CO 430 with Karen Yeats[*]

Eason Li

2026 W

---

[*] Class has voted for oral midterm and written final.

Contents

1 Formal Power Series

Lecture 1 - Monday, January 05

We will work with integral domain in the course.

Definition 1.1. **[Ring of Formal Power Series]**

$R[[x]]$, the **ring of formal power series of** R, is the set of expansions of the form

$$\sum_{n\geq 0} a_n x^n$$

with the obvious addition and multiplication operations. We view $R \subseteq R[[x]]$ where

$$r \quad \mapsto \quad rx^0 + 0x^1 + 0x^2 + \cdots$$

Definition 1.2. **[Valuation]**

Let $A(x) = \sum_{n\geq 0} a_n x^n$. Then, the **coefficient of** x^n is

$$[x^n]A(x) = a_n$$

Also, the **valuation of** A is

$$\mathrm{val}_x A(x) = \begin{cases} \min\{n : [x^n]A(x) \neq 0\} & \text{if } A(x) \neq 0 \\ \infty & \text{if } A(x) = 0 \end{cases}$$

Proposition 1.1. Properties of val

Let $A(x), B(x) \in R[[x]]$. Then,

1. $\mathrm{val}_x A(x) = \infty$ if and only if $A(x) = 0$;
2. $\mathrm{val}_x (A(x) + B(x)) \geq \min\{\mathrm{val}_x A(x), \mathrm{val}_x B(x)\}$;
3. $\mathrm{val}_x (A(x)B(x)) = \mathrm{val}_x A(x) + \mathrm{val}_x B(x)$.

Proof. (1) is trivial;
(2) If $a_n = b_n = 0$, then $a_n + b_n = 0$, so

$$\mathrm{val}_x (A(x) + B(x)) \not< \min\{\mathrm{val}_x A(x), \mathrm{val}_x B(x)\}$$

If $a_n = -b_n$, then it is possible for the strict greater than to occur. We do not always have equality.
(3) If either $A(x)$ or $B(x)$ is 0, it is clear that we get equality. If neither are 0, we look at the $\mathrm{val}_x A(x) + \mathrm{val}_x B(x)$'th coefficient of $A(x)B(x)$. Simple computation gives $\mathrm{val}_x (A(x)B(x))$ is smaller or equal to that.

Then, looking at smaller coefficients, using the definition of $\text{val}_x A(x)$ and $\text{val}_x B(x)$, we will get that in

$$\sum_{k=0}^{i} a_{i-k} b_k$$

one of $a_{i=k}$ or b_k is 0, so the sum is 0, so we have equality. ☐

Proposition 1.2. If R is an integral domain, then $R[[x]]$ is an integral domain.

Proof. It suffices to show that if R is commutative, then $R[[x]]$ is commutative (this is obvious) and that if R has nonzero zero divisors, then so does $R[[x]]$. This is also obvious by the previous proposition on valuation. ☐

Definition 1.3.

We write

$$\frac{A(x)}{B(x)} = C(x)$$

when $A(x) = B(x)C(x)$.

Comment 1.1. If you dislike the definition of symbols indexed by counting sequences, you can also define $R[[x]]$ as the x-adic completion of $R[x]$. In this spirit, we will use val_x to define a topology on $R[[x]]$, which will be exactly the x-adic metric topology.

1.1 Convergence of Sequences of FPS

Lecture 2 - Wednesday, January 07

We now use val to define a topology on $R[[x]]$. Take $0 < \varepsilon < 1$ and define $\|A(x)\|_\varepsilon$ as

$$\|A(x)\|_\varepsilon = \begin{cases} \varepsilon^{\text{val}_x(A(x))} & \text{for } A(x) \neq 0 \\ 0 & \text{for } A(x) = 0 \end{cases}$$

also define

$$d_\varepsilon(A(x), B(x)) = \|A(x) - B(x)\|_\varepsilon$$

There are some nice properties this “norm” has.

Discovery 1.1. We obset that

$\|A(x)\|_\varepsilon \geq 0$	YES
$\|A(x)\|_\varepsilon = 0 \quad \implies \quad A(x) = 0$	YES
Triangle inequality	YES

Verification of triangle inequality. We have

$$\begin{aligned}\|A(x)+B(x)\|_\varepsilon &= \varepsilon^{\mathrm{val}_x(A(x)+B(x))} \\ &\leq \varepsilon^{\min(\mathrm{val}_x(A(x)),\mathrm{val}_x(B(x)))} \\ &= \max\left(\varepsilon^{\mathrm{val}_x(A(x))}, \varepsilon^{\mathrm{val}_x(B(x))}\right) \\ &= \max\left(\|A(x)\|_\varepsilon, \|B(x)\|_\varepsilon\right) \qquad (*) \\ &\leq \|A(x)\|_\varepsilon + \|B(x)\|_\varepsilon\end{aligned}$$

as desired. $\square$

Note 1.1. (*) tells use that $\| \bullet \|_\varepsilon$ is a nonarchimedean norm.

However, this is not a vector space norm because we cannot pull out scalars. This is still okay because we still get a metric. i.e.,

- $d_\varepsilon(A(x), A(x)) = 0$;
- If $A(x) \neq B(x)$, then $d_\varepsilon(A(x), B(x)) > 0$;
- $d_\varepsilon(A(x), B(x)) = d_\varepsilon(B(x), A(x))$;
- $d_\varepsilon(A(x), B(x)) \leq d_\varepsilon(A(x), C(x)) + d_\varepsilon(C(x), B(x))$

The point is that we now have a notion of convergence.

Definition 1.4. **[Convergence of Sequences of FPS]**

For a sequence of formal power series $A_i(x) \in R[[x]]$, we define

$$\lim_{n\to\infty} A_n(x) = A(x)$$

if and only if $\lim_{n\to\infty} \mathrm{val}_x(A_n(x) - A(x)) = \infty$ in $\mathbb{R}$. This is equivalent to $\lim_{n\to\infty} d_\varepsilon(A_n(x), A(x)) = 0$.

Comment 1.2. More informally, the sequence converges if and only if the sequence of coefficients $[x^n]A_i(x)$ eventually stablizes for all n, so in this context, $1/2^n \to 0$ is not true xddx.

Example 1.1. We have the following examples:

1. $A_i(x) = x^i$, then $\lim_{n\to\infty} A_i(x) = 0$;
2. $A_i(x) = \frac{1-x^i}{1-x}$, then $\lim_{n\to\infty} A_i(x) = \frac{1}{1-x}$;
3. $A_i(x) = \frac{1}{1-x/i}$, then $\lim_{n\to\infty} A_i(x)$ does not exist.

This makes precise that $R[[x]]$ is the complete of $R[x]$.

Also now we know when an operation on formal power series is valid by asking if it converges.

Example 1.2. We can define infinite porduct a infinite sum of formal power series by limit of partial sum/product specifically

$$\sum_{n=0}^{\infty} A_n(x) = \lim_{N\to\infty} \sum_{n=0}^{N} A_n(x)$$

$$\prod_{n=0}^{\infty} A_n(x) = \lim_{N\to\infty} \prod_{n=0}^{N} A_n(x)$$

Discovery 1.2. Compositions of formal power series is a special case of sum

$$A(B(x)) = \sum_{n=0}^{\infty} a_n B(x)^n \quad \text{where } A(x) = \sum_{n\geq 0} a_n x^n$$

which converges if and only if

$$\mathrm{val}_x(a_n B(x)^n) \longrightarrow_{n\to\infty} = 0$$

which happens if and only if a_n eventually zero (when $A(x)$ is a polynomial) or $\mathrm{val}_x(B(x)) > 0$ (when $B(x)$'s constant term is zero).

Comment 1.3. We have a notation for the set of FPS with zero constant term:

$$R[[x]]_+ = \{A(x) \in R[[x]] : \mathrm{val}_x(A(x)) > 0\} = xR[[x]]$$

Note 1.2. Multivariate situation has some subtlety. As rings,

$$R[[x,y]] = R[[x]][[y]] = R[[y]][[x]]$$

but they have different vals. For $R[[x,y]]$, we have

$$\mathrm{val}_{x,y}(A(x,y)) = \begin{cases} \min\{m+n : [x^m y^n]A(x,y) \neq 0\} & A(x,y) \neq 0 \\ \infty & A(x,y) = 0 \end{cases}$$

while for $R[[x]][[y]]$ we have val_y and for $R[[y]][[x]]$ we have val_x.

1.2 Some things you should know

Let $A(x), B(x) \in R[[x]]$, say $A(x) = \sum_{n\geq 0} a_n x^n$

Definition 1.5. **[Compositional Inverse]**

We say that $A(x), B(x)$ are compositional inverses if and only if $A(B(x)) = B(A(x)) = x$;

- Formal derivative:
$$\frac{d}{dx}A(x) = A'(x) = \sum_{n\geq 1} n a_n x^{n-1}$$
- Formal integral: If $\mathbb{Q} \subseteq R$, then
$$\int_x A(x) = \sum_{n\geq 0} \frac{a_n}{n+1} x^{n+1}$$
- We have
$$\exp(x) = \sum_{n\geq 0} \frac{1}{n!} x^n \in \mathbb{Q}[[x]]$$
- We have
$$L(x) = \log(1+x) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} x^n \in \mathbb{Q}[[x]]$$
- We have
$$B(x,y) = (1+x)^y = \sum_{n\geq 0} \binom{y}{n} x^n \in \mathbb{Q}[y][[x]]$$

Comment 1.4. See more on Kevin's videos (here is the link), More FPS operations and Special FPS. Also see more on A1.

1.3 Hensel's Lemma

Proposition 1.3. Hensel's Lemma

Let $F(t,x) \in R[[t,x]]$ and let $F'(t,x) = \frac{d}{dx}F(t,x)$. Suppose $F(0,0) := [t^0x^0]F(t,x) = 0$ and $F'(0,0)$ is invertible. Then there exists a unique $f(t) \in R[[t]]_+$ such that $F(t, f(t)) = 0$.

Proof. Build it with Newton's method in this context.
Existence: Let $f_0(t) = 0$, let
$$f_{n+1}(t) = f_n(t) - \frac{F(t, f_n(t))}{F'(t, f_n(t))}$$
We first want to show that $f_i(t) \in R[[t]]_+$ for all i, which is equivalent to showing that $f_i(t)$ has zero constant coefficient. This can be done fairly easily by induction ($f_n(t) \in R[[t]]_+$, so the fraction in the above is well-defined). Recall linear approximations:

For $A(x) \in R[[x]]$,
$$A(u) = A(v) + (u-v)A'(v) \mod (u-v)^2$$

Now we use linear approximation to show $\lim_{n\to\infty} f_n(t)$ exists:

$$F(t, f_n(t)) = \underbrace{F(t, f_{n-1}(t)) + (f_n(t) - f_{n-1}(t))F'(t, f_{n-1}(t))}_{=0 \text{ by definition of } f_n(t)} \mod (f_n(t) - f_{n-1}(t))^2$$

so $(f_n(t) - f_{n-1}(t))^2$ divides $F(t, f_n(t))$. Also by definition of f_n and invertibility of F', we know $F(t, f_n(t))$ divides $(f_n(t) - f_{n-1}(t))$. Take val,

$$\begin{aligned} & \mathrm{val}_t(f_n(t) - f_{n-1}(t))^2 \le \mathrm{val}_t F(t, f_n(t)) \le \mathrm{val}_t(f_n(t) - f_{n-1}(t)) \\ \implies \quad & 2 \cdot \mathrm{val}_t(f_n(t) - f_{n-1}(t)) \le \mathrm{val}_t F(t, f_n(t)) \le \mathrm{val}_t(f_n(t) - f_{n-1}(t)) \end{aligned}$$

where the second inequality comes from divisibility. Hence $\lim_{n\to\infty} \mathrm{val}_t(f_n(t) - f_{n-1}(t)) = \infty$ since $f_n(t) = R[[t]]_+$ (so val > 0). Therefore, the sequence converges.

> **Note 1.3.** We point out that in general, $\lim_{n\to\infty} d(x_n, x_{n-1}) = 0$ does not imply convergence (intuition: consider $f(x) = \sqrt{x}$ as an example). In our case, convergence is implied, this is a special property of this specific topology.

Now let

$$f(t) = \lim_{n\to\infty} f_n(t)$$

which exists. Now take limit of definition of f_n,

$$\begin{aligned} \lim_{n\to\infty} f_{n+1}(t) &= \lim_{n\to\infty} f_n(t) - \frac{F(t, f_n(t))}{F'(t, f_n(t))} \\ \implies \quad f(t) &= f(t) \quad - \quad \frac{F(t, f(t))}{F'(t, f(t))} \end{aligned}$$

which forces $F(t, f(t)) = 0$ as desired. For uniqueness, suppose

$$F(t, f(t)) = F(t, g(t)) = 0$$

Apply linear approximation

$$F(t, g(t)) = F(t, f(t)) + (g(t) - f(t))F'(t, f(t)) \mod (g(t) - f(t))^2$$

so $(g(t) - f(t))$ is divisible by $(g(t) - f(t))^2$. We know that $(g(t) - f(t))$ is not a unit in $R[[t]]$, so it is 0. ☐

Lecture 3 - Monday, January 12

> **Lemma 1.1.** Let $A(x) \in R[[x]]$ and $B \in R[[x]]$ be given. If $A(B(x)) = x$, then $B(A(x)) = x$.

Proof. Given $C(x) \in R[[x]]_+$, define

$$\begin{aligned} \mathrm{ev}_C : R[[x]] &\to R[[x]] \\ A(x) &\mapsto A(C(x)) \end{aligned}$$

then ev_C is a ring homomorphism and if $C \neq 0$, then it is injective since if $\sum_{n\geq 0} a_n C(x)^n = \sum_{n\geq 0} d_n C(x)^n$ and there is a minimal m such that $a_m \neq d_m$, then

$$\mathrm{val}_x(A(X(c)) - D(C(x))) = m\mathrm{val}_x(C(x)) < \infty$$

so returning to the problem at hand,

$$A(B(x)) = x$$

so we have

$$B(A(B(x))) = B(x)$$

but since ev_B is injective, so $B(A(x)) = x$. $\square$

1.3.1 One Use of Hensel's Lemma

Proposition 1.4. $A(x) \in R[[x]]_+$ has a compositional inverse if and only if $[x]A(x)$ is invertible in R.

Proof. [$\Longleftarrow$] Let $F(t,x) := t - A(x)$ and recall we define $F'(t,x) := \frac{\partial}{\partial x}F(t,x)$. Then

$$\begin{aligned} F(0,0) &= 0 + A(0) = 0 \\ F'(0,0) &= A'(0) \quad \text{is invertible} \end{aligned}$$

Apply Hensel's Lemma (1.3) to get $B(t) \in R[[t]]_+$ such that $F(t, B(t)) = t - A(B(t)) = 0$, so $A(B(t)) = t$. By the above lemma, we know that A and B are compositional inverses.

[$\Longrightarrow$] Given, suppose the compositional inverse of A is B, so $B(A(x)) = x$, and hence

$$[x]B(A(x)) = b_1 a_1 = 1$$

and this implies that a_1 is invertible. $\square$

2 Other Formal Series

2.1 Formal Laurent Series

Definition 2.1. **[Ring of Formal Laurent Series]**

The **ring of formal Laurent series** $R((x))$ over R is

$$R((x)) = \left\{ \sum_{n \geq N} a_n x^n : N \in \mathbb{Z}, \quad a_n \in R \right\}$$

with addition and multiplication as you would expect:

$$\sum_{n \geq N_1} a_n x^n + \sum_{n \geq N_2} b_n x^n = \sum_{n \geq \min(N_1, N_2)} (a_n + b_n) x^n$$

where by convention $a_n = 0$ for $n < N_1$ and likewise for b_is. For multiplication:

$$\left(\sum_{n \geq N_1} a_n x^n \right) \left(\sum_{n \geq N_2} b_n x^n \right) = \sum_{n \geq N_1 + N_2} \left(\sum_{i \in \mathbb{Z}} a_i b_{n-i} \right) x^n$$

note that the sum in the parenthesis on the RHS is indeed a finite sum.

Coefficient extraction works just as before, but there is a special one:

Definition 2.2. **[Formal Residue]**

For $A(x) \in R((x))$, we define the **formal residue operator** as

$$[x^{-1}]A(x)$$

Comment 2.1. val also still works in the ring of Laurent series, but note that it can now be negative.

As a result, $\|\bullet\|_\varepsilon$ and d_ε also works just as before.

Note 2.1. If F is a field, then so is $F((x))$.

Discovery 2.1. Two cautions:

1. The reason why we bound the negative degree is because it's badly behaved if you don't:

$$\begin{aligned} \sum_{n \in \mathbb{Z}} x^n &= \sum_{n \geq 0} x^n + \sum_{n < 0} x^n \\ &= \frac{1}{1-x} + \frac{x^{-1}}{1 - x^{-1}} = 0 \end{aligned}$$

Discovery 2.2. Second caution:

2. We note
$$R[[x]][[y]] = R[[y]][[x]]$$
but in Laurent context the different vals really matter:
$$R((x))((y)) \neq R((y))((x))$$
Here is an example: In $\mathbb{Q}((x))((y))$
$$\begin{aligned}(x+y)^{-1} = \frac{\frac{1}{x}}{1+\frac{y}{x}} &= \frac{1}{x}\sum_{n\geq 0}\left(-\frac{y}{x}\right)^n \\ &= \sum_{n\geq 0}(-1)^n y^n x^{-n-1}\end{aligned}$$
but in $\mathbb{Q}((y))((x))$, we have
$$(x+y)^{-1} = \sum_{n\geq 0}(-1)^n x^n y^{-n-1}$$
which are completely different.

2.1.1 Lagrange Implicit Function Theorem

With all of this in hand, we can not state the Lagrange implicit function theorem, also known as LIFT.

Theorem 2.1. Lagrange Implicit Function Theorem (LIFT)

Suppose $\mathbb{Q} \subseteq R$ and $\phi(\lambda) \in R[[\lambda]]$ is invertible. Then there exists a unqiue $A(x) \in R[[x]]_+$ such that

- $A(x) = x\phi(A(x))$;
- $[x^n]A(x) = \frac{1}{n}[\lambda^{n-1}]\phi(\lambda)^n$ for $n \geq 1$;
- for any $f(\lambda) \in R((\lambda))$,
$$[x^n]f(A(x)) = \frac{1}{n}[\lambda^{n-1}]f'(\lambda)\phi(\lambda)^n$$
for $n \neq 0$. When $n = 0$,
$$[x^0]f(A(x)) = [\lambda^0]f(\lambda) + [\lambda^{-1}]f'(\lambda)\log\left(\frac{\phi(\lambda)}{\phi(0)}\right)$$
where the second term in the sum is 0 if $\text{val}_\lambda f(\lambda) \geq 0$.

Comment 2.2. See 330 notes for a proof using formal Laurent series and the formal residue operator. See notes from next week (see appendix B) for a combinatorial proof.

2.2 Formal Dirichlet Series

Definition 2.3. **[Ring of Formal Dirichlet Series]**

The **ring of formal Dirichlet series** over R is the set of expressions of the form

$$\sum_{n\geq 1} a_n n^{-s}$$

for some constant s, with addition and multiplication:

$$\sum_{n\geq 1} a_n n^{-s} + \sum_{n\geq 1} b_n n^{-s} = \sum_{n\geq 1} (a_n + b_n) n^{-s}$$
$$\left(\sum_{n\geq 1} a_n n^{-s}\right) + \left(\sum_{n\geq 1} b_n n^{-s}\right) = \sum_{n\geq 1} \left(\sum_{k|n} a_k b_{n/k}\right) n^{-s}$$

Discovery 2.3. Formal Dirichlet series are useful for counting things where the weight of a product is the product of the weights in contrast to formal power series as generating series where weight of product is the sum of weights.

2.2.1 Motivating Example: Riemann Zeta Function

Recall that the Riemann zerta function is

$$\begin{aligned}\zeta(s) &= \sum_{n\geq 1} n^{-s} \\ &= \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}\end{aligned}$$

This is the Dirichlet generating series of positive integers.

2.3 Other Examples of Kinds of Formal Series

Other examples of kinds of formal series including series in non-commuting variables. There, you get words as your monomials.

3 Ordinary Generating Series (Ordinary Generating Functions)

Definition 3.1. **[Combinatorial Class]**

A **combinatorial class** $\mathcal{A}$ is a finite or countably infinite set with a weight function

$$w : \mathcal{A} \to \mathbb{Z}_{\geq 0}$$

such that $\mathcal{A}_n = \{a \in \mathcal{A} : w(a) = n\}$ is finite for all n. The counting sequence of $\mathcal{A}$ is

$$a_0, a_1, a_2, \ldots$$

where $a_i = |\mathcal{A}_i|$.

Definition 3.2. **[Ordinary Generating Series (OGF)]**

The **ordinary generating series** of a combinatorial class $\mathcal{A}$ is

$$A(x) = \sum_{a \in \mathcal{A}} x^{w(a)} = \sum_{n \geq 0} a_n x^n \in \mathbb{Z}[[x]]$$

We can also have a multivariate series. Suppose we have a number of weight functions

$$w_1 : \mathcal{A} \to \mathbb{Z}_{\geq 0}, \qquad w_2 : \mathcal{A} \to \mathbb{Z}_{\geq 0}, \qquad \text{etc.}$$

then the ordinary generating series is

$$A(x_1, x_2, \ldots) = \sum_{a \in \mathcal{A}} x_1^{w_1(a)} x_2^{w_2(a)} \cdots$$

Question 3.1.

What finiteness condition do I need have?

We need to have

$$\mathcal{A}_{n_1, n_2, \ldots} \coloneqq \{a \in \mathcal{A} : w_i(a) = n_i\}$$

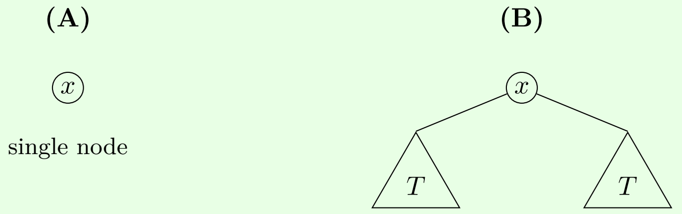
is finite for all $n_1, n_2, \ldots$. We also need that for each a, there are only finitely many nonzero weights (usually suffices to have finitely many variables).

Discovery 3.1. We have a general framework for enumeration (symbolic method):

- point down $\mathcal{A}$ and w;
- find a way to decompose $\mathcal{A}$ so that the elements of $\mathcal{A}$ are broken into smaller or simpler pieces (these decompositions must be bijetive and weight preserving);
- if weight behaves appropriately across out standard operators, then connect decompositions into equation for generating series;
- solve/ extract coefficients/ manipulate etc. to solve problem at hand.

Binary Trees

Example 3.1. Let the class of binary trees be denoted as $\mathcal{T}$. Recall that binary trees are rooted trees, and every vertex in a binary tree has 0 or 2 children (denoted as L, R), each of which is also a binary tree. For weight w, we use the number of vertices. We observe that a binary tree is either a single node (the root), or a root with two children of non-empty binary trees,



single node

root + two (ordered) non-empty subtrees

hence we know that

$$\mathcal{T} \coloneqq \{\bullet\} \sqcup \{\bullet\} \times \mathcal{T} \times \mathcal{T}$$

which yields us the following equation related to its ordinary generating series:

$$T(x) = x + xT(x)^2$$

Lecture 4 - Wednesday, January 14

3.1 Operations on Combinatorial Classes Part 1

Let $\mathcal{A}$ and $\mathcal{B}$ be combinatorial classes with weight functions $w_{\mathcal{A}}$ and $w_{\mathcal{B}}$ respectively.

Disjoint Union: The disjoint union $\mathcal{A} \cup \mathcal{B}$ (with $\mathcal{A} \cap \mathcal{B} = \varnothing$) is a combinatorial class with weight function

$$w : \mathcal{A} \cup \mathcal{B} \to \mathbb{Z}_{\geq 0}$$

ginen by $w|_{\mathcal{A}} = w_{\mathcal{A}}$ and $w|_{\mathcal{B}} = w_{\mathcal{B}}$ and with ordinary generating series $A(x) + B(x)$.

Cartesian Product: The Cartesian product $\mathcal{A} \times \mathcal{B}$ is a combinatorial class with weight function

$$w : \mathcal{A} \times \mathcal{B} \to \mathbb{Z}_{\geq 0}$$

given by $w((a, b)) = w_{\mathcal{A}}(a) + w_{\mathcal{B}}(b)$ and with ordinary generating series $A(x)B(x)$.

Sequence (Kleene Star): The sequence operator is defined as

$$\mathcal{A}^* = \bigcup_{k \geq 0} \mathcal{A}^k$$

which is a combinatorial class with weight function $w((a_1, a_2, \ldots, a_k)) = w_{\mathcal{A}}(a_1) + \cdots + w_{\mathcal{A}}(a_k)$, and the ordinary generating series is $\dfrac{1}{1 - A(x)}$.

Note 3.1. We must have $\mathcal{A}_0 = \varnothing$ to take that star of $\mathcal{A}$.

Example 3.2. Binary Strings

Recall that the class of binary strings is $\mathcal{B} = \{0,1\}^*$ where the weight of a binary string is the length of it. We have

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

Example 3.3. Binary Strings

Suppose $\mathcal{T}$ is the class of binary trees, recall that we have

$$T(x) = x + xT(x)^2$$

Remember that the set up for LIFT 2.1 is $A(x) = x\phi(A(x))$. In this specific case, we have $\phi(\lambda) = 1+\lambda^2$, and so by LIFT 2.1,

$$[x^n]T(x) = \frac{1}{n}[\lambda^{n-1}](1+\lambda^2)^n = \begin{cases} \frac{1}{n}\binom{n}{(n-1)/2} & n \text{ is odd} \\ 0 & n \text{ is even} \end{cases}$$

Example 3.4. Plane Trees

Recall that plane trees are ordered rooted trees, i.e., at each vertex, children are ordered. Hence

$$P \leftrightharpoons \{\bullet\} \times P^*$$

and so the ordinary generating series is $P(x) = \dfrac{x}{1-P(x)}$. Recall that the set up for LIFT 2.1 is $A(x) = x\phi(A(x))$. Here we have $\phi(\lambda) = \dfrac{1}{1-\lambda}$, so

$$\begin{aligned} [x^n]P(x) &= \frac{1}{n}[\lambda^{n-1}]\left(\frac{1}{1-\lambda}\right)^n && \text{for } n \geq 1 \\ &= \frac{1}{n}\binom{2n-2}{n-1} && \text{for } n \geq 1 \end{aligned}$$

which is Catalan!

Example 3.5. Let's consider the class of binary trees again, but this time we are counting the number of leaves. We have the same bijection, but translating to a different ordinary generating series:

$$\tilde{T}(x) = x + x^0 \cdot \tilde{T}(x)^2$$

We notice that this is the same ordinary general series as the class of plane trees.

Exercise 3.1. Since the counting sequences of plane trees counted by number of vertices and binary trees counted by number of leaves are the same, there must be a weight preserving bijection between them. Find (one of) the bijection(s).

We can also use a 2-variable generating series to count plane trees, denoted as $\mathcal{P}$, with respect to both the number of vertices and the number of leaves,

$$P(x, y) = \sum_{t \in \mathcal{P}} x^{\#\text{vers}(t)} y^{\#\text{leaf}(t)}$$

We know that we have

$$\begin{aligned} \mathcal{P} &\risingdotseq \{\bullet\} \times \mathcal{P}^* \\ &\risingdotseq \{\bullet\} \cup \{\bullet\} \times \mathcal{P} \times \mathcal{P}^* \end{aligned}$$

> **Comment 3.1.** The reason why we prefer the second form is because we do not know whether the root is a leaf in the first form.

Therefore, we have

$$P(x, y) = xy + x \cdot \frac{P(x, y)}{1 - P(x, y)}$$

Apply LIFT 2.1 in x we obtain

$$[x^n]P(x, y) = \frac{1}{n}[\lambda^{n-1}] \left(y + \frac{\lambda}{1 - \lambda}\right)^n \qquad \text{for } n \geq 1$$

Furthermore,

$$\begin{aligned} [x^n y^k]P(x, y) &= [y^k] \left[\frac{1}{n}[\lambda^{n-1}] \left(y + \frac{\lambda}{1 - \lambda}\right)^n\right] & \text{for } n \geq 1 \\ &= \frac{1}{n}[\lambda^{n-1}] \left[[y^k] \left(y + \frac{\lambda}{1 - \lambda}\right)^n\right] & \text{for } n \geq 1 \\ &= \frac{1}{n}[\lambda^{n-1}] \binom{n}{k} \left(\frac{\lambda}{1 - \lambda}\right)^{n-k} & \text{for } n \geq 1 \\ &= \frac{1}{n}\binom{n}{k}[\lambda^{n-1}] \frac{\lambda^{n-k}}{1 - \lambda^{n-k}} \\ &= \frac{1}{n}\binom{n}{k}[\lambda^{k-1}](1 - \lambda)^{-n+k} = \frac{1}{n}\binom{n}{k}\binom{n-2}{k-1} \end{aligned}$$

3.1.1 Counting with Dirichlet Generating Series

Recall the definition of Formal Dirichlet series: Definition 2.3.

> **Comment 3.2.** There is no object of weight 0 in the Dirichlet generating series.

The Dirichlet generating series of a combinatorial class $\mathcal{A}$ is

$$\sum_{a \in \mathcal{A}} w(a)^{-s} = \sum_{n \geq 1} a_n n^{-s}$$

Lemma 3.1. Suppose $\mathcal{A}$ has a multiplicative weight function and unique factorization into indecomposible elements, let $\mathcal{P}$ denote the set of indecomposibles, then

$$\sum_{n\geq 1} a_n n^{-s} = \prod_{n\geq 2}(1-n^{-s})^{-\mathcal{P}_n}$$

Proof. We define

$$\mathrm{val}_s\left(\sum_{n\geq 1} a_n n^{-s}\right) = \{m : a_m \neq 0\}$$

then can define convergence as before (and again the intuition is that convergence is when the coefficients stabilize). Then the product is convergent. Then multiply out and interpret (in finite partial product and take limit if you like)

$$\prod_{n\geq 2}(1-n^{-s})^{-\mathcal{P}_n} = \prod_{p\in\mathcal{P}}(1-w(p)^{-s})^{-1}$$

where $(1-w(p)^{-s})^{-1}$ is the Dirichlet generating series for $\{p\}^*$. Hence

$$\begin{aligned}\prod_{n\geq 2}(1-n^{-s})^{-\mathcal{P}_n} &= \prod_{p\in\mathcal{P}}(1-w(p)^{-s})^{-1}\\ &= \sum_{p_1^{k_1}p_2^{k_2}\cdots,p_i\in\mathcal{P}} \left(w(p_1)^{k_1}w(p_2)^{k_2}\cdots\right)^{-s}\\ &= \sum_{p_1^{k_1}p_2^{k_2}\cdots,p_i\in\mathcal{P}} \left(w(p_1^{k_1}p_2^{k_2}\cdots)\right)^{-s}\\ &= \sum_{a\in\mathcal{A}} w(a)^{-s}\end{aligned}$$

by unique factorization. □

Example 3.6. Suppose $\mathcal{A} = \mathbb{Z}_{\geq 1}$ and $\mathcal{P}$ is the set of primes, then the above tells us that

$$\zeta(s) \coloneqq \sum_{n\geq 1} n^{-s} = \prod_{p\text{ prime}}(1-p^{-s})^{-1}$$

Example 3.7. Consider finite abelian groups whose weight is their sizes. Let a_n be the number of them of size n. The indecomposible finite cyclic groups are for each power of prime p^n, so

$$\sum_{n\geq 1} a_n n^{-s} = \prod_{p\text{ prime},n\geq 1}(1-(p^n)^{-s})^{-1}$$

3.2 More Operations for Combinatorial Classes

3.2.1 Multiset

Suppose I don't want my rooted trees to have any order structure on the subtrees at a vertex.

Question 3.2.

How do I count them?

Instead of a sequence of subtrees at each vertex, I have a *multiset* of subtrees at each vertex. Hence the operation we want to build is multiset, denoted as MSet.

The Euler product construction gives the idea, but now in a formal power series context.

For $\mathsf{MSet}(\mathcal{A})$, we choose the convenience order so that all elemetns of weight 0 comes first, then order elements of weight 1, and then weight 2, etc.

Discovery 3.2. Therefore, for class $\mathcal{A}$ such that $\mathcal{A}_0 = \varnothing$ (we will see why below), we have

$$\mathsf{MSet}(\mathcal{A}) \rightleftharpoons \prod_{a \in \mathcal{A}} \{a\}^*$$

where the convenience order is implicit in the product.

Let $B(x)$ be the generating series of

$$\mathcal{B} = \mathsf{MSet}(\mathcal{A})$$

so we know that

$$\begin{aligned} B(x) &= \prod_{a \in \mathcal{A}} \frac{1}{1 - x^{w_A(a)}} \\ &= \prod_{n \geq 1} \left(\frac{1}{1 - x^n} \right)^{a_n} \end{aligned}$$

Note 3.2. Since the denominator is $1 - x^n$, we must have that $\mathcal{A}_0 = \varnothing$, and hence the index in the above product start at $n = 1$.

Rewrit the generating series a bit more,

$$\begin{aligned} B(x) = \prod_{n \geq 1} \frac{1}{(1 - x^n)^{a_n}} &= \exp\left(\log\left(\prod_{n \geq 1} \frac{1}{(1 - x^n)^{a_n}} \right) \right) \\ &= \exp\left(\sum_{n \geq 1} a_n \log((1 - x^n)^{-1}) \right) \\ &= \exp\left(\sum_{n \geq 1} a_n \sum_{i \geq 1} \frac{x^{ni}}{i} \right) \\ &= \exp\left(\sum_{i \geq 1} \frac{1}{i} \sum_{n \geq 1} a_n x^{in} \right) = \boxed{\exp\left(\sum_{i \geq 1} \frac{1}{i} A(x^i) \right)} \end{aligned}$$

which is the MSet formula.

Example 3.8. Going back to the question in hand. Let $\mathcal{T}$ be the class of rooted trees without the plane structure, so we have

$$\mathcal{T} = \{\bullet\} \times \mathsf{MSet}(\mathcal{T})$$

and hence the generating series for the class is

$$T(x) = x \cdot \exp\left(\sum_{i\geq 1} \frac{T(x^i)}{i}\right)$$

3.2.2 Set

Lecture 7 - Wednesday, January 28

What if we want to take a set (of distinct elements) from a class $\mathcal{A}$, that is,

$$\mathcal{B} = \mathsf{Set}(\mathcal{A}), \qquad \mathcal{B} \leftrightharpoons \prod_{a\in\mathcal{A}} \{\varepsilon, a\}$$

As a result, we obtain the generating series of $\mathcal{B}$ is

$$\begin{aligned} B(x) &= \prod_{a\in\mathcal{A}} (1 + x^{w(a)}) \\ &= \prod_{n=1}^{\infty} (1 + x^n)^{a_n} \\ &= \exp\left(\log\left(\prod_{n\geq 1} (1+x^n)^{a_n}\right)\right) \\ &= \exp\left(\sum_{n\geq 1} a_n \sum_{i\geq 1} (-1)^{i-1} \frac{x^{ni}}{i}\right) = \boxed{\exp\left(\sum_{i\geq 1} (-1)^{i-1} \cdot \frac{1}{i} \cdot A(x^i)\right)} \end{aligned}$$

3.2.3 Cycle

If we wanted a cycle of objects from $\mathcal{A}$, say $\mathcal{B} = \mathsf{Cyc}(\mathcal{A})$, then

$$B(x) = \sum_{i\geq 1} \frac{\phi(i)}{i} \log\left(\frac{1}{1 - A(x^i)}\right)$$

where ϕ is the Euler ϕ function. We will prove this (optional) in assignment 2.

3.2.4 Pointing (Rooting)

Another operation is called pointing, or sometimes called rooting[1].

Suppose we have a combinatorial class $\mathcal{C}$ which is built of atoms. i.e., it has a decomposition in terms of operations we have seen, involving recursive appearances of itself and explicit elements of weight 0 and 1. The ones of weight 1 are the atoms. Note that a system of equations is also possible where they recursively refer to each other and elements of weight 0 and/or 1.

Example 3.9. Any classes of trees we have seen are in this form, where we consider the weight to be the number of vertices.

Example 3.10. Binary strings, and the atom-counting weight is the length of the string.

We define $\mathcal{C}^\bullet$ a combinatorial class whose elements are pairs of an element of $\mathcal{C}$ and a choice of an atom from that element. We usually write this by drawing the element of $\mathcal{C}$ and marking the selected atom with an arrow or circled, coloured.

Theorem 3.1.

For $\mathcal{C}^\bullet$ defined as above, we have

$$C^\bullet(x) = \frac{x\mathrm{d}}{\mathrm{d}x}C(x)$$

Proof. We know that

$$C^\bullet(x) = \sum_{(c,z):c\in\mathcal{C},z\in c} x^{w(c)} = \sum_{c\in\mathcal{C}} w(c)x^{w(c)} = \frac{x\mathrm{d}}{\mathrm{d}x}C(x)$$

as simple as it is. □

Comment 3.3. As on LIFT activity, composing with $C(x)$, e.g., $C(A(x))$, is putting a $\mathcal{C}$-structure on your objects. This is refered as "wigilization".

---

[1] We will revisit the formulation here when we get to species.

3.3 Some Concrete Examples

3.3.1 Non-crossing Rooted Chord Diagrams[2]

Let the class of non-crossing rooted chord diagram counted by the number of chords be $\mathcal{N}$. We decompose the chord diagram into two parts, one enclosed under the chord on the rooted node, and one outside. Algebraically,

$$\mathcal{N} \leftrightharpoons \{\text{the chord}\} \times \mathcal{N} \times \mathcal{N} \cup \{\varepsilon\}$$

Hence the generating series of the class is

$$N(x) = 1 + xN(x)^2$$

> **Note 3.3.** This is Catalan (again).

3.3.2 Permuations

Consider a permutation of $\{1, \dots, n\}$ in the 1-line notation. For instance,

$$4123$$

We can view this as selecting a slot for inserting n into the 1-line notation and a permutation on $\{1, \dots, n-1\}$. Note, we can see this as pointing, but there is something subtle, that is, there are one more slots than the number of atoms. This suggests that we need to adjust by adding an extra "dummy" atom at the end, then we have one slot per atom, so

$$\mathcal{P} \leftrightharpoons \{\varepsilon\} \cup (\mathcal{P} \times \{\bullet\})^{\bullet}$$

Note that this extra atom also accounts for the weight of the inserted number. In terms of the generating series, we have

$$P(x) = 1 + \frac{x\mathrm{d}}{\mathrm{d}x}(xP(x))$$

Equivalently, take coefficient of x^n for $n > 0$ we have

$$P_n = nP_{n-1}$$

where $P_0 = 1$. This is factorial!, as we already knew.

---

[2] They are really just matchings on $\{1, \dots, 2n\}$.

3.4 Transfer Matrix Mathod

Oddly enought, this starts with a little bit of algebraic graph theory. Suppose we have a digraph D on v vertices and a weight associated to each arc (use 1 if you don't like weights). Its adjacency matrix is a matrix A whose entries $a_{ij} =$ (the weight from node i to node j), 0 if there is no arc.

Proposition 3.1. The (i, j)th entry of A^n is the sum of the weights of all directed walks of n arcs from vertex i to vertex j, where the weight of a walk is the product of the weights of its arcs.
In particular, with all weights set to 1, this is just the number of directed walks of n arcs from i to j.

Proof. The proof is essentially just definition of matrix multiplication. We compute

$$(A^n)_{ij} = \sum_{1 \le i_1, \ldots, i_{n-1} \le n} A_{ii_1} A_{i_1 i_2} \cdots A_{i_{n-1} j}$$

If no walk between $i, i_1, i_2, \ldots, i_{n-1}, j$, then at least one of the $A_{i_\ell i_{\ell+1}}$ term is 0. Else if there is such a walk, then the product is just the product of the weights of the arcs in the walk. The result follows. ☐

The above result is nice for enumeration. Let $D_{ij}(x)$ be the ordinary generating series for the walk from i to j counted by length. Then

$$\begin{aligned} D_{ij}(x) &= \sum_{n \ge 0} (A^n)_{ij} x^n \\ &= \left(\sum_{n \ge 0} A^n x^n \right)_{ij} \\ &= \left((1 - Ax)^{-1} \right)_{ij} \end{aligned}$$

Question 3.3.

Why does the last transition make sense?

Proof. We verify:

$$(1 - Ax) \sum_{n=0}^{N} A^n x^n = 1 - A^{N+1} x^{N+1}$$

so $\mathrm{val}_x \to \infty$, so the limit is 1, and so $1 - Ax$ is invertible. ☐

Lecture 8 - Monday, February 02

We notice that the above is a well set up for Cramer's Rule:

$$\begin{aligned} D_{ij}(x) &= \left((1 - Ax)^{-1} \right)_{ij} \\ &= \boxed{(-1)^{i+j} \frac{\det(1 - Ax; i, j)}{\det(1 - Ax)}} \end{aligned}$$

where $\det(1 - Ax; i, j)$ is the matrix $1 - Ax$ removing the ith row and jth column.

Note 3.4. The denominator is the characteristic polynomial. This tells us that D_{ij} is a rational function of x, and further as characteristic polynomials, we know things about them from eigenvalues.

Note 3.5. Closed walks, i.e., walks whose $i = j$, are particularly nice as $D_{ii}(x)$ is a diagonal entry, so all closed walks regardless of the start point is a sum of diagonal entries, i.e., a trace.

Let $C(x)$ be the generating series of closed walks (excluding the empty walks), so

$$\begin{aligned} C(x) &= \sum_{n\geq 1}\sum_{i=1}^{v} A_{ii}^{n} x^{n} \\ &= \sum_{n\geq 1} (\mathrm{tr} A^{n}) x^{n} \end{aligned}$$

If $\lambda_1, \ldots, \lambda_d$ are the non-zero eigenvalues of A, then

$$\mathrm{tr}(A^n) = \lambda_1^n + \cdots + \lambda_d^n$$

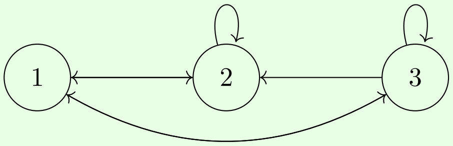
As a result, we have

$$C(x) = \sum_{n\geq 1} (\lambda_1^n + \cdots + \lambda_d^n) x^n = \frac{\lambda_1 x}{1-\lambda_1 x} + \cdots + \frac{\lambda_d x}{1-\lambda_d x}$$

Put them on a common denominator,

$$\boxed{C(x) = \frac{-\frac{x\mathrm{d}}{\mathrm{d}x}\det(1-Ax)}{\det(1-Ax)}}$$

Example 3.11. Ternary string on $\{1, 2, 3\}$ with no 11 or 23



$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Now we can compute any of these generating series we want (just small matrix determinants). For instance, the generating series of all walks is

$$\sum_{1\leq i,j\leq 3} D_{ij}(x) = \frac{\sum_{1\leq i,j\leq 3} (-1)^{i+j} \det(1-Ax; i, j)}{\det(1-Ax)}$$

we could compute this directly, which would yield us

$$\frac{3+x-x^2}{1-2x-x^2+x^3}$$

but easier, we could first calculate the characteristic polynomial $\det(1-Ax)$, then note that the degree of the numerator is at most one less than the degree of the denominator (2 in this case), so we could put a generic polynomial of that degree in the numerator and compute the initial terms.

Comment 3.4. If you like, you can have a fake start vertex. Then just turning the crank, we do not need the sum in computing the generating series for all walks anymore as the series for that vertex naturally computes it.

Exercise 3.2. What if we counted closed walks on D, what kind of strings do these generate.

Solution. If we count the bit as we hit the vertex, we get ternary strings starting and ending with the same bit and avoiding 11, 23. If instead, we interpret as counting the bit as we traverse the edge, putting on the number as we leave the vertex. As a subexample, say we start at 3, and then do end at 3, then the place we left to get back to 3 at the end cannot be 2 because there is no 23 edge. Therefore, there is no strings of the form

$$3 \dots 2,$$

and likewise, there exist no strings of the form

$$1 \dots 1.$$

These words cyclically avoid 11 and 23, so the generating series is

$$\frac{-\frac{x\mathrm{d}}{\mathrm{d}x}\det(1-Ax)}{\det(1-Ax)} = \frac{2x+2x^2-3x^3}{1-2x-x^2+x^3}$$

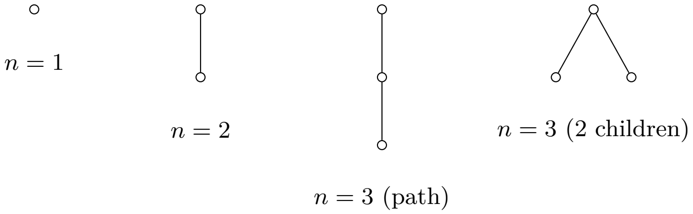
☐

Anytime your objects can be built step by step, where constant number of previous steps determines what's possible next, then you can use the transfer matrix to attack.

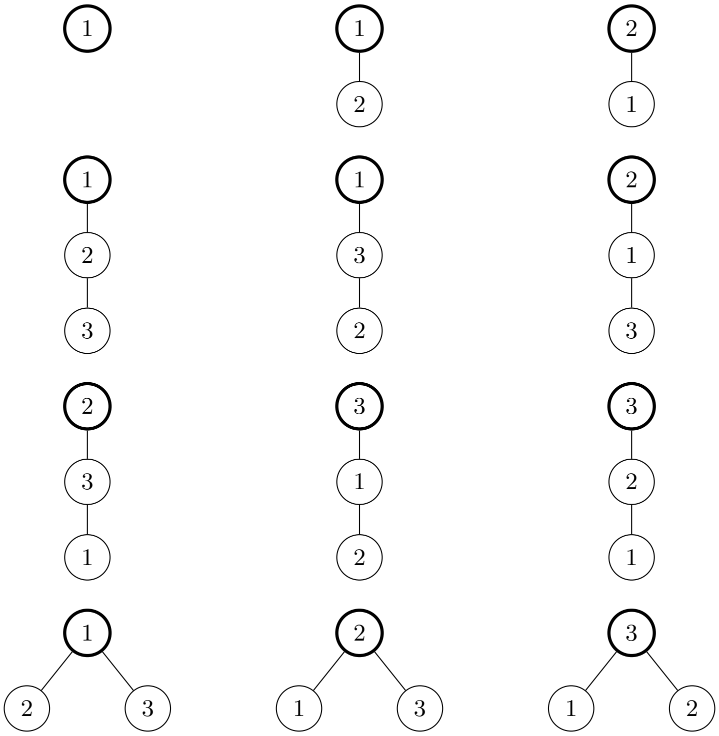
4 Exponential Generating Series

4.1 Labelled Objects Naively

Let's first consider rooted trees with no order or plane structure. Here are some examples,

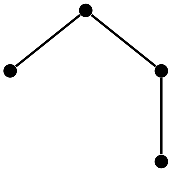


You may also have encountered labelled rooted trees:



so each vertex gets a label from 1 to # of vertices such that no two vertices get the same label. Here are some questions we want to ask:

How many labellings of the following tree are there?



Answer. $4! = 24$. □

What is the maximum number of labellings of a rooted tree on n vertices?

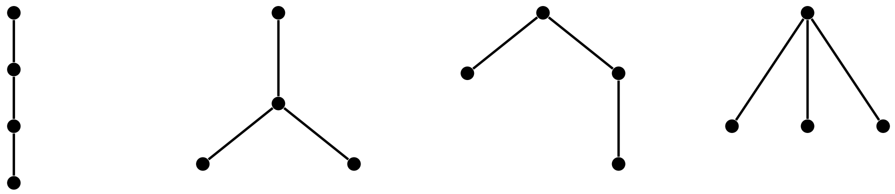
Answer. $n!$. □

What is the minimum number of labellings of a rooted tree on n vertices?

Answer. n, the special case is when there is one root and all other vertices are leaves, and you cannot do any better than this because the root is always distinguishable. □

How many labelled rooted trees on 4 vertices?

Answer. The following are the four "different" types of trees on four vertices we could put labels on:



We have $24 + 12 + 24 + 4 = 64$. □

We can also have less obvious representation of objects we are familiar with as labelled objects, for example, permutations. The correct notion of generating series for labelled objects is the exponential generating series.

Definition 4.1. **[Exponential Generating Series]**

Let $\mathcal{A}$ be a labelled combinatorial class, then

$$A(x) = \sum_{a \in \mathcal{A}} \frac{x^{w(a)}}{w(a)!} = \sum_{n \geq 0} \frac{a_n}{n!} x^n$$

which is a **exponential generating series**.

Note 4.1. In the above exponential generating series, the coefficients could be rational, i.e.,

$$A(x) \in \mathbb{Q}[[x]]$$

Example 4.1. Let $\mathcal{S}$ be the class of permutations as labelled objects, then

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

which is convenient.

Example 4.2. Let $\mathcal{C}$ be directed (non-empty) cycles as labelled objects, how many labelled cycles on n nodes are there? We could easily find that the answer is $n!/n$. Hence

$$C(x) = \sum_{n\geq 1} \frac{(n-1)!}{n!} x^n = \sum_{n\geq 1} \frac{x^n}{n} = \log\left(\frac{1}{1-x}\right)$$

Note 4.2. We notice that $S(x) = \exp(C(x))$. This is not a coincidence, we will revisit this later.

The above was pass 1, for pass 2, let us return to "built of atoms". In this context, we label the atoms. More formally speaking, we define labelling as following:

Definition 4.2. **[Labelling]**

Given $\mathcal{A}$ an unlabelled combinatorial class that is built of atoms, a **labelling** of $a \in \mathcal{A}$ is a bijection from the atoms of a to $\{1, \ldots, w(a)\}$.

4.1.1 Operations on Labelled Classes

Lecture 9 - Wednesday, February 04

Let's think about operations on labelled classes now.

Proposition 4.1. Disjoint union for labelled classes gives sum of exponential generating series, i.e.,

$$\sum_{n\geq 0} \frac{a_n}{n!} x^n + \sum_{n\geq 0} \frac{b_n}{n!} x^n = \sum_{n\geq 0} \frac{a_n + b_n}{n!} x^n$$

What about the product? We don't want Cartesian product of labelled classes. Suppose $\mathcal{A}$ is the class of labelled rooted trees and $\mathcal{B}$ is the class of permutations, so an example element from $\mathcal{A} \times \mathcal{B}$ would be

$$\left(\ \boxed{1}\!-\!\boxed{2},\ \ 132\ \right)$$

which has 5 atoms but the labelling is not in bijection with $\{1, 2, 3, 4, 5\}$ because of the existence of two copies of 1 and two copies of 2. Hence we need labels for (a, b) from $w(a) + w(b)$, so we need to distribute the labels between a and b in such a way as to preserve the relative order of the labels in a, and in b, separately encoded. We say

- a **weak labelling** of $a \in \mathcal{A}$ is an injection from the atoms of a to $\mathbb{Z}_{\geq 1}$;
- an **order preserving relabelling** of a labelled object $a \in \mathcal{A}$ is a weak labelling of a so that for any two atoms z_1, z_2 of a,

$$z_1 \leq z_2 \texttt{ in original labelling} \quad \Longleftrightarrow \quad z_1 \leq z_2 \texttt{ in the new weak labelling}$$

- Given $S \subseteq \mathbb{Z}_{\geq 1}$ with $|S| = w(a)$, there exists a unique order preserving relabelling of a using the labels of S.

Definition 4.3. **[Labelled Product]**

Let $\mathcal{A}$ and $\mathcal{B}$ be labelled classes, then the **labelled product** is

$$\mathcal{A} * \mathcal{B} = \bigcup_{n \geq 0} \left(\bigcup_{S \subseteq \{1, \ldots, n\}} \left\{ (a_S, b_{[n] \setminus S}) : a \in \mathcal{A}_{|S|}, b \in \mathcal{B}_{n-|S|} \right\} \right)$$

and these unions are disjoint.

Comment 4.1. These are objects who form the union

$$\bigcup_{\substack{a \in \mathcal{A} \\ b \in \mathcal{B}}} (a * b)$$

whose size is $|(a * b)| = w(a) + w(b)$.

Proposition 4.2. The exponential generating series of $\mathcal{C} := \mathcal{A} * \mathcal{B}$ is

$$C(x) = A(x)B(x)$$

Proof. We have

$$\begin{aligned}
\sum_{n \geq 0} \sum_{S \subseteq [n]} \sum_{(a_S, b_{[n] \setminus S})} \frac{x^{w(a)+w(b)}}{(w(a)+w(b))!} &= \sum_{n \geq 0} \sum_{k=0}^{n} \sum_{\substack{S \subseteq [n] \\ |S| = k}} \sum_{(a_S, b_{[n] \setminus S})} \frac{x^n}{n!} \\
&= \sum_{n \geq 0} \sum_{k=0}^{n} \frac{x^n}{n!} \binom{n}{k} a_k b_{n-k} \\
&= \sum_{n \geq 0} \sum_{k=0}^{n} \frac{x^k x^{n-k}}{n!} \cdot \frac{n!}{k!(n-k)!} a_k b_{n-k} \\
&= \left(\sum_{n \geq 0} \frac{a_n}{n!} x^n \right) \left(\sum_{n \geq 0} \frac{b_n}{n!} x^n \right)
\end{aligned}$$

as desired. ☐

Therefore, labelled product of labelled combinatorial classes gives the product of exponential generating series. Now we think about sequence. We get sequence (Kleene star) as before, but with labelled product at each step:

$$\mathcal{A}^* = \{\varepsilon\} \cup \mathcal{A} \cup \mathcal{A} * \mathcal{A} \cup \mathcal{A} * \mathcal{A} * \mathcal{A} \cup \cdots$$

Proposition 4.3. If the labelled class $\mathcal{B} \coloneqq \mathcal{A}^*$, then

$$B(x) = \frac{1}{1 - A(x)}$$

provided $\mathcal{A}_0 = \varnothing$.

Moreover, Set is much easier than before because labelling means no possibility of repetition:

$$\mathsf{Set}(\mathcal{A}) = \{\varepsilon\} \cup \mathcal{A} \cup \frac{\mathcal{A} * \mathcal{A}}{2!} \cup \frac{\mathcal{A} * \mathcal{A} * \mathcal{A}}{3!} \cup \cdots$$

giving that the expoenential generating series of $\mathsf{Set}(A)$ is

$$1 + A(x) + \frac{A(x)^2}{2} + \frac{A(x)^3}{3!} + \cdots$$

Proposition 4.4. If the labelled class $\mathcal{B} = \mathsf{Set}(\mathcal{A})$, then

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

provided $\mathcal{A}_0 = \varnothing$.

Discovery 4.1. We saw an example of this: permutations as labelled Set of class of labelled cycles.

4.2 Combinatorial Species

Does built of atoms seem like a kluge? Does the arbitrariness of labelling with $\{1, \dots, n\}$ bother you? Is order preserving relabelling kind of uneasy? Are you you bothered that there is no set of all graphs?

Why do we not have a set of all finite graphs? This is because the class is too big since we didn't say "up to isomorphism", so the vertices can be from any set. How do we fix this? We take graphs with a fixed vertex set X. We take this as the motivation for species.

Definition 4.4. **[Species]**

A **species** is a rule that assigns

1. to each finite set X a finite set $\mathcal{A}_X$ called the **set of $\mathcal{A}$-structure on** X;
2. to each bijection of finite sets $f : X \to Y$ a bijection $f_* : \mathcal{A}_X \to \mathcal{A}_Y$ called **transportation of $\mathcal{A}$-structure along** f such that
   (a) if $X \neq Y$, then $\mathcal{A}_X \cap \mathcal{A}_Y = \varnothing$;
   (b) $(\mathrm{Id}_X)_* : \mathcal{A}_X \to \mathcal{A}_X$ is the identity;
   (c) If $f : X \to Y$, $g : Y \to Z$ are bijections, then $(g \circ f)_* = g_* \circ f_*$.

Note 4.3. This is a category theoretic definiton. Let $\mathfrak{Set}$ be the category of finite sets where the morphism are bijections, then a species is a functor

$$\mathcal{A} : \mathfrak{Set} \to \mathfrak{Set}$$

for any finite set $X \in \mathfrak{Set}$, $\mathcal{A}_X = A[x]$ the set of $\mathcal{A}$-structures, and if $f : X \to Y$ is a morphism in $\mathfrak{Set}$, then

$$f_* = \mathcal{A}[f] : \mathcal{A}[x] \to \mathcal{A}[x]$$

is the transportation along f.

Definition 4.5. **[Isomorphic]**

Let $\mathcal{A}$ be a species, $a \in \mathcal{A}_X$ and $b \in \mathcal{A}_Y$, then we say a, b are **isomorphic** if there exists bijection $f : X \to Y$ such that $f_*(a) = b$.

Note 4.4. We note that if $|X| = |Y|$, then $|\mathcal{A}_X| = |\mathcal{A}_Y|$, so we can define $a_n = |\mathcal{A}_X|$ and get the expoenential generating series

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

4.2.1 Table of Examples

Name (notation)	Structures on X	Transportation along $f : X \to Y$	Generating series
Sets, denoted $\mathcal{E}$	$\{X\} = \mathcal{E}_X$	$f_*(X) = Y$	$E(x) = \sum_{n\geq 0} \frac{1}{n!} x^n = \exp(x)$
Linear orders ($\mathcal{L}$)	$\mathcal{L}_X = \{(x_1, \dots, x_n) : X = \{x_{[n]}\}\}$	$f_*((x_1, \dots, x_n)) = (f_*(x_1), \dots, f_*(x_n))$	$L(x) = \sum_{n\geq 0} \frac{n!}{n!} x^n = \frac{1}{1-x}$
Endofunctions ($\mathcal{N}$)	$\mathcal{N}_X = \{\alpha : X \to X\}$	$f_*(\alpha) = f \circ \alpha \circ f^{-1}$	$N(x) = \sum_{n\geq 0} \frac{n^n}{n!} x^n$
Permutation ($\mathcal{S}$)	$\mathcal{S}_X = \{\text{bijection } \alpha : X \to X\}$	$f_*(\alpha) = f \circ \alpha \circ f^{-1}$	$S(x) = \sum_{n\geq 0} \frac{n!}{n!} x^n = \frac{1}{1-x}$
Cycles ($\mathcal{C}$)	$\mathcal{C}_X = \{\alpha : X \to X, \text{ bijection with exactly one cycle}\}$	$f_*(\alpha) = f \circ \alpha \circ f^{-1}$	$C(x) = \sum_{n\geq 1} \frac{(n-1)!}{n!} x^n = \log\left(\frac{1}{1-x}\right)$
Graphs ($\mathcal{G}$)	$\mathcal{G}_X = \{G : \text{simple graph on } X\}$	$f_*(G)$ preserves edge relation	
"Free" trees ($\mathcal{T}$)	$\mathcal{T}_X : \{t \in \mathcal{G}_X : \text{connected, no cycles}\}$	same as above	$T(x) = \sum_{n\geq 1} \frac{n^{n-2}}{n!} x^n$
Set partition (Π)	$\Pi_X = \{\{P_1, \dots, P_k\} \text{ set partition}\}$	$f_*(\{P_1, \dots, P_k\}) = \{f(P_1), \dots, f(P_k)\}$	$\Pi(x) = \sum_{n\geq 0} \frac{B_n}{n!} x^n$

`Table: table of examples`

Definition 4.6. **[Natural Transportation]**

Let $\mathcal{A}$ and $\mathcal{B}$ be species, a **natural transportation** $\Phi : \mathcal{A} \to \mathcal{B}$ is a rule which assigns to each finite set X a function $\Phi_X : \mathcal{A}_X \to \mathcal{B}_X$ such that if $f : X \to Y$ is a bijection of finite sets, then

$$\begin{tikzcd}
\mathcal{A}_X \arrow[r, "\Phi"] \arrow[d, "A[f]"] & \mathcal{B}_X \arrow[d, "B[f]"] \\
\mathcal{A}_Y \arrow[r, "\Phi"] & \mathcal{B}_Y
\end{tikzcd}$$

Comment 4.2. This is the categorical nature of natural transportation.

Example 4.3. $\mathcal{T}$ inside $\mathcal{G}$ gives a natural transportation $\Phi(t) = t$ seen as a graph. This works whenever you have a subspecies. Other examples are cycles in permutations and permutations in endofunctions.

Example 4.4. Forgetful transformation from any species to sets by forgetting the structure.

Example 4.5. From $\mathcal{E}$ to directed graphs by $(\alpha : X \to X) \mapsto$ digraph with arcs $x \to \alpha(x)$ for all $x \in X$.

Lecture 10 - Monday, February 09

Lemma 4.1. Let $a \in \mathcal{A}_X$ and $b \in \mathcal{A}_Y$ be isomorphic, and Φ a natural transformation from $\mathcal{A}$ to $\mathcal{B}$. Then $\Phi_X(a)$ and $\Phi_Y(b)$ are isomorphic.

Proof. By definition, there exists a bijection $\sigma : X \to Y$ such that

$$A[\sigma](a) = b.$$

Naturality of Φ means that for every bijection $\sigma : X \to Y$,

$$B[\sigma] \circ \Phi_X \; = \; \Phi_Y \circ A[\sigma].$$

Applying both sides to a gives

$$B[\sigma]\big(\Phi_X(a)\big) = \Phi_Y\big(A[\sigma](a)\big) = \Phi_Y(b).$$

Thus $\Phi_X(a)$ and $\Phi_Y(b)$ are carried to each other by the bijection σ, so they are isomorphic B-structures. ☐

Definition 4.7. **[Natural Equivalence]**

Let Φ be a natural transformation from $\mathcal{A}$ to $\mathcal{B}$, if $\Phi_X : \mathcal{A}_X \to \mathcal{B}_X$ is an isomorphism for all x in the set, then Φ is a **natural equivalence**, and we say $\mathcal{A}$ and $\mathcal{B}$ are naturally equivalent and write $A \equiv B$.

Example 4.6. The class of permutations $\mathcal{S}$ is naturally equivalent to the subspecies of directed graphs where every vertex has in- and out-degree of exactly 1.

Example 4.7. Graphs $\mathcal{G}$ is naturally equivalent to itself under graph complement operation, and of course, identity.

Example 4.8. For all species $\mathcal{A}$, there is a natural equivalence to the species of pairs of identical $\mathcal{A}$-structure.

Definition 4.8. **[Numerical Equivalence]**

We say two species $\mathcal{A}$ and $\mathcal{B}$ are **numerically equivalent** if $A(x) = B(x)$, and we write $\mathcal{A} \approx \mathcal{B}$.

Note 4.5. Natural equivalence implies numerical equivalence, but the converse is not true, the following is an example.

Example 4.9. From the table 4.2.1, we have linear orders denoted as $\mathcal{L}$, and

$$\mathcal{L}_X = \{(x_1, \dots, x_n) : |X| = n, X = \{x_1, \dots, x_n\}\}$$

we denote $\mathcal{S}$ the set of permutations, and

$$\mathcal{S}_X = \{\alpha : X \to X : \alpha \text{ is a bijection}\}$$

whose transportation is $f \circ \alpha \circ f^{-1}$. These are numerically equivalent, but not naturally equivalent. To see why, suppose they were, i.e., there is a natural equivalence $\Phi : \mathcal{L} \to \mathcal{S}$, but any two linear orders of the same size are isomorphic (just map the elements to make it work), but then by the lemma, we can take these isomorphisms through Φ and since Φ is a natural equivalence, we get that all permutations of sets of the same size are isomorphic. However, here, the transportation was $f \circ \alpha \circ f^{-1}$, i.e., conjugation, and conjugation cannot change cycle structure. We conclude now that now all permutations of the same size are isomorphic, so Φ cannot exist.

Definition 4.9. **[Connected]**

A species $\mathcal{A}$ is **connected** if $\mathcal{A}_\varnothing = \varnothing$.

Comment 4.3. The language "connected" will return in Hopf algebras.

Species very naturally tell the labelled story. $\mathcal{A}$-structure on X are objects labelled by X. They are some sort of "diagram with widgets (atoms)" that takes X-labels, and transportation is technically relabelling. The unlabelled classes are the isomorphism classes.

4.3 Species Operations

This is a categorical reformulation of the symbolic method. Let $\mathcal{A}$ and $\mathcal{B}$ be species.

Sum/ disjoint union: This works essentially the same as before. We will use the notation $\mathcal{A} + \mathcal{B}$ for species with

$$(\mathcal{A} + \mathcal{B})_X = \mathcal{A}_X \sqcup \mathcal{B}_X$$

where we force disjointness. Formally, we accomplish this by taking two distinct singletons and take

$$\mathcal{A}_X \sqcup \mathcal{B}_X = \{0\} \times \mathcal{A}_X \cup \{1\} \times \mathcal{B}_X$$

Setminus: When $\mathcal{B} \subseteq \mathcal{A}$, i.e., $\mathcal{B}_X \subseteq \mathcal{A}_X$ for all X, then we define $\mathcal{A} - \mathcal{B}$ by

$$(\mathcal{A} - \mathcal{B})_X = \mathcal{A}_X \setminus \mathcal{B}_X$$

Restrict by order/ filter: It will be handy to have notation for restricting by order (filter). For instance, $\mathcal{A}_{\text{even}}$ is defined by

$$(\mathcal{A}_{\text{even}}) = \begin{cases} \mathcal{A}_X & |X| \text{ is even} \\ \varnothing & \text{otherwise} \end{cases}$$

We can also have other conditions written in the subscript.

Example 4.10. If $\mathcal{E}$ is the species of sets, then $\mathcal{E}_{>0}$ is the species of non-empty sets.

Species product and Star: Here is the species version of labelled product, $\mathcal{A} * \mathcal{B}$ is defined by

$$(\mathcal{A} * \mathcal{B})_X = \bigsqcup_{S \subseteq X} (\mathcal{A}_X \times \mathcal{B}_{X \setminus S})$$

Since this is the best notion of product in this context, we define $\mathcal{A}^n = \underbrace{\mathcal{A} * \mathcal{A} * \cdots * \mathcal{A}}_{n \text{ times}}$ unless otherwise states. It is worth noting that

$$\mathcal{A}^0 = \mathcal{E}_0 = \{\varnothing\}$$

Moreover, we also define

$$\mathcal{A}^* = \sum_{n \geq 0} \mathcal{A}^n$$

Cartesian product: Cartesian product works when the other set is viewed as a constnat, and sometimes cartesian product with a constant set is useful. Let $\mathcal{A}$ be a species and $\mathcal{K}$ be a set, the species $\mathcal{K} \times \mathcal{A}$ is defined by

$$(\mathcal{K} \times \mathcal{A})_X = \mathcal{K} \times \mathcal{A}_X$$

note that we have seen this for defining disjoint union.

Pointing: We can define pointing $\mathcal{A}^\bullet$ as

$$(\mathcal{A}^\bullet)_X = \mathcal{A}_X \times X$$

Composition: Here we have a new one this time, which is composition. We define $\mathcal{A}[\mathcal{B}]$ where $\mathcal{B}$ is connected by

$$(\mathcal{A}[\mathcal{B}])_X = \bigsqcup_{\{x_1,\ldots,x_k\} \in \Pi_X} \mathcal{A}_{\{x_1,\ldots,x_k\}} \times \mathcal{B}_{X_1} \times \mathcal{B}_{X_2} \times \cdots \times \mathcal{B}_{X_k}$$

Example 4.11. $\mathcal{E}[\mathcal{T}]$ deotes forests, and in fact, $\mathcal{E}[\mathcal{A}] = \mathsf{Set}[\mathcal{A}]$.

Example 4.12. $\mathcal{L}[\mathcal{A}] = \mathcal{A}^*$.

4.3.1 Exponential Generating Series for Species Operations

The exponential generating series behave nicely for all of the above operations.

Condition	Species	Expoenential Generating Series
	$\mathcal{A}+\mathcal{B}$	$A(x)+B(x)$
$\mathcal{B} \subseteq A$	$\mathcal{A}-\mathcal{B}$	$A(x)-B(x)$
	$\mathcal{A}_{\text{even}}$	$1/2(A(x)+A(-x))$
	$\mathcal{A}_{\text{odd}}$	$1/2(A(x)-A(-x))$
	$\lvert\mathcal{K}\rvert \times \mathcal{A}$	$\mathcal{K}A(X)$
	$\mathcal{A}*\mathcal{B}$	$A(x)B(x)$
$\mathcal{A}$ conncted	$\mathcal{A}^*$	$\frac{1}{1-A(x)}$
	$\mathcal{A}^\bullet$	$\frac{x\mathrm{d}}{\mathrm{d}x}A(x)$
$\mathcal{B}$ connected	$\mathcal{A}[\mathcal{B}]$	$A(B(x))$

Proof. Seires manipulations (as before). ☐

Example 4.13. How many permutations with even cycles are there?

Solution. We write this class of permutations as species, even set of cycles:

$$\mathcal{E}_{\text{even}}[\mathcal{C}]$$

The class of cycles has exponential generating series $\log\left(\frac{1}{1-x}\right)$, moreover, we know that

$$E_{\text{even}}(x) = \frac{1}{2}(e^x + e^{-x}) = \cosh(x)$$

Hence the expoenential generating series for permutations with even cycles is

$$\cosh\left(\log\left(\frac{1}{1-x}\right)\right) = 1 + \frac{1}{2}\left(\frac{x^2}{1-x}\right)$$

so the number of permutations of $\{1,\dots,n\}$ with an even number of cycles is

$$n![x^n]\left(1+\frac{1}{2}\left(\frac{x^2}{1-x}\right)\right) = \begin{cases} \frac{1}{2}n! & n \geq 2 \\ 0 & n = 1 \\ 1 & n = 0 \end{cases}$$

☐

5 Posets

Lecture 13 - Monday, March 02

5.1 Definitions and Examples

Definition 5.1. **[Poset]**

A **partially ordered set** P (or **poset**) is a set and a binary relation $\leq$ ($\leq_P$ if disambiguation needed) satisfying

- $x \leq x$ for all $x \in P$
- $x \leq y$ and $y \leq x$ implies $x = y$ for all $x, y \in P$
- $x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in P$

Comment 5.1. We write $x < y$ for $x \leq y$ and $x \neq y$.

Definition 5.2. **[Comparable]**

For $x, y \in P$, if $x \leq y$ or $y \leq x$, then we say x and y are **comparable**.

Definition 5.3. **[Incomparable]**

We say x and y are **incomparable** if they are not comparable.

Definition 5.4. **[Cover]**

If $x < y$ and there is no $z \in P$ such that $x < z < y$, then we say that y **covers** x. We write

$$x \prec y \qquad \text{or} \qquad x \lessdot y$$

Note 5.1. For a finite poset, the cover relation determines the poset. Draw a finite poset by its order diagram or Hasse diagram by drawing the graph of the cover relation on P, where if $x \prec y$ then we draw x lower on the page than y (so all edges directed upwards, without drawing any arrows).

Definition 5.5. **[Maximal]**

An element $x \in P$ is **maximal** if there is no $y \in P$ with $y \geq x$. Of course, we have a analogous definition for minimal.

Definition 5.6. **[Minimum and Maximum]**

We say a poset P has a **minimum element** (**maximum element**), write it $\hat{0}$ ($\hat{1}$) if there exists $\hat{0} \in P$ ($\hat{1} \in P$) such that $x \geq \hat{0}$ ($x \leq \hat{1}$) for all $x \in P$.

Example 5.1. Positive integers with divisibility.

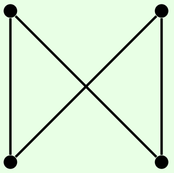
Example 5.2. $\{1, \ldots, n\}$ with usual $\leq$ (this is a total order as all pairs of elements are comparable).

Example 5.3. Anything you can order by inclusion. Subset of a fixed set gives us what's known as *Boolean poset.*

Example 5.4. Set partition by refinement.

Example 5.5. Rooted tree, the tree itself is the Hasse diagram, with the root as, say, the maximal element.

Example 5.6. The following is a poset with two maximal elements and two minimal elements:



This is called a bowtie because it looks like one.

Definition 5.7. **[Subposet]**

Given P a poset, a **subposet** of P is a $Q \subseteq P$ with $\leq$ restricted on Q.

An order preserving map between posets P and Q is some $f : P \to Q$ such that

$$x \leq_P y \quad \Rightarrow \quad f(x) \leq_Q f(y)$$

and then an isomorphism of posets is an order preserving bijection with an order preserving inverse.

Definition 5.8. **[Interval]**

An **interval** in a poset P is, for $x, y \in P$,

$$[x, y] \coloneqq \{z \in P : x \leq z \leq y\}$$

Definition 5.9. **[Locally Finite]**

A poset is **locally finite** if all intervals are finite.

Definition 5.10. **[Chain, Antichain]**

A **chain** in a poset is a set of mutually comparable elements, which must be totally ordered. On the other hand, an **antichain** in a poset is a set of mutually incomparable elements.

Discovery 5.1. There are locally finite infinite poset, take a countably infinite antichain as an example.

Definition 5.11. **[Downset (or Order Ideal)]**

A **downset** of a poset P is a set $D \subseteq P$ that is downwards closed, i.e., if $x \in D$ and $y \in P$ with $y \leq x$, then $y \in D$.

Definition 5.12. **[Upset (or Order Filter)]**

An **upset** is a set $U \subseteq P$ that is upwards closed.

The downset generated by $x_1, x_2, \ldots, x_k$ is

$$\Lambda(x_1, x_2, \ldots, x_k) = \{y \in P : \exists\, i \text{ s.t. } y \leq x_i\}$$

This is called the **inclusive past of** $x_1, x_2, \ldots, x_k$. The upset generated by $x_1, x_2, \ldots, x_k$ is

$$\mathrm{V}(x_1, x_2, \ldots, x_k) = \{y \in P : \exists\, i \text{ s.t. } y \geq x_i\}$$

This is called the **inclusive future of** $x_1, x_2, \ldots, x_k$.

Definition 5.13. **[Upper Bound and Lower Bound]**

Given $x, y \in P$, an **upper bound** of x, y is an element $z \in P$ such that $z \geq x$ and $z \geq y$. We have an analogous definition for **lower bound**.

Definition 5.14. **[Least Upper Bound]**

A **least upper bound** of x, y is a $z \in P$ such that every upper bound w of x, y satisfies $w \geq z$. We write $z = x \vee y$ and say x **join** y.

Definition 5.15. **[Greatest Lower Bound]**

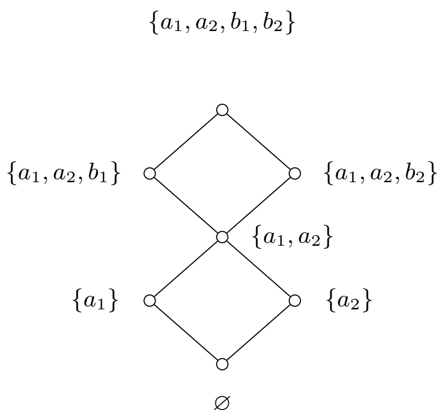
A **greatest lower bound** of x, y is a $z \in P$ such that every lower bound w of x, y satisfies $w \leq z$. We write $z = x \wedge y$ and say x **meet** y.

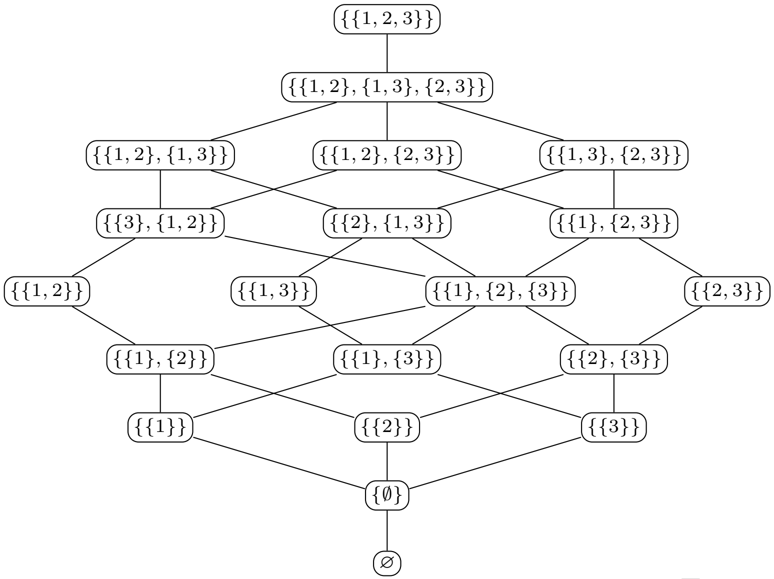
Definition 5.16. **[Lattice]**

A **lattice** is a poset for which every pair of elements has a greatest lower bound and a least upper bound.

Discovery 5.2. The set of downsets of a fixed poset P forms a poset by inclusion. Try $P_1 =$ the bow tie and $P_2 =$ the Boolean poset on $[3]$, what do you observe about the posets of downsets that you get?

Observation. We first draw the Hasse diagrams:





□

Lecture 14 - Wednesday, March 04

Example 5.7. Example and non-examples of lattices:

- A chain is a lattice;
- B_n are lattices (of course, for $n \geq 1$);
- The poset of downsets of B_3 (drawn above) is also a lattice.
- The bowtie (see example 5.6) is not a lattice because upper bound of the two minimal elements is not unique;
- The poset of downsets of the bowtie is a lattice.

Note 5.2. Some lattice fact:

1. $\wedge, \vee$ are associative, commutative, and idempotent;
2. absorption: $x \wedge (x \vee y) = x = x \vee (x \wedge y)$;
3. $x \wedge y = x \iff x \vee y = y \iff x \leq y$.

Definition 5.17. **[Distributive (Lattice)]**

A lattice L is **distributive** if for all $x, y, z \in L$:

1. $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$;
2. $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$

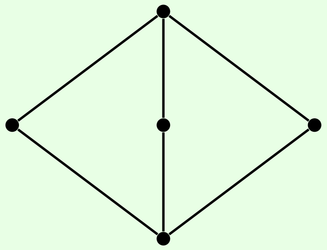
Note 5.3. We notice that in the definition above, either of them implies the other. Suppose the first is true, then

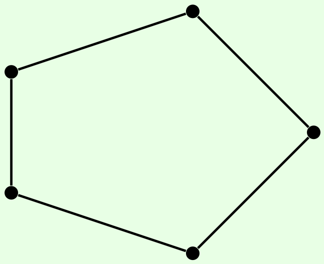
$$\begin{aligned}(x \wedge y) \vee (x \wedge z) &= ((x \wedge y) \vee x) \wedge ((x \wedge y) \vee z) && \text{given} \\ &= x \wedge ((x \wedge y) \vee z) && \text{absorption} \\ &= x \wedge ((x \vee z) \wedge (y \vee z)) && \text{given and commutative} \\ &= (x \wedge (x \vee z)) \wedge (y \vee z) && \text{associative} \\ &= x \wedge (y \vee z) && \text{absorption}\end{aligned}$$

as desired.

Example 5.8. B_n is distributive because union and intersection are distributive.

Example 5.9. Here are two non-examples:





5.2 Fundamental Theorem of Finite Distributive Lattices

This section is dedicated to proving the following theorem.

Theorem 5.1. Fundamental Theorem of Finite Distributive Lattices

Let L be a finite distributive lattice, then there is a unique (up to isomorphism) finite poset P such that L is isomorphic to the lattice of downsets of P.

To prove the above theorem, we will need a couple lemmas to assist us.

5.2.1 Lemma 1

Lemma 5.1. Let P be a finite poset, the poset of downsets of P is a finite distributive lattice.

Proof. (First note that we are working under finite poset, so everything is finite.) Union of downsets is a downset, intersection of downsets is also a downset, so the poset of downsets is a lattice. Further, union and intersection are distributive, so it is a distributive lattice. □

5.2.2 Lemma 2

Definition 5.18. **[Join-irreducible]**

Say an element x of a lattice L is **join-irreducible** if we cannot write $x = y \vee z$ with $y < x$ and $z < x$.

Lemma 5.2. Let P be a finite poset, a downset D of P is join-irreducible in the lattice of downsets if and only if D is principal, i.e., $D = \Lambda(x)$ for some $x \in P$.

Comment 5.2. As a result, we have a bijection between the join-irreducibles of the lattice of downsets of P and P itself (via generators), so we see P inside its lattice of downsets with the same poset structure.

Proof. We first show that if a downset $D = D_1 \cup D_2$ with $D_1 \subsetneq D$ and $D_2 \subsetneq D$, then D is not principal: We note that $D_1 \cap D_2$ is a downset, so there exists a maximal element of D_2 which is in $D - D_1$, call it z, likewise we take y maximal in D_1 in $D - D_2$. Then, if D were principal, $D = \Lambda(x)$, then $x \geq z$, $x \geq y$, which is impossible, so D is not principal.

On the other hand, suppose D is not principal, i.e., $D = \Lambda(x_1, \ldots, x_k)$ for x_i maximal in D, $k > 1$. Then we can take $D_1 = \Lambda(x_1)$, $D_2 = \Lambda(x_2)$, …, so D is not join-irreducible. □

Lecture 15 - Monday, March 09

Proof of Theorem 5.1. Take L to be the distributive lattice. Let P be the subposet of join irreducible elements of L. By the second lemma, it suffices to check that L is isomorphic to the lattice of downsets of P, call that $J(P)$.

Take $x \in L$, let D_x be defined as

$$D_x = \{y \in P : y \leq x\}$$

(note that x might not exist in P, so D_x is not necessarily $\Lambda(x)$). D_x is a downset of P by definition. Hence we have a map $\phi : L \to J(P)$ defined by $x \mapsto D_x$. Our goal is to show that ϕ is our bijection.

We know so far that this is an injective map, and it is order-preserving, so we just need to show that ϕ is surjective. Take $D \in J(P)$. Let $x = \bigvee_{y \in D} y$ in L, we want to show that $D = D_x$. We certainly have $D \subseteq D_x$ since all things are $\leq$ their upper bound. On the other hand, we start by noting that $\bigvee_{y \in D} y = \bigvee_{y \in D_x} y$ because of the leastness of the upper bound. Take $z \in D_x$, and take $\wedge$ of z with the above:

$$\underbrace{z \wedge \left(\bigvee_{y \in D} y \right)}_{\bigvee_{y \in D}(z \wedge y)} = \underbrace{z \wedge \left(\bigvee_{y \in D_x} y \right)}_{= \bigvee_{y \in D_x}(z \wedge y)}$$

We know that $\bigvee_{y \in D_x}(z \wedge y) = z$ since one of the terms is $z \wedge z$, and all the other terms are $\leq$ that. On the left, there is some $y \in D$ such that $y \wedge z = z$ becasue otherwise there is problem with join-irreducibility. Then $z \leq y$ and D is a downset, so $z \in D$. ☐

5.2.3 There is an Equivalence of Categories

On one side, we want category of finite posets with order preserving maps, on the other side, we want the category of finite distributive lattices with bounded lattice homomorphisms (a lattice homomorphism is a map of underlying sets preserving $\wedge$ and $\vee$, bounded means that it preserves $\hat{0}$ and $\hat{1}$).

Let **2** be the following poset:

Write $J(P)$ for lattice of downsets of P. A downset of P exhibits the same information as an order preserving map $\phi : P \to \mathbf{2}$ ($D = \phi^{-1}(0)$), so $J(P)$ is $\mathrm{Hom}(P, \mathbf{2})$. Further, order preserving map play well with this:

Let $g : Q \to P$ be an order preserving map, define

$$g^* : J(P) \to J(Q)$$

by for any downset $D \in J(P)$, we enforce

$$r_D(x \in P) = \begin{cases} 0 & x \in D \\ 1 & x \notin D \end{cases}$$

then $g^*(D) = (r_D \circ g)^{-1}(0)$ in $J(Q)$, it is the preimage of the downset D under g.

> **Comment 5.3.** We can check that this preserves $\wedge$, $\vee$, $\hat{0}$, and $\hat{1}$.

> **Note 5.4.** In category language, the above says that $J = \mathrm{Hom}(-, \mathbf{2})$ is a covariant functor giving an equivalence of categories.

5.3 Möbius Inversion

The ζ-function of a poset P is $\zeta : P \times P \to \mathbb{Z}$ defined as

$$\zeta(x,y) = \begin{cases} 1 & x \leq y \\ 0 & \text{otherwise} \end{cases}$$

It is often convenient to think of ζ as a matrix, so we fix a linear extension of P (a total order on the underlying set P so that if $x \leq_p y$ then $x \leq y$ in the total order). Then the ζ-matrix of P is

$$\zeta = (\zeta(x_i, x_j))_{i,j}$$

with $P = \{x_1, x_2, \ldots\}$ ordered by the linear extension.

Comment 5.4. This is also called the **relation matrix** of P.

Note 5.5. This is also the adjacency matrix of P as a digraph. (Not the Hasse diagram, but P itself, so an edge $x \rightarrowtail y$ whenever $x \leq y$). Just like any adjacency matrix, we know somethings about it:

$$\zeta^2(x,y) = \sum_{x \leq z \leq y} 1 = \big|[x,y]\big|$$

and

$$\zeta^k(x,y) = \sum_{x \leq x_1 \leq \cdots \leq x_{k-1} \leq y} 1 = \#\text{ of multichains of length } k \text{ from } x \text{ to } y \text{ in } P$$

where a multichain is a chain where repeats are allowed. Notice that if we worked with $(\zeta - I)^k$ instead, then we are counting honest chains.

Discovery 5.3. ζ is invertible because it is upper triangular with 1s on the diagonals, but what is its inverse explicitly?

We define $\mu : P \times P \to \mathbb{Z}$ by

$$\mu(x,y) = \begin{cases} 1 & x = y \\ -\sum_{x \leq z < y} \mu(x,z) & x < y \\ 0 & \text{otherwise} \end{cases}$$

Note 5.6. The first two can be rephrased as

$$\sum_{x \leq z \leq y} \mu(x,z) = \delta(x,y)$$

μ is well-defined on any locally finite poset by induction on size of interval, μ is called the **Möbius function**. Of course, we can also make the matrix of μ, which will be called μ :)

Proposition 5.1. Let P be a locally finite poset, then ζ and μ are inverses.

Proof. We have

$$(\mu\zeta)(x,y) = \sum_{z\in P} \mu(x,z)\zeta(z,y)$$

which is a finite sum since $\mu(x,z) \neq 0 \;\Rightarrow\; x \leq z$ and $\zeta(z,y) \neq 0 \;\Rightarrow\; z \leq y$. Further,

$$\begin{aligned}(\mu\zeta)(x,y) &= \sum_{z\in P} \mu(x,z)\zeta(z,y)\\ &= \sum_{x\leq z\leq y} \mu(x,z)\zeta(z,y)\\ &= \sum_{x\leq z\leq y} \mu(x,z) = \begin{cases} 0 & x<y\\ 1 & x=y\\ 0 & x>y\end{cases}\end{aligned}$$

by note 5.6. ☐

Theorem 5.2. Möbius Inversion

1. Let P be a poset with all downsets finite, let $f,g : P \to \mathbb{C}$, then

$$g(x) = \sum_{y\leq x} f(y) \quad \forall x\in P \quad\iff\quad f(x) = \sum_{y\leq x} \mu(y,x)g(y) \quad \forall x \in P$$

2. Let P be a poset with all upsets finite, let $f,g : P \to \mathbb{C}$, then

$$g(x) = \sum_{y\geq x} f(y) \quad \forall x\in P \quad\iff\quad f(x) = \sum_{y\geq x} \mu(x,y)g(y) \quad \forall x \in P$$

Note 5.7. The condition of "all downsets (upsets) finite" forces the sums to be finite. We can relax this if we are in a valuation ring and convergent sums.

Proof. For the first, we write f, g as column vectors (indexed by the same linear extension used by ζ and μ, so the ith entry is $f(x_i)$). Then the LHS is $g = \zeta^T f$ while the RHS is $f = \mu^T g$, which are the same by multiplying by $(\zeta^{-1})^T$ or $(\mu^{-1})^T$ since ζ and μ are inverses. ☐

Comment 5.5. In the end, Möbius inversion was a fairly trivial linear algebra fact.

Lecture 16 - Wednesday, March 11

5.3.1 Three Classic Examples

Example 5.10. Boolean Poset

Let $P = B_n$. Let $S, T \subseteq [n]$ be any two subsets, then

$$\mu(S,T) = \begin{cases} (-1)^{|S|-|T|} & S \subseteq T \\ 0 & \text{otherwise} \end{cases}$$

by Binomial Theorem essentially. This is inclusion-exclusion.

How could we use the result above?

Let $\mathcal{S}$ be a set of properties, and $\mathcal{C}$ a combinatorial object that have or don't have the properties. Let $f(U)$ be the number of elements of $\mathcal{C}$ satisfying exactly the properties of $U \subseteq \mathcal{S}$ (this is typically hard to count). Let $g(U)$ be the number of elements of $\mathcal{C}$ satisfying at least the properties of U, then

$$g(U) = \sum_{T \supseteq U} f(T)$$

By Möbius Inversion, we have

$$f(U) = \sum_{T \supseteq U} (-1)^{|T|-|U|} g(T)$$

We often want

$$f(\varnothing) = \sum_{T \subseteq S} (-1)^{|T|} g(T)$$

This is typical inclusion-exclusion. A derangements is an example of this (inclusion-exclusion). Let

$$\begin{aligned} \mathcal{C} &= \text{permutation of } [n] \\ \mathcal{S} &= \{1 \text{ is fixed point}, 2 \text{ is fixed point}, \ldots\} \end{aligned}$$

Then the number of permutations with at least $i_1, \ldots, i_k$ fixed points is $(n-k)!$ by elements counting, so the number of derangements is given by

$$\sum_{T \subseteq [n]} (-1)^{|T|} (n - |T|)! = \sum_{k=0}^{n} \binom{n}{k} (-1)^k (n-k)!$$

Example 5.11. Number Theory Example

Let P be positive integers ordered by divisibility. Note by unique factorization, a positive integer is the multiset of its primes, so P is a product poset of a total order for the powers of each prime. It turns out that

$$\mu(i,j) = \begin{cases} (-1)^m & \text{if } \frac{j}{i} \text{ is a product of distinct primes, and there are } m \text{ of them} \\ 0 & \text{otherwise} \end{cases}$$

Since only depends on $\frac{j}{i}$, we write $\mu(\frac{j}{i})$. Alternatively, we can think of the multiset of primes given by vector of multiplicities:

$$\mu((a_1, a_2, \ldots), (b_1, b_2, \ldots)) = \begin{cases} (-1)^{\sum(b_i - a_i)} & \text{if each } b_i - a_i \in \{0, 1\} \\ 0 & \text{otherwise} \end{cases}$$

Classic number theoretic Möbius inversion is

$$g(n) = \sum_{d \mid n} f(d) \quad \Longleftrightarrow \quad f(n) = \sum_{d \mid n} g(d)\, \mu\left(\frac{n}{d}\right)$$

Example 5.12. Let P be $\mathbb{Z}_{\geq 0}$ with the usual total order, then inductively,

$$\mu(i,j) = \begin{cases} 1 & i = j \\ -1 & i + 1 = j \\ 0 & \text{otherwise} \end{cases}$$

Möbius inversion tells us that

$$g(n) = \sum_{i=0}^{n} f(i) \quad \forall n \geq 0 \quad \Longleftrightarrow \quad f(n) = g(n) - g(n-1) \quad \forall n \geq 0$$

with the convention that $g(-1) = 0$. This is the discrete analogue of fundamental theorem of calculus.

6 Hopf Algebras

Product tells us how to put things together.

Example 6.1. Suppose Ω is a finite alphabet, then a product on Ω^* could be concatenation.

Example 6.2. A product on $\mathrm{Span}_{\mathbb{Q}}\Omega^*$ is shuffle:

$$ab \sqcup\!\sqcup\, cb = abcb + 2acbb + 2cabb + cbab$$

6.1 Algebra and Coalgebra

Definition 6.1. **[Algebra]**

An algebra A over K is a vector space on K with 2 linear maps

$$m : A \otimes A \to A$$
$$u : K \to A$$

(where u stands for unit) such that

$$\begin{tikzcd}
A \otimes A \otimes A \arrow[r, "m\otimes \mathrm{id}"] \arrow[d, "\mathrm{id}\otimes m"] & A \otimes A \arrow[d, "m"] \\
A \otimes A \arrow[r, "m"] & A
\end{tikzcd}$$

$$\begin{tikzcd}
K \otimes A \arrow[d, "u\otimes \mathrm{id}"] & A \arrow[l, "1\otimes a \leftarrow a"'] \arrow[r, "a\to a\otimes 1"] \arrow[d, "\mathrm{id}"] & A \otimes K \arrow[d, "\mathrm{id}\otimes u"] \\
A \otimes A \arrow[r, "m"] & A & A \times A \arrow[l, "m"']
\end{tikzcd}$$

Tensor product is a machine to turn bilinear maps into linear maps (you might have expected $m : A \times A \to A$ bilinear instead of $m : A \otimes A \to A$ linear).

Theorem 6.1. Point Universal Property

$A \otimes B$ is the vector space over K such that there exists a bilinear map

$$\iota : A \times A \to A \otimes A$$
$$(a, b) \mapsto a \otimes b$$

with the property that for any bilinear $f : A \times B \to C$, there exists a unqiue linear map $g : A \otimes B \to C$ so that

$$\begin{tikzcd}
A \times B \arrow[r, "\iota"] \arrow[d, "f"] & A \otimes B \arrow[dl, "\exists!\, g"] \\
C &
\end{tikzcd}$$

Definition 6.2. **[Coalgebra]**

A coalgebra C over K is a vector space on K with 2 linear maps

$$\Delta : C \to C \otimes C$$
$$\varepsilon : C \to K$$

(where u stands for unit) such that

$$\begin{tikzcd}
C \otimes C \otimes C & C \otimes C \arrow[l, "\Delta\otimes\mathrm{id}"] \\
C \otimes C \arrow[u, "\mathrm{id}\otimes\Delta"] & C \arrow[l, "\Delta"] \arrow[u, "\Delta"]
\end{tikzcd}$$

$$\begin{tikzcd}
K \otimes C \arrow[r, "1\otimes a\leftarrow a"] & C & C \otimes K \arrow[l, "a\to a\otimes 1"] \\
C \otimes C \arrow[u, "\varepsilon\otimes\mathrm{id}"] & C \arrow[l, "\Delta"] \arrow[u, "\mathrm{id}"] \arrow[r, "\Delta"] & C \times C \arrow[u, "\mathrm{id}\otimes\varepsilon"]
\end{tikzcd}$$

Comment 6.1. A coproduct tells you how to take one object and pull it apart.

Example 6.3. A coproduct on $\mathrm{Span}_{\mathbb{Q}}\Omega^*$ could be deconcatenation:

$$\Delta(aba) = 1 \otimes aba + a \otimes ba + ab \otimes a + aba \otimes 1$$

where 1 is the new notation for empty.

Example 6.4. A coproduct on $\mathrm{Span}_{\mathbb{Q}}\Omega^*$ could be deshuffle:

$$\Delta(aba) = 1 \otimes aba + a \otimes ba + b \otimes aa + a \otimes ab + ab \otimes a + aa \otimes b + ba \otimes a + aba \otimes 1$$

with $\varepsilon(1) = 1 \in K$ and $\varepsilon(w) = 0$ where w is a nonempty word, extended linearly.

If algebra and coalgebra are compatible, then we have a **bialgebra**. What does “compatible” mean? Compatible means that

- Δ, ε are algebra homomorphisms, or equivalently
- m, u are coalgebra morphisms.

Exercise 6.1. Show that they have the same commutative diagram.

Discovery 6.1. On $\mathrm{Span}_{\mathbb{Q}}\Omega^*$, which pairs give a bialgebra?

$\Delta\backslash m$	Concatenation	Shuffle
Deconcatenation		
Deshuffle		

Solution. **Concatenation and deconcatenation**: We first compute

$$\begin{aligned}\Delta_{decon}(m_{concat}(a,b)) &= \Delta_{decon}(ab)\\ &= 1\otimes ab + a\otimes b + ab\otimes 1\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\Delta_{decon}(a) &= 1\otimes a + a\otimes 1\\ \Delta_{decon}(b) &= 1\otimes b + b\otimes 1\end{aligned}$$

Notice that

$$(1\otimes a + a\otimes 1)(1\otimes b + b\otimes 1) = 1\otimes ab + b\otimes a + a\otimes b + ab\otimes 1$$

we find that they are NOT the same.
Shuffle and deshuffle: We first compute

$$\begin{aligned}\Delta_{desh}(m_{shuff}(a,b)) &= \Delta_{desh}(ab+ba)\\ &= 1\otimes ab + 1\otimes ba + 2a\otimes b + 2b\otimes a + ab\otimes 1 + ba\otimes 1\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\Delta_{desh}(a) &= 1\otimes a + a\otimes 1\\ \Delta_{desh}(b) &= 1\otimes b + b\otimes 1\end{aligned}$$

Notice that when we multiply with shuffle:

$$(1\otimes a + a\otimes 1)\sqcup\!\sqcup(1\otimes b + b\otimes 1) = 1\otimes ab + 1\otimes ba + b\otimes a + a\otimes b + ab\otimes 1 + ba\otimes 1$$

we find that they are NOT the same.
The other two entries can be verified to be YES.

Discovery 6.2. On $\mathrm{Span}_{\mathbb{Q}}\Omega^*$, which pairs give a bialgebra?

$\Delta\backslash m$	Concatenation	Shuffle
Deconcatenation	NO	YES
Deshuffle	YES	NO

$\square$

Lecture 17 - Monday, March 16

Say C a coalgebra, A an algebra, and

$$f, g : C \to A$$

are linear maps. We define convolution as follows:

Definition 6.3. **[Convolution]**

Convolution of f and g is

$$f * g = m \circ (f \otimes g) \circ \Delta$$

Note 6.1. This makes the space of linear maps $C \to A$ into an algebra, called the **convolution algebra**.

Definition 6.4. **[Hopf Algebra]**

A bialgebra B is a **Hopf algebra** if there exists a linear map $S : B \to B$ called antipode such that

$$S * \mathrm{id} = \mathrm{id} * S = u \circ \varepsilon$$

Comment 6.2. S is typically not an algebra homomorphism as it flips multiplication

$$S(ab) = S(b)S(a)$$

Last time we had two word bialgebras, namely the concatenation/ deshuffle and shuffle/ deconcatenation. In fact, both of these are Hopf algebras with antipode

$$S(w_1 w_2 \dots w_n) = (-1)^n w_n w_{n-1} \dots w_1$$

6.2 More Examples

Example 6.5. $K[x]$ with $\Delta(x) = x \otimes 1 + 1 \otimes x$ extend to an algebra homomorphism, so

$$\Delta(x^n) = \sum_{k=0}^{n} \binom{n}{k} x^k \otimes x^{n-k}$$

also, $S(x) = -x$ and so $S(x^n) = (-1)^n x^n$.

Example 6.6. Conner-Kreiner Hopf Algebra of Rooted Trees

Let $\mathcal{T}$ be rooted trees with the root at the top (with no empty trees). Then

$$K[\mathcal{T}] = \operatorname{span}_K(\mathcal{F})$$

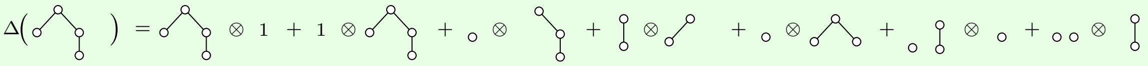
with disjoint union multiplication, where $\mathcal{F}$ are forests of rooted trees (including empty denoted $\mathbb{1}$). What is the coproduct?

$$\Delta(t) = \sum_{C \text{ antichain of } t} \Lambda(C) \otimes (t \setminus \Lambda(C))$$

and extend to an algebra homomorphism. Notice, the above is the same as

$$\Delta(t) = \sum_{D \text{downset of } f} D \otimes (t \setminus D) = \sum_{U \text{upset of } f} (t \setminus U) \otimes U$$

Here is an example:



Furthermore,

$$S(t) = -t - \sum_{\substack{C \text{ antichain} \\ C \neq \varnothing, C \neq \{\text{root}\}}} S(\Lambda(C)) \cdot (t \setminus \Lambda(C))$$

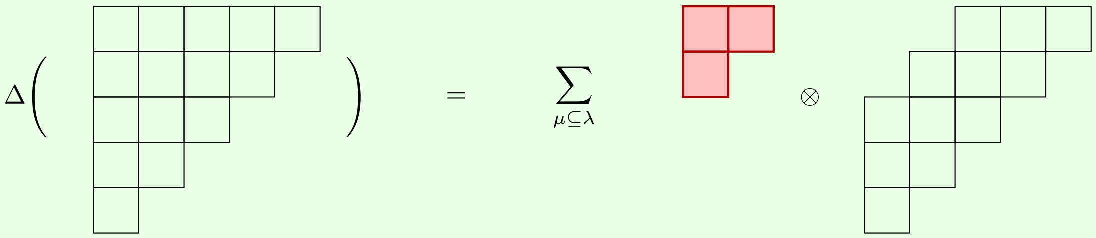
this defines S inductively.

> **Comment 6.3.** This S describes the process of subtracting subdivergences in order to renormalize in quantum field theory.

Example 6.7. A not so combinatorial example

Let G be a group and KG be a group ring – where formal sums of group elements multiplication is group operation extended linearly. We have $\Delta(g) = g \otimes g$ and $S(g) = g^{-1}$.

Example 6.8.



and then taking Schur functions we get Hopf algebra of symmetric functions,

$$\Delta(S_\lambda) = \sum_{\mu \leq \lambda} S_\mu \otimes S_{\lambda/\mu}$$

Definition 6.5. **[Graded]**

Bialgebra B is **graded** as a vector space

$$B = \bigoplus_{i=0}^{\infty} B_i$$

and all defining maps are graded.

Definition 6.6. **[Connected]**

A bialgebra B is **connected** if $B_0 = K$.

If one's bialgebra B is graded and connected, then B is automatically a Hopf algebra with S recursively defined,

$$S(x) = -x + \sum_i S(x_{i,1})x_{i,2}$$

where

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \sum_i x_{i,1} \otimes x_{i,2}$$

6.2.1 Incidence Hopf Algebras

For this course, an important example is incidence Hopf algebras:
Let P be a locally finite poset, and

$$\text{Int}(P) \coloneqq \text{ set of all intervals of } P$$

The incidence coalgebra is a coalgebra on $\text{Span}_K(\text{Int}(P))$ with

$$\Delta([x, y]) = \sum_{x \le z \le y} [x, z] \otimes [z, y]$$

$$\varepsilon([x, y]) = \begin{cases} 1 & x = y \\ 0 & \text{otherwise} \end{cases}$$

Reduced incidence coalgebra of P is the same but instead we take $\text{Int}(P)/\sim$, i.e., we take intervals up to isomorphism.

Comment 6.4. There is a more general version of the story for suitable families of intervals and suitable equivalence relations.

Note 6.2. If you take a dual (in the $\text{Hom}(-, L)$ sense), you get incidence algebra.

We make incidence Hopf algebras from the incidence coalgebra with Cartesian product of posets as the product. In the reduced case, it is Hopf algebra with antipode

$$S([x,y]) = \sum_{k\geq 0} \sum_{x=x_0<x_1<\cdots<x_k=y} (-1)^k \prod_{i=1}^{k} [x_{i-1}, x_i]$$

6.3 Hopf Algebra and Möbius Inversion

How does all those work relate to Möbius inversion? We define

$$\zeta([x,y]) = 1$$

for all $[x,y] \in \mathrm{Int}(P)$, and extend linearly to $\mathrm{Span}_K \mathrm{Int}(P)$. We define μ to be the convolution inverse of ζ, then

1. $\mu = \zeta \circ S$;
2. given $f, g :$ incidence Hopf algebra of $P \to K$,

$$f = g * \zeta \quad \iff \quad g = f * \mu$$

Proof. For (1), we have

$$\begin{aligned} \zeta * (\zeta \circ S) &= (\zeta \circ \mathrm{id}) * (\zeta \circ S) \\ & \zeta \circ (\mathrm{id} * S) \\ &= \zeta \circ (u\varepsilon) \end{aligned}$$

Now we try it on $[x,x]$:

$$(\zeta u\varepsilon)([x,x]) = \zeta(1) = 1$$

and on $[x,y]$ for $x \neq y$, we have

$$(\zeta u\varepsilon)([x,y]) = \zeta(0) = 0$$

so $\zeta * (\zeta \circ S) = u\varepsilon$ and hence $\zeta \circ S$ is the convolution inverse of ζ, so $\mu = \zeta \circ S$ by definition. For (2), we have

$$f * \mu = (g * \zeta) * u = g * (\zeta * \mu) = g * u\varepsilon = g$$

□

> **Comment 6.5.** If you have a nice formula for S, you get a nice formula for μ, because they are essentially the same.

7 Very Rapid q-Counting Overview

Here is the idea: augment a nonnegative integer into a polynomial in q so that eval at $q = 1$ gives the integer back. We do it in a nice way so that when the integer counts something, then the polynomial counts it with some parameter/ statistic.

7.1 Setup

Here are some facts:

$$\begin{aligned} [n]_q &= 1 + q + q^2 + \cdots + q^{n-1} = \frac{1-q^n}{1-q} \\ [n]_q! &= [n]_q[n-1]_q \cdots [1]_q \\ \begin{bmatrix} n \\ k \end{bmatrix}_q &= \frac{[n]_q!}{[k]_q![n-k]_q!} \end{aligned}$$

This accomplishes our first wish. Below are three examples showing why this also satisfies our second wish.

7.1.1 q-Binomial Coefficients Have Nice Properties

As stated in the title, q-binomial coefficients have nice properties: The first properties is as follows

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^k \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q$$

Second, q-binomial coefficients count a lot things, the most classical thing it counts is lattice walks by area (below the walk). If you've heard talks about major index on permutations, that is also this world. For instance,

$$\sum_{\sigma \in S_n} q^{\mathrm{inv}(\sigma)} = \sum_{\sigma \in S_n} q^{\mathrm{maj}(\sigma)} = [n]_q!$$

where $\mathrm{inv}(\sigma)$ is the number of inversion of σ, i.e., pairs of (i, j) such that $i < j$ and $\sigma(i) > \sigma(j)$, and $\mathrm{maj}(\sigma)$ is defined as

$$\mathrm{maj}(\sigma) = \sum_{i \in \mathrm{Des}(\sigma)} i$$

where $\mathrm{Des}(\sigma) = \{i : \sigma(i) > \sigma(i+1)\}$.

8 Final Preparation

8.1 Practice Problems on Species

Problem 1. Show that the hyperbolic identity $e^x = \cosh(x) + \sinh(x)$ can be "lifted" to the land of species. That is, find species F and G such that $F(x) = \cosh(x)$, $G(x) = \sinh(x)$, and $E = F + G$.

Problem 2. Recall that the power set species, or species of subsets, is defined by $\mathcal{P} = E \cdot E$. In this spirit, define the species $\mathcal{P}^{[2]}$ of subsets of cardinality 2. Derive simple expressions for the generating series

$$\mathcal{P}^{[2]}(x) \quad \text{and} \quad \widetilde{\mathcal{P}^{[2]}}(x).$$

Generalize by defining an appropriate species $\mathcal{P}^{[k]}$.

Problem 3. For $n \in \mathbb{N}$ and F a species, define the species $nF = F + F + \cdots + F$ (n terms). In the particular case $F = 1$ (the empty set species) we write $n1$.

(a) Describe the species n.

(b) Show that $nF = n \cdot F$, where the dot on the right hand side denotes species multiplication.

Problem 4. Each of n (distinguishable) telephone poles is painted red, blue or any of an additional k colors. Determine the number of ways this can be done so that an odd number are painted red and an even number blue. Try to think in terms of species when finding the solution.

Problem 5. View $e^{(1+y)x} = e^x e^{yx}$ as an identity among exponential generating series in x with coefficients in the polynomial ring $\mathbb{Q}[y]$. By identifying coefficients in this (trivial) identity what well known result follows?

Problem 6. (a) Let $f(n)$ be the number of integer partitions of n into distinct parts. Find an expression for the ordinary generating series $F(x) = \sum_{n \geq 0} f(n) x^n$.

(b) Let $g(n)$ be the number of integer partitions of n into odd parts. Find an expression for the ordinary generating series $G(x) = \sum_{n \geq 0} g(n) x^n$.

(c) Prove that $f(n) = g(n)$ for all $n \in \mathbb{N}$.

Problem 7. A permutation is a *derangement* if it has no fixed points. Let $\mathrm{Der} \subseteq \mathcal{S}$ be the corresponding species. Derive formulas for $\mathrm{Der}(x)$ and $|\mathrm{Der}[n]|$.

Problem 8. Let End and $\mathcal{A}$ be the species of endofunctions and rooted trees, respectively. Show that the type generating series for endofunctions and rooted trees are related by

$$\widetilde{\mathrm{End}}(x) = \prod_{k \geq 1} \frac{1}{1 - \widetilde{\mathcal{A}}(x^k)}$$

Problem 9. Prove that the cycle index series for the set species, E, is given by

$$Z_E(x_1, x_2, x_3, \ldots) = \exp\left(x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 + \cdots\right)$$

Hint: Work with the augmented cycle index series

$$\hat{Z}_E(x; x_1, x_2, \ldots) = \sum_{n\geq 0} \left(\sum_{\sigma \in \mathcal{S}[n]} |\operatorname{Fix} E[\sigma]|\, x_1^{c_1(\sigma)} x_2^{c_2(\sigma)} x_3^{c_3(\sigma)} \cdots \right) \frac{x^n}{n!}$$

and consider it the generating series for some weighted species.

Problem 10. The species Mob of (binary) mobiles consists of nonempty rooted binary trees where each node has exactly zero or two subtrees, and sibling subtrees are considered unordered. Thus, $\mathrm{Mob} = X + XE_2(\mathrm{Mob})$.

(a) Show that

$$Z_{\mathrm{Mob}}(x_1, x_2, x_3 \ldots) = x_1 \left(1 + \frac{1}{2} \Big(Z_{\mathrm{Mob}}(x_1, x_2, x_3, \ldots)^2 + Z_{\mathrm{Mob}}(x_2, x_4, x_6 \ldots) \Big) \right)$$

(b) Deduce a simple functional equation whose solution is the ordinary generating series $\widetilde{\mathrm{Mob}}(x)$ for unlabeled mobiles. You don't need to solve for $\widetilde{\mathrm{Mob}}(x)$.

Problem 11. Find a (virtual) species F such that $F' = \mathcal{C}$.

8.2 Practice Problems on Posets

Problem 1. Let $P = \{1, 2, 3, 5, 6, 10, 15, 30\}$ ordered by divisibility.

(a) Draw the Hasse diagram of P.

(b) Find all maximal and minimal elements. Does P have a $\hat{0}$? A $\hat{1}$?

(c) Find a chain of maximum length and an antichain of maximum size.

(d) List all elements that cover 6.

Problem 2. Let P be the poset $\{a, b, c\}$ where $a < b$ and c is incomparable to both a and b.

(a) List all downsets of P.

(b) Draw the Hasse diagram of $J(P)$, the poset of downsets ordered by inclusion.

(c) Verify directly that $J(P)$ is a distributive lattice.

Problem 3. For each poset below, determine whether it is a lattice. If not, exhibit a pair of elements with no join or no meet.

(a) The poset of all subgroups of $\mathbb{Z}/12\mathbb{Z}$, ordered by inclusion.

(b) The poset $\{1, 2, 3, 6\}$ under divisibility.

(c) The "diamond" M_3: five elements $\hat{0} < a, b, c < \hat{1}$ with a, b, c mutually incomparable.

Problem 4. The diamond M_3 from Problem 3(c) is a lattice.

(a) Show it is **not** distributive by finding x, y, z such that

$$x \wedge (y \vee z) \neq (x \wedge y) \vee (x \wedge z).$$

(b) The pentagon N_5 has elements $\hat{0} < a < b < \hat{1}$ and $\hat{0} < c < \hat{1}$ with c incomparable to both a and b. Find a distributivity failure in N_5.

(c) Prove or disprove: every sublattice of a distributive lattice is distributive.

Problem 5. Consider B_3, the Boolean lattice on $\{1, 2, 3\}$.

(a) Identify all join-irreducible elements of B_3.

(b) What is the poset P of join-irreducibles (with the inherited order)? Draw it.

(c) Verify the Fundamental Theorem of Finite Distributive Lattices: show that $J(P) \cong B_3$.

Problem 6. Let L be the chain of length 3, i.e., $L = \{\hat{0}, a, b, \hat{1}\}$ with $\hat{0} < a < b < \hat{1}$.

(a) Check that L is a distributive lattice.

(b) Find the poset P of join-irreducible elements of L.

(c) Show that $J(P) \cong L$.

8.2.1 Practice Problems on Möbius Inversion

Problem 1. Find $m, n \in \mathbb{N}$ such that they have no prime divisors other than 2 and 3, $(m,n) = 18$, $\tau(m) = 21$, and $\tau(n) = 10$.

Problem 2. Find $n \in \mathbb{N}$ such that one of the following is satisfied: $\pi(n) = 2^3 \cdot 3^6$, $\pi(n) = 3^{30} \cdot 5^{40}$, $\pi(n) = 13 \cdot 31$, or $\tau(n) = 13 \cdot 31$.

Problem 3. Define $\sigma_k(n) = \sum_{d|n} d^k$. Thus, $\sigma_0(n) = \tau(n)$ and $\sigma_1(n) = \sigma(n)$. Prove that $\sigma_k(n)$ is multiplicative for all $k \in \mathbb{N}$, and find a formula for it.

Problem 4. Show that $\tau(n)$ is odd iff n is a perfect square, and that $\sigma(n)$ is odd iff n is a perfect square or twice a perfect square.

Problem 5. If $f(n)$ is multiplicative, $f \not\equiv 0$, then show $\sum_{d|n} \mu(d) f(d) = \prod_{i=1}^{r} (1 - f(p_i))$.

Proof. we have

$$\sum_{d|n} \mu(d) f(d) = \sum_{\substack{d|n \\ d \text{ squarefree}}} \mu(d) f(d) = \sum_{I \subseteq [r]} (-1)^{|I|} \prod_{i \in I} f(p_i) = \prod_{i=1}^{r} \big(1 - f(p_i)\big).$$

as desired. □

Problem 6. If $f(n)$ is multiplicative, then show that $h(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d)$ is also multiplicative. Conclude that every multiplicative function is the sum-function of another multiplicative function.

Problem 12. Let $f(x) \in \mathbb{Z}[x]$ and let $\psi(n)$ be the number of values $f(j)$, $j = 1, 2, \ldots, n$, such that $(f(j), n) = 1$. Show that $\psi(n)$ is multiplicative and that $\psi(p^t) = p^{t-1}\psi(p)$. Conclude that

$$\psi(n) = \prod_{p|n} \psi(p)/p.$$

Problem 13. Find closed expressions for the following sums:

- $\sum_{d|n} \mu(d)\tau(d)$
- $\sum_{d|n} \mu(d)\varphi(d)$
- $\sum_{d|n} \mu(d)\sigma(d)$
- $\sum_{d|n} \mu^2(d)\varphi^2(d)$
- $\sum_{d|n} \mu(d)\tau\left(\frac{n}{d}\right)$
- $\sum_{d|n} \frac{\mu(d)}{\varphi(d)}$
- $\sum_{d|n} \mu(d)\sigma\left(\frac{n}{d}\right)$
- $\sum_{d|n} \mu\left(\frac{n}{d}\right) \ln d$
- $\sum_{d|n} \frac{\mu(d)}{d}$
- $\sum_{\substack{(t,n)=1 \\ 1 \le t < n}} t$

Problem 14. Consider the function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$, the so-called *Riemann zeta-function.* It converges for $s > 1$. In fact, one can extend it to an *analytic* function over the whole complex plane except $s = 1$. The famous Riemann Conjecture claims that all zeros of $\zeta(s)$ in the strip $0 \leq \operatorname{Re} s \leq 1$ lie on the line $\operatorname{Re} s = 1/2$. For $\zeta(s)$ show the following formal identities:

- $\zeta(s) = \prod_{p} \frac{1}{1 - p^{-s}}$
- $\zeta(s)^{-1} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$
- $\zeta(s)^2 = \sum_{n=1}^{\infty} \frac{\tau(n)}{n^s}$
- $\zeta(s)\zeta(s-1) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n^s}$

Problem 15. Suppose that we are given infinitely many tickets, each with one natural number on it. For any $n \in \mathbb{N}$, the number of tickets on which divisors of n are written is exactly n. For example, the divisors of 6, $\{1, 2, 3, 6\}$, are written in some variation on 6 tickets, and no other ticket has these numbers written on it. Prove that any number $n \in \mathbb{N}$ is written on at least one ticket.

Problem 16. Let $f(n) : \mathbb{N} \to \mathbb{N}$ be multiplicative and strictly increasing. If $f(2) = 2$, then $f(n) = n$ for all n.

8.3 Practice Problems on Hopf Algebra

Problem 1. Let $V = \mathbb{Q}^2$ with basis $\{e_1, e_2\}$.

(a) Show that $e_1 \otimes e_1 + e_2 \otimes e_2$ cannot be written as a single pure tensor $v \otimes w$ for any $v, w \in V$.

(b) Let $m : V \otimes V \to V$ be the linear map induced by the bilinear map $\mu(e_i, e_j) = e_{\min(i,j)}$. Compute $m(e_1 \otimes e_2)$, $m(e_2 \otimes e_1)$, and $m(e_1 \otimes e_1 + e_2 \otimes e_2)$.

Problem 2. Let $A = \mathbb{Q}[x]/(x^3)$ with the usual polynomial multiplication m and unit $u : \mathbb{Q} \to A$ given by $u(1) = 1$.

(a) Verify the associativity diagram for m by computing both $m \circ (m \otimes \mathrm{id})$ and $m \circ (\mathrm{id} \otimes m)$ on the element $\bar{x} \otimes \bar{x} \otimes \bar{x}$.

(b) Verify the unit diagram: show $m(u(1) \otimes \bar{x}) = \bar{x} = m(\bar{x} \otimes u(1))$.

Problem 3. Let $\Omega = \{a, b\}$ and consider the coalgebra $C = \mathrm{span}_{\mathbb{Q}}\Omega^*$ with deconcatenation coproduct and counit $\varepsilon(1) = 1$, $\varepsilon(w) = 0$ for nonempty words w.

(a) Compute $\Delta(abba)$.

(b) Verify coassociativity for the word ab, i.e., show that

$$(\Delta \otimes \mathrm{id}) \circ \Delta(ab) = (\mathrm{id} \otimes \Delta) \circ \Delta(ab).$$

(c) Verify the counit axiom for ab: show that $(\varepsilon \otimes \mathrm{id}) \circ \Delta(ab) = ab = (\mathrm{id} \otimes \varepsilon) \circ \Delta(ab)$ under the identification $\mathbb{Q} \otimes C \cong C \cong C \otimes \mathbb{Q}$.

Problem 4. Let $\Omega = \{a, b\}$ and work in $\mathrm{span}_{\mathbb{Q}}\Omega^*$.

(a) Compute the shuffle product $ab \sqcup\!\sqcup\, ba$.

(b) Compute the deshuffle coproduct $\Delta_{desh}(abc)$ where $\Omega = \{a, b, c\}$.

Problem 5. Let H be a bialgebra with multiplication m, unit u, comultiplication Δ, and counit ε.

(a) Write out explicitly what it means for Δ to be an algebra homomorphism, i.e., $\Delta \circ m = (m \otimes m) \circ (\mathrm{id} \otimes \tau \otimes \mathrm{id}) \circ (\Delta \otimes \Delta)$, where $\tau(a \otimes b) = b \otimes a$ is the swap map.

(b) Show that saying "Δ and ε are algebra homomorphisms" is equivalent to saying "m and u are coalgebra morphisms." (Draw the commutative diagrams for each statement and verify they are the same.)

Problem 6. Consider $\mathbb{Q}[x]$ with the usual multiplication and with $\Delta(x) = x \otimes 1 + 1 \otimes x$ extended as an algebra homomorphism. The counit is $\varepsilon(x^n) = 0$ for $n \geq 1$ and $\varepsilon(1) = 1$.

(a) Verify coassociativity on x^2: show $(\Delta \otimes \mathrm{id})\Delta(x^2) = (\mathrm{id} \otimes \Delta)\Delta(x^2)$.

(b) Verify the antipode axiom $m \circ (S \otimes \mathrm{id}) \circ \Delta(x^2) = u(\varepsilon(x^2))$.

Problem 7. Let C be a coalgebra and A an algebra. Let $f, g : C \to A$ be linear maps with convolution $f * g = m \circ (f \otimes g) \circ \Delta$.

(a) Show that convolution is associative: $(f * g) * h = f * (g * h)$.

(b) Show that $u \circ \varepsilon$ is the identity for convolution, i.e., $(u \circ \varepsilon) * f = f = f * (u \circ \varepsilon)$.

(c) Conclude that the set of linear maps $C \to A$ forms an algebra under convolution with identity $u \circ \varepsilon$.

Problem 8. Recall that for both word Hopf algebras (concatenation/deshuffle and shuffle/deconcatenation), the antipode is

$$S(w_1 w_2 \cdots w_n) = (-1)^n w_n w_{n-1} \cdots w_1.$$

(a) Verify the antipode axiom for the word ab in the concatenation/deshuffle bialgebra:

$$m_{concat} \circ (S \otimes \mathrm{id}) \circ \Delta_{desh}(ab) = u(\varepsilon(ab)).$$

(b) Compute $S(abc)$ and verify the antipode axiom on abc in the same bialgebra.

Problem 9. Let G be a finite group and KG the group ring with $\Delta(g) = g \otimes g$, $\varepsilon(g) = 1$, and $S(g) = g^{-1}$ for all $g \in G$.

(a) Show that Δ extends to an algebra homomorphism $KG \to KG \otimes KG$.

(b) Verify the counit axiom and the antipode axiom on a general group element g.

(c) This Hopf algebra is **not** graded connected (unless G is trivial). Explain why.

(d) Show that $\Delta(g) = g \otimes g$ makes every group element "group-like." Prove that in any Hopf algebra, a group-like element g (i.e., $\Delta(g) = g \otimes g$, $\varepsilon(g) = 1$) must satisfy $S(g) \cdot g = 1 = g \cdot S(g)$.

Problem 10. Let $B = \bigoplus_{n \geq 0} B_n$ be a graded connected bialgebra (so $B_0 = K$). For $x \in B_n$ with $n \geq 1$, write

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \sum_i x_{i,1} \otimes x_{i,2}$$

where each $x_{i,1} \in B_j$, $x_{i,2} \in B_k$ with $j, k \geq 1$ and $j + k = n$. Explain why the recursive formula $S(x) = -x - \sum_i S(x_{i,1})\, x_{i,2}$ terminates: why does computing $S(x)$ for $x \in B_n$ only require knowing S on $\bigoplus_{k=0}^{n-1} B_k$?

Problem 11. Let $\mathcal{F}$ denote the set of rooted forests (including the empty forest $\mathbb{1}$), and let $H = \mathrm{span}_K(\mathcal{F})$ with disjoint union as multiplication. The coproduct is defined on a tree t by summing over downsets:

$$\Delta(t) = \sum_{D \text{ downset of } t} D \otimes (t \setminus D).$$

Let t be the rooted tree that is a path of length 2 (three vertices: root → child → grandchild).

(a) List all downsets of t and compute $\Delta(t)$.

(b) Using the recursive formula for the antipode, compute $S(t)$.

(c) Verify the antipode axiom: $m \circ (S \otimes \mathrm{id}) \circ \Delta(t) = u(\varepsilon(t))$.

Problem 12. Let P be the poset $\{1, 2, 3, 4\}$ with the usual total order $1 < 2 < 3 < 4$.

(a) Compute $\Delta([1, 4])$ in the incidence coalgebra.

(b) Compute $\Delta([1, 3])$ and $\Delta([2, 4])$.

(c) In the reduced incidence coalgebra (intervals up to isomorphism), intervals only depend on length. Write e_k for the class of intervals of length k. Express $\Delta(e_3)$ in terms of $e_i \otimes e_j$.

(d) The ζ-function sends every interval class to 1. Use $\mu = \zeta \circ S$ and the antipode formula

$$S([x,y]) = \sum_{k \geq 0} \sum_{x = x_0 < x_1 < \cdots < x_k = y} (-1)^k \prod_{i=1}^{k} [x_{i-1}, x_i]$$

to compute $\mu(e_0)$, $\mu(e_1)$, $\mu(e_2)$, and $\mu(e_3)$ for the total order. Verify these match the Möbius function you expect.

Problem 13. Let P be a locally finite poset with incidence Hopf algebra H.

(a) Recall that $\mu = \zeta \circ S$. Starting from $\zeta * (\zeta \circ S) = u \circ \varepsilon$, derive the recursive formula for the Möbius function:

$$\mu(x,y) = \begin{cases} 1 & x = y \\ -\sum_{x \leq z < y} \mu(x,z) & x < y. \end{cases}$$

(b) Let P be the Boolean poset B_n. Using the antipode formula, show that $\mu(S,T) = (-1)^{|T|-|S|}$ for $S \subseteq T$.

A Appendix – Some Theorems

Theorem A.1. Binomial Theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k}, \qquad n \in \mathbb{Z}_{\geq 0}.$$

Theorem A.2. Negative Binomial Theorem

$$(1-x)^{-t} = \sum_{n\geq 0} \binom{n+t-1}{t-1} x^n = \sum_{n\geq 0} \binom{n+t-1}{n} x^n, \qquad (t \in \mathbb{Z}_{>0}).$$

In particular, we have

$$\binom{-t}{n} = (-1)^n \binom{n+t-1}{t-1} = (-1)^n \binom{n+t-1}{n}.$$

Theorem A.3. Trinomial Theorem

$$(a+b+c)^n = \sum_{\substack{i,j,k\geq 0\\ i+j+k=n}} \binom{n}{i,j,k} a^i b^j c^k.$$

where the trinomial coefficient is given by

$$\binom{n}{i,j,k} = \frac{n!}{i!\, j!\, k!}$$

B Combinatorial Proof of LIFT

Lecture 5 - Monday, January 19

Recall how a combinatorial proof works, we have the following two schemes:

```
Find a set $\mathcal{A}$ such that $|\mathcal{A}| = a$;
Find a set $\mathcal{B}$ such that $|\mathcal{B}| = b$;
Find a bijection between $\mathcal{A}$ and $\mathcal{B}$.
```

Algorithm 1: Prove $a = b$

```
Find a set $\mathcal{A}$ and a weight function such that $A(x) = a(x)$;
Find a set $\mathcal{B}$ and a weight function such that $B(x) = b(x)$;
Find a bijection between $\mathcal{A}$ and $\mathcal{B}$.
```

Algorithm 2: Prove $a(x) = b(x)$

B.1 Part 1 — $\phi(x) = (1-x)^{-1}$

1. If $A(x) = x\phi(A(x))$, what is the combinatorial interpretation of $[x^n]A(x)$?

We can think of $\mathcal{A}$ as the class of plane trees, so $[x^n]A(x)$ is the number of plane trees with n vertices (nodes).

2. Give an example of a set and weight function whose OGF is $\phi(x) = (1-x)^{-1}$.

A simple example is the class of non-negative integers.

3. What is the combinatorial interpretation of $[x^{n-1}]\phi(x)^n$?

Number of tuple of n non-negative inters whose entries sum up to be $n-1$.

4. What is the combinatorial interpretation of $\frac{1}{n}[x^{n-1}]\phi(x)^n$?

Number of tuple of n non-negative inters whose entries sum up to be $n-1$. However, tuples that are the same up to cyclic shift is considered as the same one. Because the size is strictly smaller than the number of parts, so the number of tuples (the value of $[x^{n-1}]\phi(x)^n$) is always a multiple of n.

Note B.1. The above works because $\gcd(n-1, n) = 1$. See the example below for a case that fails.

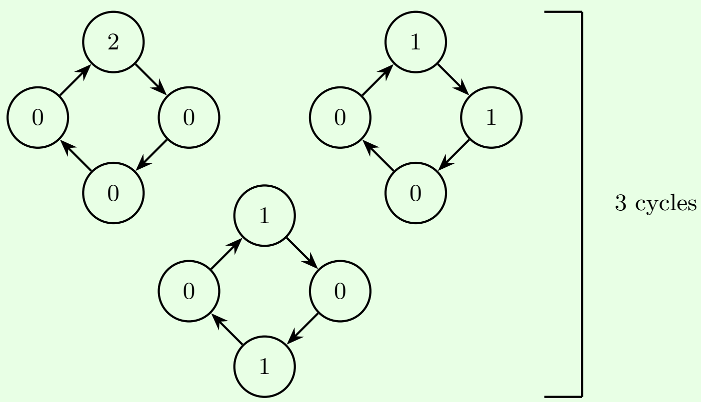
Example B.1. Sequences vs. cycles, length = 4 and sum = 2.

Sequences

10 sequences

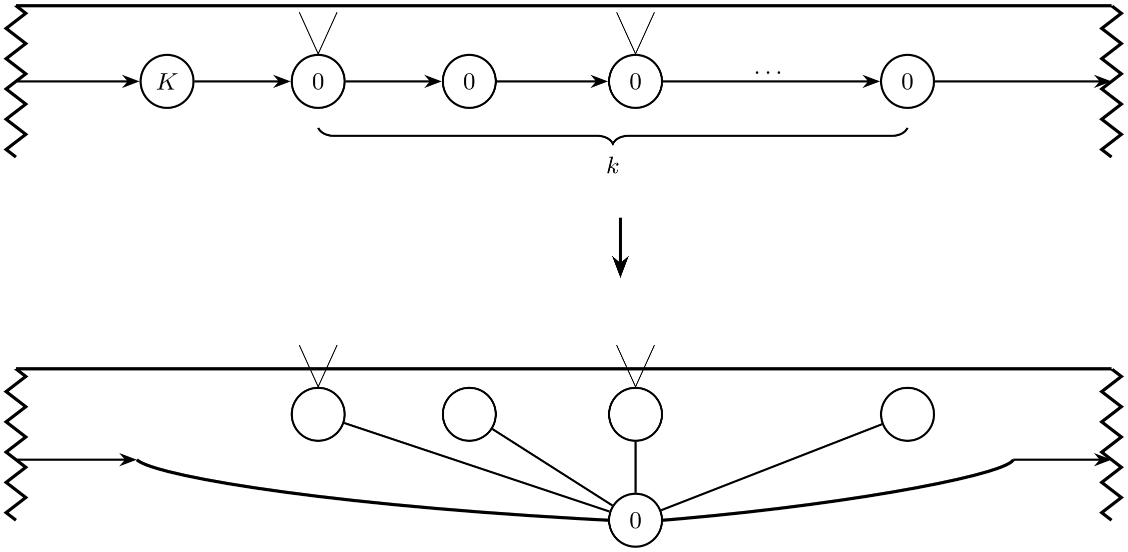
$(2,0,0,0)$
$(0,2,0,0)$
$(0,0,2,0)$
$(0,0,0,2)$
$(1,1,0,0)$
$(0,1,1,0)$
$(0,0,1,1)$
$(1,0,0,1)$
$(1,0,1,0)$
$(0,1,0,1)$

Cycles



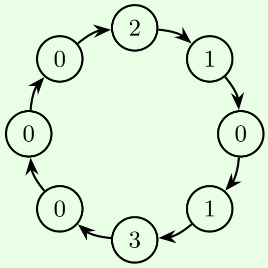
Now, to prove the $\phi(x) = (1-x)^{-1}$ case, we wish to find a bijection between the plane trees of n nodes and those "acyclic" n-tuples of size $n-1$.

Transformation rule:

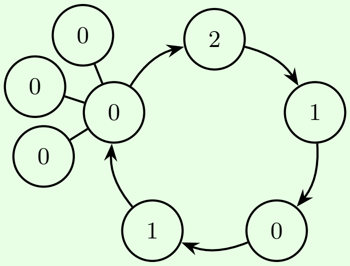


The above is the transformation rule in visual. In words, each transformation takes a node with value k and the subsequent k nodes with 0 and contract it. The new node has value 0, and there are k children attached to it. The picture below is an example of transforming a tuple (cycle) to a plane tree.

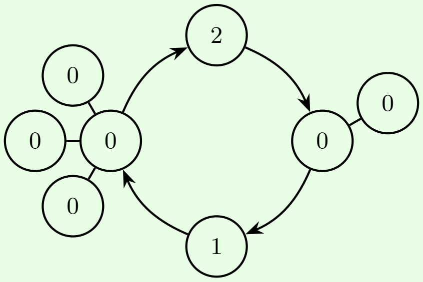
Example B.2. The picture below is a "acyclic" tuple we start with:



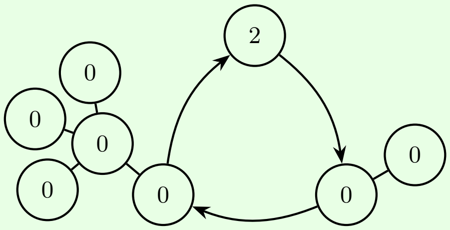
The following pictures illustrate the state after each transformation.



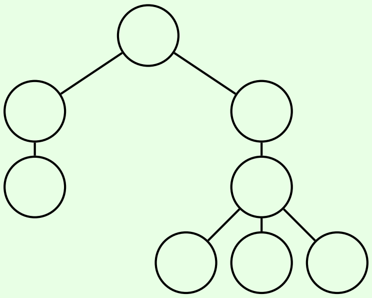
Transformation 1



Transformation 2



Transformation 3



Transformation 4

Discovery B.1. I discovered that the cycle can be obtained quickly from the tree using DFS, where the numbers in the cycle are the children of the nodes in DFS order.

B.2 Part 2 — Add Weight Function

Lecture 6 - Wednesday, January 21

5. What weight function can we add to sets Q_n and C_n so that bijection alone is weight preserving?

> **Comment B.1.** In other words, think about the parameters that go through the bijection nicely (you will get a sequence of parameters).

For Q_n, we let w_i to be the weight function that returns the number of vertices with i children, and for C_n, let w'_i be the weight fucntion that returns the number of entries of value i in a cycle.

Now we are almost done. If $\phi(x) = \sum_{n\geq 0} y_n x^n$ and $A(x) = x\phi(A(x))$. Show that

(a) The LHS of LIFT:

$$[x^n]A(x) = \sum_{q\in Q_n} y_0^{w_0(q)} y_1^{w_1(q)} \cdots$$

(b) The RHS of LIFT:

$$\frac{1}{n}[x^{n-1}]\phi(x) = \sum_{c\in C_n} y_0^{w'_0(q)} y_1^{w'_1(q)} \cdots$$

and then we are done with the proof.

C Rooted Triangulations Done Right

Just for fun – Wednesday, January 21

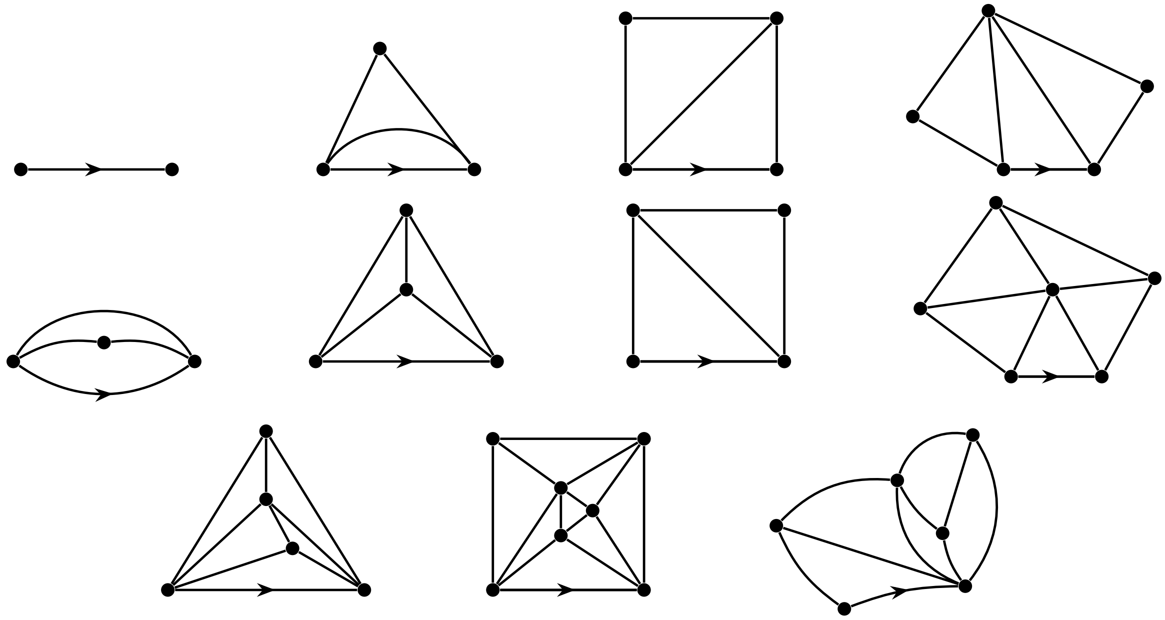
C.1 Introduction

C.1.1 Rules

We have the following rules for rooted triangulations:

- Planar drawings up to isotopy;
- All internal faces have degree 3 (outer face has degree ≥ 2);
- Parallel edges are allowed;
- Loop is not allowed, cut vertex is not allowed;
- The root edges, γ, is oriented counterclockwise on the outer face.

C.1.2 Examples



C.1.3 Notation

We use $\mathcal{T}$ to denote the class of all rooted triangulations, and we use $\mathcal{T}_n$ to denote the rooted triangulations whose outer face has degree k.

We have the following two weight functions: w_1 returns the number of edges, and w_2 returns the degree of the outer face. As a result, we have the following two generating series:

$$T(x,y) = \sum_{t \in \mathcal{T}} x^{w_1(t)} y^{w_2(t)}$$

$$T_k(x) = \sum_{t \in \mathcal{T}_k} x^{w_1(t)} = [y^k]T(x,y)$$

C.2 Rooted Triangulations

Step 1. Give a recursive decomposition of $\mathcal{T}$.

Step 2. Use the above decomposition to obtain an equation involving the ordinary generating series for $\mathcal{T}$ (this equation is quadratic in T, but it also involves T_2).

Step 3. Let R be a ring, if $a, b, c, z \in R$,

$$az^2 + bz + c = 0$$

What must be true for a, b, c?

> **Example C.1.** Prove that $y^2 + x^4$ is not a square in $\mathbb{C}[[x, y]]$.

Step 4. Find the unique series $T_2(x)$ such that

$$(y - x)^2 - 4x^2y^2(y - T_2(x))$$

is a square in $\mathbb{Q}[[x, y]]$.
Hint: Try to get to the form $g(x) = x\phi(g(x))$, $T_2(x) = f(g(x))$.

D Averages

Lecture 11 - Wednesday, February 11

Let $\mathcal{S}$ be a finite set of combinatorial objects with weight function $\phi : \mathcal{S} \to \mathbb{N}$. To find the average, the generating function method is

$$\operatorname{Avg}(\phi) = \frac{S'(1)}{S(1)}$$

where $S(x) = \sum_{s \in \mathcal{S}} x^{\phi(s)}$ is the ordinary generating function. Here is a more elementary method: Suppose $\phi(s)$ counts the number of widgets ub s, and let W denote the set of all possible widgets, then for $s \in \mathcal{S}$, $w \in \mathcal{W}$, we have

$$\delta_{s,w} = \begin{cases} 1 & \text{if object } s \text{ contains widget } w \\ 0 & \text{otherwise} \end{cases}$$

and then the average is

$$\operatorname{Avg}(\phi) = \frac{1}{|\mathcal{S}|} \cdot \sum_{s \in \mathcal{S}} \sum_{w \in \mathcal{W}} \delta_{s,w}$$

Here are two interpretations:

1. First interpretation:
$$\operatorname{Avg}(\phi) = \sum_{w \in \mathcal{W}} \frac{\#\text{ of objects containing } w}{\text{total }\#\text{ of objects}}$$
2. Second interpretation:
$$\operatorname{Avg}(\phi) = \sum_{w \in \mathcal{W}} \mathbb{P}(w)$$
where $\mathbb{P}(w)$ is the probability that w is contained in a randomly chosen object, random in the sense of a uniform distribution on $\mathcal{S}$.

More generally, we can assign different probability to choosing objects. In this case, the second interpretation above would be the expected value of ϕ:

$$\mathbb{E}(\phi) = \sum_{w \in \mathcal{W}} \mathbb{P}(w) \quad \leftarrow \text{ no longer using uniform distribution}$$

D.1 Let's Explore

1. Let B_n be the set of 01-strings of length n for $n \geq 3$. For $\sigma \in B_n$, let $\phi(\sigma)$ be the number of occurrences of 111 as a substring. Compute $\operatorname{Avg}(\phi)$.

Answer. The answer is $\frac{2^{n-3}}{2^n}(n-2) = \frac{n-2}{8}$. ☐

2. Let $\mathcal{G}_X$ be the set of graphs on X where $|X| = n$. For $\Gamma \in \mathcal{G}_X$, let $\phi(\Gamma)$ be the number of spanning trees of Γ. Compute $\operatorname{Avg}(\phi)$.

Answer. The answer is $\frac{n^{n-2} \cdot 2^{\binom{n}{2}-(n-1)}}{2^{\binom{n}{2}}} = \frac{n^{n-2}}{2^{n-1}}$. ☐

Lecture 12 - Monday, February 23

3. Let S_n be the set of permutations on $[n] \to [n]$, and we fix $k \leq n$. For $\sigma \in S_n$, let $\phi(\sigma)$ be the number of k-cycles in σ. Compute $\text{Avg}(\phi)$.

Answer. The answer is $\dfrac{\binom{n}{k}(k-1)!(n-k)!}{n!} = \dfrac{1}{k}$. ☐

4. On each square of an $n \times n$ chess board, we place a stone with probability p, what is the expected number of pairs of adjacent stons?

Answer. $\mathbb{E}(\phi) = 2n(n-1)p^2$. ☐

D.2 Generating Function Returns

In the proceeding examples, probabilities were the same for all $w \in \mathcal{W}$. When this is not the case, we can sometimes compute $\text{Avg}(\phi)$ using a generating function for w.

If there exists a weight function $wt : \mathcal{W} \to \mathbb{N}$ and $0 < p < 1$ such that $\mathbb{P}(w) = p^{wt(w)}$, then

$$\text{Avg}(\phi) = \sum_{w \in \mathcal{W}} p^{wt(w)} = W(p)$$

where $W(x)$ is the ordinary generating function for W. Let's consider this example: Let $\mathcal{G}_X$ be the set of graphs on X where $|X| = n$. For $\Gamma \in \mathcal{G}_X$, let $\phi_X(\Gamma) : \mathcal{G}_X \to \mathbb{N}$ be the number of cycles in Γ. Our goal is to compute $\text{Avg}(\phi)$. We do this in three steps:

(a) What is $\mathcal{W}_X$, the set of widgets?

Answer. The set of widgets is cycles on subsets of X,

$$\mathcal{W} = \tilde{\mathcal{C}} * \mathcal{E}$$

where $2\tilde{\mathcal{C}} = \mathcal{C} - \mathcal{C}_{1,2}$. ☐

(b) What is the weight function on $\mathcal{W}_X$?

(c) Compute the mixed generating function $W(x;t)$, hence find $\text{Avg}(\phi_X)$.

Answer. We have

$$\tilde{C}(x) = \frac{1}{2}\left(\log\left(\frac{1}{1-x}\right) - x - \frac{x^2}{2}\right)$$

and hence

$$\tilde{C}(xt) = \sum_{n \geq 3} \frac{(n-1)!}{2} \frac{x^n}{n!} t^n$$

which is a exponential generating seires in x, but an ordinary generating series in t. As a result,

$$\text{Avg}(\phi_X) = \left[\frac{x^n}{n!}\right] W\left(x, \frac{1}{2}\right)$$

as desired. ☐

E The Cyclic Sieving Phenomenon

Let $C_n = (\mathbb{Z}_n, +)$ be the cyclic group of order n and $\mathcal{Y}$ be a finite set with C_n-action. Suppose $wt : \mathcal{Y} \to \mathbb{N}$ is a weight function, then

$$Y(q) = \sum_{y \in \mathcal{Y}} q^{wt(y)}$$

is the ordinary generating function for Y, which is a polynomial with $\mathbb{N}$-coefficients. One more thing, let $\zeta = \exp\{2\pi i/n\}$ be the primitive n^{th} root of unity.

Definition E.1. **[Sieving Polynomial]**

$Y(q)$ is called a **sieving polynomial** for $(\mathcal{Y}, C_n)$ if

$$|\mathcal{Y}^c| = Y(\zeta^c) \qquad \forall c \in C_n$$

where $\mathcal{Y}^c$ is the set of c-fixed points, and the RHS is the OGF evaluated at $\zeta^c \in \mathbb{C}$.

Comment E.1. This is kind of weird, getting cardinality of a set by evaluating a polynomial at a complex number. For most weight functions on $\mathcal{Y}$, $Y(q)$ is NOT a sieving polynomial (unless $n = 1$).

Theorem E.1.

Sieving polynomials exist.

Note E.1. In practice, they are hard to determine, so we often have to guess. Here is an important clue: $Y(1) = |\mathcal{Y}|$.

Discovery E.1. What is so phenomenal? The "first thing you might guess" actually works, i.e., q-analogues.

E.1 2-Colouring of n-gon

Let $\mathcal{Y}$ be

$$\mathcal{Y} := \{\text{colourings of sides of regular } n\text{-gon, with } k \text{ reds and } n-k \text{ blues}\}$$

and let C_n act on $\mathcal{Y}$ by rotating the n-gon. Then

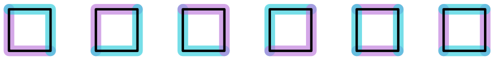
$$|\mathcal{Y}| = \underline{\qquad\qquad}$$

Guess that the sieving polynomial for $(\mathcal{Y}, C_n)$ is

$$Y(q) = \underline{\qquad\qquad}$$

E.1.1 Verify our Guess

Here, we verify our guess for $n = 4$ and $k = 2$. Hence, we have

$y = \{$  $\}$

$$Y(q) = \frac{(q^4-1)(q^3-1)}{(q^2-1)(q-1)} \qquad \text{and} \qquad \zeta = i$$

c	$\lvert\mathcal{Y}^c\rvert$	$Y(\zeta^c)$
$1, 3$		
0		
2		

E.1.2 Verify our Guess, again

Now we verify our guess for $n = 8$ and $k = 4$:

$$Y(q) = \frac{(q^8-1)(q^7-1)(q^6-1)(q^5-1)}{(q^4-1)(q^3-1)(q^2-1)(q-1)}$$

c	$\lvert\mathcal{Y}^c\rvert$	$Y(\zeta^c)$
$1, 3, 5, 7$		
0		
4		
$2, 6$		

E.2 C_n Acting with a Single Orbit

We want to study an action of C_n on a set Y which has exactly one orbit. In other words, the action is *transitive*: every element of Y can be reached from any other element by applying a suitable group element.

A standard way to obtain such an action is to let

$$Y = \mathbb{Z}_d = \{0, 1, \ldots, d-1\},$$

where $d \mid n$. Since d divides n, the cyclic group C_n has a quotient of order d, and this gives a natural transitive action on $\mathbb{Z}_d$.

We write elements of C_n additively, so that an element $c \in C_n$ acts on $k \in \mathbb{Z}_d$ by translation:

$$C_n \times Y \to Y, \qquad (c, k) \longmapsto c + k \pmod{d}.$$

Thus each group element rotates the set $\mathbb{Z}_d$ by a fixed amount modulo d.

> **Note E.2.** Because addition modulo d can move any element of $\mathbb{Z}_d$ to any other element, this action has a single orbit.

We are asked to find the *Sieving Polynomial* for this action in the following cases:

(a) $d = 1$,

(b) $d = n$.

(c) $n = 4, d = 2$.

(d) $n = 6, d = 2$.

(e) $n = 6, d = 3$.

Hint. First determine $|\mathcal{Y}^c|$ for all $c \in C_n$.

E.3 Main Result

> **Theorem E.2.**
>
> If C_n acts on $\mathcal{Y}$, then there exists a sieving polynomial.

Proof. If $\mathcal{Y}$ has a single orbit of size d, where $d|n$, then take

____

If $\mathcal{Y}$ has multiple orbits, then take

____

$\square$

Corollary E.1. Orbit Counting via the Sieving Polynomial

Suppose $Y(q)$ is a sieving polynomial for $(\mathcal{Y}, C_n)$, then

$$|\widetilde{\mathcal{Y}}| = \sum_{k\geq 0}[q^{kn}]Y(q)$$

where $|\tilde{\mathcal{Y}}|$ denotes the number of orbits.

Proof. By the orbit counting lemma,

$$|\widetilde{\mathcal{Y}}| = \frac{1}{n}\sum_{c=0}^{n-1}|Y^c| = \frac{1}{n}\sum_{c=0}^{n-1}\gamma(\xi^c).$$

Recall: if

$$\gamma(q) = \sum_{m\geq 0} f_m q^m,$$

then

$$\frac{1}{n}\sum_{c=0}^{n-1}\gamma(q\xi^c) = \sum_{k\geq 0} f_{kn}q^{kn}.$$

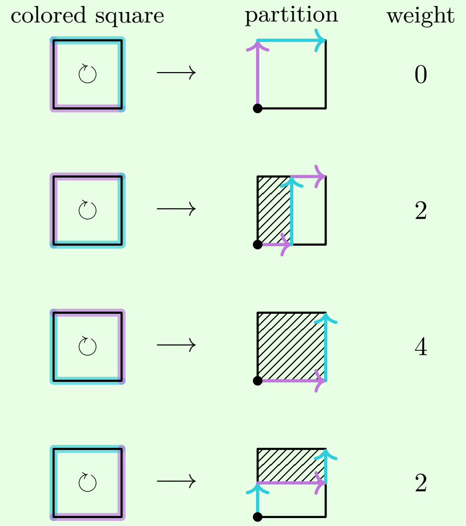
Putting $q = 1$ gives

$$\sum_{k\geq 0} f_{kn}.$$

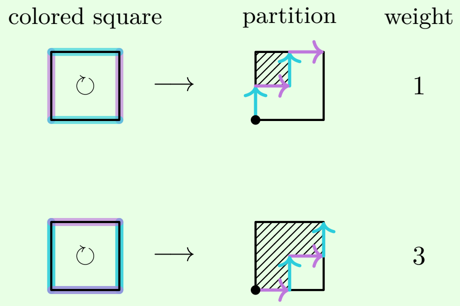
as desired. ☐

E.3.1 Examples

Example E.1. This example makes no sense. Recall that $\binom{4}{2}_q$ is the OGF for partitions inside 2×2 square. Let the weight function be $wt(\lambda) = |\lambda|$, then



First orbit: the four colored squares, their corresponding partitions, and the weights. The orbit generating function is $1 + 2q^2 + q^4$.



Second orbit: the two colored squares, their corresponding partitions, and the weights. The orbit generating function is $q + q^3$.

Example E.2. Let

$$\mathcal{S}_m = \left\{ f : \mathbb{Z}_n \to \mathbb{N} \;\middle|\; \sum_{i=0}^{n-1} f(i) = m \right\}.$$

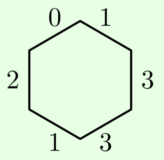
This is the set of functions assigning a natural number to each side of an n-gon, with total sum m. The cyclic group C_n acts on $\mathcal{S}_m$ by rotation.

For example, the labeling

$$0, 1, 3, 3, 1, 2$$

on a hexagon defines an element of $\mathcal{S}_{10}$, since

$$0 + 1 + 3 + 3 + 1 + 2 = 10.$$



$\in \mathcal{S}_{10}.$

We saw that

$$|\mathcal{S}_m| = [x^m](1-x)^{-n} = \binom{n+m-1}{m}.$$

Theorem E.3.

The Gaussian binomial coefficient

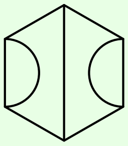
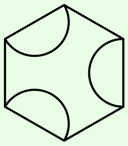
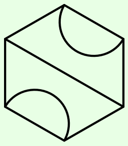
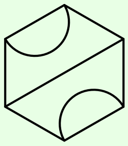
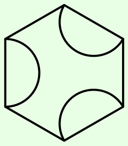
$$\binom{n+m-1}{m}_q$$

is a sieving polynomial for the action of C_n on $\mathcal{S}_m$.

Example E.3. Let

$$\mathcal{H}_n = \{\text{non-crossing matchings of the vertices of a } 2n\text{-gon}\}\,.$$

For example, when $n = 3$, the elements of $\mathcal{H}_3$ are:

We have

$$|\mathcal{H}_n| = \frac{1}{n+1}\binom{2n}{n},$$

which is the n-th Catalan number.
The cyclic group C_{2n} acts on $\mathcal{H}_n$ by rotation. Define

$$\mathrm{Cat}_n(q) = \frac{1}{[n+1]_q}\binom{2n}{n}_q,$$

the q-Catalan number, where

$$[n+1]_q = (n+1)_q = \frac{q^{n+1}-1}{q-1}.$$

Theorem E.4.

$\mathrm{Cat}_n(q)$ is a sieving polynomial for $(\mathcal{H}_n, C_{2n})$.

Theorem E.5.

$\mathrm{Cat}_n(q)$ is a sieving polynomial for $(\mathcal{J}_n, C_{n+2})$, where

$$\mathcal{J}_n = \{\text{triangulations of } (n+2)\text{-gon, no internal vertices}\}$$

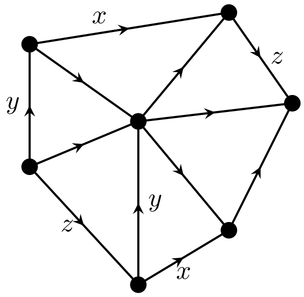
F Non-Intersecting Paths

Here is the set up: Γ is an acyclic directed graph, and R is a ring. For each $e \in E(\Gamma)$, we assign a weight $wt(e) \in R$, and for each subgraph $Q \subseteq \Gamma$, we define

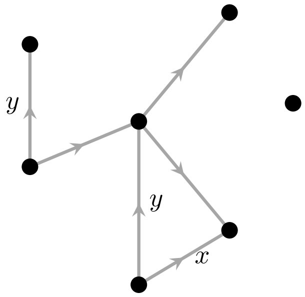
$$wt(Q) = \prod_{e \in E(Q)} wt(e)$$

F.1 Example, $R = \mathbb{Q}[x, y, z]$

Suppose our graph Γ is as follows:



where if an edge has no weight written, then by convention it has weight 1. Then the following subgraph Q has weight $\Omega(Q) = xy^2$.

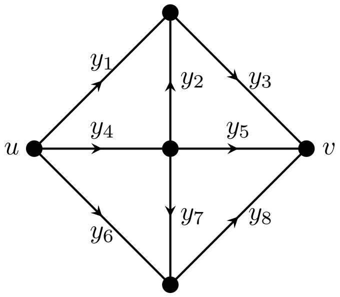


Definition F.1. **[Path Enumerator]**

For $u, v \in V(\Gamma)$, we define

$$P(u, v) = \sum_{\text{paths } Q: u \rightsquigarrow v} \Omega(Q)$$

F.2 Compute $P(u, v)$ and Calculate Determinants



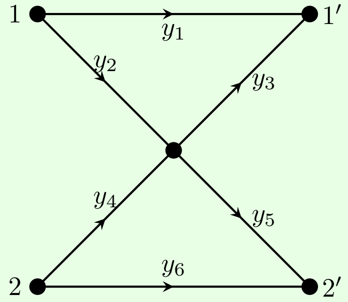
In the above graph, we have

$$P(u,v) = y_6y_8 + y_4y_5 + y_4y_2y_3 + y_4y_7y_8 + y_1y_3$$

Now we understand how the path enumerator is calculated, we will calculate (useful) determinants in the following example:

> **Comment F.1.** Take note of which terms cancel and which terms don't.

Example F.1.



Compute

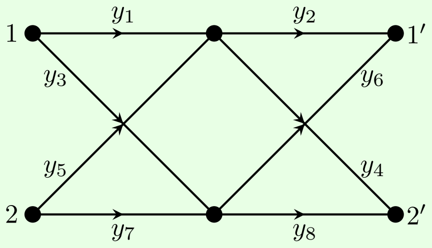
$$\det\begin{pmatrix} P(1,1') & P(1,2') \\ P(2,1') & P(2,2') \end{pmatrix}$$

Solution. We find that

$$\begin{aligned}\det\begin{pmatrix} P(1,1') & P(1,2') \\ P(2,1') & P(2,2') \end{pmatrix} &= \det\begin{pmatrix} y_1 + y_2y_3 & y_2y_5 \\ y_4y_3 & y_6 + y_4y_5 \end{pmatrix} \\ &= y_1y_6 + y_1y_4y_5 + y_2y_3y_6 + y_2y_3y_4y_5 - y_2y_5y_4y_3 \\ &= y_1y_6 + y_1y_4y_5 + y_2y_3y_6\end{aligned}$$

as desired. ☐

Example F.2.



Compute

$$\det\begin{pmatrix} P(1,1') & P(1,2') \\ P(2,1') & P(2,2') \end{pmatrix}$$

Solution. We find that

$$\det\begin{pmatrix} P(1,1') & P(1,2') \\ P(2,1') & P(2,2') \end{pmatrix} = \det\begin{pmatrix} y_1y_2 + y_3y_6 & y_1y_4 + y_3y_8 \\ y_5y_2 + y_7y_6 & y_5y_4 + y_7y_8 \end{pmatrix}$$

$$\begin{aligned} &= y_1y_2y_5y_4 + y_1y_2y_7y_8 + y_3y_6y_5y_4 + y_3y_6y_7y_8 \\ &\quad - \left(y_1y_4y_5y_2 + y_1y_4y_7y_6 + y_3y_8y_5y_2 + y_3y_8y_7y_6\right) \\ &= y_1y_2y_7y_8 + y_3y_6y_5y_4 - y_1y_4y_7y_6 - y_3y_8y_5y_2 \end{aligned}$$

as desired. ☐

F.3 Lindström–Gessel–Viennot Lemma

Lemma F.1. Let Γ be a directed graph, and let

$$1, 2, \ldots, n, \qquad 1', 2', \ldots, n'$$

be distinct vertices in $V(\Gamma)$.

An *n-path*

$$Q : (1, 2, \ldots, n) \rightsquigarrow (1', 2', \ldots, n')$$

is an $(n+1)$-tuple

$$(\sigma, Q_1, Q_2, \ldots, Q_n),$$

where $\sigma \in S_n$, and for each $i = 1, \ldots, n$,

$$Q_i : i \rightsquigarrow \sigma(i)'.$$

Its weight is defined by

$$\mathrm{wt}(Q) = \mathrm{sgn}(\sigma) \prod_{i=1}^{n} \mathrm{wt}(Q_i).$$

We say that Q is *non-intersecting* if and only if

$$Q_1, \ldots, Q_n$$

are pairwise vertex-disjoint. Let

$$\mathcal{J} = \{\text{non-intersecting } Q : (1, 2, \ldots, n) \rightsquigarrow (1', 2', \ldots, n')\}.$$

Then

$$\det\begin{pmatrix} P(1,1') & \cdots & P(1,n') \\ \vdots & \ddots & \vdots \\ P(n,1') & \cdots & P(n,n') \end{pmatrix} = \sum_{Q \in \mathcal{J}} \mathrm{wt}(Q).$$

Sketch of proof Let

$$\mathcal{S} = \{\text{all } Q : (1, 2, \ldots, n) \rightsquigarrow (1', 2', \ldots, n')\},$$

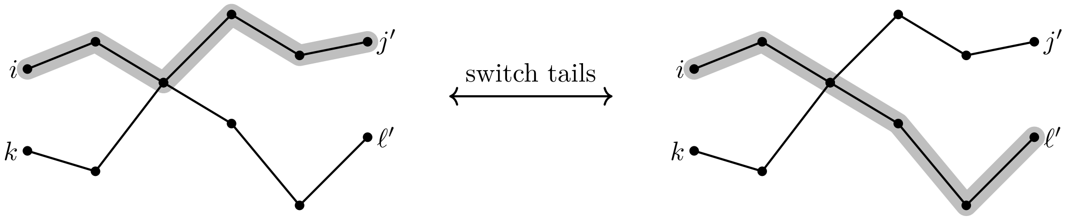
so this includes *all* n-paths, not just the non-intersecting ones. Expanding the determinant naturally yields

$$\det\bigl(P(i,j')\bigr) = \sum_{Q\in\mathcal{S}} \mathrm{wt}(Q).$$

To prove the result, we need a sign-reversing involution

$$\Phi : \mathcal{S}\setminus\mathcal{J} \longrightarrow \mathcal{S}\setminus\mathcal{J},$$

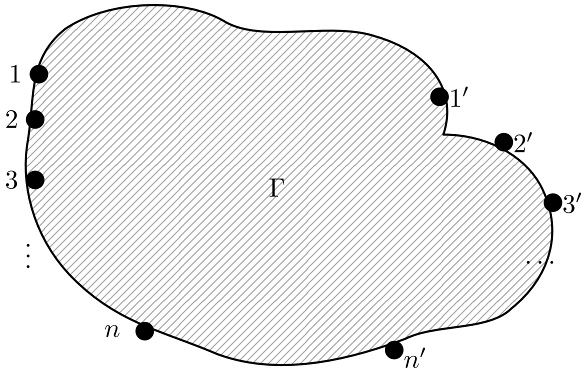
where $\mathcal{S}\setminus\mathcal{J}$ is the set of intersecting n-paths.
Main idea: switch tails at the first intersection.



This is sign-reversing. With some care, one can show that it is in fact an involution.

The Planar Case

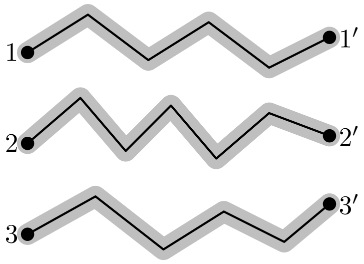
Assume Γ is embedded in the plane.



$$1, 2, \ldots, n, n', \ldots, 2', 1' \qquad \text{lie on the outer face in this order.}$$

In a planar graph, non-intersecting paths cannot cross. In this case, every $Q \in \mathcal{J}$ is of the form

$$Q = (\mathrm{id}, Q_1, \ldots, Q_n), \qquad Q_i : i \rightsquigarrow i'.$$



$$\boxed{\text{All terms in } |\mathcal{J}|_{\text{wt}} \text{ have sign } +1.}$$

F.4 Applications

Totally Non-negative Matrices

Definition F.2. **[Totally Non-negative]**

A matrix is **totally non-negative** is and only if all minors are greater or equal to zero.

Example F.3. The matrix $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}$ is totally non-negative, since all entries are ≥ 0, and

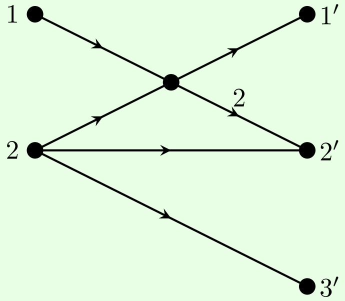
$$\det\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \geq 0, \quad \det\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \geq 0, \quad \det\begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix} \geq 0$$

Theorem F.1.

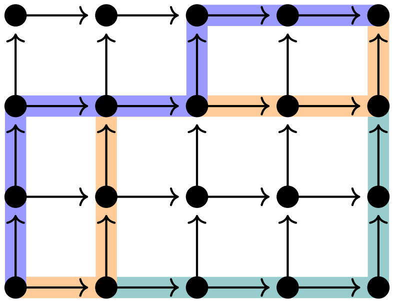
Let $M \in \texttt{Mat}_{m\times n}(\mathbb{R})$, then M is totally non-negative if and only if there exists a directed acyclic planar graph Γ, an edge weight function $wt : E(\Gamma) \to \mathbb{R}^+$, and vertices $1, 2, \ldots, m$, $1', 2', \ldots, n'$ such that

$$M_{ij} = P(i, j')$$

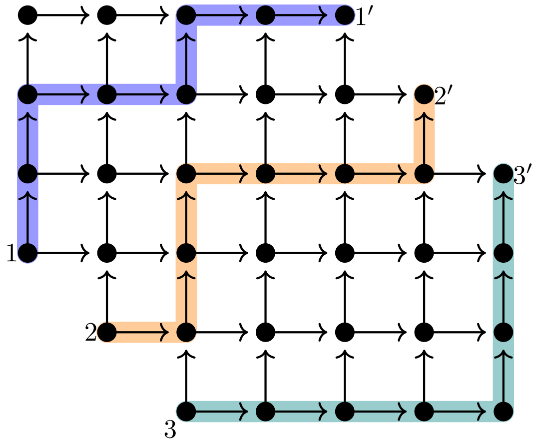
Example F.4. Recall that the matrix $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}$ is totally non-negative, here is the graph that produces this matrix:



(Weakly) Non-crossing Lattice Paths Weakly non-crossing paths can intersect, but upper path must stay weakly above lower path.



The above shows three weakly non-crossing paths. We can transform these “non-crossing” paths to “non-intersecting paths” by shifting the paths:

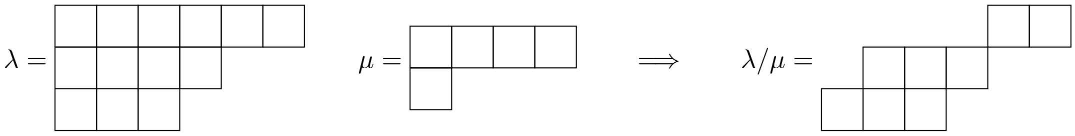


Hence in this case, the number of triples of non-crossing paths is equal to

$$\det\begin{pmatrix} \binom{7}{4} & \binom{7}{5} & \binom{7}{6} \\ \binom{7}{3} & \binom{7}{4} & \binom{7}{5} \\ \binom{7}{2} & \binom{7}{3} & \binom{7}{4} \end{pmatrix}$$

Skew Partitions A skew partition is λ/μ is a pair of partitions $\mu \leq \lambda$ depicted as the “difference of diagrams”.

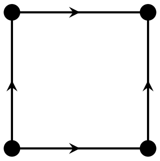
Example F.5.



Determine the number of skew partitions inside $k \times \ell$ rectangle.

F.4.1 Two Questions

Question 1. By considering n non-crossing paths in the following diagram:



compute

$$\det\begin{pmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 2 & \cdots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \end{pmatrix}$$

Question 2. How many ways are there to pack unit cubes in the corner of an $n \times n \times n$ cube?

Index