp \Omega[-uX]^\perp
\end{aligned}$$

so taking the coefficient of $(-u)^k z^m$ yields

$$e_k^\perp B_m = B_m e_k^\perp + B_{m-1} e_{k-1}^\perp$$

Now we have

$$e_k S_\lambda = e_k B_{\lambda_1} \cdots B_{\lambda_\ell} \cdot 1$$

whence it follows that

$$e_k S_\lambda = \sum_\mu S_\mu$$

where this time we *subtract* one from any k parts of λ in such a way that we end up with a valid partition. We have found a very nice duality between e_λ and $e_\lambda^\perp$, and this duality is very important, as it tells us something deep about the Schur functions and their structure. In particular, the Schur functions are determined entirely by the Pieri rule, and at the same time, the elementary symmetric functions can be generated (inductively) entirely using Schur functions and the Pieri rule. One thing we can observe relatively quickly is that when the e_k are expanded using Schur functions, all coefficients are positive. We also note that

$$\langle e_\mu, S_\lambda \rangle = e_\mu^\perp = S_\lambda$$

where $|\lambda| = |\mu|$. Note that since the e_λ functions comprise a basis for symmetric functions, knowledge of these inner products is tantamount to knowledge of all of the Schur functions.

Let's consider the dual basis $\tilde{S}_\lambda$ to the Schur functions. We first note that

$$[S_\mu] e_k S_\lambda = \begin{cases} 1 & \text{if } \mu/\lambda \text{ is a vertical } k\text{-strip} \\ 0 & \text{otherwise} \end{cases}$$

and we have that

$$\langle \tilde{S}_\lambda, e_k S_\lambda \rangle = \langle e_k^\perp \tilde{S}_\mu, S_\lambda \rangle = [\tilde{S}_\lambda] e_k^\perp \tilde{S}_\mu$$

In other words, $\tilde{S_\lambda} = S_\lambda$, so the S_λ comprise an orthonormal basis! This is a beatiful fact, because in general, having an inner product over a free $\mathbb{Z}$-module does not imply the existence of an orthonormal basis, and yet in this case we have one. The other great thing is that orthonormal bases (of $\mathbb{Z}$-modules) are unique up to reording and sign changes.

We'd now like to turn our attention to the interactions between the Bernstein operators and the complete homogeneous symmetric functions. In this case, we will require some more careful combinatorial analysis, because we won't be as lucky as we were with the elementary symmetric functions, where all of the "non-nice" terms (where we didn't get a partition after adding the 0-1 vector) vanished. This is because instead of adding 0-1 vectors, we are adding general vectors. Recall the generating function for the h_k:

$$\Omega[uX] = \sum h_k(x)u^k$$

We find that

$$\begin{aligned}\Omega[uX]\Omega[zX]\Omega[-z^{-1}X]^{\perp} &= \Omega[zX]\Omega p - z^{-1}X]^{\perp}\Omega[uX]\Omega[uz^{-1}] \\ &= \frac{\Omega[zX]\Omega[-z^{-1}X]^{\perp}\Omega[uX]}{1-uz^{-1}}\end{aligned}$$

Taking the coefficient of $u^k z^m$ gives

$$h_k B_m = \sum_{\ell=0}^{k} B_{m+\ell}h_{k-\ell}$$

In other words, unlike the case of the elementary symmetric functions, where we could only change by 0 or by 1, we can now change by any constant ℓ. This means that

$$h_k S_\lambda = \sum_\mu S_\mu$$

where $\mu = \lambda + \nu$, with ν any non-negative vector with total weight k. There are many terms of this sum in which μ is not valid partition but still does not vanish due to asymmetrization. As it turns out, it will be the case that any μ in which some part of the partition is extended beyond a part in a previous row will get cancelled by another such μ. This means that any μ with S_μ in the above some is obtained by adding a *horizontal* strip of length k to λ. This gives us the Pieri rule for the h_k:

$$h_k S_\lambda = \sum_\mu S_\mu$$

where μ/λ is a horizontal k-strip. We can offer an inductive argument for why all of the aforementioned non-partitions cancel.

We could apply a similar analysis to $h_k^{\perp}$, but we can also be clever and realize that since the h_k and the e_k are dual under ω, the rule for $h_k^{\perp}$ is dual to the rule of $e_k^{\perp}$. In other words,

$$h_k^{\perp} S_\mu = \sum S_\mu$$

where μ is such that λ/μ is a horizontal k-strip. Recall that we previously mentioned that the Kostka coefficient $K_{\lambda\mu}$ is given by $K_{\lambda\mu} = [m_\mu]S_\lambda = \langle h_\mu, S_\lambda\rangle = \langle 1, h_\mu^{\perp} S_\lambda\rangle = h_\mu^{\perp} S_\lambda$. This is nothing but the number of ways to reduce λ to nothing (i.e. the empty partition) by removing a succession of horizontal strips of sizes $\mu_\ell, \mu_{\ell-1}, \cdots, \mu_1$, which again demonstrates that $K_{\lambda\mu}$ is a positive integer (we previously demonstrated this by observing $K_{\lambda\mu}$ as the dimension of a weight space). Now, one way we can think of removing these strips is by writing the index i (as in μ_i) of the size of the strip we remove.

Having established the Schur functions as an orthonormal basis in the inner product space of symmetric functions (thanks to the Pieri rules), we can recall the Kostka coefficients:

$$K_{\lambda\mu} = [m_\mu]s_\lambda \qquad s_\lambda = \sum_\mu K_{\lambda\mu}m_\mu$$

Recalling that the m_μ functions are dual to the h_μ functions, we can write instead that

$$K_{\lambda\mu} = [s_\lambda]h_\mu$$

In particular, $K_{\lambda\mu}$ is the number of ways we can build up the partition λ by repeatedly adding horizontal strips whose lengths are given by the parts of μ. As an example, suppose we have

$$\lambda = (3, 2, 2) \qquad \mu = (2, 2, 1, 1, 1)$$

One manner of building λ as a set of horizontal strips is

3	5	
2	4	
1	1	2

We can hence view the Schur function as a generating function

$$s_\lambda = \sum_{T \in \mathrm{SSYT}(\lambda)} x^T$$

where SSYT(λ) is the set of semistandard Young tableaux of partition λ. Of note is the fact that $\omega s_\lambda = s_{\lambda^*}$, so the ω operator acts as a sort of transpose when we operate in the basis of Schur functions.

If we instead consider the difference between the diagram between a larger diagram and a smaller diagram, we get a *Skew Young diagram* (denoted λ/μ). We can recognize these diagrams as whenever we have two squares that are diagonally opposite, we must have the whole rectangle. We now look at semistandard *skew* tableaux. To verify that this construction yields symmetric functions, we need only check that consecutive variables are symmetric. We can see that this holds, as for all cells that lie in between columns in which there is a change in the height of the diagram, we may swap i and $i+1$ with no repercussions. Another way to think about this is to label (reading from left to right) all instances of i with a right parenthesis and all instances of $i+1$ with a left parenthesis. We may demonstrate symmetry by first eliminating all matching left and right parenthesis, and among the unmatched portion, swapping the number of each type of parenthesis.

Let's consider the significance of the skew Schur functions. In particular, consider

$$[m_\nu]S_{\lambda/\mu} = K_{\lambda/\mu,\nu}$$

The left-hand side can be seen to be the inner product of h_ν with $S_{\lambda/\mu}$, and we have

$$\langle h_\nu, S_{\lambda/\mu}\rangle = \langle h_\nu S_\mu, S_\lambda\rangle$$

We can then reconstruct

$$s_{\lambda/\mu} = s_\mu^\perp s_\lambda$$

In other words, taking the inner product of two Schur functions gives a skew Schur function.

As a special case of multiplying Schur functions, we consider $S_{21} \cdot S_{21}$. We choose this product since it is in some sense the smallest nontrivial example, in that it is not simply a row or a column. We consider the coefficient of S_λ in $S_{21} \cdot S_{21}$ for arbitrary λ. Note that

$$[S_\lambda]S_{21} \cdot S_{21} = \langle S_\lambda, S_{21} \cdot S_{21}\rangle = \langle S_{21}^\perp S_\lambda, S_{21}\rangle = \langle S_{\lambda/21}, S_{21}\rangle$$

This gives us that

8 Representation theory

Definition 8.1 (Coordinate free representation)
A matrix representation of a finite group G is a homomorphism $\varphi : G \mapsto GL(V)$ where V is a (finite-dimensional) vector space. We only care about a representation up to conjugation of elements in $GL(V)$, as this simply corresponds to a change in basis. Such consideration leads to coordinate free representations. As we want a characteristic 0 field that is algebraically closed, we pick $\mathbb{C}$ as our ground field.

Definition 8.2 (Character)
In our pursuit of coordinate-free properties of our representation, we define the *character* $\chi : G \mapsto \mathbb{C}$ given by

$$\chi(g) = \mathrm{tr}_V(\varphi(g))$$

This definition of character leads to several nice properties. Firstly,

$$\chi_{V\oplus W} = \chi_V + \chi_W$$

and we also have

$$\chi_{V\otimes W} = \chi_V \chi_W$$

Furthermore, as desired, the character is invariant under a change of basis. We also have that if we have a submodule $W \subset V$, we can write

$$V = W \oplus V/W$$

yielding the formula

$$\chi_V = \chi_W + \chi_{V/W}$$

We note that if a submodule exists, it must be the case that the underlying space can be decomposed into a direct sum of irreducible submodules – that is, submodules for which no invariant subspace exists.

As an example, we consider S_n acting on $\mathbb{C}^n$ by permuting coordinates. One invariant subspace (submodule) we can find is the one dimensional subspace

$$W = \mathbb{C} \cdot (1, \dots, 1)$$

The quotient (in $\mathbb{C}^n$) of this subspace

$$V/W = \{(x_1, \dots, x_n) \mid \sum_i x_i = 0\}$$

Note that we would not be able to find this quotient if our field characteristic divided the order of the group.

If $G \circlearrowright V$, then we automatically have that $G \circlearrowright V^*$ by composing φ with the dual operator. As

$$\varphi(gh)^* = \varphi(h) * \varphi(g) *$$

we must take the inverse to construct a proper homomorphism

$$\varphi((gh)^{-1}) = \varphi(g^{-1})^* \varphi(h^{-1})^*$$

Thus, we have that

$$\chi_{V^*}(g) = \chi_V(g^{-1}) = \overline{\chi_V(g)}$$

In general, if we were to consider linear maps $\mathrm{Hom}(V, W)$, we can have G act on this space by letting

$$g \cdot \alpha = g \circ \alpha \circ g^{-1}$$

Note that we can recover our previous homomorphism on the dual by considering $W = \mathbb{C}$ and letting g act trivially on W. As $\mathrm{Hom}(V, W) = V^* \otimes W$, we have by the above that

$$\chi_{\mathrm{Hom}(V,W)}(g) = \chi_{V^*} * \chi_W(g) = \chi_W(g)\chi_V(g^{-1})$$

Definition 8.3 (G-module homomorphism)
A G-module homomorphism $\psi : V \mapsto W$ has the property that

$$\psi(gv) = g\psi(v)$$

from which we have htat

$$g\psi g^{-1} = \psi$$

It follows that the space of G-module homomorphisms

$$\mathrm{Hom}_{\mathbb{C}G}(V, W) = \mathrm{Hom}_{\mathbb{C}}(V, W)^G$$

where

$$V^G := \{v \in V \mid gv = v \quad \forall g \in G\}$$

It follows that

$$\chi_V(1) = \dim V$$

We define a convenient operator known as the *Reynolds operator*

$$R := \frac{1}{|G|} \sum_{g \in G} g$$

We show that the Reynolds operator is a projection on any G-module into the subspace V^G. Formally, we have that

$$gRv = \frac{1}{|G|} \sum_{h \in G} ghv = Rv$$

Furthermore, for $v \in V^G$, we have that

$$Rv = \frac{1}{|G|} \sum_{g \in G} gv = \frac{1}{|G|} |G| \cdot v = v$$

from which it follows that R is a projection. We finally have from this that

$$\chi)V(R) = \dim V^G$$

Combining all that we have from above, we have that

$$\dim \hom_G(V, W) = \dim \hom_{\mathbb{C}}(V, W)^G = \chi_{\hom(V,W)}(R) = \frac{1}{|G|} \sum_{g \in G} \chi_W(g) \chi_V(g^{-1}) := \langle \chi_V, \chi_W \rangle$$

Thus, we have now defined an inner product of characters by considering the dimension of G-modules over linear maps from V to W.

Definition 8.4 (Irreducible)
A G-module V is irreducible if it has no nontrivial submodule W.

Lemma 8.5 (Schur's Lemma)
Suppose that V and W are irreducible G-modules. Then we have

$$\hom_G(V, W) = \begin{cases} 0 & \text{if } V \not\cong W \\ \mathbb{C} & \text{otherwise} \end{cases}$$

Proof. Consider a homomorphism φ from V to W. Suppose that φ is not injective. Then it has a nontrivial kernel $\ker\varphi \subset V$ (where inclusion is strict). But this is a nontrivial submodule of V, contradicting that V is irreducible. Hence, φ is injective. This implies that the image of φ is nontrivial, as no two elements can both map to 0. If φ is not a surjection, then $\operatorname{Im}\varphi$ is a nontrivial submodule of W, contradicting the irreducibility of W. Hence, φ is an injection and a surjection, meaning that it is a bijection, i.e. an isomorphism. If $V \not\simeq W$, then there can be no homomorphism $\varphi : V \to W$.

In the case where they are isomorphic, this problem reduces to finding $\hom_G(V, V)$ where V is irreducible. In particular, this means that we may consider the eigenvalues and eigenvectors of elements of G. Furthermore, we note that $\mathbb{C}$ is algebraically closed. Thus, if λ is an eigenvalue of φ, then we have that

$$\ker(\varphi - \lambda I) \neq 0$$

However, we also have that $\varphi - \lambda I$ is a G-module homomorphism. But by our previous argument in the other case, we have that this is not an isomorphism. Thus, we have that

$$\varphi - \lambda I = 0$$

and thus that $\varphi = \lambda I$, giving us our 1-dimensional space. □

It follows that these character functions are orthonormal under the inner product defined above.

Theorem 8.6 (Maschke's Complete Reducibility Theorem)
If V is a G-module, and W is a submodule of V, then there exists $W' \subset V$ (a submodule) such that $V = W \oplus W'$.

Proof. We aim to find an inclusion map $i : W \mapsto V$ and the projection map $p : V \mapsto W$ such that $W' = \ker(p)$ and $p \circ i - \mathrm{id}_W$. We pick a p arbitrarily from $\hom_{\mathbb{C}}(V, W)$ and "fix" it by applying the Reynolds operator and considering $Rp \in \hom_G(V, W)$. We need only show that $(Rp) \circ i = \mathrm{id}_W$. We can see this as

$$\frac{1}{|G|}\sum_{g\in G} gpg^{-1}iw = \frac{1}{|G|}\sum_{g\in G} gpig^{-1}w = \frac{1}{|G|}\sum_{g\in G} w = w$$

Note that in this proof we relied on the fact that the character of the field does not divide the order of the group. □

As a corollary, to this (and the fact that irreducible characters are orthonormal), we can state that χ_V determines V. We can further say that if

$$V \simeq \bigoplus_{V_i \, \mathrm{irr.}} V_i$$

(a form which is always possible to obtain), the multiplicities of the V_i (i.e. lumping together isomorphic V_i) are fixed. That is, a character encapsulates multiplicities of irreducible submodules in a decomposition thereinto.

We now consider $G \circlearrowright X$ where X is a finite set. While this is not a vector space, we can make it a vector space by considering elements of g as permutation matrices on indicator functions (vectors in $\mathbb{C}^{|X|}$) for set elements. We can think of this more succinctly as $G \circlearrowright \mathbb{C}X$. If we look at the trace of a permutation matrix, we get a 1 in the ith diagonal entry iff the ith set element is fixed under the permutation. Thus, we have that

$$\chi_{\mathbb{C}G} = |X^g|$$

We can in fact prove Burnside's lemma through our existing character formulas. We have that

$$\dim(\mathbb{C}X)^G = |X^G| = X_{\mathbb{C}X}(R) = \frac{1}{|G|}\sum_{g\in G} X_{\mathbb{C}X}(g) = \frac{1}{|G|}\sum_{g} |X^g|$$

As an example, we consider $G \circlearrowright G$ by left multiplication. Again, this is equivalent to considering $G \circlearrowright \mathbb{C}G$. We look at the character of this representation:

$$\chi_{\mathbb{C}G}(g) = \begin{cases} |G| & g = 1 \\ 0 & g \neq 1 \end{cases}$$

We have that

$$\chi_{\mathbb{C}G} = \sum_{V\,\text{inv.}} (\dim V)\chi_V$$

from which it follows that

$$\mathbb{C}G \simeq \bigoplus_{V\,\text{inv.}} (\dim V) \cdot V$$