Lecture Notes on Algebraic Combinatorics

Jeremy L. Martin

jmartin@math.ku.edu

December 3, 2012

Foreword

The starting point for these lecture notes was my notes from Vic Reiner's Algebraic Combinatorics course at the University of Minnesota in Fall 2003. I currently use them for graduate courses at the University of Kansas. They will always be a work in progress. Please use them and share them freely for any research purpose. I have added and subtracted some material from Vic's course to suit my tastes, but any mistakes are my own; if you find one, please contact me at `jmartin@math.ku.edu` so I can fix it. Thanks to those who have suggested additions and pointed out errors, including but not limited to: Logan Godkin, Alex Lazar, Nick Packauskas, Billy Sanders, Tony Se.

1. Posets and Lattices

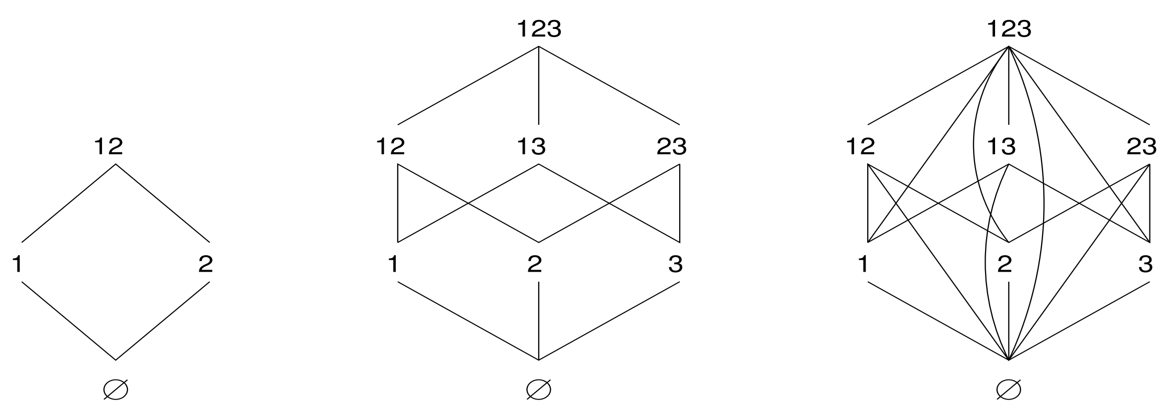
1.1. Posets.

Definition 1.1. A **partially ordered set** or **poset** is a set P equipped with a relation $\leq$ that is reflexive, antisymmetric, and transitive. That is, for all $x, y, z \in P$:

(1) $x \leq x$ (reflexivity).
(2) If $x \leq y$ and $y \leq x$, then $x = y$ (antisymmetry).
(3) If $x \leq y$ and $y \leq z$, then $x \leq z$ (transitivity).

We'll usually assume that P is finite.

Example 1.2 (**Boolean algebras**). Let $[n] = \{1, 2, \ldots, n\}$ (a standard piece of notation in combinatorics) and let $\mathscr{B}_n$ be the power set of $[n]$. We can partially order $\mathscr{B}_n$ by writing $S \leq T$ if $S \subseteq T$.



The first two pictures are **Hasse diagrams**. They don't include all relations, just the **covering relations**, which are enough to generate all the relations in the poset. (As you can see on the right, including *all* the relations would make the diagram unnecessarily complicated.)

Definition 1.3. Let P be a poset and $x, y \in P$.

- x is **covered** by y, written $x \lessdot y$, if $x < y$ and there exists no z such that $x < z < y$.
- The **interval** from x to y is

$$[x, y] := \{z \in P \mid x \leq z \leq y\}.$$

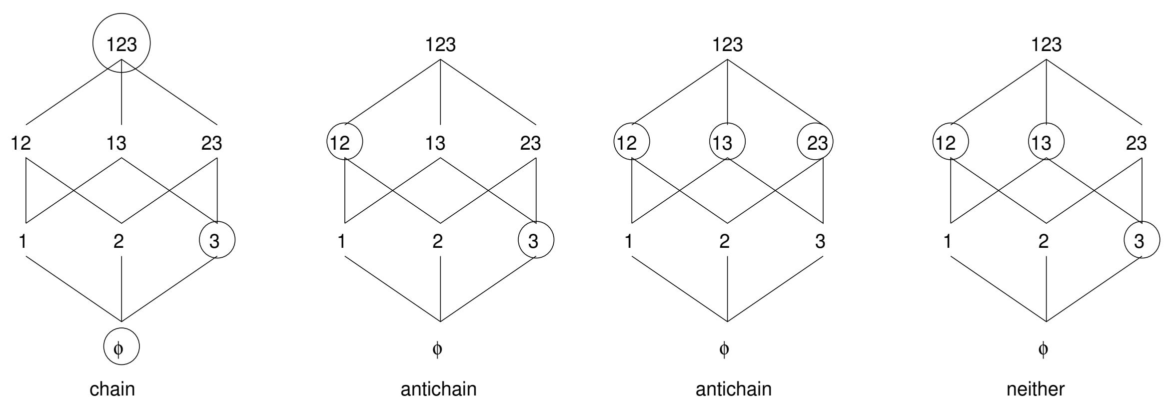
(This is nonempty if and only if $x \leq y$, and it is a singleton set if and only if $x = y$.)

The Boolean algebra $\mathscr{B}_n$ has a unique minimum element (namely $\emptyset$) and a unique maximum element (namely $[n]$). Not every poset has to have such elements, but if a poset does, we'll call them $\hat{0}$ and $\hat{1}$ respectively (or if necessary $\hat{0}_P$ and $\hat{1}_P$).

Definition 1.4. A poset that has both a $\hat{0}$ and a $\hat{1}$ is called **bounded**.[1] An element that covers $\hat{0}$ is called an **atom**, and an element that is covered by $\hat{1}$ is called a **coatom**. (For example, the atoms in $\mathscr{B}_n$ are the singleton subsets of $[n]$.)

We can make a poset P bounded: define a new poset $\hat{P}$ by adjoining new elements $\hat{0}, \hat{1}$ such that $\hat{0} < x < \hat{1}$ for every $x \in P$. Meanwhile, sometimes we have a bounded poset and want to delete the bottom and top elements.

Definition 1.5. A subset $C \subset P$ is called a **chain** if its elements are pairwise comparable. Thus every chain is of the form $C = \{x_0, \dots, x_n\}$, where $x_0 < \cdots < x_n$. An **antichain** is a subset of P in which no two of its elements are comparable.[2]



1.2. **Ranked Posets.** One of the many nice properties of $\mathscr{B}_n$ is that its elements fall nicely into horizontal slices (sorted by their cardinalities). Whenever $S \lessdot T$, it is the case that $|T| = |S| + 1$. A poset for which we can do this is called a **ranked** poset. However, it would be tautological to define a ranked poset to be a poset in which we can rank the elements! The actual definition of rankedness is a little more subtle, but makes perfect sense after a little thought.

Definition 1.6. A chain $x_0 < \cdots < x_n$ is **saturated**[3] if it is not properly contained in any other chain from x_0 to x_n; equivalently, if $x_{i-1} \lessdot x_i$ for every $i \in [n]$. In this case, the number n is the **length** of the chain A poset P is **ranked** if for every $x \in P$, all saturated chains with top element x have the same length; this number is called the **rank** of x and denoted $r(x)$. It follows that

$$x \lessdot y \implies r(y) = r(x) + 1. \tag{1.1}$$

A poset is **graded** if it is ranked and bounded.

Note:

(1) "Length" means the number of *steps*, not the number of *elements* — i.e., edges rather than vertices in the Hasse diagram.

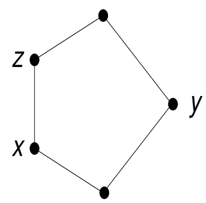
[1] This has nothing to do with the more typical metric-space definition of "bounded".

[2] To set theorists, "antichain" means something stronger: a set of elements such that no two have a common lower bound. This concept does not typically arise in combinatorics, where many posets are bounded.

[3] Sometimes called "maximal", but that word can easily be misinterpreted to mean "of maximum size".

(2) The literature is not consistent on the usage of the term "ranked". Sometimes "ranked" is used for the weaker condition that for every pair $x, y \in P$, every chain from x to y has the same length. Under this definition, the implication (1.1) fails (proof left to the reader).

(3) For any finite poset P (and some infinite ones) one can define $r(x)$ to be the supremum of the lengths of all chains with top element x — but if P is not a ranked poset, then there will be some pair a, b such that $b \gtrdot a$ but $r(y) > r(x) + 1$. For instance, in the bounded poset shown below (known as N_5), we have $\hat{1} \gtrdot y$, but $r(\hat{1}) = 3$ and $r(y) = 1$.



Definition 1.7. Let P be a ranked poset with rank function r. The **rank-generating function** of P is the formal power series

$$F_P(q) = \sum_{x \in P} q^{r(x)}.$$

Thus, for each k, the coefficient of q^k is the number of elements at rank k.

We can now say that the Boolean algebra is ranked by cardinality. In particular,

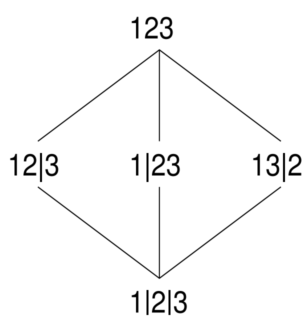
$$F_{\mathscr{B}_n}(q) = \sum_{S \subset [n]} q^{|S|} = (1+q)^n.$$

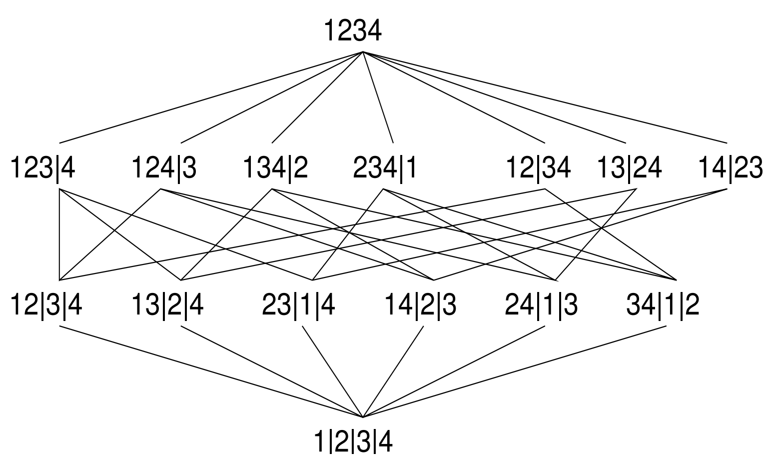
The expansion of this polynomial is palindromic, because the coefficients are a row of Pascal's Triangle. That is, $\mathscr{B}_n$ is **rank-symmetric**. In fact, much more is true. For any poset P, we can define the **dual poset** P^* by reversing all the order relations, or equivalently turning the Hasse diagram upside down. It's not hard to prove that the Boolean algebra is **self-dual**, i.e., $\mathscr{B}_n \cong \mathscr{B}_n^*$, from which it immediately follows that it is rank-symmetric.

Example 1.8 (**The partition lattice**)**.** Let Π_n be the poset of all set partitions of $[n]$. E.g., two elements of Π_5 are

$$S = \big\{\{1,3,4\},\ \{2,5\}\big\} \qquad \text{(abbr.: } S = 134|25)$$
$$T = \big\{\{1,3\},\ \{4\},\ \{2,5\}\big\} \qquad \text{(abbr.: } T = 13|4|25)$$

The sets $\{1,3,4\}$ and $\{2,5\}$ are called the *blocks* of S. We can impose a partial order on Π_n by putting $T \leq S$ if every block of T is contained in a block of S; for short, T *refines* S.



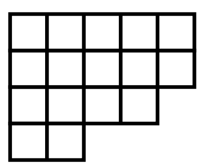


- The covering relations are of the form "merge two blocks into one".
- Π_n is graded, with $\hat{0} = 1|2|\cdots|n$ and $\hat{1} = 12\cdots n$. The rank function is $r(S) = n - |S|$.
- The coefficients of the rank-generating function of Π_n are the Stirling numbers of the second kind: $S(n,k)$ = number of partitions of $[n]$ into k blocks. That is,
$$F_n(q) = F_{\Pi_n}(q) = \sum_{k=1}^{n} S(n,k) q^{n-k}.$$
For example, $F_3(q) = 1 + 3q + q^2$ and $F_4(q) = 1 + 6q + 7q^2 + q^3$.

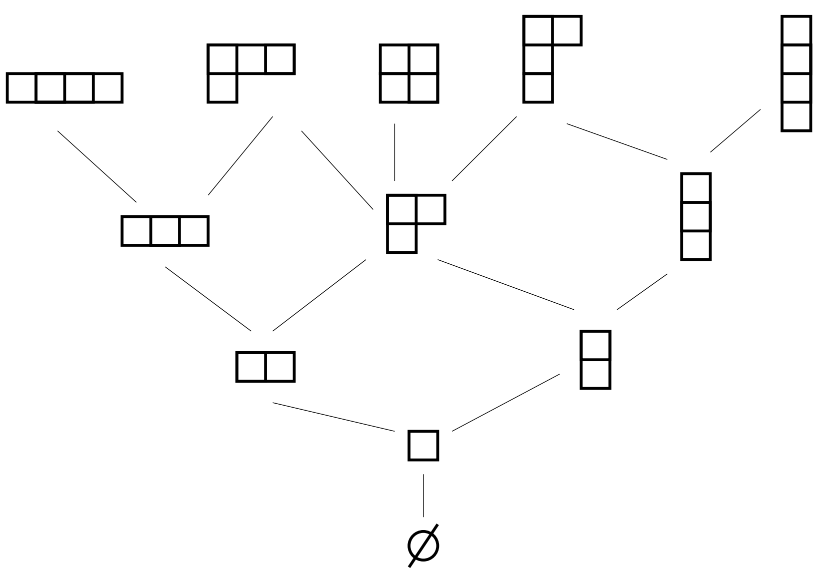
Example 1.9 (**Young's lattice**.)**.** A *partition* is a sequence $\lambda = (\lambda_1, \dots, \lambda_\ell)$ of weakly decreasing positive integers: i.e., $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. If $n = \lambda_1 + \cdots + \lambda_\ell$, we write $\lambda \vdash n$ and/or $n = |\lambda|$. For convenience, set $\lambda_i = 0$ for all $i > \ell$. Let Y be the set of all partitions, partially ordered by $\lambda \geq \mu$ if $\lambda_i \geq \mu_i$ for all $i = 1, 2, \dots$.

This is an infinite poset, but it is *locally finite*, i.e., every interval is finite. Its rank-generating function
$$\sum_{\lambda} q^{|\lambda|} = \sum_{n\geq 0} \sum_{\lambda \vdash n} q^n$$
is given by the justly celebrated formula
$$\prod_{k=1}^{\infty} \frac{1}{1-q^k}.$$

There's a nice pictorial way to look at Young's lattice. Instead of thinking about partitions as sequence of numbers, view them as their corresponding **Ferrers diagrams**: northwest-justified piles of boxes whose i^{th} row contains λ_i boxes. For example, 5542 is represented by the following Ferrers diagram:



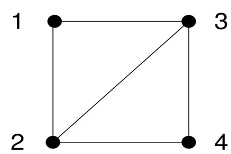
Then $\lambda \geq \mu$ if and only the Ferrers diagram of λ contains that of μ. The bottom part of the Hasse diagram of Y looks like this:



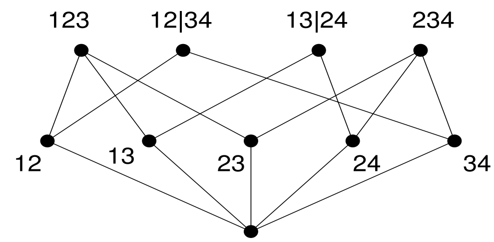
Definition 1.10. An *isomorphism* of posets P, Q is a bijection $f : P \to Q$ such that $x \leq y$ if and only if $f(x) \leq f(y)$. We say that P and Q are **isomorphic**, written $P \cong Q$, if there is an isomorphism $P \to Q$. An *automorphism* is an isomorphism from a poset to itself.

Young's lattice Y has a nontrivial automorphism $\lambda \mapsto \tilde{\lambda}$ called *conjugation*. This is most easily described in terms of Ferrers diagrams: reflect across the line $x + y = 0$ so as to swap rows and columns. It is easy to check that if $\lambda \geq \mu$, then $\tilde{\lambda} \geq \tilde{\mu}$.

Example 1.11 (**The clique poset of a graph**)**.** Let $G = (V, E)$ be a graph with vertex set $[n]$. A *clique* of G is a set of vertices that are pairwise adjacent. Let $K(G)$ be the poset consisting of set partitions all of whose blocks are cliques in G, ordered by refinement.



G



K(G)

This is a *subposet* of Π_n: a subset of Π_n that inherits its order relation. This poset is ranked but not graded, since there is not necessarily a $\hat{1}$. Notice that $\Pi_n = K(K_n)$ (the complete graph on n vertices).

1.3. **Lattices.**

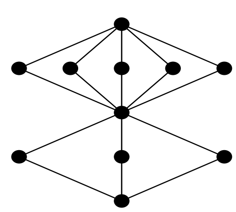
Definition 1.12. A poset L is a **lattice** if every pair $x, y \in L$ has a unique **meet** $x \wedge y$ and **join** $x \vee y$. That is,

$$x \wedge y = \max\{z \in L \mid z \leq x, y\},$$
$$x \vee y = \min\{z \in L \mid z \geq x, y\}.$$

Note that, e.g., $x \wedge y = x$ if and only if $x \leq y$. These operations are commutative and associative, so for any finite $M \subset L$, the meet $\wedge M$ and join $\vee M$ are well-defined elements of L. In particular, every finite lattice is bounded (with $\hat{0} = \wedge L$ and $\hat{1} = \vee L$). For convenience, we define $\wedge\emptyset = \hat{1}$ and $\vee\emptyset = \hat{0}$.

Example 1.13. The Boolean algebra $\mathscr{B}_n$ is a lattice, with $S \wedge T = S \cap T$ and $S \vee T = S \cup T$.

Example 1.14. The *complete graded poset* $P(a_1, \ldots, a_n)$ has $r(\hat{1}) = n + 1$ and $a_i > 0$ elements at rank i for every $i > 0$, with every possible order relation (i.e., $r(x) > r(y) \implies x > y$).



P(3,1,5)

This poset is a lattice if and only if no two consecutive a_i's are 2 or greater.

Example 1.15. The clique poset $K(G)$ of a graph G is in general not a lattice, because join is not well-defined. Meet, however, is well-defined, because the intersection of two cliques is a clique. Therefore, the clique poset is what is called a **meet-semilattice**. It can be made into a lattice by adjoining a brand-new $\hat{1}$ element. In the case that $G = K_n$, the clique poset is a lattice, namely the partition lattice Π_n.

Example 1.16. Lattices don't have to be ranked. For example, the poset N_5 is a perfectly good lattice.

Proposition 1.17 (Absorption laws). *Let L be a lattice and $x, y \in L$. Then $x \vee (x \wedge y) = x$ and $x \wedge (x \vee y) = x$. (Proof left to the reader.)*

Proposition 1.18. *Let P be a bounded poset that is a meet-semilattice (i.e., every nonempty $B \subseteq P$ has a well-defined meet $\wedge B$). Then P every finite nonempty subset of P has a well-defined join, and consequently P is a lattice.*

Proof. Let $A \subseteq P$, and let $B = \{b \in P \mid b \geq a \text{ for all } a \in A\}$. Note that $B \neq \emptyset$ because $\hat{1} \in B$. I claim that $\wedge B$ is the unique least upper bound for A. First, we have $\wedge B \geq a$ for all $a \in A$ by definition of B and of meet. Second, if $x \geq a$ for all $a \in A$, then $x \in B$ and so $x \geq \wedge B$, proving the claim. $\square$

Definition 1.19. Let L be a lattice. A *sublattice* of L is a subposet $L' \subset L$ that (a) is a lattice and (b) inherits its meet and join operations from L. That is, for all $x, y \in L'$, we have

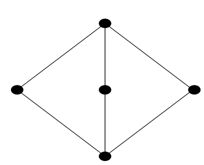
$$x \wedge_{L'} y = x \wedge_L y \quad \text{and} \quad x \vee_{L'} y = x \vee_L y.$$

A sublattice $L' \subset L$ does not have to have the same $\hat{0}$ and $\hat{1}$ elements. As an important example, every interval $L' = [x, z] \subseteq L$ (i.e., $L' = \{y \in L \mid x \leq y \leq z\}$) is a sublattice with minimum element x and maximum element z. (We might write $\hat{0}_{L'} = x$ and $\hat{1}_{L'} = z$.)

Example 1.20 (The subspace lattice). Let q be a prime power, let $\mathbb{F}_q$ be the field of order q, and let $V = \mathbb{F}_q^n$ (a vector space of dimension n over $\mathbb{F}_q$). The *subspace lattice* $L_V(q) = L_n(q)$ is the set of all vector subspaces of V, ordered by inclusion. (We could replace $\mathbb{F}_q$ with any old field if you don't mind infinite posets.)

The meet and join operations on $L_n(q)$ are given by $W \wedge W' = W \cap W'$ and $W \vee W' = W + W'$. We could construct analogous posets by ordering the (normal) subgroups of a group, or the prime ideals of a ring, or the submodules of a module, by inclusion. (However, these posets are not necessarily ranked, while $L_n(q)$ is ranked, by dimension.)

The simplest example is when $q = 2$ and $n = 2$, so that $V = \{(0,0), (0,1), (1,0), (1,1)\}$. Of course V has one subspace of dimension 2 (itself) and one of dimension 0 (the zero space). Meanwhile, it has three subspaces of dimension 1; each consists of the zero vector and one nonzero vector. Therefore, $L_2(2) \cong M_5$.



Note that $L_n(q)$ is self-dual, under the anti-automorphism $W \to W^\perp$. (An *anti-automorphism* is an isomorphism $P \to P^*$.)

Example 1.21 (Bruhat order and weak Bruhat order). Let $\mathfrak{S}_n$ be the set of permutations of $[n]$ (i.e., the symmetric group).[4] Write elements of $\mathfrak{S}_n$ as strings $\sigma_1\sigma_2\cdots\sigma_n$ of distinct digits, e.g., $47182635 \in \mathfrak{S}_8$. Impose a partial order on $\mathfrak{S}_n$ defined by the following covering relations:

(1) $\sigma \lessdot \sigma'$ if σ' can be obtained by swapping σ_i with σ_{i+1}, where $\sigma_i < \sigma_{i+1}$. For example,

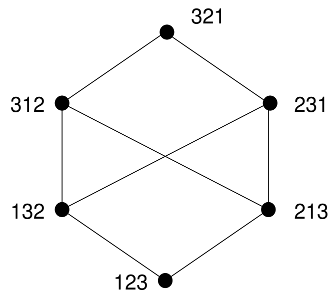
$$4718\underline{26}35 \lessdot 4718\underline{62}35 \quad \text{and} \quad 4\underline{71}82635 \gtrdot 4\underline{17}82635.$$

(2) $\sigma \lessdot \sigma'$ if σ' can be obtained by swapping σ_i with σ_j, where $i < j$ and $\sigma_j = \sigma_i + 1$. For example,

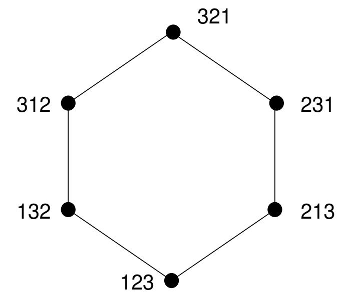
$$4718\underline{2}6\underline{3}5 \lessdot 4718\underline{3}6\underline{2}5.$$

[4]The letter S has many other uses in combinatorics: Stirling numbers of the first and second kind, Schur symmetric functions, ... The symmetric group is important enough to merit an ornate symbol.

If we only use the first kind of covering relation, we obtain the **weak Bruhat order** (or just "weak order").



Bruhat order



Weak Bruhat order

The Bruhat order is not in general a lattice (e.g., 132 and 213 do not have a well-defined join in $\mathfrak{S}_3$). The weak order actually is a lattice, though this is not so easy to prove.

A *Coxeter group* is a finite group generated by elements $s_1, \dots, s_n$, called *simple reflections*, satisfying $s_i^2 = 1$ and $(s_i s_j)^{m_{ij}} = 1$ for all $i \neq j$ and some integers $m_{ij} \geq 23$. For example, setting $m_{ij} = 3$ if $|i - j| = 1$ and $m_{ij} = 2$ if $|i - j| > 1$, we obtain the symmetric group $\mathfrak{S}_{n+1}$. Coxeter groups are fantastically important in geometric combinatorics and we could spend at least a semester on them. For now, it's enough to mention that every Coxeter group has associated Bruhat and weak orders, whose covering relations correspond to multiplying by simple reflections.

The Bruhat and weak order give graded, self-dual poset structures on $\mathfrak{S}_n$, with the same rank function, namely the number of **inversions**:

$$r(\sigma) = \Big| \big\{ \{i, j\} \mid i < j \text{ and } \sigma_i > \sigma_j \big\} \Big|.$$

(For a general Coxeter group, the rank of an element σ is the minimum number r such that σ is the product of r simple reflections.) The rank-generating function of $\mathfrak{S}_n$ is a very nice polynomial called the **q-factorial**:

$$F_{\mathfrak{S}_n}(q) = 1(1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}) = \prod_{i=1}^{n} \frac{1-q^i}{1-q}.$$

1.4. **Distributive Lattices.**

Definition 1.22. A lattice L is **distributive** if the following two equivalent conditions hold:

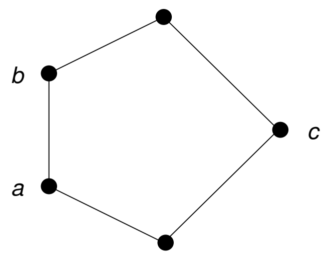
(1.2a) $$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \qquad \forall x, y, z \in L,$$

(1.2b) $$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \qquad \forall x, y, z \in L.$$

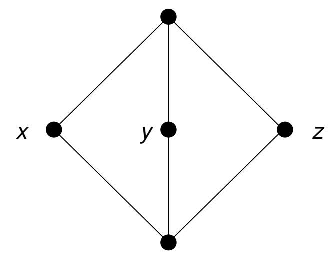
Proving that the two conditions (1.2a) and (1.2b) are equivalent is not too hard, but is not trivial (it's a homework problem). Note that replacing the equalities with $\geq$ and $\leq$ respectively gives statements that are true for all lattices.

The condition of distributivity seems natural, but in fact distributive lattices are quite special.

(1) The Boolean algebra $\mathscr{B}_n$ is a distributive lattice, because the set-theoretic operations of union and intersection are distributive over each other.
(2) M_5 and N_5 are not distributive:



N_5



M_5

$$(a \vee c) \wedge b = b \qquad\qquad (x \vee y) \wedge z = z$$

$$(a \wedge b) \vee (a \wedge c) = a \qquad\qquad (x \wedge z) \vee (y \wedge z) = \hat{0}$$

In particular, the partition lattice Π_n is not distributive for $n \geq 3$ (recall that $\Pi_3 \cong M_5$).

(3) Any sublattice of a distributive lattice is distributive. In particular, Young's lattice Y is distributive because it is locally a sublattice of a Boolean lattice.

(4) The set D_n of all positive integer divisors of a fixed integer n, ordered by divisibility, is a distributive lattice (proof for homework).

Definition 1.23. Let P be a poset. An **order ideal** of P is a set $A \subseteq P$ that is closed under going down, i.e., if $x \in A$ and $y \leq x$ then $y \in A$. The poset of all order ideals of P (ordered by containment) is denoted $J(P)$. The order ideal **generated** by $x_1, \ldots, x_n \in P$ is the smallest order ideal containing them, namely

$$\langle x_1, \ldots, x_n \rangle \; := \; \big\{ y \in P \mid y \leq x_i \text{ for some } i \big\}.$$

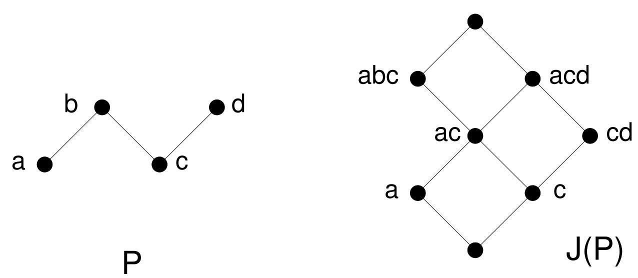


FIGURE 1. The lattice $J(P)$ of order ideals of a poset P.

There is a natural bijection between $J(P)$ and the set of antichains of P, since the maximal elements of any order ideal A form an antichain that generates it.

Proposition 1.24. *The operations $A \vee B = A \cup B$ and $A \wedge B = A \cap B$ make $J(P)$ into a distributive lattice, partially ordered by set containment.*

Sketch of proof: Check that $A \cup B$ and $A \cap B$ are in fact order ideals of P (this is fairly easy from the definition). It follows that $J(P)$ is a sublattice of the Boolean algebra on P, hence is distributive. ☐

Definition 1.25. Let L be a lattice. An element $x \in L$ is **join-irreducible** if it cannot be written as the join of two other elements. That is, if $x = y \vee z$ then either $x = y$ or $x = z$. The subposet (not sublattice!) of L consisting of all join-irreducible elements is denoted $\mathrm{Irr}(L)$.

Provided that L has no infinite descending chains (e.g., L is finite, or is locally finite and has a $\hat{0}$), every element of L can be written as the join of join-irreducibles (but not necessarily uniquely; e.g., M_5).

All atoms are join-irreducible, but not all join-irreducible elements need be atoms. An extreme (and slightly trivial) example is a chain: *every* element is join-irreducible, but there is only one atom. As a less trivial example, in the lattice below, a, b, c, d are all join-irreducible, although the only atoms are a and c.

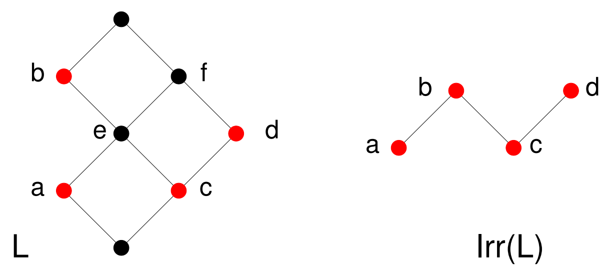


FIGURE 2. The poset of join-irreducibles of a lattice.

Theorem 1.26 (**Fundamental Theorem of Finite Distributive Lattices; Birkhoff 1933**). *Up to isomorphism, the finite distributive lattices are exactly the lattices $J(P)$, where P is a finite poset. Moreover, $L \cong J(\mathrm{Irr}(L))$ for every lattice L and $P \cong \mathrm{Irr}(J(P))$ for every poset P.*

The proof encompasses a series of lemmata.

Lemma 1.27. *Let L be a distributive lattice and let $p \in L$ be join-irreducible. Suppose that $p \leq a_1 \vee \cdots \vee a_n$. Then $p \leq a_i$ for some i.*

Proof. By distributivity we have

$$p = p \wedge (a_1 \vee \cdots \vee a_n) = (p \wedge a_1) \vee \cdots \vee (p \wedge a_n)$$

and since p is join-irreducible, it must equal $p \wedge a_i$ for some i, whence $p \leq a_i$. □

(Analogue: If a prime p divides a product of positive numbers, then it divides at least one of them. This is in fact exactly what Lemma 1.27 says when applied to the divisor lattice D_n.)

Proposition 1.28. *Let L be a distributive lattice. Then every $x \in L$ can be written uniquely as an irredundant join of join-irreducible elements.*

Proof. We have observed above that any element in a finite lattice can be written as an irredundant join of join-irreducibles, so we have only to prove uniqueness. So, suppose that we have two irredundant decompositions

$$x = p_1 \vee \cdots \vee p_n = q_1 \vee \cdots \vee q_m \tag{1.3}$$

with $p_i, q_j \in \mathrm{Irr}(L)$ for all i, j.

By Lemma 1.27, $p_1 \leq q_j$ for some j. Again by Lemma 1.27, $q_j \leq p_i$ for some i. If $i \neq 1$, then $p_1 \leq p_i$, which contradicts the fact that the p_i form an antichain. Therefore $p_1 = q_j$. Replacing p_1 with any join-irreducible appearing in (1.3) and repeating this argument, we find that the two decompositions must be identical. □

Sketch of proof of Birkhoff's Theorem. The lattice isomorphism $L \to J(\mathrm{Irr}(L))$ is given by

$$\phi(x) = \langle p \mid p \in \mathrm{Irr}(L),\ p \leq x \rangle.$$

Meanwhile, the join-irreducible order ideals in P are just the principal order ideals, i.e., those generated by a single element. So the poset isomorphism $P \to \mathrm{Irr}(J(P))$ is given by

$$\psi(y) = \langle y \rangle.$$

These facts need to be checked (as a homework problem).

Corollary 1.29. *Every distributive lattice is isomorphic to a sublattice of a Boolean algebra (whose atoms are the join-irreducibles in L).*

The atoms of the Boolean lattice $\mathscr{B}_n$ are the singleton sets $\{i\}$ for $i \in [n]$; these form an antichain.

Corollary 1.30. *Let L be a finite distributive lattice. TFAE:*

(1) *L is a Boolean algebra;*
(2) $\mathrm{Irr}(L)$ *is an antichain;*
(3) *L is atomic (i.e., every element in L is the join of atoms).*
(4) *Every join-irreducible element is an atom;*
(5) *L is* complemented. *That is, for each $x \in L$, there exists a unique element $\bar{x} \in L$ such that $x \vee \bar{x} = \hat{1}$ and $\bar{x} \wedge y = \hat{0}$.*
(6) *L is* relatively complemented. *That is, for every interval*[5] *$[x, z] \subseteq L$ and every $y \in [x, z]$, there exists a unique element $u \in [x, z]$ such that $y \vee u = z$ and $y \wedge u = x$.*

Proof. (6) $\implies$ (5): Trivial.

(5) $\implies$ (4): Suppose that L is complemented, and suppose that $z \in L$ is a join-irreducible that is not an atom. Let x be an atom in $[\hat{0}, z]$. Then Then

$$\begin{aligned}(x \vee \bar{x}) \wedge z &= \hat{1} \wedge z = z \\ &= (x \wedge z) \vee (\bar{x} \wedge z) = x \vee (\bar{x} \wedge z),\end{aligned}$$

by distributivity. Since z is join-irreducible, we must have $\bar{x} \wedge z = z$, i.e., $\bar{x} \geq z$. But then $\bar{x} > x$ and $\bar{x} \wedge x = x \neq \hat{0}$, a contradiction.

(4) $\iff$ (3): Trivial.

(4) $\implies$ (2): This follows from the observation that any two atoms are incomparable.

(2) $\implies$ (1): By FTFDL, since $L = J(\mathrm{Irr}(L))$.

(1) $\implies$ (6): If $X \subseteq Y \subseteq Z$ are sets, then let $U = X \cup (Y \setminus Z)$. Then $Y \cap U = X$ and $Y \cup U = Z$. $\square$

Dually, we could show that every element in a distributive lattice can be expressed uniquely as the meet of meet-irreducible elements. (This might be a roundabout way to show that distributivity is a self-dual condition.)

1.5. **Modular Lattices.**

Definition 1.31. A lattice L is **modular** if for every $x, y, z \in L$ with $x \leq z$, the *modular equation* holds:

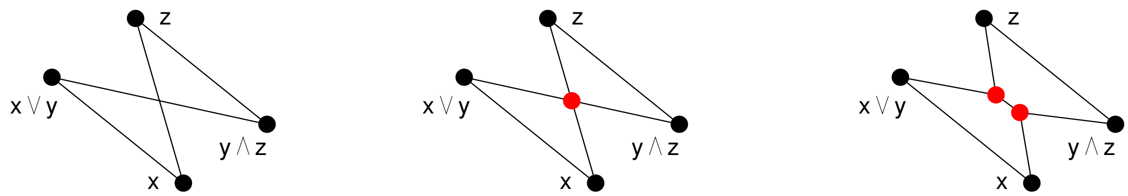
$$x \vee (y \wedge z) = (x \vee y) \wedge z. \tag{1.4}$$

Here is one way to picture modularity. Even without assuming $x \leq z$, we have

$$x \ \leq \ x \vee y \ \geq \ z \wedge y \ \leq \ z \ \geq \ x.$$

as pictured on the left. Modularity can be thought of saying that "the relations cross properly" – the intersection point of the two lines in the Hasse diagram is a unique element of the poset.

[5] Think of this interval as a sublattice with top element z and bottom element x.



Modular   Not modular

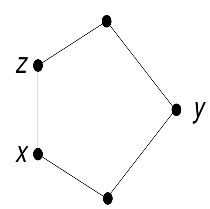
Note that for all lattices, if $x \leq z$, then $x \vee (y \wedge z) \leq (x \vee y) \wedge z$. Modularity says that, in fact, equality holds.

Some basic facts and examples:

(1) Every sublattice of a modular lattice is modular. Also, distributive lattices are modular: if L is distributive and $x \leq z \in L$, then
$$x \vee (y \wedge z) \;=\; (x \wedge z) \vee (y \wedge z) \;=\; (x \vee y) \wedge z,$$
so L is modular.

(2) The lattice L is modular if and only if its dual L^* is modular. Unlike the corresponding statement for distributivity, this is completely trivial, because the modular equation is invariant under dualization.

(3) The nonranked lattice N_5 is not modular. With the labeling below, we have $x \leq z$, but
$$x \vee (y \wedge z) = x \vee \hat{0} \;=\; x,$$
$$(x \vee y) \wedge z = \hat{1} \wedge z \;=\; z.$$



In fact, N_5 is the unique obstruction to modularity, as we will soon see.

(4) The nondistributive lattice $M_5 \cong \Pi_3$ is modular. However, Π_4 is not modular (exercise).

The etymology of the term "modular lattice" may be as follows: if R is a (not necessarily commutative) ring and M is a (left) R-submodule, then the poset $L(M)$ of (left) R-submodules of M, ordered by inclusion, is a modular lattice with operations $A \vee B = A + B$, $A \wedge B = A \cap B$, satisfies the modular condition (although it may not be a finite poset). Common examples include the lattice of subspaces of a vector space (e.g., $L_n(q)$) and the lattice of subgroups of an abelian group (i.e., of a $\mathbb{Z}$-module).

The lattice of subgroups of a nonabelian group need not be modular. For example, let $G = \mathfrak{S}_4$ and let $X = \langle(1\,2\,3)\rangle$, $Y = \langle(1\,2)\rangle$ (using cycle notation), and let Z be the alternating group $\mathfrak{A}_4$. Then $(XY) \cap Z = Z$ but $X(Y \cap Z) = Z$.

Proposition 1.32. *Let L be a lattice. TFAE:*

(1) *L is modular.*
(2) *For all $x, y, z \in L$, if $x \in [y \wedge z, z]$, then $x = (x \vee y) \wedge z$.*
(2*) *For all $x, y, z \in L$, if $x \in [y, y \vee z]$, then $x = (x \wedge z) \vee y$.*
(3) *For all $y, z \in L$, there is an isomorphism of lattices $[y \wedge z, z] \cong [y, y \vee z]$.*

Proof. (1) $\implies$ (2): If $y \wedge z \leq x \leq z$, then the modular equation (1.4) becomes $x = (x \vee y) \wedge z$.

(2) $\implies$ (1): Suppose that (2) holds. Let $a, b, c \in L$ with $a \leq c$. Then

$$b \wedge c \quad \leq \quad a \vee (b \wedge c) \quad \leq \quad c \vee c = c$$

so applying (2) with $y = b$, $z = c$, $x = a \vee (b \wedge c)$ gives

$$a \vee (b \wedge c) \quad = \quad \big((a \vee (b \wedge c)) \vee b\big) \wedge c \quad = \quad (a \vee b) \wedge c$$

as desired.

(2) $\iff$ (2*): Modularity is a self-dual condition.

Finally, for every y, z, there are functions $\alpha : [y \wedge z, z] \to [y, y \vee z]$ and $\beta : [y, y \vee z] \to [y \wedge z, z]$ given by $\alpha(q) = q \vee y$ and $\beta(q) = q \wedge z$. Conditions (2) and (2*) say respectively that $\beta \circ \alpha$ and $\alpha \circ \beta$ are the identity. Together, these are equivalent to assertion (3). □

Corollary 1.33. *Let R be a ring and M an R-module. Then the lattice $L(M)$ of R-submodules of M is modular.*

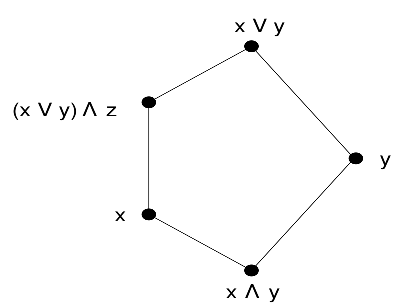
Proof. The Second Isomorphism Theorem says that $B/(A \cap B) \cong (A + B)/A$ for all $A, B \in L(M)$. Therefore $L(B/(A \cap B)) \cong L((A + B)/A)$, which says that $L(M)$ satisfies condition (3) of Prop. 1.32. □

Theorem 1.34. *Let L be a lattice.*

(1) *L is modular if and only if it contains no sublattice isomorphic to N_5.*
(2) *L is distributive if and only if it contains no sublattice isomorphic to N_5 or M_5.*

Proof. Both $\implies$ directions are easy, because N_5 is not modular and M_5 is not distributive.

Suppose that x, y, z is a triple for which modularity fails. One can check that



is a sublattice (details left to the reader).

Suppose that L is not distributive. If it isn't modular then it contains an N_5, so there's nothing to prove. If it is modular, then choose x, y, z such that

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

You can then show that

(1) this inequality is invariant under permuting x, y, z;
(2) $[x \wedge (y \vee z)] \vee (y \wedge z)$ and the two other lattice elements obtained by permuting x, y, z form a cochain; and
(3) the join (resp. meet) of any of two of x, y, z is equal.

Hence, we have constructed a sublattice of L isomorphic to M_5. □

A corollary is that every modular lattice (hence, every distributive lattice) is graded, because a non-graded lattice must contain a sublattice isomorphic to N_5. The details are left to the reader; we will eventually prove the stronger statement that every *semimodular* lattice is graded.

1.6. Semimodular Lattices.

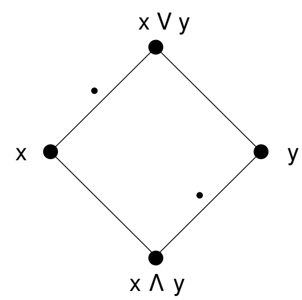
Definition 1.35. A lattice L is **(upper) semimodular** if for all $x, y \in L$,

$$x \wedge y \lessdot y \implies x \lessdot x \vee y. \tag{1.5}$$

Conversely, L is **lower semimodular** if the converse holds.

The implication (1.5) is trivial if x and y are comparable. If they are incomparable (as we will often assume), then there are several useful colloquial rephrasings of semimodularity:

- "If meeting with x merely nudges y down, then meeting with y merely nudges x up."
- In the interval $[x \wedge y, x \vee y] \subset L$ pictured below, if the southeast relation is a cover, then so is the northwest relation.



- Contrapositively, "If there is other stuff between x and $x \vee y$, then there is also other stuff between $x \wedge y$ and y."

Lemma 1.36. *If L is modular then it is upper and lower semimodular.*

Proof. If $x \wedge y \lessdot y$, then the sublattice $[x \wedge y, y]$ has only two elements. If L is modular, then by condition (3) of Proposition 1.32 we have $[x \wedge y, y] \cong [x, x \vee y]$, so $x \lessdot x \vee y$. Hence L is upper semimodular. A similar argument proves that L is lower semimodular. $\square$

In fact, upper and lower semimodularity together imply modularity. To make this more explicit, we will show that each of these three conditions on a lattice L implies that it is ranked, and moreover, for all $x, y \in L$, the rank function r satisfies

$$r(x \vee y) + r(x \wedge y) \leq r(x) + r(y) \qquad \text{if } L \text{ is upper semimodular;} \tag{1.6}$$

$$r(x \vee y) + r(x \wedge y) \geq r(x) + r(y) \qquad \text{if } L \text{ is lower semimodular;} \tag{1.7}$$

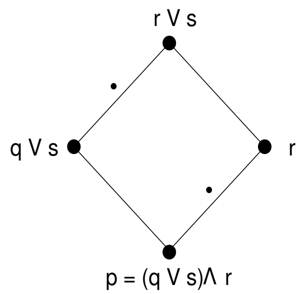
$$r(x \vee y) + r(x \wedge y) = r(x) + r(y) \qquad \text{if } L \text{ is modular.} \tag{1.8}$$

Lemma 1.37. *Suppose L is semimodular and let $q, r, s \in L$. If $q \lessdot r$, then either $q \vee s = r \vee s$ or $q \vee s \lessdot r \vee s$.*

("If it only takes one step to walk up from q to r, then it takes *at most* one step to walk from $q \vee s$ to $r \vee s$.")

Proof. Let $p = (q \vee s) \wedge r$. Note that $q \leq p \leq r$. Therefore, either $p = q$ or $p = r$.

- If $p = r$, then $q \vee s \geq r$. So $q \vee s = r \vee (q \vee s) = r \vee s$.
- If $p = q$, then $p = (q \vee s) \wedge r = q \lessdot r$. Applying semimodularity to the diamond figure below, we obtain $(q \vee s) \lessdot (q \vee s) \vee r = r \vee s$.



□

Theorem 1.38. *L is semimodular if and only if it is ranked, with rank function r satisfying $r(x \vee y) + r(x \wedge y) \leq r(x) + r(y)$ for all $x, y \in L$.*

Proof. Suppose that L is a ranked lattice with rank function r satisfying (1.6). If $x \wedge y \lessdot y$, then $x \vee y > x$ (otherwise $x \geq y$ and $x \wedge y = y$). On the other hand, $r(y) = r(x \wedge y) + 1$, so by (1.6)

$$r(x \vee y) - r(x) \leq r(y) - r(x \wedge y) = 1$$

which implies that in fact $x \vee y \gtrdot x$.

The hard direction is showing that a semimodular lattice has such a rank function. First, observe that if L is semimodular, then

$$x \wedge y \lessdot x, y \implies x, y \lessdot x \vee y. \tag{1.9}$$

Denote by $c(L)$ the maximum length[6] of a chain in L. We will show by induction on $c(L)$ that L is ranked.

Base case: If $c(L) = 0$ or $c(L) = 1$, then this is trivial.

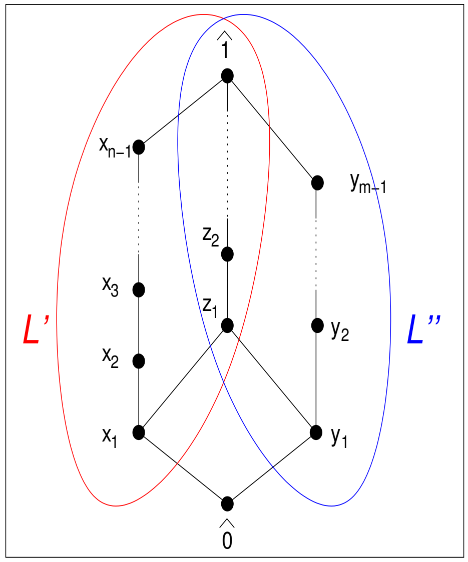
Inductive step: Suppose that $c(L) = n \geq 2$. Assume by induction that every semimodular lattice with no chain of length $c(L)$ has a rank function satisfying (1.6).

<u>First</u>, we show that L is ranked.

Let $X = \{\hat{0} = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_{n-1} \lessdot x_n = \hat{1}\}$ be a chain of maximum length. Let $Y = \{\hat{0} = y_0 \lessdot y_1 \lessdot \cdots \lessdot y_{m-1} \lessdot y_m = \hat{1}\}$ be any saturated chain in L. We wish to show that $m = n$.

Let $L' = [x_1, \hat{1}]$ and $L'' = [y_1, \hat{1}]$. By induction, these sublattices are both ranked. Moreover, $c(L') = n - 1$. If $x_1 = y_1$ then Y and X are both saturated chains in the ranked lattice L' and we are done, so suppose that $x_1 \neq y_1$. Let $z_1 = x_1 \vee y_1$. By (1.9), z_1 covers both x_1 and y_1. Let $z_1, z_2, \ldots, \hat{1}$ be a saturated chain in L (thus, in $L' \cap L''$).

[6]Remember that the length of a chain is the number of minimal relations in it, which is one less than its cardinality as a subset of L. So, for example, $c(\mathscr{B}_n) = n$, not $n + 1$.



Since L' is ranked and $z \gtrdot x_1$, the chain $z_1, \dots, \hat{1}$ has length $n-2$. So the chain $y_1, z_1, \dots, \hat{1}$ has length $n-1$.

On the other hand, L'' is ranked and $y_1, y_2, \dots, \hat{1}$ is a saturated chain, so it also has length $n-1$. Therefore the chain $\hat{0}, y_1, \dots, \hat{1}$ has length n as desired.

<u>Second</u>, we show that the rank function r of L satisfies (1.6).

Let $x, y \in L$ and take a saturated chain $x \wedge y = c_0 \lessdot c_1 \lessdot \cdots \lessdot c_{n-1} \lessdot c_n = x$. Note that $n = r(x) - r(x \wedge y)$. Then we have a chain

$$y = c_0 \vee y \leq c_1 \vee y \leq \cdots \leq c_n \vee y = x \vee y.$$

By Lemma 1.37, each $\leq$ in this chain is either an equality or a covering relation. Therefore, the *distinct* elements $c_i \vee y$ form a saturated chain from y to $x \vee y$, whose length must be $\leq n$. Hence

$$r(x \vee y) - r(y) \leq n = r(x) - r(x \wedge y)$$

and so

$$r(x \vee y) + r(x \wedge y) \leq n = r(x) + r(y).$$ □

The same argument shows that L is lower semimodular if and only if it is ranked, with a rank function satisfying the reverse inequality of (1.6)

Theorem 1.39. *L is modular if and only if it is ranked, with a rank function r satisfying*

$$r(x \vee y) + r(x \wedge y) = r(x) + r(y) \qquad \forall x, y \in L. \tag{1.10}$$

Proof. If L is modular, then it is both upper and lower semimodular, so the conclusion follows by Theorem 1.38. On the other hand, suppose that L is a lattice whose rank function r satisfies (1.10). Let $x \leq z \in L$. We already know that $x \vee (y \wedge z) \leq (x \vee y) \wedge z$, so it suffices to show that these two elements have the same rank. Indeed,

$$\begin{aligned}
r(x \vee (y \wedge z)) &= r(x) + r(y \wedge z) - r(x \wedge y \wedge z) && \text{(by (1.10))}\\
&= r(x) + r(y) + r(z) - r(y \vee z) - r(x \wedge y \wedge z) && \text{(by (1.10) again)}\\
&\geq \underline{r(x) + r(y)} + r(z) - r(x \vee y \vee z) - \underline{r(x \wedge y)} &&\\
&= r(x \vee y) + r(z) - r(x \vee y \vee z) && \text{(by (1.10) applied to the underlined terms)}\\
&= r((x \vee y) \wedge z). && \square
\end{aligned}$$

1.7. **Geometric Lattices.** The prototype of a geometric lattice is as follows. Let $\mathbb{F}$ be a field, let V be a vector space over $\mathbb{F}$, and let E be a finite subset of V. Define

$$L(E) = \{W \cap E \mid W \subseteq V \text{ is a vector subspace}\} = \{\operatorname{span}(A) \mid A \subseteq E\}.$$

This is a poset under inclusion, and is easily checked to be a lattice under the operations

$$(W \cap E) \wedge (X \cap E) = (W \cap X) \cap E, \qquad (W \cap E) \vee (X \cap E) = (W + X) \cap E.$$

The elements of $L(E)$ are called *flats*. For example, E and $\emptyset$ are both flats, because $V \cap E = E$ and $O \cap E = \emptyset$, where O means the zero subspace of V. On the other hand, if $v, w, x \in E$ with $v + w = x$, then $\{v, w\}$ is not a flat, because any vector subspace that contains both v and w must also contain x. So, an equivalent definition of "flat" is that $A \subseteq E$ is a flat if no vector in $E \setminus A$ is in the linear span of the vectors in A.

The lattice $L(E)$ is submodular, with rank function $r(A) = \dim \operatorname{span} A$. (Exercise: Check that r satisfies the submodular inequality.) It is *not* in general modular; e.g., see Example 1.44 below. On the other hand, $L(E)$ is always an *atomic* lattice: every element is the join of atoms. This is a consequence of the simple fact that $\operatorname{span}\{v_1, \dots, v_k\} = \operatorname{span}\{v_1\} + \cdots + \operatorname{span}\{v_k\}$. This motivates the following definition:

Definition 1.40. A lattice is **geometric** if it is (upper) semimodular and atomic.

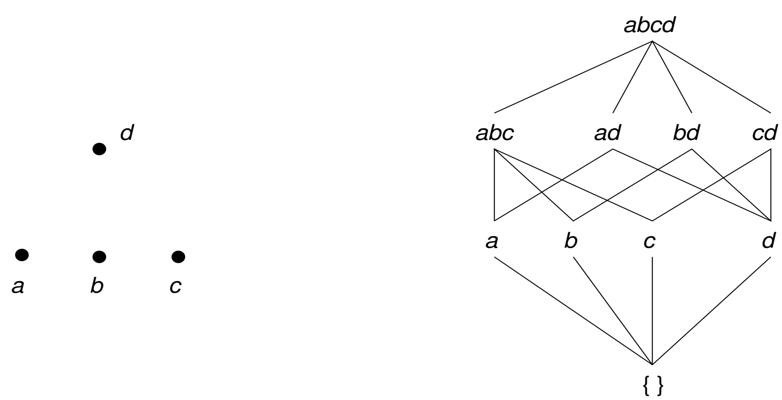
Example 1.41. Π_n is a geometric lattice. (Homework problem.)

Example 1.42. A construction closely related to $L(E)$ is the lattice

$$L^{\text{aff}}(E) = \big\{W \cap E \mid W \subseteq V \text{ is an affine subspace}\big\}.$$

(An affine subspace of V is a translate of a vector subspace; for example, a line or plane not necessarily containing the origin.) In fact, any lattice of the form $L^{\text{aff}}(E)$ can be expressed in the form $L(\hat{E})$, where $\hat{E}$ is a certain point set constructed from E (homework problem). However, the dimension of the affine span of a set $A \subseteq E$ is one less than its rank — which means that we can draw geometric lattices of rank 3 conveniently as planar point configurations.

Example 1.43. Let E be the point configuration on the left below. Then $L^{\text{aff}}(E)$ is the lattice on the right (which happens to be modular).



Example 1.44. The lattice $L(E)$ is not in general modular. For example, let $E = \{w, x, y, z\}$, where $w, x, y, z \in \mathbb{R}^3$ are *in general position*; that is, any three of them form a basis. Then $A = \{w, x\}$ and $B = \{y, z\}$ are flats. Letting r be the rank function on $L(E)$, we have

$$r(A) = r(B) = 2, \quad r(A \wedge B) = 0, \quad r(A \vee B) = 3.$$

Recall that a lattice is *relatively complemented* if, whenever $y \in [x, z] \subseteq L$, there exists $u \in [x, z]$ such that $y \wedge u = x$ and $y \vee u = z$.

Proposition 1.45. *Let L be a finite semimodular lattice. Then L is atomic (hence geometric) if and only if it is relatively complemented; that is, whenever $y \in [x, z] \subseteq L$, there exists $u \in [x, z]$ such that $y \wedge u = x$ and $y \vee u = z$.*

Here's the geometric interpretation of being relatively complemented. Suppose that V is a vector space, $L = L(E)$ for some point set $E \subseteq V$, and that $X \subseteq Y \subseteq Z \subseteq V$ are vector subspaces spanned by flats of $L(E)$. For starters, consider the case that $X = O$. Then we can choose a basis B of the space Y and extend it to a basis B' of Z, and the vector set $B' \setminus B$ spans a subspace of Z that is complementary to Y. More generally, if X is any subspace, we can choose a basis B for X, extend it to a basis B' of Y, and extend B' to a basis B'' of Z. Then $B \cup (B'' \setminus B')$ spans a subspace $U \subseteq Z$ that is relatively complementary to Y, i.e., $U \cap Y = X$ and $U + Y = Z$.

Proof. ($\Longrightarrow$) Suppose that L is atomic. Let $y \in [x, z]$, and choose $u \in [x, z]$ such that $y \wedge u = x$ (for instance, $u = x$). If $y \vee u = z$ then we are done. Otherwise, choose an atom $a \in L$ such that $a \leq z$ but $a \not\leq y \vee u$. Set $u' = u \vee a$. By semimodularity $u' \gtrdot u$. Then $u' \vee y \gtrdot u \vee y$ (by Lemma 1.37; details omitted) and in addition $u' \wedge y = x$ (why?) By repeatedly replacing u with u' if necessary, we eventually obtain a complement for y in $[x, z]$.

($\Longleftarrow$) Suppose that L is relatively complemented and let $x \in L$. We want to write x as the join of atoms. If $x = \hat{0}$ then it is the empty join; otherwise, let $a_1 \leq x$ be an atom and let x_1 be a complement for a_1 in $[\hat{0}, x]$. Then $x_1 < x$ and $x = a_1 \vee x_1$. Replace x with x_1 and repeat, getting

$$x = a_1 \vee x_1 = a_1 \vee (a_2 \vee x_2) = (a_1 \vee a_2) \vee x_2 = \cdots = (a_1 \vee \cdots \vee a_n) \vee x_n = \cdots$$

We have $x > x_1 > x_2 > \cdots$, so eventually $x_n = \hat{0}$, and $x = a_1 \vee \cdots \vee a_n$. $\square$

2. Matroids

The motivating example of a geometric lattice is the lattice of flats of a finite set E of vectors. The underlying combinatorial data of this lattice can be expressed in terms of the rank function, which says the dimension of the space spanned by every subset of E. However, there are many other equivalent ways to describe the "combinatorial linear algebra" of a set of vectors: the family of linearly independent sets; the family of sets that form bases; which vectors lie in the span of which sets; etc. All of these ways are descriptions of a **matroid** structure on E. Matroids can also be regarded as generalizations of graphs, and are important in combinatorial optimization as well.

2.1. Closure operators.

Definition 2.1. Let E be a finite set. A **closure operator on** E is a map $2^E \to 2^E$, written $A \mapsto \bar{A}$, such that *(i)* $A \subseteq \bar{A} = \bar{\bar{A}}$ and *(ii)* if $A \subseteq B$, then $\bar{A} \subseteq \bar{B}$. As a not-quite-trivial consequence,

$$\overline{\bar{A} \cap \bar{B}} = \bar{A} \cap \bar{B} \qquad \forall A, B \subseteq E, \tag{2.1}$$

because $\overline{\bar{A} \cap \bar{B}} \subset \bar{\bar{A}} = \bar{A}$.

Definition 2.2. A **matroid closure operator** on E is a closure operator satisfying in addition the *exchange axiom*:

$$\text{if } e \notin \bar{A} \text{ but } e \in \overline{A \cup f}, \text{ then } f \in \overline{A \cup e}. \tag{2.2}$$

Remark on notation: Here $A \cup e$ is short for $A \cup \{e\}$. Similarly, I will often abbreviate $A \setminus \{e\}$ bu $A \setminus e$.

Definition 2.3. A **matroid** M is a set E (the "ground set") together with a matroid closure operator. A closed subset of M (i.e., a set that is its own closure) is called a **flat** of M. The matroid is called **simple** if the empty set and all singleton sets are closed.

Example 2.4. *Vector matroids.* Let V be a vector space over a field $\mathbb{F}$, and let $E \subseteq V$ be a finite set. Then

$$A \mapsto \bar{A} := (\operatorname{span} A) \cap E$$

is a matroid closure operator on E. It is easy to check the conditions for a closure operator. To check condition (2.2), if $e \in \overline{A \cup \{f\}}$, then we have a linear equation

$$e = c_f f + \sum_{a \in A} c_a a, \qquad c_f, c_a \in \mathbb{F}.$$

If $e \notin \bar{A}$, then $c_f \neq 0$, so we can solve for f to express it as a linear combination of the vectors in $A \cup \{e\}$, obtaining (2.2). A matroid arising in this way (or, more generally, isomorphic to such a matroid) is called a *vector matroid*, *vectorial matroid*, *linear matroid* or *representable matroid* (over $\mathbb{F}$).

What does it mean if the matroid is simple? The condition $\bar{\emptyset} = O$ says that none of the vectors can be the zero vector, the condition that all singleton sets are closed says that no two vectors are scalar multiples of each other. If we want to study linear independence of a set of vectors, these are reasonable conditions to impose.

A vector matroid records information about linear dependence (i.e., which vectors belong to the linear spans of other sets of vectors) without having to worry about the actual coordinates of the vectors. More generally, a matroid can be thought of as a combinatorial, coordinate-free abstraction of linear dependence and independence.

2.2. **Matroids and geometric lattices.** The following theorem says that simple matroids and geometric lattices are essentially the same things.

Theorem 2.5. *1. Let M be a simple matroid with finite ground set E. Let $L(M)$ be the poset of flats of M, ordered by inclusion. Then $L(M)$ is a geometric lattice, under the operations $A \wedge B = A \cap B$, $A \vee B = \overline{A \cup B}$.*

2. Let L be a geometric lattice and let E be its set of atoms. Then the function $\bar{A} = \{e \in E \mid e \leq \bigvee A\}$ is a matroid closure operator on E.

Proof. Step 1a: Show that $L(M)$ is an atomic lattice. The intersection of flats is a flat (an easy exercise), so the operation $A \wedge B = A \cap B$ makes $L(M)$ into a meet-semilattice. It's bounded (with $\hat{0} = \bar{\emptyset}$ and $\hat{1} = E$), so it's a lattice by Proposition 1.18. Meanwhile, $\overline{A \cup B}$ is by definition the smallest flat containing $A \cup B$, so it is the meet of all flats containing both A and B. (Note that this argument shows that any closure operator, not necessarily matroidal, gives rise to a lattice.)

By definition of a simple matroid, the singleton subsets of E are atoms in $L(M)$. Every flat is the join of the atoms corresponding to its elements, so $L(M)$ is atomic.

Step 1b: Show that $L(M)$ is semimodular. First, I claim that if $F \in L(M)$ and $a \notin F$, then $F \lessdot F \vee \{a\}$. Indeed, let G be a flat such that

$$F \subsetneq G \subseteq F \vee \{a\} = \overline{F \cup \{a\}}.$$

For any $b \in G \setminus F$, we have $b \in \overline{F \cup \{a\}}$ so $a \in \overline{F \cup \{b\}}$ by the exchange axiom (2.2), which implies $F \vee \{a\} \subseteq F \vee \{b\} \subseteq G$. So the $\subseteq$ above is actually an equality.

On the other hand, if $F \lessdot G$ then $G = F \vee \{a\}$ for any atom $a \in G \setminus F$. So the covering relations are exactly the relations of this form.

Suppose now that F and G are incomparable and that $F \gtrdot F \wedge G$. Then Then $F = (F \wedge G) \vee \{a\}$ for some $a \in M$. We must have $a \not\leq G$ (otherwise $F \leq G$; the easiest way to prove this is using the atomic property), so $G < G \vee \{a\}$, and by the previous assertion this must be a cover. We have just proved that $L(M)$ is semimodular. In particular, it is ranked, with rank function

$$r(F) = \min\left\{|B| : \ B \subseteq E, \ F = \bigvee B\right\}.$$

(Such a set B is called a *basis* of F.)

For assertion (2), it is easy to check that $A \mapsto \bar{A}$ is a closure operator, and that $\bar{A} = A$ for $|A| \leq 1$. So the only nontrivial part is to establish the exchange axiom (2.2).

Recall that if L is semimodular and $x, e \in L$ with e an atom and $x \not\geq e$, then $x \vee e \gtrdot x$ (because $r(x \vee e) - r(x) \leq r(e) - r(x \wedge e) = 1 - 0 = 1$).

Suppose that e, f are atoms and A is a set of atoms such that $e \notin \bar{A}$ but $e \in \overline{A \cup f}$. We wish to show that $f \in \overline{A \cup e}$. Let $x = \bigvee A \in L$. Then $x \lessdot x \vee f$ and $x < x \vee e \leq x \vee f$. Together, this implies that $x \vee f = x \vee e$. In particular $f \leq x \vee e$, i.e., $f \in \overline{A \cup e}$ as desired. $\square$

In view of this bijection, we can describe a matroid on ground set E by the function $A \mapsto r(\bar{A})$, where r is the rank function of the associated geometric lattice. It is standard to abuse notation by calling this function r also. Formally:

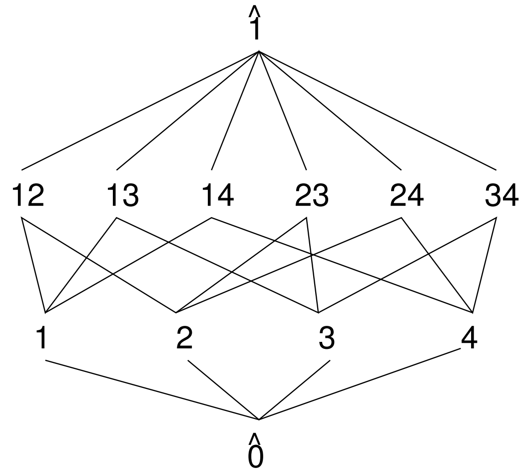
Definition 2.6. A *matroid rank function* on E is a function $r : 2^E \to \mathbb{N}$ satisfying the following conditions for all $A, B \subseteq E$:

(1) $r(A) \leq |A|$.
(2) If $A \subseteq B$ then $r(A) \leq r(B)$.
(3) $r(A) + r(B) \geq r(A \cap B) + r(A \cup B)$ (the *rank submodular inequality*).

Observe that

- If r is a matroid rank function on E, then the corresponding matroid closure operator is given by
$$\overline{A} = \{e \in E : \ r(A \cup e) = r(A)\}.$$
Moreover, this closure operator defines a *simple* matroid if and only if $r(A) = |A|$ whenever $|A| \leq 2$.
- If $A \mapsto \bar{A}$ is a matroid closure operator on E, then the corresponding matroid rank function r is
$$r(A) = \min\{|B| : \ \bar{B} = \bar{A}\}.$$

Example 2.7. Let $n = |E|$ and $0 \leq k \leq E$, and define
$$r(A) = \min(k, |A|).$$
This clearly satisfies the conditions of a matroid rank function (Definition 2.6). The corresponding matroid is called the *uniform matroid* $U_k(n)$, and has closure operator
$$\bar{A} = \begin{cases} A \text{ if } |A| < k, \\ E \text{ if } |A| \geq k. \end{cases}$$
So the flats of M of the sets of cardinality $< k$, as well as (of course) E itself. Therefore, the lattice of flats looks like a Boolean algebra $\mathscr{B}_n$ that has been truncated at the k^{th} rank. For $n = 3$ and $k = 2$, this lattice is M_5; for $n = 4$ and $k = 3$, the Hasse diagram is as follows (see Example 1.44).



If S is a set of n points in general position in $\mathbb{F}^k$, then the corresponding matroid is isomorphic to $U_k(n)$. This sentence is tautological, in the sense that it can be taken as a definition of "general position". Indeed, if $\mathbb{F}$ is infinite and the points are chosen randomly (in some reasonable analytic or measure-theoretic sense), then $L(S)$ will be isomorphic to $U_k(n)$ with probability 1. On the other hand, $\mathbb{F}$ must be sufficiently large (in terms of n) in order for $\mathbb{F}^k$ to have n points in general position.

As for "isomorphic", here is a precise definition.

Definition 2.8. Let M, M' be matroids on ground sets E, E' respectively. We say that M and M' are **isomorphic**, written $M \cong M'$, if there is a bijection $f : E \to E'$ meeting any (hence all) of the following conditions:

(1) There is a lattice isomorphism $L(M) \cong L(M')$;
(2) $r(A) = r(f(A))$ for all $A \subseteq E$. (Here $f(A) = \{f(a) \mid a \in A\}$.)
(3) $\overline{f(A)} = f(\bar{A})$ for all $A \subseteq E$.

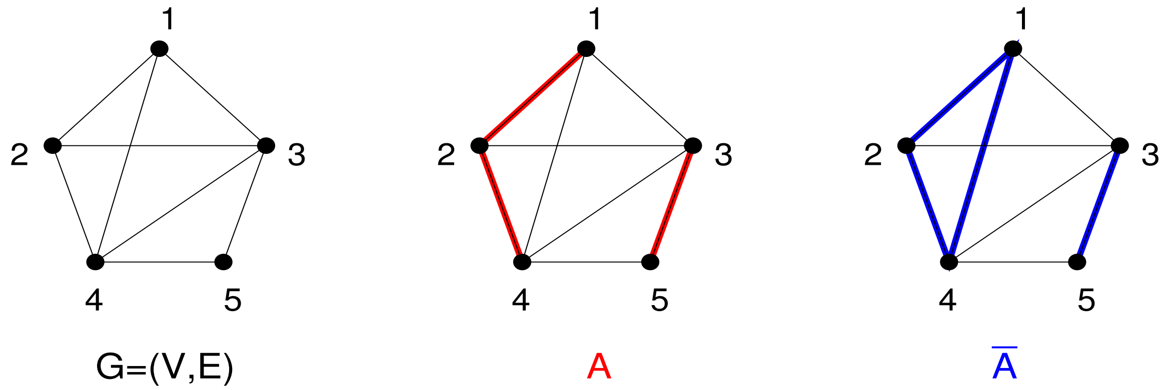
In general, every equivalent definition of "matroid" (and there are several more coming) will induce a corresponding equivalent notion of "isomorphic".

2.3. **Graphic Matroids.** One important application of matroids is in graph theory. Let G be a finite graph with vertices V and edges E. For convenience, we'll write $e = xy$ to mean "e is an edge with endpoints x, y"; this should not be taken to exclude the possibility that e is a loop (i.e., $x = y$) or that some other edge might have the same pair of endpoints.

Definition 2.9. For each subset $A \subseteq E$, the corresponding *induced subgraph* of G is the graph $G|_A$ with vertices V and edges A. The *graphic matroid* or *complete connectivity matroid* $M(G)$ on E is defined by the closure operator

$$\bar{A} = \{e = xy \in E \mid x, y \text{ belong to the same component of } G|_A\}. \tag{2.3}$$

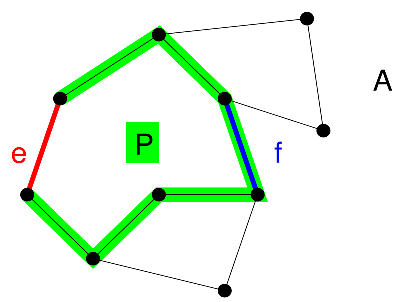
Equivalently, an edge $e = xy$ belongs to $\bar{A}$ if there is a path between x and y consisting of edges in A (for short, an *A-path*). For example, in the following graph, $14 \in \bar{A}$ because $\{12, 24\} \subset A$.



Proposition 2.10. *The operator $A \mapsto \bar{A}$ defined by* (2.3) *is a matroid closure operator.*

Proof. It is easy to check that $A \subseteq \bar{A}$ for all A, and that $A \subseteq B \implies \bar{A} \subseteq \bar{B}$. If $e = xy \in \bar{\bar{A}}$, then x, y can be joined by an $\bar{A}$-path P, and each edge in P can be replaced with an A-path, giving an A-path between x and y.

Finally, suppose $e = xy \notin \bar{A}$ but $e \in \overline{A \cup f}$. Let P be an $(A \cup f)$-path; in particular, $f \in P$. Then $P \cup f$ is a cycle, from which deleting f produces an $(A \cup e)$-path between the endpoints of f. $\square$



The rank function of the graphic matroid is given by

$$r(A) = \min\{|B| : \ B \subseteq A, \ \overline{B} = \overline{A}\}.$$

Such a subset B is called a *spanning forest* of A (or of $G|_A$). They are the bases of the graphic matroid $M(G)$.

Theorem 2.11. *Let $B \subseteq A$. Then any two of the following conditions imply the third (and characterize spanning forests of A):*

1. *$r(B) = r(A)$;*
2. *B is acyclic;*
3. *$|B| = |V| - c$, where c is the number of connected components of A.*

The flats of $M(G)$ correspond to the subgraphs whose components are all *induced subgraphs* of G. For $W \subseteq V$, the induced subgraph $G[W]$ is the graph with vertices W and edges $\{xy \in E \mid x, y \in W\}$.

Example 2.12. If G is a *forest* (a graph with no cycles), then no two vertices are joined by more than one path. Therefore, every edge set is a flat, and $M(G)$ is a Boolean algebra.

Example 2.13. If G is a cycle of length n, then every edge set of size $< n - 1$ is a flat, but the closure of a set of size $n - 1$ is the entire edge set. Therefore, $M(G) \cong U_{n-1}(n)$.

Example 2.14. If $G = K_n$ (the complete graph on n vertices), then a flat of $M(G)$ is the same thing as an equivalence relation on $[n]$. Therefore, $M(K_n)$ is naturally isomorphic to the partition lattice Π_n.

2.4. **Equivalent Definitions of Matroids.** In addition to rank functions, lattices of flats, and closure operators, there are many other equivalent ways to define a matroid on a finite ground set E. In the fundamental example of a linear matroid M, some of these definitions correspond to linear-algebraic notions such as linear independence and bases.

Definition 2.15. A **(matroid) independence system** $\mathscr{I}$ is a family of subsets of E such that

$$\emptyset \in \mathscr{I}; \tag{2.4a}$$

$$\text{if } I \in \mathscr{I} \text{ and } I' \subseteq I, \text{ then } I' \in \mathscr{I}; \quad \text{and} \tag{2.4b}$$

$$\text{if } I, J \in \mathscr{I} \text{ and } |I| < |J|, \text{ then there exists } x \in J \setminus I \text{ such that } I \cup x \in \mathscr{I}. \tag{2.4c}$$

Note: A family of subsets satisfying (2.4a) and (2.4b) is called a *(abstract) simplicial complex* on E.

If E is a finite subset of a vector space, then the linearly independent subsets of E form a matroid independence system. Conditions (2.4a) and (2.4b) are clear. For condition (2.4c), the span of J has greater dimension than that of I, so there must be some $x \in J$ outside the span of I, and then $I \cup x$ is linearly independent.

A matroid independence system records the same combinatorial structure on E as a matroid rank function:

Proposition 2.16. *Let E be a finite set.*

(1) *If r is a matroid rank function on E, then*
$$\mathscr{I} = \{A \subset E \mid r(A) = |A|\}$$
is an independence system.
(2) *If $\mathscr{B}$ is an independence system on E, then*
$$r(A) = \max\{|I \cap A| \mid I \in \mathscr{B}\}$$
is a matroid rank function.
(3) *These constructions are mutual inverses.*

If $M = M(G)$ is a graphic matroid, the associated independence system is the family of *acyclic* edge sets in G. To see this, notice that if A is a set of edges and $e \in A$, then $r(A \setminus e) < r(A)$ if and only if deleting e breaks a component of $G|_A$ into two smaller components (so that in fact $r(A \setminus e) = r(A) - 1$). This is equivalent to the condition that e belongs to no cycle in A. Therefore, if A is acyclic, then deleting its edges one by one gets you down to $\emptyset$ and decrements the rank each time, so $r(A) = |A|$. On the other hand, if A contains a cycle, then deleting any of its edges won't change the rank, so $r(A) < |A|$.

Here's what the "donation" condition (2.4c) means in the graphic setting. Suppose that $|V| = n$, and let $c(H)$ denote the number of components of a graph H. If I, J are acyclic edge sets with $|I| < |J|$, then
$$c(G|_I) = n - |I| \ > \ c(G|_J) = n - |J|,$$
and there must be some edge $e \in J$ whose endpoints belong to different components of $G|_I$; that is, $I \cup e$ is acyclic.

Which abstract simplicial complexes are matroid independence complexes? The following answer is useful in combinatorial commutative algebra. First, a simplicial complex is called **pure** if all its maximal faces have the same cardinality. The donation condition implies that matroid complexes are pure, but in fact being matroidal is much stronger than being pure, to wit:

Proposition 2.17. *Let Δ be an abstract simplicial complex on E. The following are equivalent:*

(1) *Δ is a matroid independence complex.*
(2) *For every $F \subseteq E$, the* ***induced subcomplex*** *$\Delta|_F = \{\sigma \in E \mid \sigma \subseteq F\}$ is Cohen-Macaulay.*

(3) *For every $F \subseteq E$, the induced subcomplex $\Delta|_F$ is pure.*

What "Cohen-Macaulay" means and why it is important is a long story, beyond the scope of this course. Suffice it here to say that the condition is a fundamental link between commutative algebra and combinatorics. The standard references are Bruns and Herzog [3] (which has a relatively algebraic point of view) and Stanley [13] (affectionately known as "Stanley's Green Book"; relatively combinatorial). Many important recent developments can be found in Miller and Sturmfels [8]. Algebraists may know that the implications (1) $\implies$ (2) $\implies$ (3) are not too hard (the first can be proved using shellability; the second is more or less the statement that Cohen-Macaulay ideals are equidimensional). The equivalence of (1) and (3) is purely combinatorial, to wit:

Proof of (1) $\implies$ (3). Suppose that Δ is a matroid independence complex and $F \subseteq E$. Let I, J be faces of $\Delta|_F$ with $|I| < |J|$. Donation says that there is some $x \in J \setminus I$ such that $I \cup \{x\} \in \Delta$, and certainly $I \cup \{x\} \subset F$. So I is not a maximal face of $\Delta|_F$. Hence Δ_F is pure. $\square$

Proof of (3) $\implies$ (1). On the other hand, suppose that $\Delta|_F$ is pure for every $F \subseteq E$. Let $I, J \in \Delta$ with $|I| < |J|$; then I is not a maximal face of $\Delta|_{I \cup J}$. So there is some $x \in (I \cup J) \setminus I = J \setminus I$ such that $I \cup \{x\} \in \Delta|_F \subseteq \Delta$, which is precisely the donation condition. $\square$

The maximal independent sets — that is, bases — provide another way of axiomatizing a matroid.

Definition 2.18. A **(matroid) basis system on** E is a family $\mathscr{B} \subseteq 2^E$ such that for all $B, B' \in \mathscr{B}$,

$$|B| = |B'|; \quad \text{and} \tag{2.5a}$$

$$\text{for all } e \in B \setminus B', \text{ there exists } e' \in B' \setminus B \text{ such that } B \setminus e \cup e' \in \mathscr{B}. \tag{2.5b}$$

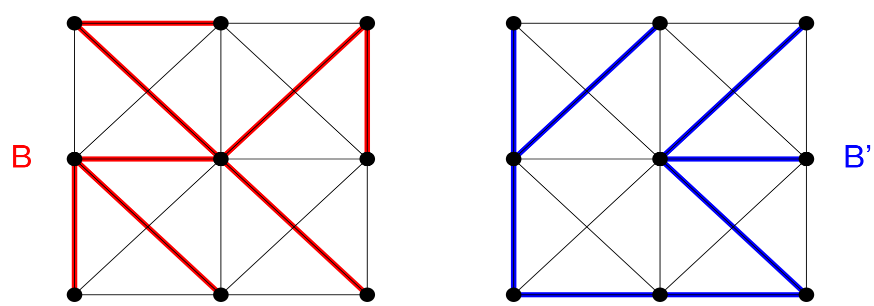
Given (2.5a), the condition (2.5b) can be replaced with

$$\text{for all } e \in B \setminus B', \text{ there exists } e' \in B' \setminus B \text{ such that } B' \setminus e' \cup e \in \mathscr{B}, \tag{2.5c}$$

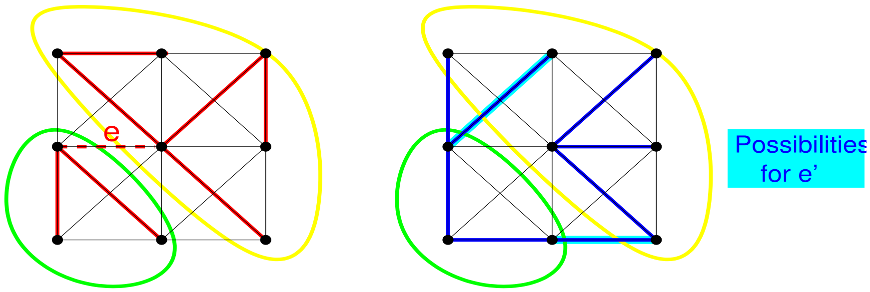
although this requires some proof (homework!).

For example, if S is a finite set of vectors spanning a vector space V, then the subsets of S that are bases for V all have the same cardinality (namely $\dim V$) and satisfy the basis exchange condition (2.5b).

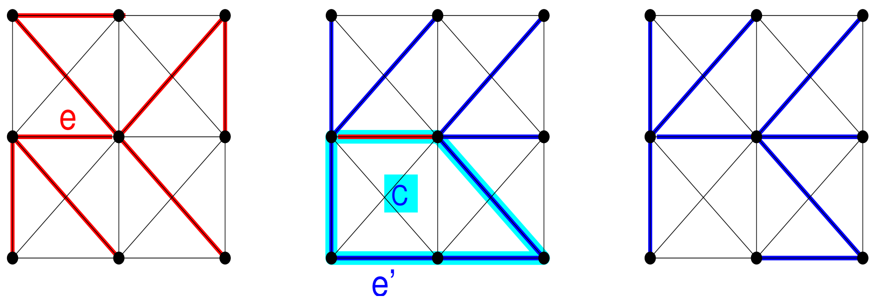
If G is a graph, then the bases of $M(G)$ are its **spanning forests**, i.e., its maximal acyclic edge sets. If G is connected (which, as we will see, we may as well assume when studying graphic matroids) then the bases of $M(G)$ are spanning trees.



Here is the graph-theoretic interpretation of (2.5b). If G is connected, B, B' are spanning trees, and $e \in B \setminus B'$, then $B \setminus e$ has two connected components. Since B' is connected, it must have some edge e' with one endpoint in each of those components, and then $B \setminus e \cup e'$ is a spanning tree.



As for (2.5c), if $e \in B \setminus B'$, then $B' \cup e$ must contain a unique cycle C (formed by e together with the unique path in B' between the endpoints of e). Deleting any edge $e' \in C \setminus e$ will produce a spanning tree.



If G is a graph with edge set E and $M = M(G)$ is its graphic matroid, then

$$\mathscr{I} = \{A \subseteq E \mid A \text{ is acyclic}\},$$
$$\mathscr{B} = \{A \subseteq E \mid A \text{ is a spanning forest of } G\}.$$

If S is a set of vectors and $M = M(S)$ is the corresponding linear matroid, then

$$\mathscr{I} = \{A \subseteq S \mid A \text{ is linearly independent}\},$$
$$\mathscr{B} = \{A \subseteq S \mid A \text{ is a basis for } \operatorname{span}(S)\}.$$

Proposition 2.19. *Let E be a finite set.*

(1) *If $\mathscr{I}$ is an independence system on E, then the family of maximal elements of $\mathscr{I}$ is a basis system.*
(2) *If $\mathscr{B}$ is a basis system, then $\mathscr{I} = \bigcup_{B \in \mathscr{B}} 2^B$ is an independence system.*
(3) *These constructions are mutual inverses.*

The proof is left as an exercise. We already have seen that an independence system on E is equivalent to a matroid rank function. So Proposition 2.19 asserts that a basis system provides the same structure on E. Bases turn out to be especially convenient for describing fundamental operations on matroids such as duality, direct sum, and deletion/contraction (all of which are coming soon).

One last way of defining a matroid (there are many more!):

Definition 2.20. A **(matroid) circuit system on** E is a family $\mathscr{C} \subseteq 2^E$ such that, for all $C, C' \in \mathscr{C}$,

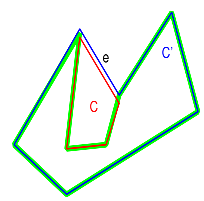
(2.6a) $\quad C \not\subseteq C'$; and

(2.6b) $\quad$ for all $e \in C \cap C'$, $C \cup C' \setminus e$ contains an element of $\mathscr{C}$.

In a linear matroid, the circuits are the minimal dependent sets of vectors. Indeed, if C, C' are such sets and $e \in C \cap C'$, then we can find two expressions for e as nontrivial linear combinations of vectors in C and in

C', and equating these expressions and eliminating e shows that $C \cup C' \setminus e$ is dependent, hence contains a circuit.

In a graph, if two cycles C, C' meet in an edge $e = xy$, then $C \setminus e$ and $C' \setminus e$ are paths between x and y, so concatenating them forms a closed path, which must contain some cycle.



Proposition 2.21. *Let E be a finite set.*

(1) *If $\mathscr{I}$ is an independence system on E, then $\{C \subseteq E \mid C \notin \mathscr{I}$ and $C' \in \mathscr{I}\ \forall C' \subsetneq C\}$ is a circuit system.*
(2) *If $\mathscr{C}$ is a circuit system, then $\{I \subseteq E \mid C \not\subseteq I\ \forall C \in \mathscr{C}\}$ is an independence system.*
(3) *These constructions are mutual inverses.*

Proof: Exercise.

The following definition of a matroid is different from what has come before, and gives a taste of the importance of matroids in combinatorial optimization

Let E be a finite set and let Δ be an abstract simplicial complex on E (see Definition 2.15). Let $w : E \to \mathbb{R}_{\geq 0}$ be a function, which we regard as assigning weights to the elements of E, and for $A \subseteq E$, define $w(A) = \sum_{e \in A} w(e)$. Consider the problem of maximizing $w(A)$ over all subsets $A \in \Delta$ (also known as *facets* of Δ). A naive approach to try to produce such a set A, which may or may not work for a given Δ and w, is the following *greedy algorithm*:

(1) Let $A = \emptyset$.
(2) If A is a facet of Δ, stop.
Otherwise, find $e \in E \setminus A$ of maximal weight such that $A \cup \{e\} \in \Delta$ (if there are several such e, pick one at random), and replace A with $A \cup \{e\}$.
(3) Repeat step 2 until A is a facet of Δ.

Proposition 2.22. *Δ is a matroid independence system if and only if the greedy algorithm produces a facet of maximal weight for* every *weight function w.*

The proof is left as an exercise, as is the construction of a simplicial complex and a weight function for which the greedy algorithm does not produce a facet of maximal weight.

Summary of Matroid Axiomatizations

- *Geometric lattice*: a lattice that is atomic and semimodular. Onlysimple matroids can be described this way.
- *Rank function*: function $r : 2^E \to \mathbb{N}$ such that $r(A) \leq |A|$ and $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$. Simple if $r(A) = |A|$ whenever $|A| \leq 1$.
- *Closure operator*: function $2^E \to 2^E$, $A \mapsto \bar{A}$ such that $A \subseteq \bar{A} = \bar{\bar{A}}$; $A \subseteq B \implies \bar{A} \subseteq B$; and $x \notin \bar{A}$, $x \in \overline{A \cup y} \implies y \in \overline{A \cup x}$. Simple if $\bar{A} = A$ whenever $|A| \leq 1$.
- *Independence system*: set family $\mathscr{I} \subseteq 2^E$ such that $\emptyset \in \mathscr{I}$; $I \in \mathscr{I}$, $I' \subseteq I \implies I' \in \mathscr{I}$; and $I, J \in \mathscr{I}$, $|I| < |J| \implies \exists x \in J \setminus I : I \cup x \in \mathscr{I}$. Simple if $A \in \mathscr{I}$ whenever $|A| \leq 2$.
- *Basis system*: set family $\mathscr{I} \subseteq 2^E$ such that $\emptyset \in \mathscr{I}$; $I \in \mathscr{I}$, $I' \subseteq I \implies I' \in \mathscr{I}$; and $I, J \in \mathscr{I}$, $|I| < |J| \implies \exists x \in J \setminus I : I \cup x \in \mathscr{I}$. Simple if every element and every pair of elements belong to some basis.
- *Circuit system:* set family $\mathscr{C} \subseteq 2^E$ such that no element contains any other, and $C.C' \in \mathscr{C}$, $e \in C \cap C' \implies \exists C'' \in \mathscr{C} : C'' \subseteq C \cup C' \setminus e$. Simple if all elements have size at least 3.
- *Greedy algorithm:* simplicial complex Δ on E such that the greedy algorithm successfully constructs a maximum-weight facet for every weight function $w : E \to \mathbb{R}_{\geq 0}$.

2.5. **Representability and Regularity.** The motivating example of a matroid is a finite collection of vectors in $\mathbb{R}^n$ — but what if we work over a different field? What if we turn this question on its head by specifying a matroid M purely combinatorially and then asking which fields give rise to vector sets whose matroid is M?

Definition 2.23. Let M be a matroid and V a vector space over a field $\mathbb{F}$. A set of vectors $S \subset V$ **represents** or **realizes** M **over** $\mathbb{F}$ if the linear matroid $M(S)$ associated with S is isomorphic to M.

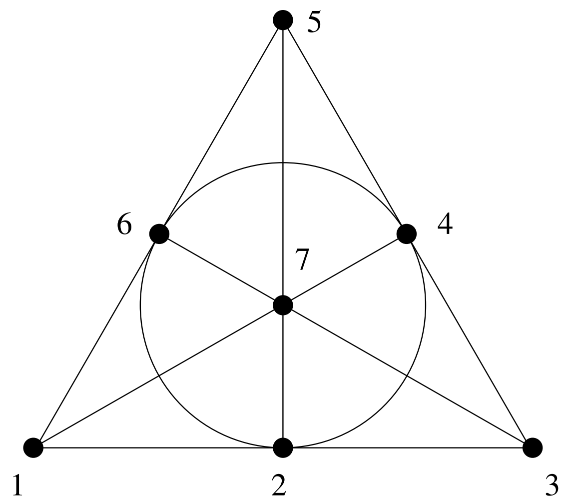
For example:

- The matroid $U_2(3)$ is representable over $\mathbb{F}_2$ (in fact, over any field): we can take $S = \{(1,0),\ (0,1),\ (1,1)\}$, and any two of these vectors form a basis of $\mathbb{F}_2^2$.
- If $\mathbb{F}$ has at least three elements, then $U_2(4)$ is representable, by, e.g., $S = \{(1,0),\ (0,1),\ (1,1),\ (1,a)\}$. where $a \in \mathbb{F} \setminus \{0,1\}$. Again, any two of these vectors form a basis of $\mathbb{F}^2$.
- On the other hand, $U_2(4)$ is not representable over $\mathbb{F}_2$, because $\mathbb{F}_2^2$ doesn't contain four nonzero elements.

More generally, suppose that M is a simple matroid with n elements (i.e., the ground set E has $|E| = n$) and rank r (i.e., every basis of M has size r) that is representable over the finite field $\mathbb{F}_q$ of order q. Then each element of E must be represented by some nonzero vector in $\mathbb{F}_q^n$, and no two vectors can be scalar multiples of each other. Therefore,

$$n \leq \frac{q^r - 1}{q - 1}.$$

Example 2.24. *The Fano plane.* Consider the affine point configuration with 7 points and 7 lines (one of which looks like a circle), as shown:

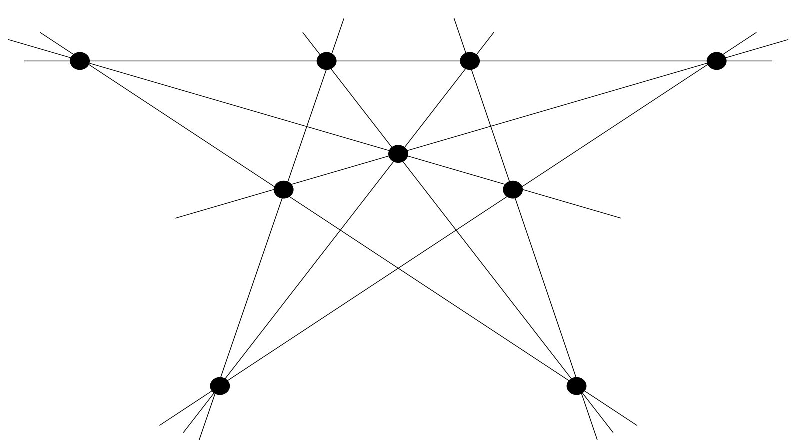


This point configuration can't be represented over $\mathbb{R}$ — if you try to draw seven non-collinear points such that the six triples 123, 345, 156, 147, 257, 367 are each collinear, then 246 will not be collinear (and in fact this is true over any field of characteristic $\neq 2$) — but it can be represented over $\mathbb{F}_2$, for example by the columns of the matrix

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \in (\mathbb{F}_2)^{3\times 7}.$$

Viewed as a matroid, the Fano plane has rank 3. Its bases are the $\binom{7}{3} - 7 = 28$ noncollinear triples of points. Its circuits are the seven collinear triples and their complements (known as **ovals**). For instance, 4567 is an oval: it is too big to be independent, but on the other hand every three-element subset of it forms a basis (in particular, is independent), so it is a circuit.

Representability is a tricky issue. As we have seen, $U_2(4)$ can be represented over any field other than $\mathbb{F}_2$, while the Fano plane is representable only over fields of characteristic 2. The point configuration below gives a rank-3 matroid that is representable over $\mathbb{R}$ but not over $\mathbb{Q}$ [7, pp. 93–94].

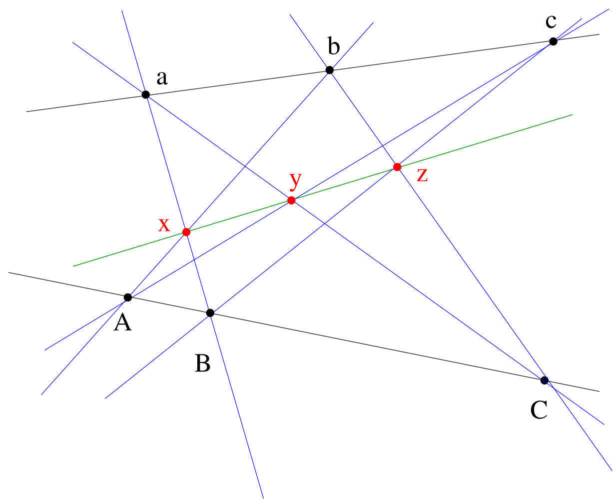


A **regular** matroid is one that is representable over every field. (For instance, we will see that graphic matroids are regular.) For some matroids, the choice of field matters. For example, every uniform matroid is representable over every infinite field, but $U_k(n)$ can be represented over $\mathbb{F}_q$ only if $k \leq q^n - 1$ (so that there are enough nonzero vectors in $\mathbb{F}_q^n$), although this condition is not sufficient. (For example, $U_2(4)$ is not representable over $\mathbb{F}_2$.)

There exist matroids that are not representable over *any* field. The best-known example is arguably the **non-Pappus matroid**, has a ground set of size 9 (see Example 2.25).

Example 2.25. Pappus' Theorem from Euclidean geometry says the following:

Let a, b, c, A, B, C be distinct points in $\mathbb{R}^2$ such that a, b, c and A, B, C are collinear. Then the three points $x = \overline{aB} \cap \overline{Ab}$, $y = \overline{aC} \cap \overline{Ac}$, $z = \overline{bC} \cap \overline{Bc}$ are collinear.



Accordingly, there is a rank-3 simple matroid on ground set $E = \{a, b, c, A, B, C, x, y, z\}$ whose flats are

$$\{\emptyset, E\} \ \cup\ \big\{\{x\} \mid x \in E\big\} \ \cup\ \{abc,\ ABC,\ aBx,\ Abx,\ aCy,\ Acy,\ bCz,\ Bcz,\ xyz\}.$$

It turns out that deleting $\{x, y, z\}$ produces the family of closed sets of a matroid closure operator. Since Pappus' theorem can be proven using analytic geometry, and the equations that say that x, y, z are collinear work over any field, it follows that the corresponding matroid is not representable over any field.

The smallest matroids not representable over any field have ground sets of size 8; one of these is the rank-4 *Vámos matroid* V_8 [10, p. 511].

2.6. **Operations on Matroids.**

2.6.1. *Duality.*

Definition 2.26. Let M be a matroid with basis system $\mathscr{B}$. The **dual matroid** M^* (also known as the **orthogonal matroid** to M and denoted $M^\perp$) has basis system

$$\mathscr{B}^* = \{E \setminus B \mid B \in \mathscr{B}\}.$$

Note that (2.5a) is clearly invariant under complementation, and complementation swaps (2.5b) and (2.5c) — so if you believe that those conditions are equivalent then you also believe that $\mathscr{B}^*$ is a matroid basis system. Also, it is clear that $(M^*)^* = M$.

What does duality mean for a vector matroid? Let $S = \{v_1, \ldots, v_n\} \subset \mathbb{F}^r$, and let $M = M(S)$. We may as well assume that S spans $\mathbb{F}^r$. That is, $r \leq n$, and the $r \times n$ matrix X with columns v_i has full rank r. Let

Y be any $(n-r) \times n$ matrix with

$$\text{rowspace}(Y) = \text{nullspace}(X).$$

That is, the rows of Y span the orthogonal complement of rowspace(X) according to the standard inner product. Then the columns of Y represent M^*. To see this, first, note that $\text{rank}(Y) = \dim \text{nullspace}(X) = n-r$, and second, check that a set of columns of Y spans its column space if and only if the complementary set of columns of X has full rank.

Example 2.27. Let $S = \{v_1, \dots, v_5\}$ be the set of column vectors of the following matrix (over $\mathbb{R}$, say):

$$X = \begin{bmatrix} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Notice that X has full rank (it's in row-echelon form, after all), so it represents a matroid of rank 3 on 5 elements. We could take Y to be the matrix

$$Y = \begin{bmatrix} 0 & 0 & -2 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{bmatrix}.$$

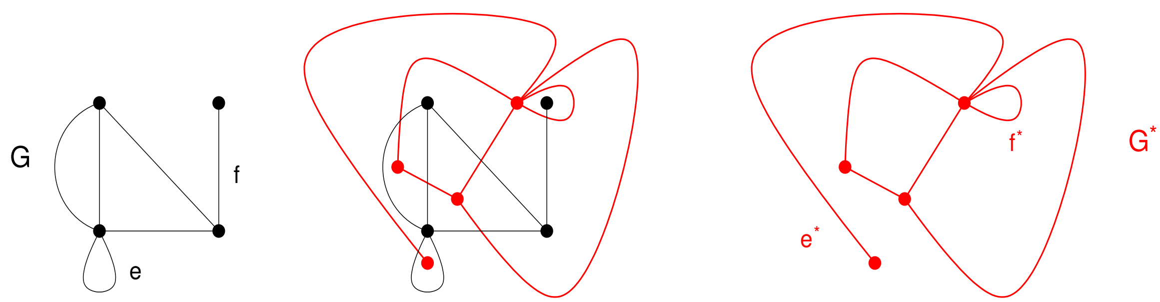
Then Y has rank 2. Call its columns $\{v_1^*, \dots, v_5^*\}$; then the pairs of columns that span its column space are

$$\{v_1^*, v_3^*\},\ \{v_1^*, v_4^*\},\ \{v_2^*, v_3^*\},\ \{v_2^*, v_4^*\},\ \{v_3^*, v_4^*\},$$

whose (unstarred) complements are precisely those triples of columns of X that span *its* column space.

In particular, every basis of M contains v_5, which corresponds to the fact that no basis of M^* contains v_5^*.

Example 2.28. Let G be a connected *planar* graph, i.e., one that can be drawn in the plane with no crossing edges. The **planar dual** is the graph G^* whose vertices are the regions into which G divides the plane, with two vertices of G^* joined by an edge e^* if the corresponding faces of G are separated by an edge e of G. (So e^* is drawn *across* e in the construction.)



Some facts to check about planar duality:

- $A \subset E$ is acyclic if and only if $E^* \setminus A^*$ is connected.
- $A \subset E$ is connected if and only if $E^* \setminus A^*$ is acyclic.
- G^{**} is naturally isomorphic to G.
- e is a loop (bridge) if and only if e^* is a bridge (loop).

Definition 2.29. Let M be a matroid on E. A **loop** is an element of E that does not belongs to any basis of M. A **coloop** is an element of E that belongs to every basis of M. An element of E that is neither a loop nor a coloop is called **ordinary**.

In a linear matroid, a loop is a copy of the zero vector, while a coloop is a vector that is not in the span of all the other vectors.

A **cocircuit** of M is by definition a circuit of the dual matroid M^*. Set-theoretically, a cocircuit is a minimal set not contained in any basis of M^*, so it is a minimal set that meets every basis of M. For a connected graph G, the cocircuits of the graphic matroid $M(G)$ are the *bonds* of G: the minimal edge sets K such that $G - K$ is not connected. A matroid can be described by its cocircuit system, which satisfy the same axioms as those for circuits (Definition 2.20).

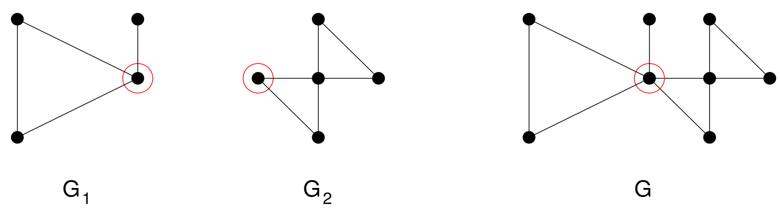
2.6.2. *Direct Sum.* Let E_1, E_2 be disjoint sets, and let $\mathscr{B}_i$ be a basis system for a matroid M_i on E_i. The **direct sum** $M_1 \oplus M_2$ is the matroid on $E_1 \cup E_2$ with basis system

$$\mathscr{B} = \{B_1 \cup B_2 \mid B_1 \in \mathscr{B}_1,\ B_2 \in \mathscr{B}_2\}.$$

(I'll omit the routine proof that this is a basis system.)

If M_1, M_2 are linear matroids whose ground sets span vector spaces V_1, V_2 respectively, then $M_1 \oplus M_2$ is the matroid you get by regarding the vectors as living in $V_1 \oplus V_2$: the linear relations have to come either from V_1 or from V_2.

If G_1, G_2 are graphs with disjoint vertex sets, then $M(G_1) \oplus M(G_2) \cong M(G_1 + G_2)$, where $+$ denotes the disjoint union. Actually, something more is true: you can identify any vertex of G_1 with any vertex of G_2 and still get a graph whose associated graphic matroid is $M(G_1) \oplus M(G_2)$ (such as G in the following figure).



Corollary: Every graphic matroid arises from a *connected* graph.

Direct sum is additive on rank functions: for $A_1 \subseteq E_1$, $A_2 \subseteq E_2$, we have

$$r_{M_1 \oplus M_2}(A_1 \cup A_2) = r_{M_1}(A_1) + r_{M_2}(A_2).$$

The geometric lattice of a direct sum is a (Cartesian) product of posets:

$$L(M_1 \oplus M_2) \cong L(M_1) \times L(M_2),$$

subject to the order relations $(F_1, F_2) \leq (F_1', F_2')$ iff $F_i \leq F_i'$ in $L(M_i)$ for each i.

As you should expect from an operation called "direct sum", and as the last two equations illustrate, pretty much all of the properties of $M_1 \oplus M_2$ can be deduced easily from those of its summands.

Definition 2.30. A matroid that cannot be written nontrivially as a direct sum of two smaller matroids is called **connected** or **indecomposable**.[7]

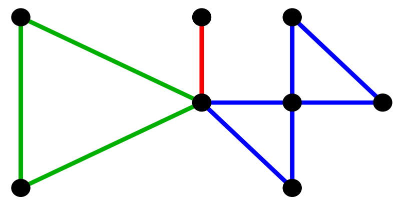
Proposition 2.31. *Let $G = (V, E)$ be a loopless graph. Then $M(G)$ is indecomposable if and only if G is* ***2-connected*** *— i.e., not only is it connected, but so is every subgraph obtained by deleting a single vertex.*

The "only if" direction is immediate: the discussion above implies that

$$M(G) = \bigoplus_H M(H)$$

where H ranges over all the **blocks** (maximal 2-connected subgraphs) of G.

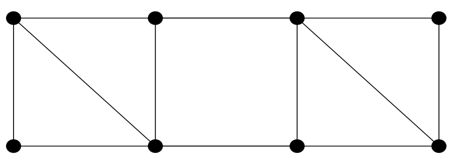
[7]The first term is more common among matroid theorists, but I prefer "indecomposable" to avoid potential confusion with the graph-theoretic meaning of "connected".

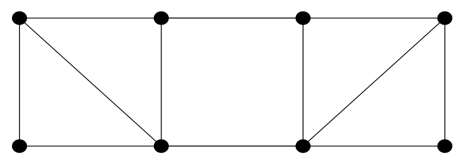


We'll prove the other direction later.

Remark: If $G \cong H$ as graphs, then clearly $M(G) \cong M(H)$. The converse is not true: if T is any tree (or even forest) on n vertices, then every set of edges is acyclic, so the independence complex is the Boolean algebra $\mathscr{B}_n$ (and, for that matter, so is the lattice of flats).

In light of Proposition 2.31, it is natural to suspect that every 2-connected graph is determined up to isomorphism by its graphic matroid, but even this is not true; the 2-connected graphs below are not isomorphic, but have isomorphic matroids.





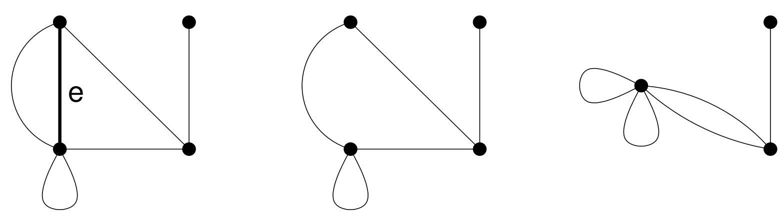
More on this later.

2.6.3. *Deletion and Contraction.*

Definition 2.32. Let M be a matroid on E with basis system $\mathscr{B}$, and let $e \in E$.

(1) If e is not a coloop, then the **deletion** of e is the matroid $M - e$ (or $M \setminus e$) on $E \setminus \{e\}$ with bases
$$\{B \mid B \in \mathscr{B},\ e \notin B\}.$$
(2) If e is not a loop, then the **contraction** of e is the matroid M/e (or $M : e$) on $E \setminus \{e\}$ with bases
$$\{B \setminus \{e\} \mid B \in \mathscr{B},\ e \in B\}.$$

Again, the terms come from graph theory. Deleting an edge of a graph means what you think it means, while contracting an edge means to throw it away and to glue its endpoints together.



G G – e G / e

Notice that contracting can cause some edges to become parallel, and can cause other edges (namely, those parallel to the edge you're contracting) to become loops. In matroid language, deleting an element from a simple matroid always yields a simple matroid, but the same is not true for contraction.

How about the linear setting? Let V be a vector space over a field $\mathbb{F}$, let $E \subset V$ be a set of vectors with linear matroid M, and let $e \in E$. Then $M - e$ is just the linear matroid on $E \setminus \{e\}$, while M/e is what you get by projecting $E \setminus \{e\}$ onto the quotient space $V/(\mathbb{F}e)$. (For example, if e is the i^{th} standard basis vector, then erase the i^{th} coordinate of every vector in $E \setminus \{e\}$.)

Deletion and contraction are interchanged by duality:

$$(M-e)^* \cong M^*/e \quad \text{and} \quad (M/e)^* \cong M^* - e. \tag{2.7}$$

Example 2.33. If M is the uniform matroid $U_k(n)$, then $M-e \cong U_k(n-1)$ and $M/e \cong U_{k-1}(n-1)$ for every ground set element e.

Many invariants of matroids can be expressed recursively in terms of deletion and contraction. The following fact is immediate from Definition 2.32.

Proposition 2.34. *Let M be a matroid on ground set E, and let $b(M)$ denote the number of bases of M. For every $e \in E$, we have*

$$b(M) = \begin{cases} b(M-e) & \text{if } e \text{ is a loop;} \\ b(M/e) & \text{if } e \text{ is a coloop;} \\ b(M-e) + b(M/e) & \text{otherwise.} \end{cases}$$

Example 2.35. If $M \cong U_k(n)$, then $b(M) = \binom{n}{k}$, and the recurrence of Proposition 2.34 is just the Pascal relation

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n}{k-1}.$$

Deletion and contraction can be described nicely in terms of the independence complex. If $\Delta = \mathscr{I}(M)$ is the independence complex of $M-e$ and M/e, then

$$\begin{aligned} \mathscr{I}(M-e) &= \operatorname{del}_\Delta(e) = \{\sigma \in \Delta \mid e \notin \sigma\}, \\ \mathscr{I}(M/e) &= \operatorname{lk}_\Delta(e) = \{\sigma \in \Delta \mid e \notin \sigma \text{ and } \sigma \cup e \in \Delta\}, \end{aligned}$$

subcomplexes known respectively as the *deletion* and *link* of e in Δ.

Every simplicial complex on vertex set V can be regarded as a topological space: imagine the vertices as the standard basis vectors in $\mathbb{R}^{|V|}$ and each face as the convex hull of its vertices (so a set of k vertices is realized as a $(k-1)$-dimensional simplex). This is called the *standard geometric realization* of Δ and should be denoted $|\Delta|$ or $[\Delta]$, but combinatorialists get used to ignoring this technical distinction. The reduced Euler characteristic of Δ is thus

$$\tilde{\chi}(\Delta) = \sum_{\sigma \in \Delta} (-1)^{\dim \sigma} = \sum_{\sigma \in \Delta} (-1)^{|\sigma|-1}. \tag{2.8}$$

This important topological invariant also satisfies a deletion-contraction recurrence. For any $e \in V$, we have

$$\begin{aligned} \tilde{\chi}(\Delta) &= \sum_{\sigma \in \Delta:\ e \notin \sigma} (-1)^{\dim \sigma} + \sum_{\sigma \in \Delta:\ e \in \sigma} (-1)^{\dim \sigma} \\ &= \sum_{\sigma \in \operatorname{del}_\Delta(e)} (-1)^{\dim \sigma} + \sum_{\tau \in \operatorname{lk}_\Delta(e)} (-1)^{1+\dim \tau} \\ &= \tilde{\chi}(\operatorname{del}_\Delta(e)) - \tilde{\chi}(\operatorname{lk}_\Delta(e)). \end{aligned} \tag{2.9}$$

These observations are the tip of an iceberg.

3. The Tutte Polynomial

The Tutte polynomial is a vitally important matroid invariant: it simultaneously encodes all invariants that can be obtained from a deletion-contraction recurrence.

3.1. **Definitions.**

Definition 3.1. Let M be a matroid with ground set E and let $e \in E$. The **Tutte polynomial** $T(M) = T(M;\, x, y)$ is computed recursively as follows:

- **(T1)** If $E = \emptyset$, then $T(M) = 1$.
- **(T2a)** If $e \in E$ is a loop, then $T(M) = y \cdot T(M - e)$.
- **(T2b)** If $e \in E$ is a coloop, then $T(M) = x \cdot T(M/e)$.
- **(T3)** If $e \in E$ is ordinary, then $T(M) = T(M - e) + T(M/e)$.

If $M = M(G)$ is a graphic matroid, we may write $T(G)$ instead of $T(M(G))$. Also, I will probably frequently slip up and write T_M instead of $T(M)$.

This is really an algorithm rather than a definition, and at this point, it is not even clear that $T(M)$ is well-defined, because the formula seems to depend on the order in which we pick elements to delete and contract. However, a miracle occurs: it doesn't! We will soon prove this by giving a closed formula for $T(M)$ that does not depend on any such choice.

In the case that M is a uniform matroid, then it *is* clear at this point that $T(M)$ is well-defined by Definition 3.1, because, up to isomorphism, $M - e$ and M/e are independent of the choices of $e \in E$.

Example 3.2. Suppose that $M \cong U_n(n)$, that is, every element of E is a coloop. By induction, $T(M)(x, y) = x^n$. Dually, if $M \cong U_0(n)$ (i.e., every element of E is a loop), then $T(M)(x, y) = y^n$.

Example 3.3. Let $M \cong U_1(2)$ (the graphic matroid of the "digon", two vertices joined by two parallel edges). Let $e \in E$; then

$$\begin{aligned} T(M) &= T(M - e) + T(M/e) \\ &= T(U_1(1)) + T(U_0(1)) \;=\; x + y. \end{aligned}$$

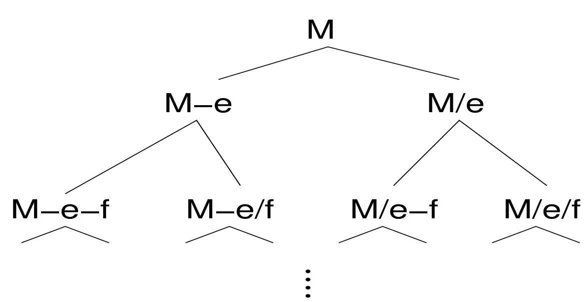
Example 3.4. Let $M \cong U_2(3)$ (the graphic matroid of K_3, as well as the matroid associated with the geometric lattice $\Pi_3 \cong M_5$). Applying Definition 3.1 for any $e \in E$ gives

$$T(U_2(3)) \;=\; T(U_2(2)) + T(U_1(2)) \;=\; x^2 + x + y.$$

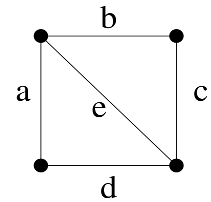
On the other hand,

$$T(U_1(3)) \;=\; T(U_1(2)) + T(U_0(2)) \;=\; x + y + y^2.$$

In general, we can represent a calculation of $T(M)$ by a binary tree in which moving down corresponds to deleting or contracting:



Example 3.5. Here is a non-uniform example. Let G be the graph below.



One possibility is to recurse on edge a (or equivalently on b, c, or d). When we delete a, the edge d becomes a coloop, and contracting it produces a copy of K_3. Therefore

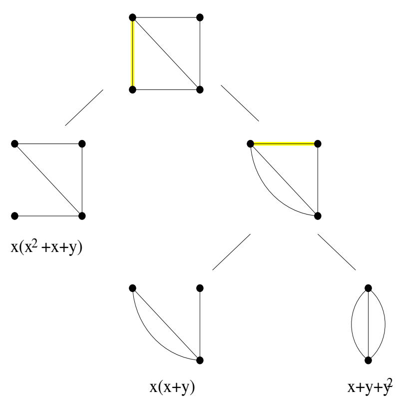
$$T(G-a) \;=\; x(x^2+x+y)$$

by Example 3.4. Next, apply the recurrence to the edge b in G/a. The graph $G/a-b$ has a coloop c, contracting which produces a digon. Meanwhile, $M(G/a/b) \cong U_1(3)$. Therefore

$$T(G/a-b) \;=\; x(x+y) \quad \text{and} \quad T(G/a/b) \;=\; x+y+y^2.$$

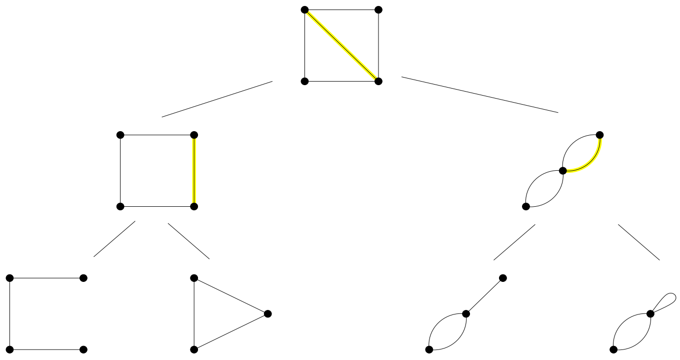
Putting it all together, we get

$$\begin{aligned} T(G) \;&=\; T(G-a)+T(G/a) \\ &=\; T(G-a)+T(G/a-b)+T(G/a/b) \\ &=\; x(x^2+x+y) \;+\; x(x+y) \;+\; (x+y+y^2) \\ &=\; x^3+2x^2+2xy+x+y+y^2. \end{aligned}$$



On the other hand, we could have recursed first on e, getting

$$\begin{aligned} T(G) \;&=\; T(G-e)+T(G/e) \\ &=\; T(G-e-c)+T(G-e/c)+T(G/e-c)+T(G/e/c) \\ &=\; x^3 \;+\; (x^2+x+y) \;+\; x(x+y) \;+\; y(x+y) \\ &=\; x^3+2x^2+2xy+x+y+y^2. \end{aligned}$$



We can actually see the usefulness of $T(M)$ even before proving that it is well-defined!

Proposition 3.6. *$T(M;\,1,1)$ equals the number of bases of M.*

Proof. Let $b(M) = T(M;\,1,1)$. Then Definition 3.1 gives

$$b(M) = \begin{cases} 1 & \text{if } E = \emptyset, \\ b_{G/e} & \text{if } e \text{ is a coloop} \\ b_{G-e} & \text{if } e \text{ is a loop} \\ b(M-e) + b(M/e) & \text{otherwise} \end{cases}$$

which is identical to the recurrence for $|\mathscr{B}(M)|$ observed in Proposition 2.34. □

Many other invariants of M can be found in this way by making appropriate substitutions for the indeterminates x, y in $T(M)$. In particular, if we let $c(M) = T(M;\,2,2)$, then

$$c(M) = \begin{cases} 1 & \text{if } E = \emptyset, \\ 2c_{G/e} & \text{if } e \text{ is a coloop} \\ 2c_{G-e} & \text{if } e \text{ is a loop} \\ c(M-e) + c(M/e) & \text{otherwise} \end{cases}$$

so $c(M) = 2^{|E|}$. This suggests that $T(M)$ ought to have a closed formula as a sum over subsets $A \subseteq E$, with each summand becoming 1 upon setting $x = 1$ and $y = 1$—for example, with each summand a product of powers of $x-1$ and $y-1$. In fact, this is the case.

Theorem 3.7. *Let r be the rank function of the matroid M. Then*

$$T(M;\,x,y) \;=\; \sum_{A \subseteq E} (x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)}. \tag{3.1}$$

Proof coming shortly. The quantity $r(E) - r(A)$ is the **corank** of A; it is the minimum number of elements one needs to add to A to obtain a spanning set of M. Meanwhile, $|A| - r(A)$ is the **nullity** of A: the minimum number of elements one needs to remove from A to obtain an acyclic set. Accordingly, (3.1) is referred to as the the **corank-nullity generating function**.

Proof of Theorem 3.7. Let M be a matroid on ground set E. The definitions of rank function, deletion, and contraction imply that for any $e \in E$ and $A \subseteq E \setminus \{e\}$:

(1) If e is not a coloop, then $r_{M-e}(A) = r_M(A)$.
(2) If e is not a loop, then $r_{M/e}(A) = r_M(A \cup e) - 1$.

Let $\tilde{T}(M) = \tilde{T}(M;\, x, y)$ denote the generating function on the right-hand side of (3.1). We will prove by induction on $n = |E|$ that $\tilde{T}(M)$ obeys the recurrence of Definition 3.1 for every ground set element e, hence equals $T(M)$. Let r' and r'' denote the rank functions on $M - e$ and M/e respectively. To save space, we will also set

$$X = x - 1, \quad Y = y - 1.$$

For **(T1)**, if $E = \emptyset$, then (3.1) gives $\tilde{T}(M) = 1 = T(M)$.

For **(T2a)**, let e be a loop. Then

$$\begin{aligned}
\tilde{T}(M) &= \sum_{A \subseteq E} X^{r(E)-r(A)} Y^{|A|-r(A)} \\
&= \sum_{A \subseteq E:\ e \notin A} X^{r(E)-r(A)} Y^{|A|-r(A)} + \sum_{B \subseteq E:\ e \in B} X^{r(E)-r(B)} Y^{|A|-r(B)} \\
&= \sum_{A \subseteq E \backslash e} X^{r'(E\backslash e)-r'(A)} Y^{|A|-r'(A)} + \sum_{A \subseteq E \backslash e} X^{r'(E\backslash e)-r'(A)} Y^{|A|+1-r'(A)} \\
&= (1+Y) \sum_{A \subseteq E \backslash e} X^{r'(E\backslash e)-r'(A)} Y^{|A|-r'(A)} \qquad = y\tilde{T}(M-e).
\end{aligned}$$

For **(T2b)**, let e be a coloop. Then

$$\begin{aligned}
\tilde{T}(M) &= \sum_{A \subseteq E} X^{r(E)-r(A)} Y^{|A|-r(A)} \\
&= \sum_{e \notin A \subseteq E} X^{r(E)-r(A)} Y^{|A|-r(A)} + \sum_{e \in B \subseteq E} X^{r(E)-r(B)} Y^{|B|-r(B)} \\
&= \sum_{A \subseteq E \backslash e} X^{(r''(E\backslash e)+1)-r''(A)} Y^{|A|-r''(A)} + \sum_{A \subseteq E \backslash e} X^{(r''(E\backslash e)+1)-(r''(A)+1)} Y^{|A|+1-(r''(A)+1)} \\
&= \sum_{A \subseteq E \backslash e} X^{r''(E\backslash e)+1-r''(A)} Y^{|A|-r''(A)} + \sum_{A \subseteq E \backslash e} X^{r''(E\backslash e)-r''(A)} Y^{|A|-r''(A)} \\
&= (X+1) \sum_{A \subseteq E \backslash e} X^{r''(E\backslash e)-r''(A)} Y^{|A|-r''(A)} \qquad = x\tilde{T}(M/e).
\end{aligned}$$

For **(T3)**, suppose that e is ordinary. Then

$$\begin{aligned}
\tilde{T}(M) &= \sum_{A \subseteq E} X^{r(E)-r(A)} Y^{|A|-r(A)} \\
&= \sum_{A \subseteq E \backslash e} \left[X^{r(E)-r(A)} Y^{|A|-r(A)}\right] + \left[X^{r(E)-r(A\cup e)} Y^{|A\cup e|-r(A\cup e)}\right] \\
&= \sum_{A \subseteq E \backslash e} \left[X^{r'(E\backslash e)-r'(A)} Y^{|A|-r'(A)}\right] + \left[X^{(r''(E)+1)-(r''(A)+1)} Y^{|A|+1-(r''(A)-1)}\right] \\
&= \sum_{A \subseteq E \backslash e} X^{r'(E\backslash e)-r'(A)} Y^{|A|-r'(A)} + \sum_{A \subseteq E \backslash e} X^{r''(E\backslash e)-r''(A)} Y^{|A|-r''(A)} \qquad = \tilde{T}(M-e) + \tilde{T}(M/e). \quad \square
\end{aligned}$$

Example 3.8. For the graph G of Example 3.5, the formula (3.1) gives

A	$\lvert A\rvert$	$r(A)$	$3-r(A)$	$\lvert A\rvert-r(A)$	contribution to (3.1)	
1 empty set	0	0	3	0	$1(x-1)^3(y-1)^0$	$=x^3-3x^2+3x-1$
5 singletons	1	1	2	0	$5(x-1)^2(y-1)^0$	$=5x^2-10x+5$
10 doubletons	2	2	1	0	$10(x-1)^1(y-1)^0$	$=10x-10$
2 triangles	3	2	1	1	$2(x-1)^1(y-1)^1$	$=2xy-2x-2y+2$
8 spanning trees	3	3	0	0	$8(x-1)^0(y-1)^0$	$=8$
5 quadrupletons	4	3	0	1	$5(x-1)^0(y-1)^1$	$=5y-5$
1 whole	5	3	0	2	$1(x-1)^0(y-1)^2$	$=y^2-2y+1$
Total	$x^3+2x^2+x+2xy+y^2+y$					

which is the same result as deletion-contraction.

As consequences of Theorem 3.7, we can obtain several invariants of a matroid easily from its Tutte polynomial.

Corollary 3.9. *For every matroid M, we have*

(1) $T(M;\,0,0)$ *= 1 if $E=\emptyset$, 0 otherwise;*
(2) $T(M;\,1,1)$ *= number of bases of M;*
(3) $T(M;\,2,2) = |E|$*;*
(4) $T(M;\,2,1)$ *= number of independent sets of M;*
(5) $T(M;\,1,2)$ *= number of spanning sets of M.*

Proof. For (1), plugging in $x=y=1$ gives $(-1)^{r(E)}\sum_A(-1)^{|A|}$. If $E\neq\emptyset$ then the sum vanishes; otherwise we get $(-1)^0(-1)^0=1$. We've already proved (2) and (3), but they also follow from the corank-nullity generating function. Plugging in $x=2,y=2$ will change every summand to 1. Plugging in $x=1$ and $y=1$ will change every summand to 0, *except* for those sets A that have corank and nullity both equal to 0 — that is, those sets that are both spanning and independent. The verifications of (4) and (5) are analogous. $\square$

A little more generally, we can use the Tutte polynomial to enumerate independent and spanning sets by their cardinality:

$$\sum_{A\subseteq E\text{ independent}} q^{|A|} = q^{r(M)}T(1/q+1,1); \tag{3.2}$$

$$\sum_{A\subseteq E\text{ spanning}} q^{|A|} = q^{r(M)}T(1,1/q+1). \tag{3.3}$$

I In particular,

$$T(0,1) = (-1)^{r(M)}\sum_{A\subseteq E\text{ independent}}(-1)^{|A|} = -(-1)^{r(M)}\tilde{\chi}(\mathscr{I}(\mathcal{M})),$$

the reduced Euler characteristic of the independence complex of M (see (2.8) and (2.9)).

Another easy fact is that $T(M)$ is multiplicative on direct sums:

$$T(M_1\oplus M_2)=T(M_1)T(M_2).$$

Moreover, duality interchanges x and y, i.e.,

$$T(M;\,x,y)=T(M^*;\,y,x). \tag{3.4}$$

This can be deduced either from the deletion-contraction recurrence (since duality interchanges deletion and contraction; see (2.7)) or from the corank-nullity generating function (by expressing r^* in terms of r).

The recursive Definition 3.1 implies that every coefficient of T_M is a nonnegative integer (i.e., $T(M;x,y)\in\mathbb{N}[x,y]$), a property which is not at all obvious from the closed formula (3.1).

3.2. **The Chromatic Polynomial.**

Definition 3.10. Let $G = (V, E)$ be a graph. A *k-coloring* of G is a function $f : V \to [k]$; the coloring is *proper* if $f(v) \neq f(w)$ whenever vertices v and w are adjacent.

The function
$$p(G;\ k) = \text{number of proper } k\text{-colorings of } G$$
is called the *chromatic polynomial* of G. Technically, at this point, we should call it the "chromatic function." But in fact one of the first things we will prove is that $p(G;\ k)$ is a polynomial function of k for every graph G.

Example 3.11. If G has a loop, then its endpoints automatically have the same color, so it's impossible to color G properly and $p(G;\ k) = 0$.

If $G = K_n$, then all vertices must have different colors. There are k choices for $f(1)$, $k-1$ choices for $f(2)$, etc., so $p(K_n;\ k) = k(k-1)(k-2)\cdots(k-n+1)$.

At the other extreme, let $G = \overline{K_n}$, the graph with n vertices and no edges. Then $p(\overline{K_n};\ k) = k^n$.

If T_n is a tree with n vertices, then pick any vertex as the root; this imposes a partial order on the vertices in which the root is $\hat{1}$ and each non-root vertex v is covered by exactly one other vertex $p(v)$ (its "parent"). There are k choices for the color of the root, and once we know $f(p(v))$ there are $k-1$ choices for $f(v)$. Therefore $p(T_n;\ k) = k(k-1)^{n-1}$.

If G has connected components $G_1, \dots, G_s$, then $p(G;\ k) = \prod_{i=1}^{s} p(G_i;\ k)$. Equivalently, $p(G + H;\ k) = p(G;\ k)p(H;\ k)$, where $+$ denotes disjoint union of graphs.

Theorem 3.12. *For every graph G we have*
$$p(G;\ k) \ = \ (-1)^{n-c}k^c \cdot T(G,\ 1-k, 0)$$
where n is the number of vertices of G and c is the number of components. In particular, $p(G; k)$ is a polynomial function of k.

Proof. First, we show that the chromatic function satisfies the recurrence

$$p(G;\ k) = k^n \qquad \text{if } E = \emptyset; \tag{3.5}$$
$$p(G;\ k) = 0 \qquad \text{if } G \text{ has a loop;} \tag{3.6}$$
$$p(G;\ k) = (k-1)p(G/e;\ k) \qquad \text{if } e \text{ is a coloop;} \tag{3.7}$$
$$p(G;\ k) = p(G-e;\ k) - p(G/e;\ k) \qquad \text{otherwise.} \tag{3.8}$$

If $E = \emptyset$ then every one of the k^n colorings of G is proper, and if G has a loop then it has no proper colorings, so (3.5) and (3.6) are easy.

Suppose $e = xy$ is not a loop. Let f be a proper k-coloring of $G - e$. If $f(x) = f(y)$, then we can identify x and y to obtain a proper k-coloring of G/e. If $f(x) \neq f(y)$, then f is a proper k-coloring of G. So (3.8) follows.

This argument applies even if e is a coloop. In that case, however, the component H of G containing e becomes two components H' and H'' of $G - e$, whose colorings can be chosen independently of each other. So the probability that $f(x) = f(y)$ in any proper coloring is $1/k$, implying (3.7).

(A corollary, by induction on $|V|$, is that $p(G;\ k)$ is a polynomial in k, and thus has the right to be called the *chromatic polynomial* of G.)

The graph $G - e$ has n vertices and either $c+1$ or c components, according as e is or is not a coloop. Meanwhile, G/e has $n-1$ vertices and c components. By the recursive definition of the Tutte polynomial

$$\begin{aligned}
\tilde{p}(G;\,k) &= (-1)^{n-c}k^c T(G,\,1-k,0)\\
&= \begin{cases} k^n & \text{if } E=\emptyset,\\ 0 & \text{if } e \text{ is a loop,}\\ (1-k)(-1)^{n+1-c}k^c T(G/e,\,1-k,0) & \text{if } e \text{ is a coloop,}\\ (-1)^{n-c}k^c\left(T(G-e,\,1-k,0)+T(G/e,\,1-k,0)\right) & \text{otherwise}\end{cases}\\
&= \begin{cases} k^n & \text{if } E=\emptyset,\\ 0 & \text{if } e \text{ is a loop,}\\ (k-1)p(G/e;\,k) & \text{if } e \text{ is a coloop,}\\ p(G-e;\,k)-p(G/e;\,k) & \text{otherwise}\end{cases}
\end{aligned}$$

which is exactly the recurrence satisfied by the chromatic polynomial. This proves the theorem. $\square$

This result raises the question of what this specialization of $T(M)$ means in the case that M is a an arbitrary (not necessarily graphic) matroid. Stay tuned!

3.3. **Acyclic Orientations.** An *orientation* D of a graph $G=(V,E)$ is an assignment of a direction to each edge $xy \in E$ (either $\vec{xy}$ or $\vec{yx}$). A *directed cycle* is a sequence $(x_0, x_1, \ldots, x_{n-1})$ of vertices such that $\vec{x_i x_{i+1}}$ is a directed edge for every i. (Here the indices are taken modulo n.)

An orientation is *acyclic* if it has no directed cycles. Let $A(G)$ be the set of acyclic orientations of G, and let $a(G) = |A(G)|$.

Theorem 3.13 (Stanley 1973)**.** *For every graph G on n vertices, we have*

$$a(G) = T(G;\,2,0) = (-1)^{n-1}p(G;\,-1).$$

Stanley actually proved a stronger result that gives a combinatorial interpretation of $p(G,k)$ for every negative integer k. This is the prototyical *combinatorial reciprocity theorem*; see, e.g., the forthcoming book by Beck and Sanyal.

Proof. The second equality is a consequence of Theorem 3.12. Plugging $x=2$ and $y=0$ into the Definition of the Tutte polynomial, we obtain the recurrence we need to establish in order to prove the first equality:

- **(A1)** If $E=\emptyset$, then $a(G)=1$.
- **(A2a)** If $e \in E$ is a loop, then $a(G)=0$.
- **(A2b)** If $e \in E$ is a coloop, then $a(G)=2a(G/e)$.
- **(A3)** If $e \in E$ is ordinary, then $a(G)=a(G-e)+a(G/e)$.

(A1) holds because the number of orientations of G is $2^{|V|}$, and any orientation of a forest (in particular, an edgelesss graph) is acyclic.

For **(A2a)**, note that if G has a loop then it cannot possibly have an acyclic orientation.

If G has a coloop e, then e doesn't belong to any cycle of G, so any acyclic orientation of G/e can be extended to an acyclic orientation of G by orienting e in either direction, proving **(A2b)**.

The trickiest part is **(A3)**. Fix an edge $e = xy \in E(G)$. For each orientation D of G, let $\tilde{D}$ be the orientation produced by reversing the direction of e, and let

$$A_1 = \{D \in A(G) \mid \tilde{D} \in A(G)\},$$
$$A_2 = \{D \in A(G) \mid \tilde{D} \notin A(G)\}.$$

Clearly $a(G) = |A_1| + |A_2|$.

Let D be an acyclic orientation of $G - e$. If D has a path from x to y (for short, an "x, y-path") then it cannot have a y, x-path, so directing e as $\vec{xy}$ (but not $e = \vec{yx}$) produces an acyclic orientation of G; all this is true if we reverse the roles of x and y. We get every orientation in A_2 in this way. On the other hand, if D does not have either an x, y-path or a y, x-path, then we can orient e in either direction to produce an orientation in A_1. Therefore

$$a(G - e) \;=\; \frac{1}{2}|A_1| + |A_2|. \tag{3.9}$$

Now let D be an acyclic orientation of G/e, and let $\hat{D}$ be the corresponding acyclic orientation of $G - e$. I claim that $\hat{D}$ can be extended to an acyclic orientation of G by orienting e in either way. Indeed, if it were impossible to orient e as $\vec{xy}$, then the reason would have to be that $\hat{D}$ contained a path from y to x, but y and x are the same vertex in D and D wouldn't be acyclic. Therefore, there is a bijection between $A(G/e)$ and matched pairs $\{D, \tilde{D}\}$ in $A(G)$, so

$$a(G/e) \;=\; \frac{1}{2}|A_1|. \tag{3.10}$$

Now combining (3.9) and (3.10) proves **(A3)**. $\square$

Some other related graph-theoretic invariants you can find from the Tutte polynomial:

- The number of *totally cyclic orientations*, i.e., orientations in which every edge belongs to a directed cycle (HW problem).
- The *flow polynomial* of G, whose value at k is the number of edge-labelings $f : E \to [k-1]$ such that the sum at every vertex is 0 mod k.
- The *reliability polynomial* $f(p)$: the probability that the graph remains connected if each edge is deleted with independent probability p.
- The "enhanced chromatic polynomial", which enumerates all q-colorings by "improper edges":
$$\tilde{\chi}(q, t) = \sum_{f: V \to [q]} t^{\#\{xy \in E \;\mid\; f(x) = f(y)\}}.$$
  This is essentially Crapo's *coboundary polynomial*, and provides the same information as the Tutte polynomial.
- And more; the canonical source for all things Tutte is T. Brylawski and J. Oxley, "The Tutte polynomial and its applications," Chapter 6 of *Matroid applications*, N. White, editor (Cambridge Univ. Press, 1992).

3.4. **Basis Activities.** We know that $T(M; x, y)$ has nonnegative integer coefficients and that $T(M; 1, 1)$ is the number of bases of M. These observations suggest that we should be able to interpret the Tutte polynomial as a generating function for bases: that is, there should be combinatorially defined functions $i, e : \mathscr{B}(M) \to \mathbb{N}$ such that

$$T(M; x, y) \;=\; \sum_{B \in \mathscr{B}(M)} x^{i(B)} y^{e(B)}. \tag{3.11}$$

In fact, this is the case. The tricky part is that $i(B)$ and $e(B)$ must be defined with respect to a total order on the ground set E, so they are not really invariants of B itself. However, another miracle occurs: the right-hand side of (3.11) does not depend on this choice of total order.

Index the ground set of E as $\{e_1, \ldots, e_n\}$, and totally order the elements of E by their subscripts.

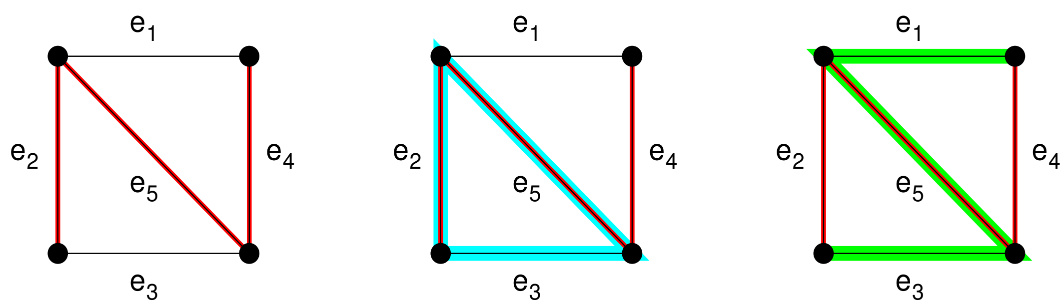
Definition 3.14. Let B be a basis of M.

- Let $e_i \in B$. The **fundamental cocircuit** $C^*(e_i, B)$ is the unique cocircuit in $(E \setminus B) \cup e_i$. That is,
$$C^*(e_i, B) = \{e_j \mid B \setminus e_i \cup e_j \in \mathscr{B}\}.$$
We say that e_i is **internally active** with respect to B if e_i is the minimal element of $C(e_i, B)$.

- Let $e_i \notin B$. The **fundamental circuit** $C(e_i, B)$ is the unique circuit in $B \cup e_i$. That is,
$$C(e_i, B) = \{e_j \mid B \setminus e_j \cup e_i \in \mathscr{B}\}.$$
We say that e_i is **externally active** with respect to B if e_i is the minimal element of $C(e_i, B)$.

- Finally, we let $e(B)$ and $i(B)$ denote respectively the number of externally active and internally active elements with respect to B.

Here's an example. Let G be the graph with edges labeled as shown below, and let B be the spanning tree $\{e_2, e_4, e_5\}$ shown in red. The middle figure shows $C(e_1, B)$, and the right-hand figure shows $C^*(e_5, B)$.



Then

$$\begin{aligned}
C(e_1, B) &= \{e_1, e_4, e_5\} && \text{so } e_1 \text{ is externally active;}\\
C(e_3, B) &= \{e_2, e_3, e_5\} && \text{so } e_3 \text{ is not externally active;}\\
C^*(e_2, B) &= \{e_2, e_3\} && \text{so } e_1 \text{ is internally active;}\\
C^*(e_4, B) &= \{e_1, e_4\} && \text{so } e_3 \text{ is not internally active;}\\
C^*(e_5, B) &= \{e_1, e_3, e_5\} && \text{so } e_3 \text{ is not internally active.}
\end{aligned}$$

Theorem 3.15. *Let M be a matroid on E. Fix a total ordering of E and let $e(B)$ and $i(B)$ denote respectively the number of externally active and internally active elements with respect to B. Then* (3.11) *holds.*

Thus, in the example above, the spanning tree B would contribute the monomial $xy = x^1y^1$ to $T(G; x, y)$.

The proof is omitted; it requires careful bookkeeping but is not hard. It's a matter of showing that the generating function on the right-hand side of (3.11) satisfies a deletion-contraction recurrence. Note in particular that if e is a loop (resp. coloop), then $e \notin B$ (resp. $e \in B$) for every basis B, and $C(e, B) = \{e\}$ (resp. $C^*(e, B) = \{e\}$), so e is externally (resp. internally) active with respect to B, so the activity generating function (3.14) is divisible by y (resp. x).

3.5. **The Tutte polynomial and linear codes.**

Definition 3.16. A **linear code** $\mathscr{C}$ is a subspace of $(Ff_q)n$, where q is a prime power and $\mathbb{F}_q$ is the field of order q. The number n is the **length** of $\mathscr{C}$. The elements $c = (c_1, \dots, c_n) \in \mathscr{C}$ are called **codewords**. The **support** of a codeword is $\operatorname{supp}(c) = \{i \in [n] \mid c_i \neq 0\}$, and its **weight** is $\operatorname{wt}(c) = |\operatorname{supp}(c)|$.

The **weight enumerator** of $\mathscr{C}$ is the polynomial

$$W_{\mathscr{C}}(t) = \sum_{c \in \mathscr{C}} t^{\operatorname{wt}(c)}.$$

For example, let $\mathscr{C}$ be the subspace of $(\mathbb{F}_2)^3$ generated by the rows of the matrix

$$X = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \in (\mathbb{F}_2)^{3\times 2}.$$

So $\mathscr{C} = \{000, 101, 011, 110\}$, and $W_{\mathscr{C}}(t) = 1 + 3t^2$.

The **dual code** $\mathscr{C}^\perp$ is the orthogonal complement under the standard inner product. This inner product is nondegenerate, i.e., $\dim \mathscr{C}^\perp = n - \dim \mathscr{C}$. (On the other hand, a subspace and its orthogonal complement can intersect nontrivially — this does not happen over $\mathbb{R}$, where the inner product is not only nondegenerate but also positive-definite; "positive" does not make sense over a finite field.) In this case, $\mathscr{C}^\perp = \{000, 111\}$ and $W_{\mathscr{C}^\perp}(t) = 1 + 3t^2$.

Theorem 3.17 (Curtis Greene, 1976)**.** *Let $\mathscr{C}$ be a linear code of length n and dimension r over $\mathbb{F}_q$, and let M be the matroid represented by the columns of a matrix X whose rows are a basis for $\mathscr{C}$. Then*

$$W_{\mathscr{C}}(t) = t^{n-r}(1-t)^r T\left(M;\ \frac{1+(q-1)t}{1-t},\ \frac{1}{t}\right)$$

The proof is again a deletion-contraction argument and is omitted. As an example, if $\mathscr{C}$ and $\mathscr{C}^\perp$ are as above (so $q = 2$), then the matroid M is $U_2(3)$. Its Tutte polynomial is $x^2x + y$, and Greene's theorem gives

$$\begin{aligned}
W_{\mathscr{C}}(t) &= t(1-t)^2 T\left(M;\ \frac{1+t}{1-t},\ \frac{1}{t}\right) \\
&= t(1-t)^2 \left(\frac{1+t}{1-t}\right)^2 + \left(\frac{1+t}{1-t}\right) + \frac{1}{t} \\
&= t(1+t)^2 + t(1+t)(1-t) + (1-t)^2 \\
&= (t + 2t^2 + t^3) + (t - t^3) + (1 - 2t + t^2) \\
&= 1 + 3t^2.
\end{aligned}$$

If $X^\perp$ is a matrix whose rows are a basis for the dual code, then the corresponding matroid $M^\perp$ is precisely the dual matroid to M. We know that $T(M; x, y) = T(M^\perp; y, x)$ (see (3.4)), so setting $s = (1-t)/(1+(q-1)t)$ (so $t = (1-s)/(1+(q-1)s)$; isn't that convenient?) gives

$$\begin{aligned}
W_{\mathscr{C}^\perp}(t) &= t^r(1-t)^{n-r} T\left(M;\ \frac{1+(q-1)s}{1-s},\ \frac{1}{s}\right) \\
&= t^r(1-t)^{n-r} s^{r-n}(1-s)^{-r} W_{\mathscr{C}}(s),
\end{aligned}$$

or rewriting in terms of t,

$$W_{\mathscr{C}^\perp}(t) = \frac{1+(q-1)t^n}{q^r} W_{\mathscr{C}}\left(\frac{1-t}{1+(q-1)t}\right)$$

which is known as the *MacWilliams identity* and is important in coding theory.

4. Poset Algebra

4.1. **The Incidence Algebra of a Poset.** Many enumerative properties of posets can be expressed in terms of a ring called the *incidence algebra* of the poset. This looks weird and abstract at first, but actually it's an extremely convenient framework to work with once you get used to it.

Let P be a poset and let $\mathrm{Int}(P)$ denote the set of intervals of P, i.e., the sets

$$[x, y] := \{z \in P \mid x \leq z \leq y\}.$$

Definition 4.1. Suppose that P is **locally finite**, i.e., every interval is finite. The **incidence algebra** $I(P)$ is the set of functions $f : \mathrm{Int}(P) \to \mathbb{C}$. I'll abbreviate $f([x, y])$ by $f(x, y)$. For convenience, we set $f(x, y) = 0$ if $x \not\leq y$. This is a $\mathbb{C}$-vector space with pointwise addition, subtraction and scalar multiplication. It can be made into an associative algebra by the *convolution product*:

$$(f * g)(x, y) = \sum_{z \in [x,y]} f(x, z)g(z, y).$$

Note that the locally-finite assumption is necessary and sufficient for convolution to be well-defined.

Convolution is not commutative, but it is associative:

$$\begin{aligned}
[(f * g) * h](x, y) &= \sum_{z \in [x,y]} (f * g)(x, z) \cdot h(z, y) \\
&= \sum_{z \in [x,y]} \left(\sum_{w \in [x,z]} f(x, w)g(w, z) \right) h(z, y) \\
&= \sum_{w \in [x,y]} f(x, w) \left(\sum_{z \in [w,y]} g(w, z)h(z, y) \right) \\
&= \sum_{w \in [x,y]} f(x, w) \cdot (g * h)(w, y) \\
&= [f * (g * h)](x, y).
\end{aligned}$$

The identity element of $I(P)$ is the Kronecker delta function:

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$$

Therefore, we might just write 1 for δ.

Proposition 4.2. *A function $f \in I(P)$ has a left/right/two-sided convolution inverse if and only if $f(x, x) \neq 0$ for all x.*

Proof. Let g be a convolution inverse of f. I.e., $f(x, x) = g(x, x)^{-1}$ for all x (so the condition is necessary) and $\sum_{z:\ z \in [x,y]} g(x, z)f(z, y) = 0$; in particular $g(x, z)f(z, z) = -\sum_{z:\ x \leq z < y} g(x, z)f(z, y)$, so

$$g(x, z) = -f(z, z)^{-1} \sum_{z:\ x \leq z < y} g(x, z)f(z, y).$$

This can be used as a formula for the convolution inverse of f. □

The *zeta function* and *eta function* of P are defined as

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{if } x \not\leq y, \end{cases} \qquad \eta(x, y) = \begin{cases} 1 & \text{if } x < y, \\ 0 & \text{if } x \not< y, \end{cases}$$

i.e., $\eta = \zeta - 1$.

These trivial-looking incidence functions are useful because their convolution powers count important things: multichains and chains in P. Specifically,

$$\begin{aligned}
\zeta^2(x,y) &= \sum_{z\in[x,y]} \zeta(x,z)\zeta(z,y) = \sum_{z\in[x,y]} 1 \\
&= \Big|\{z:\ x \leq z \leq y\}\Big|; \\
\zeta^3(x,y) &= \sum_{z\in[x,y]} \sum_{w\in[z,y]} \zeta(x,z)\zeta(z,w)\zeta(w,y) = \sum_{x\leq z\leq w\leq y} 1 \\
&= \Big|\{z,w:\ x \leq z \leq w \leq y\}\Big|; \\
\zeta^k(x,y) &= \Big|\{x_1,\dots,x_{k-1}:\ x \leq x_1 \leq x_2 \leq \cdots \leq x_{k-1} \leq y\}\Big|.
\end{aligned}$$

That is, $\zeta^k(x,y)$ counts the number of **multichains** of length k between x and y. If we replace ζ with η, then the calculations go the same way, except that all the $\leq$'s are replaced with $<$'s, and we get

$$\eta^k(x,y) = \Big|\{x_1,\dots,x_{k-1}:\ x < x_1 < x_2 < \cdots < x_{k-1} < y\}\Big|,$$

the number of chains of length k between x and y. In particular, if the chains of P are bounded in length, then $\eta^n = 0$ for $n \gg 0$.

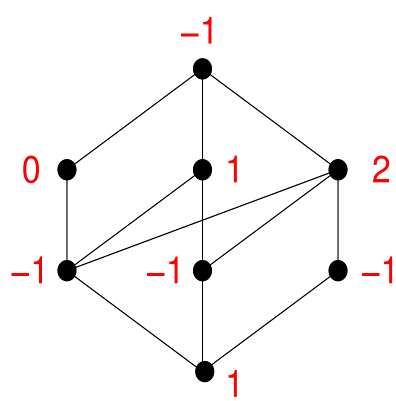
4.2. **The Möbius Function.** Let P be a poset. We are going to define an incidence function $\mu = \mu_P \in I(P)$, called the *Möbius function* of P. The definition is recursive:

(1) $\mu_P(x,x) = 1$ for all $x \in P$.
(2) If $x \not\leq y$, then $\mu_P(x,y) = 0$.
(3) If $x < y$, then $\mu_P(x,y) = -\sum_{z:\ x\leq z<y} \mu_P(x,z)$.

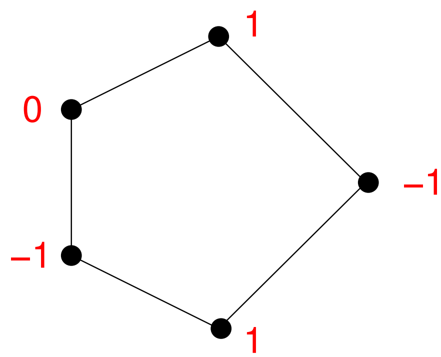
Thus the Möbius function is the unique function in $I(P)$ satisfying the equations

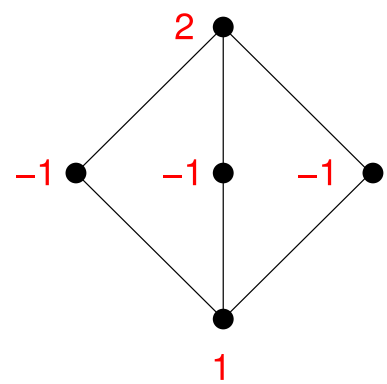
$$\sum_{z:\ x\leq z\leq y} \mu_P(x,z) = \delta(x,y).$$

Example 4.3. In the diagram of the following poset P, the red numbers indicate $\mu_P(\hat{0},x)$.



Example 4.4. In these diagrams of the posets M_5 and N_5, the red numbers indicate $\mu_P(\hat{0},x)$.





Example 4.5. *The Möbius function of a boolean algebra.* Let $\mathscr{B}_n$ be the boolean algebra of rank n and let $A \in \mathscr{B}_n$. Then $\mu(\hat{0}, A) = (-1)^{|A|}$. To prove this, induct on $|A|$. The case $|A| = 0$ is clear. For $|A| > 0$, we have

$$\begin{aligned}
\mu(\hat{0}, A) = -\sum_{B \subsetneq A} (-1)^{|B|} &= -\sum_{k=0}^{|A|-1} (-1)^k \binom{|A|}{k} \quad \text{(by induction)} \\
&= (-1)^{|A|} - \sum_{k=0}^{|A|} (-1)^k \binom{|A|}{k} \\
&= (-1)^{|A|} - (1-1)^{|A|} \;=\; (-1)^{|A|}.
\end{aligned}$$

More generally, if $B \subseteq A$, then $\mu(B, A) = (-1)^{|B \setminus A|}$, because every interval of $\mathscr{B}_n$ is a Boolean algebra.

Even more generally, suppose that P is a product of n chains of lengths $a_1, \ldots, a_n$. That is,

$$P = \{x = (x_1, \ldots, x_n) \mid 0 \leq x_i \leq a_i \text{ for all } i \in [n]\},$$

ordered by $x \leq y$ iff $x_i \leq y_i$ for all i. Then

$$\mu(\hat{0}, x) = \begin{cases} 0 & \text{if } x_i \geq 2 \text{ for at least one } i; \\ (-1)^s & \text{if } x \text{ consists of } s \text{ 1's and } n-s \text{ 0's.} \end{cases}$$

(The Boolean algebra is the special case that $a_i = 2$ for every i.) This conforms to the definition of Möbius function that you saw in Math 724. This formula is sufficient to calculate $\mu(y, x)$ for all $x, y \in P$, because every interval $[y, \hat{1}] \subset P$ is also a product of chains.

Example 4.6. *The Möbius function of the subspace lattice.* Let $L = L_n(q)$ be the lattice of subspaces of $\mathbb{F}_q^n$. Note that if $X \subset Y \subset \mathbb{F}_q^n$ with $\dim Y - \dim X = m$, then $[X, Y] \cong L_m(q)$. Therefore, it suffices to calculate

$$f(q, n) := \mu(0, \mathbb{F}_q^n).$$

Let $g_q(k, n)$ be the number of k-dimensional subspaces of $\mathbb{F}_q^n$. Clearly $g(q, 1) = -1$.

If $n = 2$, then $g_q(1, 2) = \dfrac{q^2 - 1}{q - 1} = q + 1$, so $f(q, 2) = -1 + (q + 1) = \boxed{q.}$

If $n=3$, then $g_q(1,3)=g_q(2,3)=\dfrac{q^3-1}{q-1}=q^2+q+1$, so

$$\begin{aligned}
f(q,3)=\mu(\hat{0},\hat{1}) &= -\sum_{V\subsetneq\mathbb{F}_q^3}\mu(\hat{0},V)\\
&= -\sum_{k=0}^{2} g_q(k,3)f(q,k)\\
&= -1-(q^2+q+1)(-1)-(q^2+q+1)(q) \;=\; \boxed{-q^3.}
\end{aligned}$$

For $n=4$:

$$\begin{aligned}
f(q,4) &= -\sum_{k=0}^{3} g_q(k,4)f(q,k)\\
&= -1-\frac{q^4-1}{q-1}(-1)-\frac{(q^4-1)(q^3-1)}{(q^2-1)(q-1)}(q)-\frac{q^4-1}{q-1}(-q^3) \;=\; \boxed{q^6.}
\end{aligned}$$

It is starting to look like

$$f(q,n)=(-1)^n q^{\binom{n}{2}}$$

in general, and indeed this is the case. We could prove this by induction now, but there is a slicker proof coming soon.

Why is the Möbius function useful?

- It is the inverse of ζ in the incidence algebra (check this!)
- It implies a more general version of inclusion-exclusion called *Möbius inversion.*
- It behaves nicely under poset operations such as product.
- It has geometric and topological applications. Even just the single number $\mu_P(\hat{0},\hat{1})$ tells you a lot about a bounded poset P; it is analogous to the Euler characteristic of a topological space.

Theorem 4.7 (Möbius inversion formula). *Let P be a poset in which every principal order ideal is finite, and let $f,g:P\to\mathbb{C}$. Then*

$$g(x)=\sum_{y:\,y\le x} f(y)\quad \forall x\in P \iff f(x)=\sum_{y:\,y\le x}\mu(y,x)g(y)\quad \forall x\in P, \tag{4.1a}$$

$$g(x)=\sum_{y:\,y\ge x} f(y)\quad \forall x\in P \iff f(x)=\sum_{y:\,y\ge x}\mu(x,y)g(y)\quad \forall x\in P. \tag{4.1b}$$

Proof. "A trivial observation in linear algebra" —Stanley.

We regard the incidence algebra as acting $\mathbb{C}$-linearly on the vector space V of functions $f:P\to\mathbb{Z}$ by

$$\begin{aligned}
(f\cdot\alpha)(x) &= \sum_{y:\,y\le x}\alpha(y,x)f(y),\\
(\alpha\cdot f)(x) &= \sum_{y:\,y\ge x}\alpha(x,y)f(y).
\end{aligned}$$

for $\alpha\in I(P)$. In terms of these actions, formulas (4.1a) and (4.1b) are respectively just the "trivial" observations

$$g=f\cdot\zeta \iff f=g\cdot\mu, \tag{4.2a}$$

$$g=\zeta\cdot f \iff f=\mu\cdot g. \tag{4.2b}$$

We just have to prove that these actions are indeed actions, i.e.,

$$[\alpha * \beta] \cdot f = \alpha \cdot [\beta \cdot f] \quad \text{and} \quad f \cdot [\alpha * \beta] = [f \cdot \alpha] \cdot \beta.$$

Indeed, this is straightforward:

$$\begin{aligned}
(f \cdot [\alpha * \beta])(y) &= \sum_{x\colon x \leq y} (\alpha * \beta)(x, y) f(x) \\
&= \sum_{x\colon x \leq y} \sum_{z\colon z \in [x,y]} \alpha(x, z) \beta(z, y) f(x) \\
&= \sum_{z\colon z \leq y} \left(\sum_{x\colon x \leq z} \alpha(x, z) f(x) \right) \beta(z, y) \\
&= \sum_{z\colon z \leq y} (f \cdot \alpha)(z) \beta(z, y) \quad = \quad ((f \cdot \alpha) \cdot \beta)(y).
\end{aligned}$$

and the other verification is a mirror image of this one. □

In the case of $\mathscr{B}_n$, the proposition says that

$$g(x) = \sum_{B \subseteq A} f(B) \quad \forall A \subseteq [n] \qquad \iff \qquad f(x) = \sum_{B \subseteq A} (-1)^{|B \setminus A|} g(B) \qquad \forall A \subseteq [n]$$

which is just the inclusion-exclusion formula. So Möbius inversion can be thought of as a generalized form of inclusion-exclusion that applies to an arbitrary locally finite poset P. If we know the Möbius function of P, then knowing a combinatorial formula for either f or g allows us to write down a formula for the other one.

Example 4.8. Here's an oldie-but-goodie. A **derangement** is a permutations $\sigma \in \mathfrak{S}_n$ with no fixed points. If D_n is the set of derangements, then $|D_1| = 0$, $|D_2| = 1$, $|D_3| = |\{231, 312\}| = 2$, $|D_4| = 9$, …What is $|D_n|$?

For $S \subset [n]$, let

$$\begin{aligned}
f(S) &= \{\sigma \in \mathfrak{S}_n \mid \sigma(i) = i \text{ iff } i \in S\}, \\
g(S) &= \{\sigma \in \mathfrak{S}_n \mid \sigma(i) = i \text{ if } i \in S\}.
\end{aligned}$$

Thus $D_n = f(\emptyset)$.

It's easy to count $g(S)$ directly. If $s = |S|$, then a permutation fixing the elements of S is equivalent to a permutation on $[n] \setminus S$, so $g(S) = (n - s)!$.

It's hard to count $f(S)$ directly. However,

$$g(S) = \sum_{R \supseteq S} f(R).$$

Rewritten in the incidence algebra $I(\mathscr{B}_n)$, this is just $g = \zeta \cdot f$. Thus $f = \mu \cdot g$, or in terms of the Möbius inversion formula (4.1b),

$$f(S) \;=\; \sum_{R \supseteq S} \mu(S, R) g(R) \;=\; \sum_{R \supseteq S} (-1)^{|R|-|S|} (n - |R|)! \;=\; \sum_{r=s}^{n} \binom{n}{r} (-1)^{r-s} (n - r)!$$

The number of derangements is then $f(\emptyset)$, which is given by the well-known formula

$$\sum_{r=0}^{n} \binom{n}{r} (-1)^r (n - r)!$$

Example 4.9. You can also use Möbius inversion to compute the Möbius function itself. In this example, we'll do this for the lattice $L_n(q)$. As a homework problem, you can use a similar method to compute the Möbius function of the partition lattice.

Let $V = \mathbb{F}_q^n$, let $L = L_n(q)$, let r be the rank function of L (i.e., $r(W) = \dim W$) and let X be a $\mathbb{F}_q$-vector space of *cardinality* x (yes, cardinality, not dimension!) Define

$$g(W) = \#\{\mathbb{F}_q\text{-linear maps } \phi : V \to X \mid \ker\phi \supset W\}.$$

Then $g(W) = x^{n-\dim W}$. (Choose a basis B for W and extend it to a basis B' for V. Then ϕ must send every element of B to zero, but can send each of the $n - \dim W$ elements of $B' \setminus B$ to any of the x elements of X.) Let

$$f(W) = \#\{\mathbb{F}_q\text{-linear maps } \phi : V \to X \mid \ker\phi = W.$$

Then $g(W) = \sum_{U \supset W} f(U)$, so by Möbius inversion

$$f(W) = \sum_{U:\ V \supseteq U \supseteq W} \mu_L(W,U) x^{n-\dim U}.$$

In particular, if we take W to be the zero subspace $0 = \hat{0}$, we obtain

$$(4.3)\quad f(\hat{0}) \;=\; \sum_{U \in L} \mu_L(\hat{0},U) x^{n-r(U)} \;=\; \#\{\text{1-1 linear maps } V \to X\} \;=\; (x-1)(x-q)(x-q^2)\cdots(x-q^{n-1}).$$

For this last count, choose an ordered basis $\{v_1,\dots,v_n\}$ for V, and send each v_i to a vector in X not in the linear span of $\{\phi(v_1),\dots,\phi(v_{i-1})\}$; there are $x - q^{i-1}$ such vectors. The identity (4.3) holds for infinitely many values of x and is thus an identity of polynomials in the ring $\mathbb{Q}[x]$. Therefore, it remains true upon setting x to 0 (even though no vector space can have cardinality zero!), which gives $\mu_{L_n(q)}(\hat{0},\hat{1}) = (-1)^n q^{\binom{n}{2}}$.

4.3. **The Characteristic Polynomial.**

Definition 4.10. Let P be a finite graded poset with rank function r, and suppose that $r(\hat{1}) = n$. The *characteristic polynomial* of P is defined as

$$\chi(P;x) \;=\; \sum_{z \in P} \mu(\hat{0},z) x^{n-r(z)}.$$

This is an important invariant of a poset, particularly if it is a lattice.

- We have just seen that
$$\chi(L_n(q);x) \;=\; (x-1)(x-q)(x-q^2)\cdots(x-q^{n-1}).$$
- The Möbius function is multiplicative on direct products of posets; therefore, so is the characteristic polynomial. For instance, if P is a product of n chains, then $\chi(P;x) = \sum_{k=0}^n (-1)^k \binom{n}{k} = (x-1)^n$.
- Π_n has a nice characteristic polynomial, which you will see soon.

The characteristic polynomial of a geometric lattice is a specialization of the Tutte polynomial of the corresponding matroid.

Theorem 4.11. *Let L be a geometric lattice with atoms E. Let M be the corresponding matroid on E, and r its rank function. Then*

$$\chi(L;x) \;=\; (-1)^{r(M)} T(M;\, 1-x, 0).$$

Proof. For $A \subset E$, let $\bar{A}$ denote the closure of A in M. We have

$$\begin{aligned}(-1)^{r(M)} T(M;\,1-x,0) &= (-1)^{r(M)} \sum_{A \subseteq E} (-x)^{r(M)-r(A)} (-1)^{|A|-r(A)} \\ &= \sum_{A \subseteq E} x^{r(M)-r(A)} (-1)^{|A|} \\ &= \sum_{K \in L} \underbrace{\left(\sum_{A \subseteq E:\ \bar{A}=K} (-1)^{|A|} \right)}_{f(K)} x^{r(M)-r(K)},\end{aligned}$$

so it suffices to check that $f(K) = \mu_L(\hat{0}, K)$. To do this, we use Möbius inversion on L. For $K \in L$, let

$$g(K) \;=\; \sum_{A \subseteq E:\ \bar{A} \subseteq K} (-1)^{|A|}$$

so $g = f \cdot \zeta$ and $f = g \cdot \mu$ in $I(L)$, and by Möbius inversion (this time, (4.1a)) we have

$$f(K) \;=\; \sum_{J \leq K} \mu(J,K) g(J) \;=\; \mu(\hat{0}, K)$$

On the other hand, $g(\hat{0}) = 1$, but if $J \neq \hat{0}$ then $g(J) = 0$. So the previous equation gives $f(K) = \mu(\hat{0}, K)$. □

Example 4.12. Let G be a simple graph with n vertices and c components so that its graphic matroid $M(G)$ has rank $n-c$. Let L be the geometric lattice corresponding to M. The flats of L are the *(vertex-)induced subgraphs* of G: the subgraphs H such that if $e = xy \in E(G)$, and x, y are in the same component of H, then $e \in E(H)$. We have seen before that the chromatic polynomial of G is

$$p(G;k) = (-1)^{n-c}\, k^c\, T(G,\, 1-k, 0).$$

Combining this with Theorem 4.11, we see that

$$p(G;k) = k^c\, \chi(L;k).$$

Here is a useful connection between the characteristic polynomial and chain-counting, with topological applications (among others).

Theorem 4.13 (Philip Hall's Theorem)**.** [15, Prop. 3.8.5] *Let P be a chain-finite, bounded poset, and let*

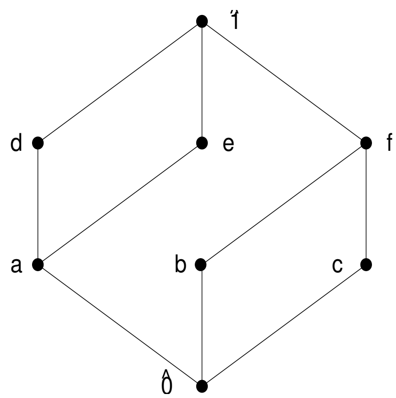
$$c_k = \left| \{\hat{0} = x_0 < x_1 < \cdots < x_k = \hat{1}\} \right|$$

be the number of chains of length i between $\hat{0}$ and $\hat{1}$. Then

$$\mu_P(\hat{0}, \hat{1}) = \sum_k (-1)^k c_k.$$

Example 4.14. For the poset P shown below, we have $\mu_P(\hat{0}, \hat{1}) = 1$ (check it). Meanwhile,

$$c_0 = 0, \quad c_1 = 1, \quad c_2 = 6, \quad c_3 = 4, \qquad c_0 - c_1 + c_2 - c_3 = 1.$$



Proof. Here's why the incidence algebra is so convenient: it makes the proof almost trivial.

Recall that $c_k = \eta^k(\hat{0},\hat{1}) = (\zeta-\delta)^k(\hat{0},\hat{1})$. The trick is to use the geometric series expansion $1/(1+h) = 1-h+h^2-h^3+h^4-\cdots$. Clearing both denominators and replacing h with η, we get

$$(\delta+\eta)\left(\sum_{k=0}^{\infty}(-1)^k\eta^k\right) = \delta$$

where 1 means δ (the multiplicative unit in $I(P)$). Since sufficiently high powers of η vanish, this is a perfectly good equation of polynomials in $I(P)$. Therefore,

$$(\delta+\eta)^{-1} = \left(\sum_{k=0}^{\infty}(-1)^k\eta^k\right)$$

and

$$\sum_{k=0}^{\infty}(-1)^k c_k \;=\; \sum_{k=0}^{\infty}(-1)^k\eta^k(\hat{0},\hat{1}) \;=\; (\delta+\eta)^{-1}(\hat{0},\hat{1}) \;=\; \zeta^{-1}(\hat{0},\hat{1}) \;=\; \mu(\hat{0},\hat{1}).$$ □

Example 4.15. Alternating sums, like those that appear in Möbius inversion and inclusion/exclusion, arise in topology as Euler characteristics. For any poset P, we can define its *order complex* $\Delta = \Delta(P)$ to be the simplicial complex on vertices P whose faces are the chains of P. (Note that this is a simplicial complex because every subset of a chain is a chain.) Specifically, $c_k(P)$ is the number of k-dimensional faces of Δ (often denoted $f_k(\Delta)$), so $\mu_{\hat{P}}(\hat{0},\hat{1}) = \tilde{\chi}(\Delta)$, the reduced Euler characteristic of Δ.

4.4. **Möbius Functions of Lattices.**

Theorem 4.16. *The Möbius function of a geometric lattice alternates in sign. I.e., if L is a geometric lattice and r is its rank function, then $(-1)^{r(x)}\mu_L(\hat{0},x) \geq 0$ for all $x \in L$.*

Proof. Let L be a geometric lattice with atoms E. Let M be the corresponding matroid on E, and r its rank function. Substituting $x = 0$ in the definition of the characteristic polynomial and in the formula of Theorem 4.11 gives

$$\mu(L) \;=\; \chi(L;\,0) \;=\; (-1)^{r(M)}T(M;\,1,0).$$

But $T(M;\,1,0) \geq 0$ for every matroid M, because $T(M;\,x,y) \in \mathbb{N}[x,y]$. Meanwhile, every interval $[\hat{0},z] \subset L$ is a geometric lattice, so the sign of $\mu(\hat{0},z)$ is the same as that of $(-1)^{r(z)}$ (or zero). □

In fact, more is true: the Möbius function of any *semimodular* (not necessarily atomic) lattice alternates in sign. This can be proven algebraically using tools we're about to develop (Stanley, Prop. 3.10.1) or combinatorially, by intepreting $(-1)^{r(M)}\mu(L)$ as enumerating *R-labellings* of L; see Stanley, §§3.12–3.13.

It is easier to compute the Möbius function of a lattice than of an arbitrary poset. The main technical tool is the following ring.

Definition 4.17. Let L be a lattice. The **Möbius algebra** $A(L)$ is the vector space of formal $\mathbb{C}$-linear combinations of elements of L, with multiplication given by the meet operation and extended linearly. (In particular, $\hat{1}$ is the multiplicative unit of $A(L)$.)

In general, the elements of L form a vector space basis of $A(L)$ consisting of *idempotents* (since $x \wedge x = x$ for all $x \in L$). For example, if $L = \mathscr{B}_n$ then $A(L) \cong \mathbb{C}[x_1, \dots, x_n]/(x_1^2 - x_1, \dots, x_n^2 - x_n)$.

It looks like $A(L)$ could have a complicated structure, but actually...

Proposition 4.18. *$A(L) \cong \mathbb{C}^{|L|}$ as rings.*

Proof. This is an application of Möbius inversion. For $x \in L$, define

$$\varepsilon_x = \sum_{y \leq x} \mu(y, x) y.$$

By Möbius inversion

$$x = \sum_{y \leq x} \varepsilon_y. \tag{4.4}$$

For $x \in L$, let $\mathbb{C}_x$ be a copy of $\mathbb{C}$ with unit 1_x, so we can identify $\mathbb{C}^{|L|}$ with $\prod_{x \in L} \mathbb{C}_x$.

Define a $\mathbb{C}$-linear map $\phi : A(L) \to \mathbb{C}^{|L|}$ by $\varepsilon_x \mapsto 1_x$. This is a vector space isomorphism, and by (4.4) we have

$$\phi(x)\phi(y) \;=\; \phi\left(\sum_{w \leq x} \varepsilon_w\right) \phi\left(\sum_{z \leq y} \varepsilon_z\right) \;=\; \left(\sum_{w \leq x} 1_w\right)\left(\sum_{z \leq y} 1_z\right) \;=\; \sum_{v \leq x \wedge y} 1_v \;=\; \phi(x \wedge y)$$

so in fact ϕ is a ring isomorphism. ☐

Question: *Suppose more generally that R is a $\mathbb{C}$-algebra of finite dimension n as a $\mathbb{C}$-vector space, and that $x_1, \dots, x_n$ are linearly independent idempotents (i.e., $x_i^2 = x_i$ for all i). Does it follow that $R \cong \mathbb{C}^n$ as rings? That is, does there exist a vector space basis $\{e_1, \dots, e_n\}$ such that $e_i^2 = e_i$ and $e_i e_j = 0$ for $i \neq j$? If so, there should be an explicit construction of the e_i's, perhaps through linear algebra.*

Note in particular that

$$\varepsilon_x \varepsilon_y = \begin{cases} \varepsilon_x & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases} \tag{4.5}$$

The reason the Möbius algebra is useful is that it lets us compute $\mu(x, y)$ more easily by summing over a cleverly chosen *subset* of $[x, y]$, rather than the entire interval.

Proposition 4.19. *Let L be a finite lattice with at least two elements. Then for every $a \in L \setminus \{\hat{1}\}$ we have*

$$\sum_{x : x \wedge a = \hat{0}} \mu(x, \hat{1}) = 0.$$

Proof. On the one hand

$$a \varepsilon_{\hat{1}} \;=\; \left(\sum_{b \leq a} \varepsilon_b\right) \varepsilon_{\hat{1}} \;=\; 0.$$

On the other hand

$$a\varepsilon_{\hat{1}} \;=\; a\left(\sum_{x\in L}\mu(x,\hat{1})x\right) \;=\; \sum_{x\in L}\mu(x,\hat{1})x\wedge a.$$

Now take the coefficient of $\hat{0}$. $\square$

A corollary of Proposition 4.19 is the useful formula

$$\mu(L) \;=\; \mu_L(\hat{0},\hat{1}) \;=\; -\sum_{\substack{x\neq\hat{0}:\\ x\wedge a=\hat{0}}}\mu(x,\hat{1}) \tag{4.6}$$

Note that this formula does not apply to posets that are not lattices (because the summation requires a well-defined meet).

Example 4.20 (Möbius function of the partition lattice Π_n)**.** Let $a = \{[n-1],\ \{n\}\} \in \Pi_n$. Then the partitions x that show up in the sum of (4.6) are the atoms whose non-singleton block is $\{i,n\}$ for some $i\in[n-1]$. For each such x, the interval $[x,\hat{1}]\subset\Pi_n$ is isomorphic to Π_{n-1}, so (4.6) gives

$$\mu(\Pi_n) \;=\; -(n-1)\mu(\Pi_{n-1})$$

from which it follows by induction that

$$\mu(\Pi_n) = (-1)^{n-1}(n-1)!.$$

(Wasn't that easy?)

Example 4.21 (Möbius function of the subspace lattice $L_n(q)$)**.** Let $L = L_n(q)$, and let $A = \{(v_1,\dots,v_n)\in\mathbb{F}_q^n \mid v_n = 0\}$. This is a codimension-1 subspace in $\mathbb{F}_q^n$, hence a coatom in L. If X is a nonzero subspace such that $X\cap A = 0$, then X must be a line spanned by some vector $(x_1,\dots,x_n)$ with $x_n\neq 0$. We may as well assume $x_n = 1$ and choose $x_1,\dots,x_{n-1}$ arbitrarily, so there are q^{n-1} such lines. Moreover, the interval $[X,\hat{1}]\subset L$ is isomorphic to $L_{n-1}(q)$. Therefore

$$\mu(L_n(q)) \;=\; -q^{n-1}\mu(L_{n-1}(q))$$

and by induction

$$\mu(L_n(q)) \;=\; (-1)^n q^{\binom{n}{2}}.$$

4.5. **Crosscuts.**

Definition 4.22. Let L be a lattice.
— An **upper crosscut** of L is a set $X\subset L\setminus\{\hat{1}\}$ such that if $y\in L\setminus X\setminus\{\hat{1}\}$, then $y<x$ for some $x\in X$. (Equivalently X contains all coatoms of L.)
— A **lower crosscut** of L is a set $X\subset L\setminus\{\hat{0}\}$ such that if $y\in L\setminus X\setminus\{\hat{0}\}$, then $y>x$ for some $x\in X$. (Equivalently X contains all atoms of L.)

The reason we don't simply define "crosscut" as "a set that contains all the coatoms" is that sometimes it's useful to have a more general notion of crosscut to apply the following theorem.

Proposition 4.23 (Rota's Crosscut Theorem)**.** *Let L be a finite lattice and let X be an upper crosscut. Then*

$$\mu(L) = \sum_{Y\subseteq X:\ \bigwedge Y=\hat{0}}(-1)^{|Y|}. \tag{4.7a}$$

Dually, if X is a lower crosscut, then

$$\mu(L) = \sum_{Y\subseteq X:\ \bigvee Y=\hat{1}}(-1)^{|Y|}. \tag{4.7b}$$

Proof. We will prove (4.7a). Suppose X is an upper crosscut. For any $x \in L$, we have the following equation in the Möbius algebra of L:

$$\hat{1} - x \;=\; \left(\sum_{y \in L} \varepsilon_y\right) - \left(\sum_{y \leq x} \varepsilon_y\right) \;=\; \left(\sum_{y \not\leq x} \varepsilon_y\right).$$

Therefore

$$\prod_{x \in X} (\hat{1} - x) \;=\; \prod_{x \in X} \left(\sum_{y \not\leq x} \varepsilon_y\right) \;=\; \sum_{y \in Y} \varepsilon_y$$

where $Y = \{y \in L \mid y \not\leq x \text{ for all } x \in X\}$ (expand the sum and apply (4.5)). But $Y = \{\hat{1}\}$ because X is an upper crosscut. That is,

$$\prod_{x \in X} (\hat{1} - x) \;=\; \varepsilon_{\hat{1}} \;=\; \sum_{y \in L} \mu(y, \hat{1}) y. \tag{4.8}$$

On the other hand

$$\prod_{x \in X} (\hat{1} - x) \;=\; \sum_{A \subseteq X} (-1)^{|A|} \bigwedge A. \tag{4.9}$$

Now extracting the coefficient of $\hat{0}$ on the right-hand sides of (4.8) and (4.9) yields (4.7a). The proof of (4.7b) is similar. □

Corollary 4.24. *Let L be a lattice in which $\hat{1}$ is not a join of atoms. Then $\mu(L) = 0$.*

In particular, $\mu(L) = 0$ if L is distributive and not Boolean (because then L is not atomic, so there is some join-irreducible non-atom b, so $b \not\leq a$ for all atoms $a \in A$ (where A is the set of atoms), so $b \not\leq \bigvee A$, so $\hat{1} \not\leq \bigvee A$, so $\bigvee A \neq \hat{1}$).

A topological application is the following result due to J. Folkman (1966), whose proof used the crosscut theorem.

Theorem 4.25. *Let L be a geometric lattice of rank r, and let $P = L \setminus \{\hat{0}, \hat{1}\}$. Then*

$$\tilde{H}_i(\Delta(P), \mathbb{Z}) \cong \begin{cases} \mathbb{Z}^{|\mu(L)|} & \text{if } i = r - 2, \\ 0 & \text{otherwise} \end{cases}$$

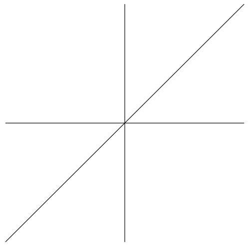
where $\tilde{H}_i$ denotes reduced simplicial homology. That is, $\Delta(P)$ has the homology type of the wedge of $\mu(L)$ spheres of dimension $r - 2$.

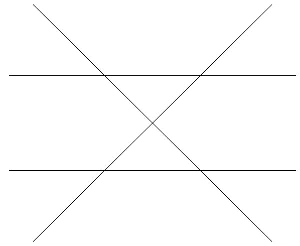
The crosscut theorem will also be useful in studying hyperplane arrangements.

5. Hyperplane Arrangements

Definition 5.1. Let $\mathbb{F}$ be a field, typically either $\mathbb{R}$ or $\mathbb{C}$, and let $n \geq 1$. A **linear hyperplane** in $\mathbb{F}^n$ is a vector subspace of codimension 1. An **affine hyperplane** is a translate of a linear hyperplane. A **hyperplane arrangement** $\mathcal{A}$ is a finite collection of (distinct) hyperplanes. The number n is called the **dimension** of $\mathcal{A}$.

Example 5.2. The left-hand arrangement $\mathcal{A}_1$ is linear; it consists of the lines $x = 0$, $y = 0$, and $x = y$. The right-hand arrangement $\mathcal{A}_2$ is affine; it consists of the four lines $x = y$, $x = -y$, $y = 1$ and $y = -1$.

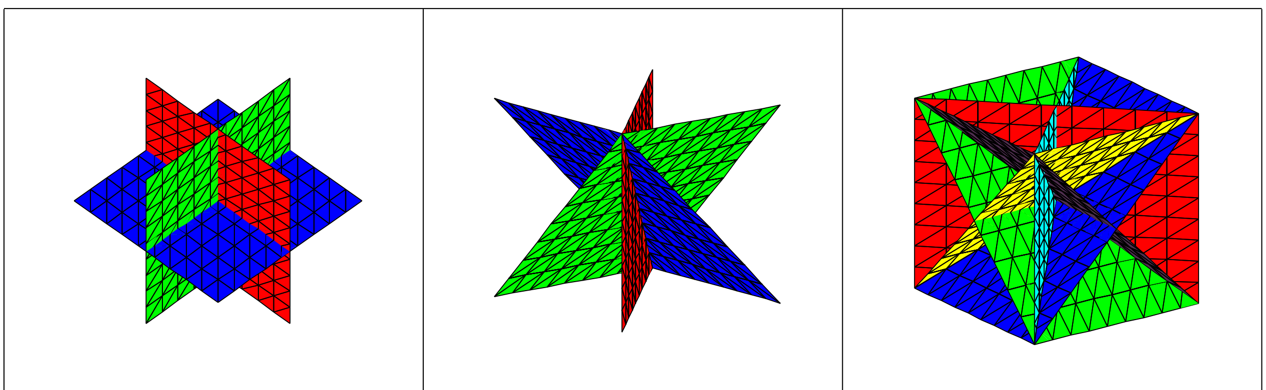




Each hyperplane is the zero set of some linear form, so their union is the zero set of the product of those s linear forms. We can specify an arrangement concisely by that product, called the **defining polynomial** of $\mathcal{A}$ (as an algebraic variety, in fact). For example, the defining polynomials of $\mathcal{A}_1$ and $\mathcal{A}_2$ are $xy(x-y)$ and $xy(x-y-1)$ respectively.

Example 5.3. Here are some 3-D arrangements (pictures produced using Maple). The *Boolean arrangement* $\mathscr{B}_n$ consists of the coordinate hyperplanes in n-space, so its defining polynomial is $x_1x_2\dots x_n$.

The *braid arrangement* $\mathscr{B}_n$ consists of the $\binom{n}{2}$ hyperplanes $x_i - x_j$ in n-space. In particular, every hyperplane in $\mathscr{B}_n$ contains the line $x_1 = x_2 = \cdots = x_n$. Projecting $\mathbb{R}^4$ along that line allows us to picture $\mathscr{B}_4$ as an arrangement $\mathrm{ess}(Br_4)$ in $\mathbb{R}^3$ ("ess" means *essentialization*, to be explained soon).



$\mathscr{B}_3$ | Br_3 | $\mathrm{ess}(Br_4)$

5.1. **The Intersection Poset.**

Definition 5.4. Let $\mathcal{A} \subset \mathbb{F}^n$ be an arrangement. Its **intersection poset** $L(\mathcal{A})$ is the poset of all intersections of subsets of $\mathcal{A}$, ordered by reverse inclusion.

- $L(\mathcal{A})$ always has a $\hat{0}$ element, namely $\mathbb{F}^n$.
- $L(\mathcal{A})$ is ranked by codimension in $\mathbb{F}^n$, because each covering relation $Y \lessdot Z$ comes from intersecting an affine linear subspace Y with a hyperplane H that neither contains nor is disjoint from Y — and in this case $\dim(Y \cap H) = \dim(Y) - 1$.
- $L(\mathcal{A})$ has a $\hat{1}$ element if and only if $\bigcap_{H\in\mathcal{A}} H \neq \emptyset$. Such an arrangement is called **central**.

Proposition 5.5. *Let $\mathcal{A} \subset \mathbb{F}^n$ be an arrangement. The following are equivalent:*

- *$\mathcal{A}$ is central.*
- *$\mathcal{A}$ is a translation of a linear arrangement.*
- *$L(\mathcal{A})$ is a geometric lattice.*

Proof. Linear arrangements are central because every hyperplane contains $\vec{0} \in \mathbb{F}^n$. Conversely, if $\mathcal{A}$ is central and $p \in \bigcap_{H\in\mathcal{A}} H$, then translating by $-p$ produces a linear arrangement.

If $\mathcal{A}$ is central, then $L(\mathcal{A})$ is bounded. It is a join-semilattice, with join given by intersection, hence it is a lattice. Indeed, it is a geometric lattice: it is clearly atomic, and it is submodular because it is a sublattice of the chain-finite modular lattice $L(\mathbb{F}^n)^*$, (i.e., the lattice of all subspaces of $\mathbb{F}^n$ ordered by reverse inclusion). $\square$

When $\mathcal{A}$ is central (we may as well assume linear), the matroid associated with $L(\mathcal{A})$ is naturally represented by the normal vectors to the hyperplanes in $\mathcal{A}$. Therefore, all of the tools we have developed for looking at lattices and matroids can be applied to study hyperplane arrangements. (We will remove the requirement of centrality in Section 5.5.)

The dimension of an arrangement cannot be inferred from the intersection poset. For example, if $\mathcal{A}_1$ is as above, then $L(\mathcal{A}_1) \cong L(Br_3)$ but $\dim \mathcal{A}_1 = 2$ and $\dim Br_3 = 3$. A more useful invariant of $\mathcal{A}$ is its **rank** $\operatorname{rank} \mathcal{A}$, defined as the rank of $L(\mathcal{A})$. Equivalently, define $W \subset \mathbb{F}^n$ to be the subspace spanned by the normal vectors v_i. Then $\operatorname{rank} \mathcal{A} = \dim W$.

Definition 5.6. An arrangement $\mathcal{A} \subset \mathbb{F}^n$ is **essential** if $\operatorname{rank} \mathcal{A} = \dim \mathcal{A}$. In general, the **essentialization** $\operatorname{ess}(\mathcal{A})$ is the arrangement

$$\{H \cap W \mid H \in \mathcal{A}\} \;\subset\; W$$

where again W is the subspace of $\mathbb{F}^n$ spanned by all normals to hyperplanes in $\mathcal{A}$.

Equivalently, if $V = W^\perp = \bigcap_{H\in\mathcal{A}} H$, then $\operatorname{rank} \mathcal{A} = n - \dim V$, and it is combinatorially equivalent to define the essentialization of $\mathcal{A}$ as the quotient

$$\{H/V \mid H \in \mathcal{A}\} \;\subset\; \mathbb{F}^n/V.$$

Observe that $\operatorname{ess}(\mathcal{A})$ is essential, and that $L(\mathcal{A})$ is naturally isomorphic to $L(\operatorname{ess}(\mathcal{A}))$.

5.2. **Counting Regions of Hyperplane Arrangements: Zaslavsky's Theorems.** Let $\mathcal{A} \subset \mathbb{R}^n$ be a real hyperplane arrangement. The **regions** of $\mathcal{A}$ are the connected components of $\mathbb{R}^n \setminus \mathcal{A} = \mathbb{R}^n \setminus \bigcup_{H\in\mathcal{A}} H$. Each component is the interior of a (bounded or unbounded) polyhedron; in particular, it is homeomorphic to $\mathbb{R}^n$. We write

$$r(\mathcal{A}) = \text{ number of regions of } \mathcal{A}.$$

We'd also like to count the number of bounded regions. However, we must be careful, because if $\mathcal{A}$ is not essential then every region is unbounded. Accordingly, call a region **relatively bounded** if the corresponding region in $\operatorname{ess}(\mathcal{A})$ is bounded, and define

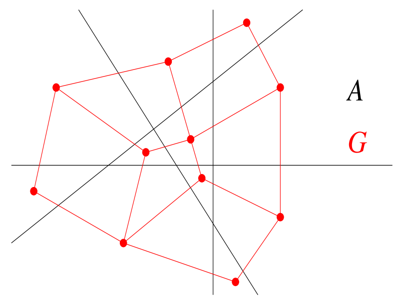
$$b(\mathcal{A}) = \text{ number of relatively bounded regions of } \mathcal{A}.$$

Note that $b(\mathcal{A}) = 0$ if and only if $\operatorname{ess}(\mathcal{A})$ is central.

Example 5.7. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be the 2-dimensional arrangements shown on the left and right of the figure below, respectively. Then $r(\mathcal{A}_1) = 6$, $b(\mathcal{A}_1) = 0$, $r(\mathcal{A}_2) = 10$, $b(\mathcal{A}_2) = 2$.

Example 5.8. The Boolean arrangement $\mathscr{B}_n$ consists of the n coordinate hyperplanes in $\mathbb{R}^n$. The complement $\mathbb{R}^n \setminus \mathscr{B}_n$ is $\{(x_1, \dots, x_n) \mid x_i \neq 0 \text{ for all } i\}$, and the connected components are the open orthants, specified by the signs of the n coordinates. Therefore, $r(\mathscr{B}^n) = 2^n$.

Example 5.9. Let $\mathcal{A}$ consist of m lines in $\mathbb{R}^2$ in *general position*: that is, no two lines are parallel and no three are coincident. Draw the *dual graph* G: the graph whose vertices are the regions of $\mathcal{A}$, with an edge between every two regions that share a common border.



Let $r = r(\mathcal{A})$ and $b = b(\mathcal{A})$, and let v, e, f denote the numbers of vertices, edges and faces of G, respectively. Each bounded region contains exactly one point where two lines of $\mathcal{A}$ meet, and each unbounded face has four sides. Therefore

$$v = r, \tag{5.1a}$$

$$f = 1 + \binom{m}{2} = \frac{m^2 - m + 2}{2} \tag{5.1b}$$

$$4(f-1) = 2e - (r-b). \tag{5.1c}$$

Moreover, the number $r - b$ of unbounded regions is just $2m$. (Take a walk around a very large circle. You will enter each unbounded region once, and will cross each line twice.) Therefore, from (5.1c) and (5.1b) we obtain

$$e = m + 2(f-1) = m^2. \tag{5.1d}$$

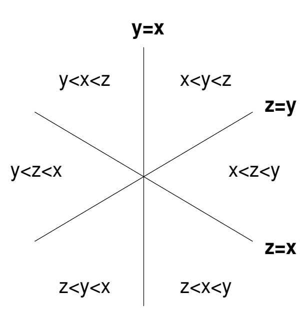
Now, Euler's formula for planar graphs says that $v - e + f = 2$. Substituting in (5.1a), (5.1b) and (5.1d) and solving for r gives

$$r = \frac{m^2 + m + 2}{2}$$

and therefore

$$b = r - 2m = \frac{m^2 - 3m + 2}{2} = \binom{m-1}{2}.$$

Example 5.10. The braid arrangement Br_n consists of the $\binom{n}{2}$ hyperplanes $x_i = x_j$ in $\mathbb{R}^n$. The complement $\mathbb{R}^n \setminus Br_n$ consists of all vectors in $\mathbb{R}^n$ with no two coordinates equal, and the connected components of this set are specified by the ordering of the set of coordinates as real numbers:



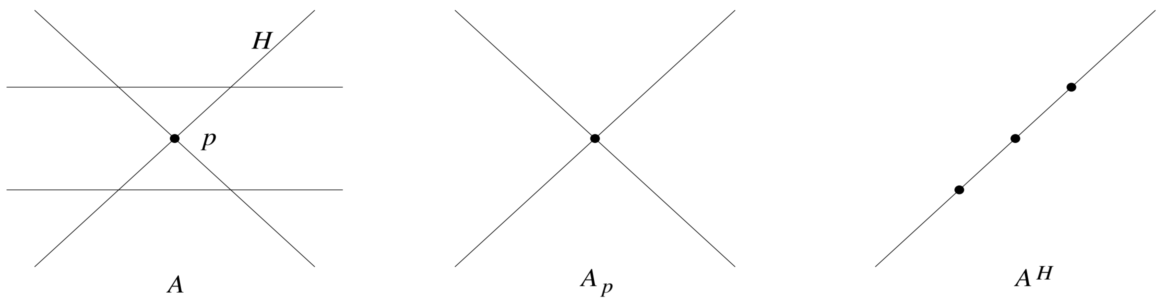
Therefore, $r(Br_n) = n!$

Our next goal is to prove Zaslavsky's theorems that the numbers $r(\mathcal{A})$ and $b(\mathcal{A})$ can be obtained as simple evaluations of the characteristic polynomial of the intersection poset $L(\mathcal{A})$.

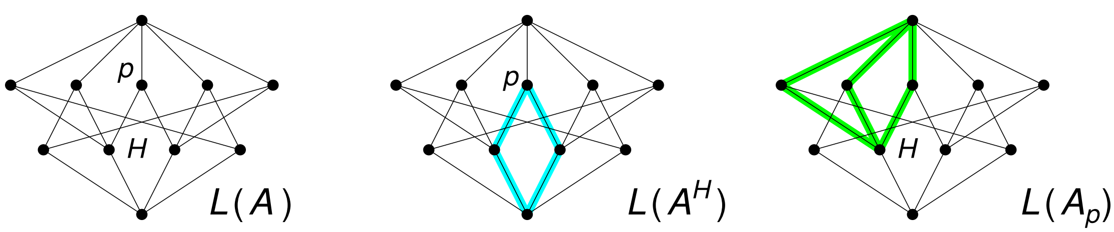
Let $x \in L(\mathcal{A})$; recall that this means that x is an affine space formed by the intersection of some subset of $\mathcal{A}$. Define arrangements

$$\mathcal{A}_x = \{H \in \mathcal{A} \mid H \supseteq x\},$$
$$\mathcal{A}^x = \{W \mid W = H \cap x,\ H \in \mathcal{A} \setminus \mathcal{A}_x\}.$$

Example 5.11. Let $\mathcal{A}$ be the 2-dimensional arrangement shown on the left, with the line H and point p as shown. Then $\mathcal{A}_p$ and $\mathcal{A}^H$ are shown on the right.



The reason for this notation is that $L(\mathcal{A}_x)$ and $L(\mathcal{A}^x)$ are isomorphic respectively to the principal order ideal and principal order filter generated by x in $L(\mathcal{A})$.



Both $\mathcal{A}_x$ and $\mathcal{A}^x$ describe what part of the arrangement x "sees", but in different ways: $\mathcal{A}_x$ is obtained by deleting the hyperplanes not containing x, while $\mathcal{A}^x$ is obtained by restricting $\mathcal{A}$ to x so as to get an arrangement whose ambient space is x itself.

Proposition 5.12. *Let $\mathcal{A}$ be a real arrangement and $H \in \mathcal{A}$. Let $\mathcal{A}' = \mathcal{A} \setminus \{H\}$ and $\mathcal{A}'' = \mathcal{A}^H$. Then*

$$r(\mathcal{A}) = r(\mathcal{A}') + r(\mathcal{A}'') \tag{5.2}$$

and

$$b(\mathcal{A}) = \begin{cases} b(\mathcal{A}') + b(\mathcal{A}'') & \text{if } \operatorname{rank} \mathcal{A} = \operatorname{rank} \mathcal{A}', \\ 0 & \text{if } \operatorname{rank} \mathcal{A} = \operatorname{rank} \mathcal{A}' + 1. \end{cases} \tag{5.3}$$

Notice that $\operatorname{rank} \mathcal{A}'$ equals either $\operatorname{rank} \mathcal{A} - 1$ or $\operatorname{rank} \mathcal{A}$, according as the normal vector $\vec{n}_H$ is or is not a coloop in the matroid $M(\mathcal{A})$ represented by all normal vectors.

Proof. Consider what happens when we add H to $\mathcal{A}'$ to obtain $\mathcal{A}$. Some regions of $\mathcal{A}'$ will remain the same, while others will be split into two regions. Say S and U are the number of split and unsplit regions.

The unsplit regions each count once in both $r(\mathcal{A})$ and $r(\mathcal{A}')$. The split regions in the second category each contribute 2 to $r(\mathcal{A})$, but they also correspond bijectively to the regions of $\mathcal{A}''$. So

$$r(\mathcal{A}') = S + U, \qquad r(\mathcal{A}) = 2S + U, \qquad r(\mathring{A}'') = S$$

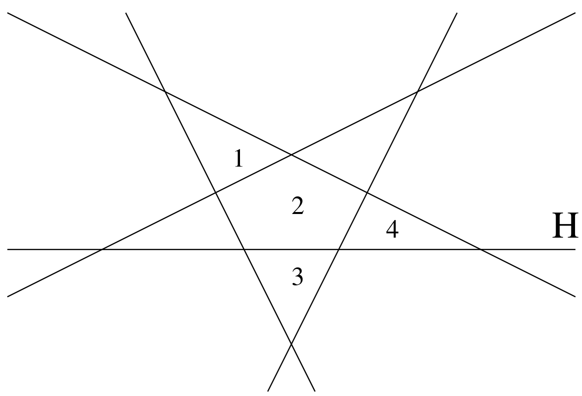
proving (5.2).

By the way, if (and only if) H is a coloop then it borders *every* region of $\mathcal{A}$, so $r(\mathcal{A}) = 2r(\mathcal{A}')$ in this case.

Now, what about bounded regions? If H is a coloop, then $\mathcal{A}$ has no bounded regions — every region of $\mathcal{A}'$ will contain a line parallel to $\vec{n}_H$, so every region of $\mathcal{A}$ will contain a ray. Otherwise, the bounded regions of $\mathcal{A}$ come in three flavors:

First, the regions not bordered by H (e.g., #1 below) correspond bijectively to bounded regions of $\mathcal{A}'$ through which H does not pass.

Second, for each region R of $\mathcal{A}$ bordered by H, the region $\bar{R} \cap H$ is bounded in $\mathcal{A}''$ (where $\bar{R}$ denotes the topological closure). Moreover, R comes from a bounded region in $\mathcal{A}'$ if and only if walking from R across H gets you to a bounded region of $\mathcal{A}$. (Yes in the case of the pair #2 and #3, which together contribute two to each side of (5.3); no in the case of #4, which contributes one to each side of (5.3).)



Therefore, we can count the bounded regions of $\mathcal{A}$ by starting with $b(\mathcal{A}')$, then adding one for each bounded region of $\mathcal{A}''$ (either by splitting a bounded region into two bounded regions, or by cutting off a piece of an unbounded region). So $b(\mathcal{A}) = b(\mathcal{A}') + b(\mathcal{A}'')$. $\square$

This looks a lot like a Tutte polynomial deletion/contraction recurrence. However, we only have a matroid to work with when $L(\mathcal{A})$ is a geometric lattice, that is, when $\mathcal{A}$ is central (otherwise, $L(\mathcal{A})$ is not even a bounded poset). On the other hand, $L(\mathcal{A})$ is certainly ranked (by codimension) for every arrangement, so we can work instead with its characteristic polynomial, which as you recall is defined as

$$\chi_{\mathcal{A}}(k) = \chi(L(\mathcal{A}); k) \;=\; \sum_{x \in L(\mathcal{A})} \mu(\hat{0}, x) k^{\dim x}. \tag{5.4}$$

In order to establish a recurrence for the characteristic polynomial, we first find a closed form for it.

Proposition 5.13 (Whitney's formula)**.** *For any real hyperplane arrangement $\mathcal{A}$, we have*

$$\chi_{\mathcal{A}}(k) \;=\; \sum_{\text{central } \mathcal{B} \subseteq \mathcal{A}} (-1)^{|\mathcal{B}|} k^{\dim \mathcal{A} - \operatorname{rank} \mathcal{B}}.$$

Proof. Consider the interval $[\hat{0}, x]$. The atoms in this interval are the hyperplanes of $\mathcal{A}$ containing x, and they form a lower crosscut of $[\hat{0}, x]$. Therefore, Rota's crosscut theorem (Proposition 4.23) says that

$$\mu(\hat{0}, x) \;=\; \sum_{Y \subset \mathcal{A}:\; x = \bigcap Y} (-1)^{|Y|}. \tag{5.5}$$

(To clarify, the sum is over all sets Y of hyperplanes of $\mathcal{A}$ such that x is (exactly) the intersection of the elements of Y.) Plugging (5.5) into the definition of the characteristic polynomial, we get

$$\begin{aligned}
\chi_{\mathcal{A}}(k) \;&=\; \sum_{x \in L(\mathcal{A})} \sum_{Y \subset \mathcal{A}:\; x = \bigcap Y} (-1)^{|Y|} k^{\dim x} \\
&=\; \sum_{Y \subset \mathcal{A}:\; \bigcap Y \neq 0} (-1)^{|Y|} k^{\dim(\bigcap Y)} \\
&=\; \sum_{\text{central } \mathcal{B} \subseteq \mathcal{A}} (-1)^{|\mathcal{B}|} k^{\dim \mathcal{A} - \operatorname{rank} \mathcal{B}}
\end{aligned}$$

as desired. □

Proposition 5.14 (Deletion/Restriction)**.** *Let $\mathcal{A} \subset \mathbb{R}^n$ be a real arrangement and $H \in \mathcal{A}$. Let $\mathcal{A}' = \mathcal{A} \setminus \{H\}$ and $\mathcal{A}'' = \mathcal{A}^H$. Then*

$$\chi_{\mathcal{A}}(k) \;=\; \chi_{\mathcal{A}'}(k) - \chi_{\mathcal{A}''}(k). \tag{5.6}$$

Sketch of proof. Split the sum in Whitney's formula into two pieces, depending on whether $H \in \mathcal{B}$. First,

$$\sum_{\substack{\text{central } \mathcal{B} \subseteq \mathcal{A} \\ H \notin \mathcal{B}}} (-1)^{|\mathcal{B}|} k^{n - \operatorname{rank} \mathcal{B}} \;=\; \sum_{\text{central } \mathcal{B} \subseteq \mathcal{A}'} (-1)^{|\mathcal{B}|} k^{n - \operatorname{rank} \mathcal{B}} \;=\; \chi_{\mathcal{A}'}(k). \tag{5.7}$$

Second, we wish to show that

$$\sum_{\mathcal{B}} (-1)^{|\mathcal{B}|} k^{n - \operatorname{rank} \mathcal{B}} \;=\; -\chi_{\mathcal{A}''}(k). \tag{5.8}$$

where the sum runs over all central arrangements $\mathcal{B} \subseteq \mathcal{A}$ such that $H \in \mathcal{B}$. This is a little trickier, because many different $\mathcal{B}$'s may give rise to the same central arrangement in $\mathcal{A}''$, so we need some careful bookkeeping.

Impose an equivalence relation $\sim$ on $\mathcal{A} \setminus \{H\}$ by setting $K \sim K'$ if $K \cap H = K' \cap H$, and let the equivalence classes be $\mathcal{A}_1, \dots, \mathcal{A}_s$. For each arrangement $\mathcal{B} \subseteq \mathcal{A}$, let $S(\mathcal{B}) = \{j \in [s] \mid \mathcal{A}_j \cap \mathcal{B} \neq \emptyset\}$; call this the *index set* of $\mathcal{B}$. Then two central arrangements with the same index set give rise to the same central subarrangement of $\mathcal{A}''$, and in fact the index sets themselves are in bijection with, and may be regarded as, just those central subarrangements of $\mathcal{A}''$. Accordingly, we see that

$$\operatorname{rank} S = \operatorname{rank} \bigcup_{j \in S} \mathcal{A}_j = \dim H - \dim \bigcap_{j \in S} \bigcap_{K \in \mathcal{A}_j} K \cap H$$

and

$$\begin{aligned}
\sum_{\mathcal{B}} (-1)^{|\mathcal{B}|} k^{n - \operatorname{rank} \mathcal{B}} \;&=\; \sum_{S} \sum_{\mathcal{B}:\, S(\mathcal{B}) = S} (-1)^{|\mathcal{B}|} k^{n - \operatorname{rank} \mathcal{B}} \\
&=\; \sum_{S} k^{\dim H - \operatorname{rank} S} \sum_{\substack{(\mathcal{C}_j)_{j \in S}: \\ \emptyset \subsetneq \mathcal{C}_j \subseteq \mathcal{A}_j}} (-1)^{\sum |\mathcal{C}_j|} \\
&=\; \sum_{S} k^{\dim H - \operatorname{rank} S} \prod_{j \in S} \left(\sum_{\emptyset \subsetneq \mathcal{C}_j \subseteq \mathcal{A}_j} (-1)^{|C_j|} \right) \;=\; \sum_{S} k^{\dim H - \operatorname{rank} S} \prod_{j \in S} (-1) \\
&=\; \sum_{S} k^{\dim H - \operatorname{rank} S} (-1)^{|S|} \;\;=\;\; -\chi_{\mathcal{A}''}(k).
\end{aligned}$$

as desired. Now the desired recurrence follows from Proposition 5.13, together with (5.7) and (5.8). □

Theorem 5.15 (Zaslavsky 1975)**.** *Let $\mathcal{A}$ be a real hyperplane arrangement. Then*

$$r(\mathcal{A}) = (-1)^{\dim \mathcal{A}} \chi_{\mathcal{A}}(-1), \tag{5.9}$$

$$b(\mathcal{A}) = (-1)^{\operatorname{rank} \mathcal{A}} \chi_{\mathcal{A}}(1). \tag{5.10}$$

Sketch of proof. Compare the recurrences for r and c with those for these evaluations of the characteristic polynomial (from Proposition 5.14). $\square$

Corollary 5.16. *Let $\mathcal{A} \subset \mathbb{R}^n$ be a central, essential hyperplane arrangement, so that $L(\mathcal{A})$ is a geometric lattice. Let M be the corresponding matroid. Then*

$$r(\mathcal{A}) = T(M;\, 2, 0), \qquad b(\mathcal{A}) = T(M;\, 0, 0) = 0.$$

Proof. Combine Zaslavsky's theorem with the formula $\chi_{\mathcal{A}}(k) = (-1)^n T(M;\, 1-k, 0)$. $\square$

Example 5.17. Let $m \geq n$, and let $\mathcal{A}$ be an arrangement of m linear hyperplanes in general position in $\mathbb{R}^n$. The corresponding matroid M is $U_n(m)$, whose rank function is

$$r(A) = \min(n, |A|)$$

for $A \subseteq [m]$. Therefore

$$\begin{aligned}
r(\mathcal{A}) \;=\; T(M;\, 2, 0) \;&=\; \sum_{A \subseteq [m]} (2-1)^{n-r(A)} (0-1)^{|A|-r(A)} \\
&=\; \sum_{A \subseteq [m]} (-1)^{|A|-r(A)} \\
&=\; \sum_{k=0}^{m} \binom{m}{k} (-1)^{k-\min(n,k)} \\
&=\; \sum_{k=0}^{n} \binom{m}{k} \;+\; \sum_{k=n+1}^{m} \binom{m}{k} (-1)^{k-n} \\
&=\; \sum_{k=0}^{n} \binom{m}{k} (1-(-1)^{k-n}) \;+\; \sum_{k=0}^{m} \binom{m}{k} (-1)^{k-n} \\
&=\; \sum_{k=0}^{n} \binom{m}{k} (1-(-1)^{k-n}) \\
&=\; 2\left(\binom{m}{n-1} + \binom{m}{n-3} + \binom{m}{n-5} + \cdots \right).
\end{aligned}$$

For instance, if $n = 3$ then

$$r(\mathcal{A}) = 2\left(\binom{m}{2} + \binom{m}{0} \right) = m^2 - m + 2.$$

Notice that this is *not* the same as the formula we obtained last time for the number of regions formed by m *affine* lines in general position in $\mathbb{R}^2$.

5.3. **Arrangements over $\mathbb{F}_q$.** Let $\mathbb{F} = \mathbb{F}_q$ be the finite field of order q, and let $\mathcal{A} \subset \mathbb{F}^n$ be a hyperplane arrangement. The "regions" of $\mathbb{F}^n \setminus \mathcal{A}$ are just its points (assuming, if you wish, that we endow $\mathbb{F}^n$ with the discrete topology). The following very important result is implicit in the work of Crapo and Rota (1970) and was stated explicitly by Athanasiadis (1996):

Proposition 5.18. $|\mathbb{F}_q^n \setminus \mathcal{A}| = \chi_{\mathcal{A}}(q)$.

Proof. By inclusion-exclusion, we have

$$|\mathbb{F}_q^n \setminus \mathcal{A}| \;=\; \sum_{\mathcal{B}\subseteq\mathcal{A}} (-1)^{|\mathcal{B}|} \left|\bigcap \mathcal{B}\right|.$$

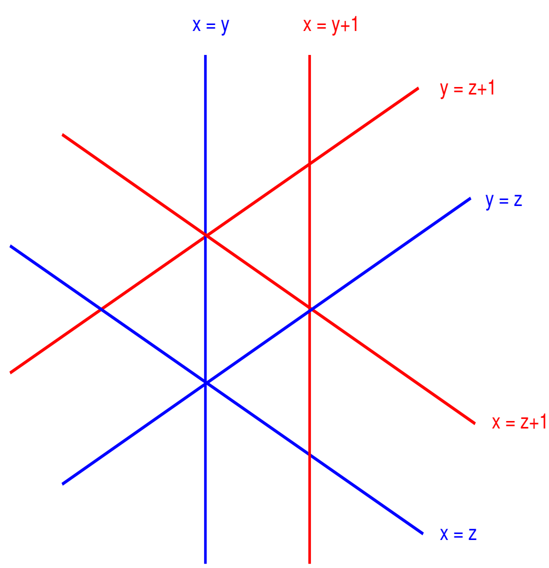
If $\mathcal{B}$ is not central, then by definition $|\bigcap \mathcal{B}| = 0$. Otherwise, $|\bigcap \mathcal{B}| = q^{n-\operatorname{rank}\mathcal{B}}$. So the sum becomes Whitney's formula for $\chi_{\mathcal{A}}(q)$ (Prop. 5.13). $\square$

This fact has a much more general application, which was systematically mined by Athanasiadis (1996). Let $\mathcal{A} \subset \mathbb{R}^n$ be an arrangement defined over the integers (i.e., such that the normal vectors to its hyperplanes lie in $\mathbb{Z}^n$). For a prime p, let $\mathcal{A}_p \subset \mathbb{F}_p^n$ be the arrangement defined by regarding the coordinates of the normal vectors as numbers modulo p. If p is sufficiently large, then it will be the case that $L(\mathcal{A}_p) \cong L(\mathcal{A})$. In this case we say that $\mathcal{A}$ **reduces correctly modulo** p. But that means that we can compute the characteristic polynomial of $\mathcal{A}$ by counting the points of $\mathcal{A}_p$ as a function of p, for large enough p.

Here is an application. The *Shi arrangement* is the arrangement of $n(n-1)$ hyperplanes in $\mathbb{R}^n$ defined by

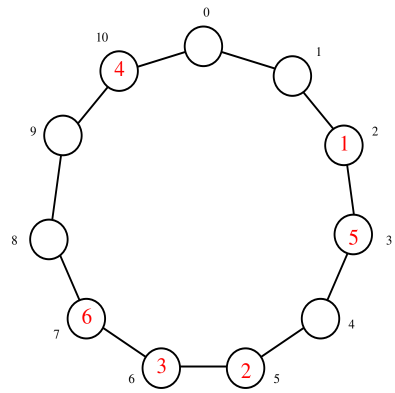
$$\mathcal{S}_n = \{x_i = x_j,\ x_i = x_j + 1 \mid 1 \le i < j \le n\}.$$

In other words, take the braid arrangement, clone it, and nudge each of the cloned hyperplanes a little bit. The Shi arrangement is not central, but every hyperplane in it contains a line parallel to the all-ones vector, so we may project along that line (just as for the braid arrangement) to obtain a combinatorially identical arrangement in $\mathbb{R}^{n-1}$. Doing this to $\mathcal{S}_2$ produces the arrangement of two points on a line; the arrangement $\mathcal{S}_3$ looks like this.



Proposition 5.19. *The characteristic polynomial of the Shi arrangement is* $\chi_{\mathcal{S}_n}(q) = q(q-n)^{n-1}$. *In particular, the numbers of regions and bounded regions are respectively* $r(\mathcal{S}_n) = |\chi(-1)| = (n+1)^{n-1}$ *and* $b(\mathcal{S}_n) = |\chi(1)| = (n-1)^{n-1}$.

Proof. It suffices to count the points in $\mathbb{F}_q^n \setminus \mathcal{S}_n$ for a large enough prime q. Let $x = (\times_1, \ldots, x_n) \in \mathbb{F}_q^n \setminus \mathcal{S}_n$. Draw a necklace with q beads labeled by the elements $0, 1, \ldots, q-1 \in \mathbb{F}_q$, and for each $i \in [n]$, put a big red $\mathbf{i}$ on the x_i bead. For example, if $q = 6$ and $x = (2, 5, 6, 10, 3, 7)$, the picture looks like this:



The condition that x avoids the hyperplanes $x_i = x_j$ says that the red numbers are all on different beads. If we read the red numbers counterclockwise, starting at the red **1** and putting in a divider sign | for each bead without a red number, we get

$$15|236||4|$$

which can be regarded as the *weak ordered set partition*

$$15, 236, \emptyset, 4, \emptyset$$

that is, a $(q-n)$-tuple $B_1, \dots, B_{q-n}$, where the B_i are pairwise disjoint sets (possibly empty) whose union is $[n]$, and $1 \in B_1$. Avoiding the hyperplanes $x_i = x_j + 1$ says that each contiguous block of beads has its red numbers in strictly increasing order counterclockwise. This correspondence is bijective (given a weak set partition, write out each block in increasing order, with bars between successive blocks). Moreover, it is easy to count weak ordered set partitions; they are just functions $f : [2, n] \to [q-n]$, where $f(i)$ is the index of the block containing i (note that $f(1)$ must equal 1), and there are $(q-n)^{n-1}$ such things. Since there is also a choice about where the red **1** goes, we obtain $|\mathbb{F}_q^n \setminus \mathcal{S}_n| = q(q-n)^{n-1}$ as desired. □

The number $(n+1)^{n-1}$ should look very suspicious — by Cayley's formula, that is the number of spanning trees of the complete graph K_{n+1}.

5.4. **Arrangements over $\mathbb{C}$.** What if $\mathcal{A} \subset \mathbb{C}^n$ is a *complex* hyperplane arrangement? Since the hyperplanes of $\mathcal{A}$ have codimension 2 as real vector subspaces, the complement $X = \mathbb{C}^n \setminus \mathcal{A}$ is connected, but not simply connected. Instead of counting regions, we can look at the topology of the complement.

Theorem 5.20 (Brieskorn 1971 [2])**.** *The homology groups $H_i(X, \mathbb{Z})$ are free abelian, and the Poincáre polynomial of X is the characteristic polynomial backwards:*

$$\sum_{i=0}^{n} \operatorname{rank}_{\mathbb{Z}} H_{2i}(X, \mathbb{Z}) q^i \;=\; (-q)^n \chi_{L(\mathcal{A})}(-1/q).$$

Orlik and Solomon [9] strengthened Brieskorn's result by giving a presentation of the cohomology ring $H^*(X, \mathbb{Z})$ in terms of $L(\mathcal{A})$, thereby proving that the cohomology is a combinatorial invariant of $\mathcal{A}$. (Brieskorn's theorem says only that the additive structure of $H^*(X, \mathbb{Z})$ is a combinatorial invariant.) By the way, the homotopy type of X is *not* a combinatorial invariant; Rybnikov [11] constructed arrangements with isomorphic lattices of flats but different fundamental groups.

5.5. **Projectivization and Coning.**

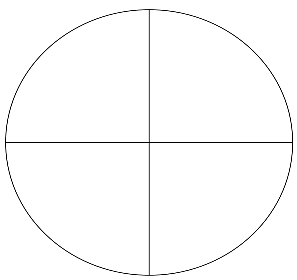
Definition 5.21. Let $\mathbb{F}$ be a field and $n \geq 1$. The set of lines through the origin in $\mathbb{F}^n$ is called n-**dimensional projective space** over $\mathbb{F}$ and denoted by $\mathbb{P}^{n-1}\mathbb{F}$.

If $\mathbb{F} = \mathbb{R}$, we can regard $\mathbb{P}^{n-1}\mathbb{R}$ as the unit sphere S^{n-1} with opposite points identified. (In particular, it is an $(n-1)$-dimensional manifold, although it is orientable only if n is even.)

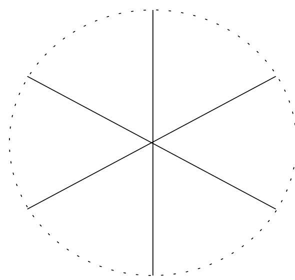
Algebraically, write $x \sim y$ if x and y are nonzero scalar multiples of each other. Then $\sim$ is an equivalence relation on $\mathbb{F}^n \setminus \{\vec{0}\}$, and $\mathbb{P}^{n-1}$ is the set of equivalence classes.

Linear hyperplanes in $\mathbb{F}^n$ correspond to affine hyperplanes in $\mathbb{P}^{n-1}\mathbb{F}$. Thus, given a central arrangement $\mathcal{A} \subset \mathbb{F}^n$, we can construct its **projectivization** $\operatorname{proj}(\mathcal{A}) \subset \mathbb{P}^{n-1}\mathbb{F}$.

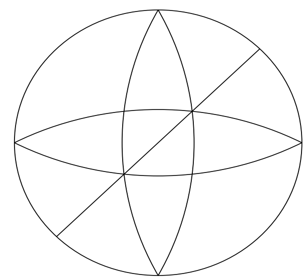
Projectivization supplies a nice way to draw central 3-dimensional real arrangements. Let S be the unit sphere, so that $H \cap S$ is a great circle for every $H \in \mathcal{A}$. Regard $H_0 \cap S$ as the equator and project the northern hemisphere into your piece of paper.



Boole3

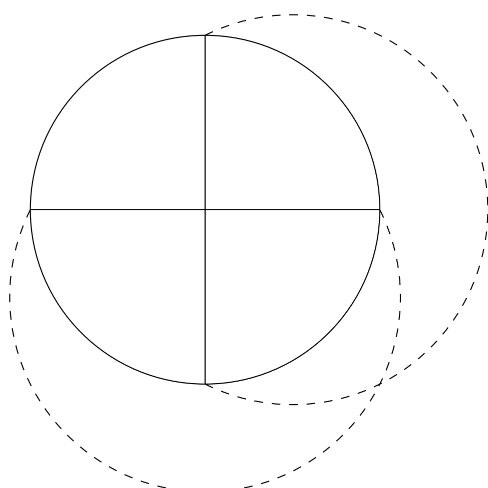


Braid3



Braid4

Of course, a diagram of $\operatorname{proj}(\mathcal{A})$ only shows the "upper half" of $\mathcal{A}$. We can recover $\mathcal{A}$ from $\operatorname{proj}(\mathcal{A})$ by "reflecting the interior of the disc to the exterior" (Stanley); e.g., for the Boolean arrangement $\mathcal{A} = \mathscr{B}_3$, the picture is as shown below. In general, $r(\operatorname{proj}(\mathcal{A})) = \frac{1}{2}r(\mathcal{A})$.



We now look at an operation that is sort of the inverse of projectivization, that lets us turn a non-central arrangement into a central arrangement (at the price of increasing the dimension by 1).

Definition 5.22. Let $\mathcal{A} \subseteq \mathbb{F}^n$ be a hyperplane arrangement, not necessarily central. The **cone** $c\mathcal{A}$ is the central arrangement in $\mathbb{F}^{n+1}$ defined as follows:

- *Geometrically:* Make a copy of $\mathcal{A}$ in $\mathbb{F}^{n+1}$, choose a point p not in any hyperplane of $\mathcal{A}$, and replace each $H \in \mathcal{A}$ with the affine span H' of p and H (which will be a hyperplane in $\mathbb{F}^{n+1}$). Then, toss in one more hyperplane containing p and in general position with respect to every H'.

- *Algebraically:* For $H = \{x \mid L(x) = a_i\} \in \mathcal{A}$ (with L a homogeneous linear form on $\mathbb{F}^n$ and $a_i \in \mathbb{F}$), construct a hyperplane $H' = \{(x_1, \ldots, x_n, y) \mid L(x) = a_i y\} \subset \mathbb{F}^{n+1}$ in $c\mathcal{A}$. Then, toss in the hyperplane $y = 0$.

For example, if $\mathcal{A}$ consists of the points $x = 0$, $x = -3$ and $x = 5$ in $\mathbb{R}^1$, then $c\mathcal{A}$ consists of the lines $x = y$, $x = -3y$, $x = 5y$, and $y = 0$ in $\mathbb{R}^2$.

Proposition 5.23. $\chi_{c\mathcal{A}}(k) = (k-1)\chi_{\mathcal{A}}(k)$.

5.6. **Graphic Arrangements.**

Definition 5.24. Let G be a simple graph on vertex set $[n]$. The **graphic arrangement** $\mathcal{A}_G \subset \mathbb{F}^n$ consists of the hyperplanes $x_i = x_j$, where ij is an edge of G.

The arrangement $\mathcal{A}_G$ is central (but not essential), so $L(\mathcal{A}_G)$ is a geometric lattice. The corresponding matroid is naturally isomorphic to the graphic matroid of G. In particular, $r(\mathcal{A}_G) = T(G; 2, 0)$ equals the number of acyclic orientations of G.

For instance, if $G = K_n$ and $\mathbb{F} = \mathbb{R}$, then $\mathcal{A} = Br_n$, which we have seen has $n!$ regions. On the other hand, the acyclic orientations of K_n are in bijection with total orderings of its vertices.

Moreover, the chromatic polynomial of G equals the characteristic polynomial of $L(\mathcal{A}_G)$.

This last fact has a concrete combinatorial interpretation. Regard a point $(x_1, \ldots, x_n) \in \mathbb{F}_q^n$ as a q-coloring of G that assigns color x_i to vertex i. Then the proper q-colorings are precisely the points of $\mathbb{F}_q^n \setminus \mathcal{A}_G$. The number of such colorings is $p(G;q)$ (the chromatic polynomial of G evaluated at q); on the other hand, by Proposition 5.18, it is also the characteristic polynomial $\chi_{\mathcal{A}_G}(q)$. Since $p(G;q) = \chi_{\mathcal{A}_G}(q)$ for infinitely many q (namely, all integer prime powers), the polynomials must be equal.

For some graphs, the chromatic polynomial factors into linear terms. Recall from Example 3.11 that

$$p(K_n;q) = q(q-1)(q-2)\cdots(q-n+1) \qquad \text{and} \qquad p(F_n;q) = q^c(q-1)^{n-c}$$

where F_n denotes any forest with n vertices and c components.

Example 5.25. Let $G = C_4$ (a cycle with four vertices and four edges), and let $\mathcal{A} = \mathcal{A}_G$. Then $L(\mathcal{A})$ is the lattice of flats of the matroid $U_3(4)$; i.e.,

$$L = \{F \subseteq [4] : \ |F| \neq 3\}$$

with $r(F) = \min(|F|, 3)$. Since the Möbius function of an element of L depends only on its rank, it is easy to check that

$$\chi_L(k) \ = \ k^3 - 4k^2 + 6k - 3 = (k-1)(k^2 - 3k + k).$$

Multiplying by $k^{\dim \mathcal{A}_L - \operatorname{rank} \mathcal{A}_L} = k^{4-3}$ gives the characteristic polynomial of $\mathcal{A}_L$, which is the chromatic polynomial of C_4:

$$\chi_{C_4}(k) \ = \ k(k-1)(k^2 - 3k + k).$$

5.7. **Supersolvable Lattices.** For which graphs does the chromatic polynomial factor into linear terms? More generally, for which arrangements $\mathcal{A}$ does the characteristic polynomial $\chi_{\mathcal{A}}(k)$ factor? A useful sufficient condition is that the intersection poset be a *supersolvable lattice*. We first need the notion of a modular element of a lattice.

Let L be a lattice. Recall from (1.4) that L is *modular* if it is ranked, and its rank function r satisfies

$$r(x) + r(y) = r(x \vee y) + r(x \wedge y)$$

for every $x, y \in L$. This is not how we first defined modular lattices, but we proved that it is an equivalent condition (Theorem 1.39).

Definition 5.26. An element $x \in L$ is a **modular element** if $r(x) + r(y) = r(x \vee y) + r(x \wedge y)$ holds for every $y \in L$. Thus L is modular if and only if every element of L is modular.

- The elements $\hat{0}$ and $\hat{1}$ are clearly modular in any lattice.
- If L is geometric, then every atom x is modular. Indeed, for $y \in L$, if $y \geq x$, then $y = x \vee y$ and $x = x \wedge y$, while if $y \not\geq x$ then $y \wedge x = \hat{0}$ and $y \vee x \gtrdot y$.
- The coatoms of a geometric lattice, however, need not be modular. Let $L = \Pi_n$; recall that Π_n has rank function $r(\pi) = n - |\pi|$. Let $x = 12|34$, $y = 13|24 \in \Pi_4$. Then $r(x) = r(y) = 2$, but $r(x \vee y) = r(\hat{1}) = 3$ and $r(x \wedge y) = r(\hat{0}) = 0$. So x is not a modular element.

Proposition 5.27. *The modular elements of Π_n are exactly the partitions with at most one nonsingleton block.*

Proof. Suppose that $\pi \in \Pi_n$ has one nonsingleton block B. For $\sigma \in \Pi_n$, let

$$X = \{C \in \sigma \mid C \cap B \neq \emptyset\}, \qquad Y = \{C \in \sigma \mid C \cap B = \emptyset\}.$$

Then

$$\pi \wedge \sigma = \Big\{C \cap B \mid C \in X\Big\} \cup \Big\{\{i\} \mid i \notin B\Big\},$$
$$\pi \vee \sigma = \left\{\bigcup_{C \in X} C\right\} \cup Y$$

so

$$\begin{aligned} |\pi \wedge \sigma| + |\pi \vee \sigma| &= (|X| + n - |B|) + (1 + |Y|) \\ &= (n - |B| + 1) + (|X| + |Y|) \;=\; |\pi| + |\sigma|, \end{aligned}$$

proving that π is a modular element.

For the converse, let B, C be nonsingleton blocks of π, then let σ have the two nonsingleton blocks $\{i,k\}, \{j,\ell\}$, where $i, j \in B$ and $k, \ell \in C$. Then $r(\sigma) = 2$ and $r(\pi \wedge \sigma) = r(\hat{0}) = 0$, but

$$r(\pi \vee \sigma) = r(\pi) + 1 < r(\pi) + r(\sigma) - r(\pi \wedge \sigma)$$

so π is not a modular element. $\square$

The usefulness of a modular element is that if one exists, we can factor the characteristic polynomial of L.

Theorem 5.28. *Let L be a geometric lattice of rank n, and let $z \in L$ be a modular element. Then*

$$\chi_L(k) \;=\; \chi_{[\hat{0},z]}(k) \cdot \left[\sum_{y:\ y \wedge z = \hat{0}} \mu_L(\hat{0}, y) k^{n - r(z) - r(y)}\right]. \tag{5.11}$$

I'll skip the proof, which uses calculation in the Möbius algebra; see Stanley, HA, pp. 50–52.

Corollary 5.29. *Let L be a geometric lattice, and let $a \in L$ be an atom. Then*
$$\chi_L(k) \;=\; (k-1) \sum_{x:\; x \not\geq a} \mu_L(\hat{0}, x) k^{r(L)-1-r(x)}.$$

(We already knew that $k-1$ had to be a factor of $\chi_L(k)$, because $\chi_L(1) = \sum_{x \in L} \mu_L(\hat{0}, x) = 0$. Still, it's nice to see it another way.)

Corollary 5.30. *Let L be a geometric lattice, and let $z \in L$ be a coatom that is a modular element. Then*
$$\chi_L(k) \;=\; (k-e)\chi_{[\hat{0},z]}(k),$$
where e is the number of atoms $a \in L$ such that $a \not\leq z$.

Example 5.31. Corollary 5.30 provides another way of calculating the characteristic polynomial of Π_n. Let z be the coatom with blocks $[n-1]$ and $\{n\}$, which is a modular element by Proposition 5.27. There are $n-1$ atoms $a \not\leq z$, namely the partitions whose nonsingleton block is $\{i, n\}$ for some $i \in [n-1]$, so we obtain
$$\chi_{\Pi_n}(k) \;=\; (k-n+1)\chi_{\Pi_{n-1}}(k)$$
and by induction
$$\chi_{\Pi_n}(k) \;=\; (k-1)(k-2)\cdots(k-n+1).$$

Let L be a geometric lattice with atoms A. Recall from (5.11) that if z is a modular element of L, then the characteristic polynomial of L factors:
$$\chi_L(k) \;=\; \chi_{[\hat{0},z]}(k) \cdot \left[\sum_{y:\; y \wedge z = \hat{0}} \mu_L(\hat{0}, y) k^{n-r(z)-r(y)} \right].$$

Of course, we can always apply this for an atom z (Corollary 5.29). But, as we've seen with Π_n, something even better happens if z is a *coatom*: we can express $\chi_L(k)$ as the product of a linear form (the bracketed sum) with the characteristic polynomial of a smaller geometric lattice, namely $[\hat{0}, z]$.

If we are extremely lucky, L will have a saturated chain of modular elements
$$\hat{0} = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_{n-1} \lessdot x_n = \hat{1}.$$
In this case, we can apply Corollary 5.30 successively with $z = x_{n-1}$, $z = x_{n-2}$, $\ldots$, $z = x_1$ to split the characteristic polynomial completely into linear factors:
$$\begin{aligned}
\chi_L(k) &= (k-e_{n-1})\chi_{[\hat{0},x_{n-1}]}(k)\\
&= (k-e_{n-1})(k-e_{n-2})\chi_{[\hat{0},x_{n-2}]}(k)\\
&= \ldots\\
&= (k-e_{n-1})(k-e_{n-2})\cdots(k-e_0),
\end{aligned}$$
where
$$\begin{aligned}
e_i &= \#\{\text{atoms } a \text{ of } [\hat{0}, x_{i+1}] \mid a \not\leq x_i\}\\
&= \#\{a \in A \mid a \leq x_{i+1},\; a \not\leq x_i\}.
\end{aligned}$$

Definition 5.32. A geometric lattice L is **supersolvable** if it has a *modular saturated chain*, that is, a saturated chain $\hat{0} = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_n = \hat{1}$ such that every x_i is a modular element. A central hyperplane arrangement $\mathcal{A}$ is called supersolvable if $L(\mathcal{A})$ is supersolvable.

- Any modular lattice is supersolvable, because every saturated chain is modular.
- Π_n is supersolvable. because we can take x_i to be the partition whose unique nonsingleton block is $[i+1]$. Thus the braid arrangement Br_n is supersolvable.

- Let $G = C_4$ (a cycle with four vertices and four edges), and let $\mathcal{A} = \mathcal{A}_G$. Then $L(\mathcal{A})$ is the lattice of flats of the matroid $U_3(4)$; i.e.,
$$L = \{F \subseteq [4] : \ |F| \neq 3\}$$
with $r(F) = \min(|F|, 3)$. This lattice is not supersolvable, because no element at rank 2 is modular. For example, let $x = 12$ and $y = 34$; then $r(x) = r(y) = 2$ but $r(x \vee y) = 3$ and $r(x \wedge y) = 0$.

Theorem 5.33. *Let $G = (V, E)$ be a simple graph. Then $\mathcal{A}_G$ is supersolvable if and only if the vertices of G can be ordered $v_1, \ldots, v_n$ such that for every $i > 1$, the set*
$$C_i \ := \ \{v_j \mid j \leq i, \ v_i v_j \in E\}$$
forms a clique in G.

I'll omit the proof, which is not too hard; see Stanley, pp. 55–57. An equivalent condition is that G is a **chordal graph**: if $C \subseteq G$ is a cycle of length ≥ 4, then some pair of vertices that are not adjacent in C are in fact adjacent in G.

By the way, it is easy to see that if G satisfies the condition of Theorem 5.33, then the chromatic polynomial $\chi(G; k)$ splits into linear factors. Consider what happens when we color the vertices in order. When we color vertex v_i, it has $|C_i|$ neighbors that have already been colored, and they all have received different colors because they form a clique. Therefore, there are $k - |C_i|$ possible colors available for v_i, and we see that
$$\chi(G; k) \ = \ \prod_{i=1}^{n} (k - |C_i|).$$

6. More Topics

6.1. **Oriented Matroids.** Let $\mathcal{A} = \{H_1, \ldots, H_n\}$ be a hyperplane arrangement in $\mathbb{R}^d$. For each i, let ℓ_i be an affine linear functional on $\mathbb{R}^n$ such that $H_i = \{\mathbf{x} \in \mathbb{R}^d \mid \ell_i(\mathbf{x}) = 0\}$. (Equivalently, $\ell_i(x) = \|\operatorname{proj}_{\mathbf{n}_i}(\mathbf{x})\|$, where $\mathbf{n}_i$ is a normal vector to H_i.)

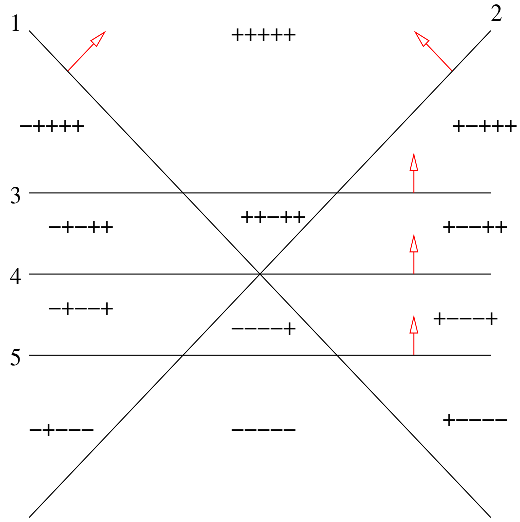
The intersections of hyperplanes in $\mathcal{A}$, together with its regions, decompose $\mathbb{R}^d$ as a *cell complex* — a disjoint union of subspaces isomorphic to Euclidean spaces of various dimensions. We can encode each cell by recording whether the linear functionals $\ell_1, \ldots, \ell_n$ are positive, negative or zero on it. Specifically, for $c = (c_1, \ldots, c_n) \in \{+, -, 0\}^n$, let
$$F = F(c) = \left\{ \mathbf{x} \in \mathbb{R}^d \;\middle|\; \begin{array}{lll} \ell_i(\mathbf{x}) > 0 & \text{if} & c_i = + \\ \ell_i(\mathbf{x}) < 0 & \text{if} & c_i = - \\ \ell_i(\mathbf{x}) = 0 & \text{if} & c_i = 0 \end{array} \right\}$$

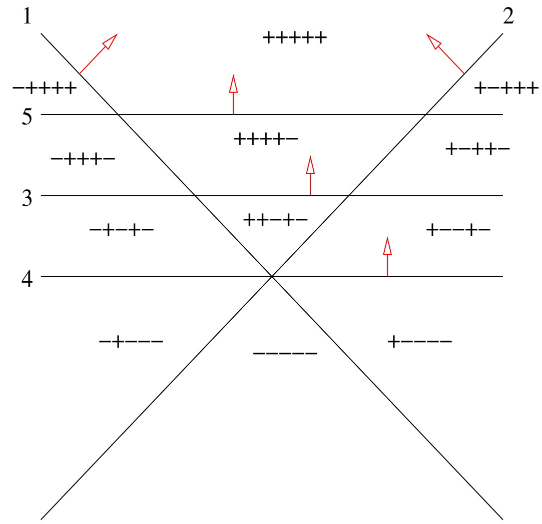
If $F \neq \emptyset$ then it is called a **face** of $\mathcal{A}$, and $c = c(F)$ is the corresponding **covector**. The set of all faces is denoted $\mathscr{F}(\mathcal{A})$. The poset $\hat{\mathscr{F}}(\mathcal{A}) = \mathscr{F}(\mathcal{A}) \cup \{\hat{0}, \hat{1}\}$ (ordered by $F \leq F'$ if $\bar{F} \subseteq \bar{F}'$) is a lattice, called the **big face lattice of** $\mathcal{A}$. This is a finer invariant than the intersection lattice.

Consider the linear forms ℓ_i that were used in representing each face by a covector. Specifying ℓ_i is equivalent to specifying a normal vector v_i to the hyperplane H_i (with $\ell_i(x) = v_i \cdot x$. As we know, the vectors v_i represent a matroid whose lattice of flats is precisely $L(\mathcal{A})$.

Scaling v_i (equivalently, ℓ_i) by a nonzero constant $\lambda \in \mathbb{R}$ has no effect on the matroid represented by the v_i's, but what does it do to the covectors? If $\lambda > 0$, then nothing happens, but if $\lambda < 0$, then we have to switch $+$ and $-$ signs in the i^{th} position of every covector. So, in order to figure out the covectors, we need not just the normal vectors v_i, but an orientation for each one.

Example: Let's go back to the two arrangements considered at the start. Their regions are labeled by the following covectors:





Now, you should object that the oriented normal vectors are the same in each case. Yes, but this couldn't happen if the arrangements were *central*, because two vector subspaces of the same space cannot possibly be parallel. In fact, if $\mathcal{A}$ is a central arrangement, then the oriented normals determine $\mathscr{F}(\mathcal{A})$ uniquely.

Proposition 6.1. *The covectors of $\mathcal{A}$ are preserved under the operation of* negation *(changing all $+$'s to $-$'s and vice versa) if and only if $\mathcal{A}$ is central. In fact, the maximal covectors that can be negated are exactly those that correspond to bounded regions.*

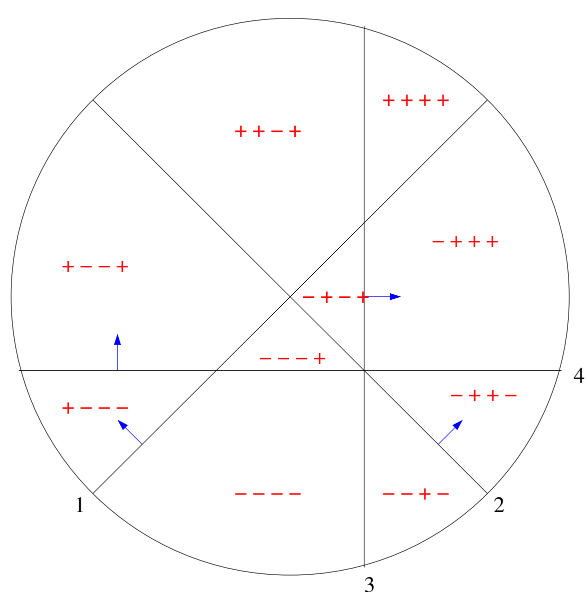
Example 6.2. Consider the central arrangement $\mathcal{A}$ whose hyperplanes are the zero sets of the linear forms

$$\ell_1 = x + y, \qquad \ell_2 = x - y, \qquad \ell_3 = x - z, \qquad \ell_1 = y + z.$$

The corresponding normal vectors are $V = \{v_1, \ldots, v_4\}$, where

$$v_1 = (1, -1, 0), \qquad v_2 = (1, 1, 0), \qquad v_3 = (1, 0, 1), \qquad v_4 = (0, 1, -1).$$

The projectivization $\operatorname{proj}(\mathcal{A})$ looks like this:



Each region F that borders the equator has a polar opposite $-F$ such that $c(-F) = -c(F)$.

The regions with covectors $---+$ and $-+-+$ do not border the equator, i.e., they are bounded in $\mathrm{proj}(\mathcal{A})$. Since they do not border the equator, neither do their opposites in $\mathcal{A}$, so those opposites do not occur in $\mathrm{proj}(\mathcal{A})$.

In the figure of Example 6.2, consider the point $p = \ell_2 \cap \ell_3 \cap \ell_4$. That three lines intersect at p means that there is a linear dependence among the corresponding normal vectors.

$$v_2 - v_3 + v_4 = 0,$$

or on the level of linear forms,

$$\ell_2 - \ell_3 + \ell_4 = 0. \tag{6.1}$$

Of course, knowing which subsets of V are linearly dependent is equivalent to knowing the matroid M represented by V. Indeed, $\{v_2, v_3, v_4\}$ is a circuit of M.

However, (6.1) tells us more than that: there exists no $x \in \mathbb{R}^3$ such that

$$\ell_2(x) > 0, \qquad \ell_3(x) < 0, \quad \text{and} \quad \ell_4(x) > 0.$$

That is, $\mathcal{A}$ has no covector of the form $*+-+$ (for any $* \in \{+,-,0\}$). We say that $0+-+$ is the corresponding **oriented circuit**.

For $c \in \{+,-,0\}^n$, write

$$c_+ = \{i \mid c_i = +\}, \qquad c_- = \{i \mid c_i = -\}.$$

Definition: Let n be a positive integer. A **circuit system for an oriented matroid** is a collection $\mathscr{C}$ of n-tuples $c \in \{+,-,0\}^n$ satisfying the following properties:

(1) $00\cdots 0 \notin \mathscr{C}$.
(2) If $c \in \mathscr{C}$, then $-c \in \mathscr{C}$.
(3) If $c, c' \in \mathscr{C}$ and $c \neq c'$, then it is not the case that both $c_+ \subset c'_+$ and $c_- \subset c'_-$
(4) If $c, c' \in \mathscr{C}$ and $c \neq c'$, and there is some i with $c_i = +$ and $c'_i = -$, then there exists $d \in \mathscr{C}$ with $d_i = 0$, and, for all $j \neq i$, $d_+ \subset c_+ \cup c'_+$ and $d_- \subset c_- \cup c'_-$.

Again, the idea is to record not just the linearly dependent subsets of a set $\{\ell_i, \ldots, \ell_n\}$ of linear forms, but also the sign patterns of the corresponding linear dependences, or "syzygies".

Condition (1) says that the empty set is linearly independent.

Condition (2) says that multiplying any syzygy by -1 gives a syzygy.

Condition (3), as in the definition of the circuit system of an (unoriented) matroid, must hold if we want circuits to record syzygies with minimal support.

Condition (4) is the oriented version of circuit exchange. Suppose that we have two syzygies

$$\sum_{j=1}^{n} \gamma_j \ell_j \;=\; \sum_{j=1}^{n} \gamma'_j \ell_j \;=\; 0,$$

with $\gamma_i > 0$ and $\gamma_i' < 0$ for some i. Multiplying by positive scalars if necessary (hence not changing the sign patterns), we may assume that $\gamma_i = -\gamma_i'$. Then

$$\sum_{j=1}^{n} \delta_j \ell_j \;=\; 0,$$

where $\delta_j = \gamma_j + \gamma_j'$. In particular, $\delta_i = 0$, and δ_j is positive (resp., negative) if and only if at least one of γ_j, γ_j' is positive (resp., negative).

- The set

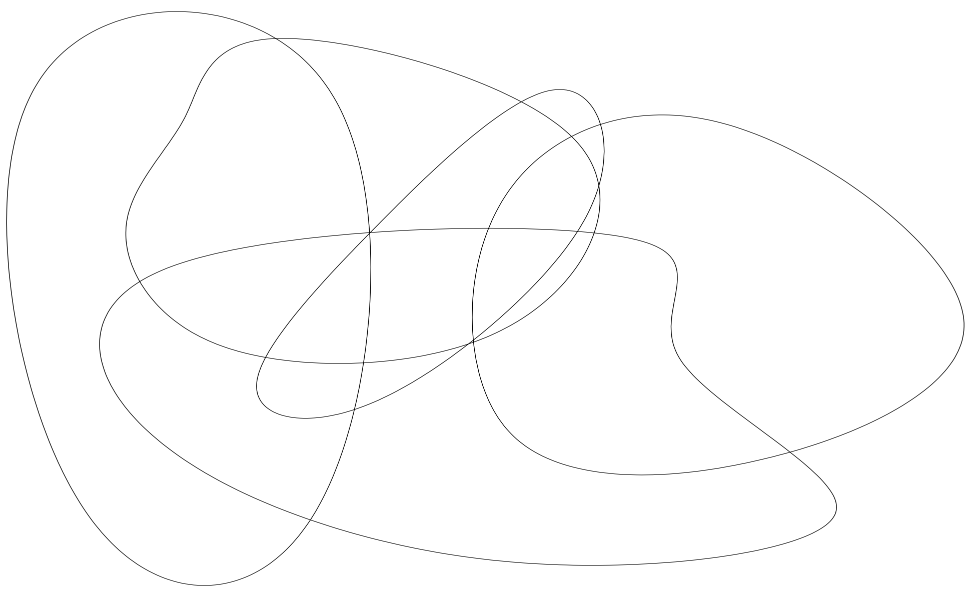
$$\{c_+ \cup c_- \mid c \in \mathscr{C}\}$$

forms a circuit system for an (ordinary) matroid.

- Just as every graph gives rise to a matroid, any loopless directed graph gives rise to an oriented matroid (homework problem!)

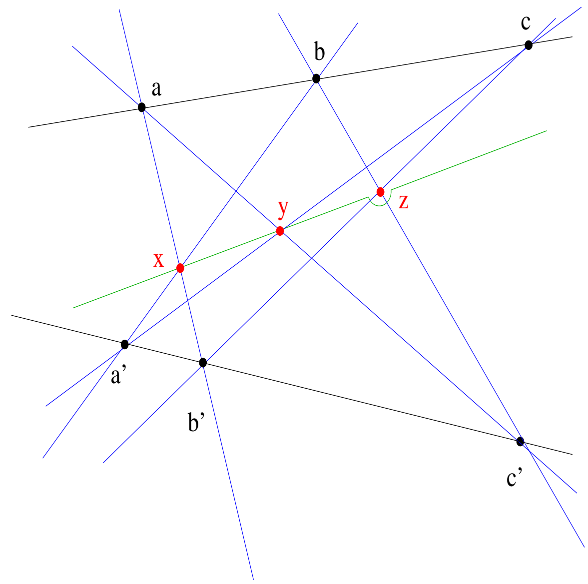
As in the unoriented setting, the circuits of an oriented matroid represent minimal obstructions to being a covector. That is, for every real hyperplane arrangement $\mathcal{A}$, we can construct a circuit system $\mathscr{C}$ for an oriented matroid such that if k is a covector of $\mathcal{A}$ and c is a circuit, then it is not the case that $k_+ \supseteq c_+$ and $k_- \supseteq c_-$.

More generally, we can construct an oriented matroid from any real *pseudosphere arrangement*, i.e., a collection of homotopy $d-1$-spheres embedded in $\mathbb{R}^n$ such that the intersection of the closures of the spheres in any subcollection is connected or empty. Here is an example of a pseudocircle arrangement in $\mathbb{R}^2$:



In fact, the Topological Representation Theorem of Folkman and Lawrence (1978) says that every oriented matroid can be represented by such a pseudosphere arrangement. However, there exist (lots of!) oriented matroids that cannot be represented as hyperplane arrangements.

Example 6.3. Recall the construction of the non-Pappus matroid (Example 2.25). If we perturb the line xyz a little bit so that it meets x and y but not z, we obtain a pseudoline arrangement whose oriented matroid $\mathcal{M}$ cannot be represented by means of a line arrangement.



6.2. **Min-Max Theorems on Posets.** A *chain cover* of a poset P is a collection of chains whose union is P. The minimum size of a chain cover is called the *width* of P.

Theorem 6.4 (Dilworth's Theorem)**.** *Let P be a finite poset. Then*

$$\mathit{width}(P) \;=\; \max\left\{s \mid P \text{ has an antichain of size } s\right\}.$$

Dilworth's Theorem can be proven as a consequence of the max-flow/min-cut theorem (one of the basic results in combinatorial optimization), but instead, here is a self-contained poset-theoretic proof.

Proof. The "$\geq$" direction is clear, because if A is an antichain, then no chain can meet A more than once, so P cannot be covered by fewer than $|A|$ chains.

For the more difficult "$\leq$" direction, we induct on $n = |P|$. The result is trivial if $n = 1$ or $n = 2$.

Let Y be the set of all minimal elements of P, and let Z be the set of all maximal elements. Note that Y and Z are both antichains. First, suppose that no set other than Y and Z is an antichain of maximum size. Dualizing if necessary, we may assume Y is maximum. Let $y \in Y$ and $z \in Z$ with $y \leq z$. Then the maximum size of an antichain in $P' = P - \{y, z\}$ is $|Y| - 1$, so by induction it can be covered with $|Y| - 1$ chains, and tossing in the chain $\{y, z\}$ gives a chain cover of P of size $|Y|$.

Now, suppose that A is an antichain of maximum size that contains neither Y nor Z as a subset. Define

$$P^+ = \{x \in P \mid x \geq a \text{ for some } a \in A\},$$
$$P^- = \{x \in P \mid x \leq a \text{ for some } a \in A\}.$$

Then

- P^+, $P^- \neq \emptyset$ (otherwise A equals Z or Y).
- $P^+ \cup P^- = P$ (otherwise A is contained in some larger antichain).
- $P^+ \cap P^- = A$ (otherwise A isn't an antichain).

So P^+ and P^- are posets smaller than P, each of which has A as a maximum antichain. By induction, each has a chain cover of size $|A|$. So for each $a \in A$, there is a chain $C_a^+ \subset P^+$ and a chain $C_a^- \subset P^-$ with

$a \in C_a^+ \cap C_a^-$, and

$$\{C_a^+ \cap C_a^- \mid a \in A\}$$

is a chain cover of P of size $|A|$. □

If we switch "chain" and "antichain", then Dilworth's theorem remains true; in fact, it becomes (nearly) trivial.

Proposition 6.5 (Trivial Proposition)**.** *In any finite poset, the minimum size of an antichain cover equals the maximum size of an chain.*

This is much easier to prove than Dilworth's Theorem.

Proof. For the $\geq$ direction, if C is a chain and $\mathcal{A}$ is an antichain cover, then no antichain in $\mathcal{A}$ can contain more than one element of C, so $|\mathcal{A}| \geq |C|$. On the other hand, let

$$A_i = \{x \in P \mid \text{ the longest chain headed by } x \text{ has length } i\};$$

then $\{A_i\}$ is an antichain cover whowe cardinality equals the length of the longest chain in P. □

These theorems have graph-theoretic consequences.

The *chromatic number* $\chi(G)$ of a graph G is the smallest number k such that G has a proper k-coloring. The *clique number* $\omega(G)$ is the largest size of a clique in G (a set of pairwise adjacent vertices). Since each vertex in a clique must be assigned a different color, it follows that

$$\chi(G) \geq \omega(G). \tag{6.2}$$

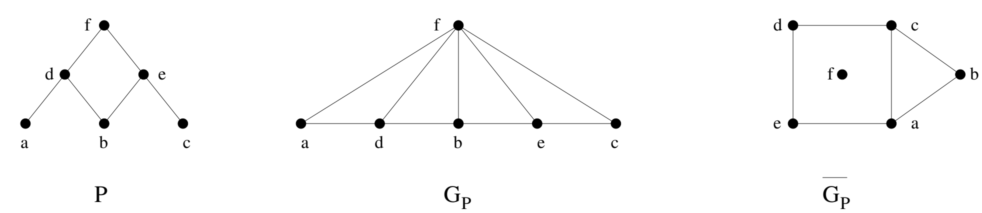
always; however, equality need not hold (for instance, for a cycle of odd length). The graph G is called **perfect** if $\omega(H) = \chi(H)$ for every induced subgraph $H \subseteq G$.

Definition 6.6. Let P be a finite poset. Its *comparability graph* G_P to be the graph G with vertices P and edges

$$\{xy \mid x \leq y \text{ or } x \geq y\}.$$

Equivalently, G_P is the underlying undirected graph of the transitive closure of the Hasse diagram of P. The *incomparability graph* $\overline{G_P}$ is the complement of G_P; that is, x, y are adjacent if and only if they are incomparable.

For example, if P is the poset whose Hasse diagram is shown on the left, then G_P is P plus the edges



A chain in P corresponds to a clique in G_P and to a coclique in $\overline{G_P}$. Likewise, an antichain in P corresponds to a coclique in G_P and to a clique in $\overline{G_P}$.

Observe that a covering of the vertex set of a graph by cocliques is exactly the same thing as a proper coloring. Therefore, the Trivial Proposition and Dilworth's Theorem say respectively that

Theorem 6.7. *Comparability and incomparability graphs of posets are perfect.*

Theorem 6.8 (Perfect Graph Theorem; Lovász 1972)**.** *Let G be a finite graph. Then G is perfect if and only if $\bar{G}$ is perfect.*

Theorem 6.9 (Strong Perfect Graph Theorem; Seymour/Chudnovsky 2002)**.** *Let G be a finite graph. Then G is perfect if and only if it has no "obvious bad counterexamples", i.e., induced subgraphs of the form C_r or $\bar{C}_r$, where $r \geq 5$ is odd.*

6.3. **The Greene-Kleitman Theorem.** There is a wonderful generalization of Dilworth's theorem due to Curtis Greene and Daniel Kleitman [6, 5].

Theorem 6.10. *Let P be a finite poset. Define two sequences of positive integers*

$$\lambda = (\lambda_1, \lambda_2, \dots), \lambda_\ell), \qquad \mu = (\mu_1, \mu_2, \dots, \mu_m)$$

by

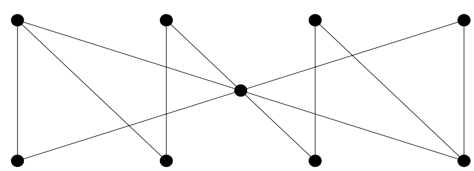
$$\lambda_1 + \cdots + \lambda_k = \max\left\{|C_1 \cup \cdots \cup C_k| :\ C_i \subseteq P \text{ chains } \right\},$$
$$\mu_1 + \cdots + \mu_k = \max\left\{|A_1 \cup \cdots \cup A_k| :\ A_i \subseteq P \text{ disjoint antichains } \right\}.$$

Then:

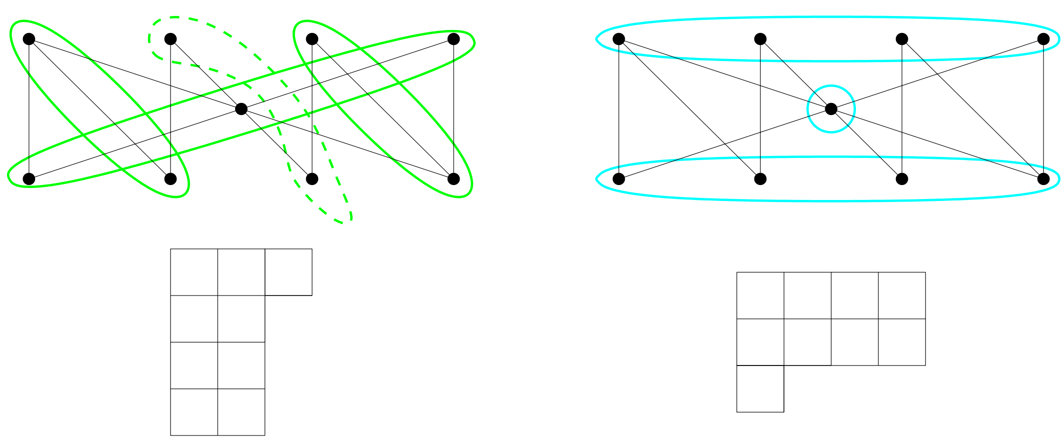
(1) *λ and μ are both* partitions *of $|P|$, i.e., weakly decreasing sequences whose sum is $|P|$.*
(2) *λ and μ are* conjugates*, i.e.,*

$$\mu_i = \#\{j \mid \lambda_j \geq i\}.$$

For example, consider the following poset:



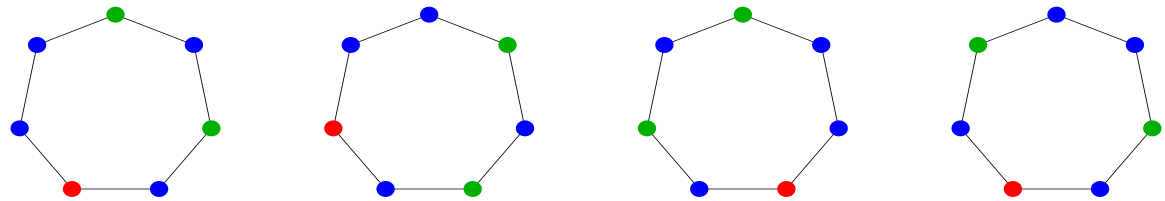
Then $\lambda = (3, 2, 2, 2)$ and $\mu = (4, 4, 1)$:



Dilworth's Theorem is now just the special case $\mu_1 = \ell$.

6.4. **Group Actions and Polyá Theory.** How many different necklaces can you make with four blue, two green, and one red bead?

It depends what "different" means. The second necklace can be obtained from the first by rotation, and the third by reflection, but the fourth one is honestly different from the first two.



If we just wanted to count the number of ways to permute four blue, two green, and one red beads, the answer would be the multinomial coefficient

$$\binom{7}{4,2,1} = \frac{7!}{4!\,2!\,1!} = 105.$$

However, what we are really trying to count is *orbits under a group action*.

Let G be a group and X a set. An **action** of G on X is a group homomorphism $\alpha : G \to \mathfrak{S}_X$, the group of permutations of X.

Equivalently, an action can also be regarded as a map $G \times X \to X$, sending (g, x) to gx, such that

- $1_G x = x$ for every $x \in X$ (where 1_G denotes the identity element of G);
- $g(hx) = (gh)x$ for every $g, h \in G$ and $x \in X$.

The *orbit* of $x \in X$ is the set

$$O_x = \{gx \mid g \in G\} \subset X$$

and its *stabilizer* is

$$S_x = \{g \in G \mid gx = x\} \subset G,$$

which is a subgroup of G.

To go back to the necklace problem, we now see that "same" really means "in the same orbit". In this case, X is the set of all 105 necklaces, and the group acting on them is the dihedral group D_7 (the group of symmetries of a regular heptagon). The number we are looking for is the number of orbits of D_7.

Lemma 6.11. *For every $x \in X$, we have $|O_x||S_x| = |G|$.*

Proof. The element gx depends only on which coset of S_x contains g, so $|O_x|$ is the number of cosets, which is $|G|/|S_x|$. □

Proposition 6.12 (Burnside's Theorem)**.** *The number of orbits of the action of G on X equals the average number of fixed points:*

$$\frac{1}{|G|} \sum_{g \in G} \#\{x \in X \mid gx = x\}$$

Proof. For a sentence P, let $\chi(P) = 1$ if P is true, or 0 if P is false (the "Garsia chi function"). Then

$$\begin{aligned}
\text{Number of orbits} &= \sum_{x \in X} \frac{1}{|O_x|} \;=\; \frac{1}{|G|} \sum_{x \in X} |S_x| \\
&= \frac{1}{|G|} \sum_{x \in X} \sum_{g \in G} \chi(gx = x) \\
&= \frac{1}{|G|} \sum_{g \in G} \sum_{x \in X} \chi(gx = x) \;=\; \frac{1}{|G|} \sum_{g \in G} \#\{x \in X \mid gx = x\}.
\end{aligned}$$

□

Typically, it is easier to count fixed points than to count orbits directly.

Example 6.13. We can apply this technique to the necklace example above.

- The identity of D_7 has 105 fixed points.
- Each of the seven reflections in D_7 has three fixed points (the single bead lying on the reflection line must be red, and then the two green beads must be equally distant from it, one on each side).
- Each of the six nontrivial rotations has no fixed points.

Therefore, the number of orbits is

$$\frac{105 + 7 \cdot 3}{|D_7|} = \frac{126}{14} = 9,$$

which is much more pleasant than trying to count them directly.

Example 6.14. Suppose we wanted to find the number of orbits of 7-bead necklaces with 3 colors, without specifying how many times each color is to be used.

- The identity element of D_7 has $3^7 = 2187$ fixed points.
- Each reflection fixes one bead, which can have any color. There are then three pairs of beads flipped, and we can specify the color of each pair. Therefore, there are $3^4 = 81$ fixed points.
- Each rotation acts by a 7-cycle on the beads, so it has only three fixed points (all the beads must have the same color).

Therefore, the number of orbits is

$$\frac{2187 + 7 \cdot 81 + 6 \cdot 3}{14} = 198.$$

More generally, the number of inequivalent 7-bead necklaces with k colors allowed is

$$\frac{k^7 + 7k^4 + 6k}{14}. \tag{6.3}$$

As this example indicates, it is helpful to look at the cycle structure of the elements of G, or more precisely on their images $\alpha(g) \in \mathfrak{S}_X$.

Proposition 6.15. *Let X be a finite set, and let $\alpha : G \to \mathfrak{S}_X$ be a group action. Color the elements of X with k colors, so that G also acts on the colorings.*

1. For $g \in G$, the number of fixed points of the action of g is $k^\ell(g)$, where $\ell(g)$ is the number of cycles in the disjoint-cycle representation of $\alpha(g)$.

2. Therefore,

$$\#\textit{equivalence classes of colorings} = \frac{1}{|G|} \sum_{g \in G} k^{\ell(g)}. \tag{6.4}$$

Let's rephrase Example 6.14 in this notation. The identity has cycle-shape 1111111 (so $\ell = 7$); each of the six reflections has cycle-shape 2221 (so $\ell = 4$); and each of the seven rotations has cycle-shape 7 (so $\ell = 1$). Thus (6.3) is an example of the general formula (6.4).

Example 6.16. How many ways are there to k-color the vertices of a tetrahedron, up to moving the tetrahedron around in space?

Here X is the set of four vertices, and the group G acting on X is the alternating group on four elements. This is the subgroup of $\mathfrak{S}_4$ that contains the identity, of cycle-shape 1111; the eight permutations of cycle-shape 31; and the three permutations of cycle-shape 22. Therefore, the number of colorings is

$$\frac{k^4 + 11k^2}{12}.$$

7. Combinatorial Algebraic Varieties

A standard reference for everything in this section is Fulton [4].

Part of the motivations for the combinatorics of partitions and tableaux comes from classical enumerative geometric questions like this:

Problem 7.1. *Let there be given four lines L_1, L_2, L_3, L_4 in $\mathbb{R}^3$ in general position. How many lines M meet each of L_1, L_2, L_3, L_4 nontrivially?*

To a combinatorialist, "general position" means "all pairs of lines are skew, and the matroid represented by four direction vectors is $U_3(4)$." To a probabilist, it means "choose the lines randomly according to some reasonable measure on the space of all lines." So, what does the space of all lines look like?

7.1. Grassmannians.

In general, if V is a vector space over a field $\mathbb{F}$ (which we will henceforth take to be $\mathbb{R}$ or $\mathbb{C}$), and $0 \leq k \leq \dim V$, then the space of all k-dimensional vector subspaces of V is called the *Grassmannian* (short for *Grassmannian variety*) and denoted by $\mathrm{Gr}(k, V)$ or $\mathrm{Gr}_{\mathbb{F}}(k, n)$ (warning: this notation varies considerably from source to source). As we'll see, $\mathrm{Gr}(k, V)$ has a lot of nice properties:

- It is a smooth manifold of dimension $k(n-k)$ over $\mathbb{F}$.
- It can be decomposed into pieces, called *Schubert cells*, each of which is naturally diffeomorphic to $\mathbb{F}^j$, for some appropriate j.
- Here's where combinatorics comes in: the Schubert cells correspond to the interval $Y_{n,k} := [\emptyset, k^{n-k}]$ in Young's lattice. (Here $\emptyset$ means the empty partition and k^{n-k} means the partition with $n-k$ parts, all of size k, so that the Ferrers diagram is a rectangle.) That is, for each partition λ there is a corresponding Schubert cell X_λ of dimension $|\lambda|$ (the number of boxes in the Ferrers diagram).
- How these cells fit together topologically is described by $Y_{n,k}$ in the following sense: the closure of X_λ is given by the formula
$$\overline{X_\lambda} = \bigcup_{\mu \leq \lambda} X_\mu$$
where $\leq$ is the usual partial order on Young's lattice (i.e., containment of Ferrers diagrams).
- Consequently, the Poincaré polynomial of $\mathrm{Gr}_{\mathbb{C}}(k, n)$ (i.e., the Hilbert series of its cohomology ring)[8] is the rank-generating function for the graded poset $Y_{n,k}$ — namely, the q-binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_q$.

To accomplish all this, we need some way to describe points of the Grassmannian. For as long as possible, we won't worry about the ground field.

[8] If these terms don't make sense, here's what you need to know. Some of you will recognize that I have omitted lots of technical details from the explanation that is about to follow — that's exactly the point.
The cohomology ring $H^*(X) = H^*(X; \mathbb{Q})$ of a space X is just some ring that is a topological invariant of X. If X is a reasonably civilized space — say, a compact finite-dimensional real or complex manifold, or a finite simplicial complex — then $H^*(X)$ is a graded ring $H^0(X) \oplus H^1(X) \oplus \cdots \oplus H^d(X)$, where $d = \dim X$, and each graded piece $H^i(X)$ is a finite-dimensional $\mathbb{Q}$-vector space. The Poincaré polynomial records the dimensions of these vector spaces as a generating function:
$$\mathrm{Poin}(X, q) = \sum_{i=0}^{d} (\dim_{\mathbb{Q}} H^i(X)) q^i.$$
For lots of spaces, this polynomial has a nice combinatorial formula. For instance, take $X = \mathbb{R}P^d$ (real projective d-space). It turns out that $H^*(X) \cong \mathbb{Q}[z]/(z^{n+1})$. Each graded piece $H^i(X)$, for $0 \leq i \leq d$, is a 1-dimensional $\mathbb{Q}$-vector space (generated by the monomial x^i), and $\mathrm{Poin}(X, q) = 1 + q + q^2 + \cdots + q^d = (1 - q^{d+1})/(1-q)$. In general, if X is a compact orientable manifold, then Poincaré duality implies (among other things) that $\mathrm{Poin}(X, q)$ is a palindrome.

Let $W \in \mathrm{Gr}(k, \mathbb{F}^n)$; that is, W is a k-dimensional subspace of $V = \mathbb{F}^n$. We can describe W as the column space of a $n \times k$ matrix M of full rank:

$$M = \begin{bmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nk} \end{bmatrix}.$$

Is the Grassmannian therefore just the space $\mathbb{F}^{n\times k}$ of all such matrices? No, because many different matrices can have the same column space. Specifically, any invertible column operation on M leaves its column space unchanged. On the other hand, every matrix whose column space is W can be obtained from M by some sequence of invertible column operations; that is, by multiplying on the right by some invertible $k \times k$ matrix. Accordingly, we can write

$$\mathrm{Gr}(k, \mathbb{F}^n) = \mathbb{F}^{n\times k}/GL_k(\mathbb{F}). \tag{7.1}$$

That is, the k-dimensional subspaces of $\mathbb{F}^n$ can be identified with the orbits of $\mathscr{F}^{n\times k}$ under the action of the general linear group $GL_k(\mathbb{F})$.

(In fact, as one should expect from (7.1),

$$\dim \mathrm{Gr}(k, \mathbb{F}^n) = \dim \mathbb{F}^{n\times k} - \dim GL_k(\mathbb{F}) = nk - k^2 = k(n-k)$$

where "dim" means dimension as a manifold over $\mathbb{F}$. Technically, this dimension calculation does not follow from (7.1) alone; you need to know that the action of $GL_k(\mathbb{F})$ on $\mathbb{F}^{n\times k}$ is suitably well-behaved. Nevertheless, we will soon be able to calculate the dimension of $\mathrm{Gr}(k, \mathbb{F}^n)$ more directly.)

Is there a canonical representative for each $GL_k(\mathbb{F})$-orbit? In other words, given $W \in \mathrm{Gr}(k, \mathbb{F}^n)$, can we find some "nicest" matrix whose column space is W? Yes: it's called reduced column-echelon form. Basic linear algebra says that we can pick any matrix with column space W and perform Gauss-Jordan elimination on its columns. We will end up with a uniquely determined matrix $M = M(W)$ with the following properties:

- $\operatorname{colspace} M = W$.
- The top nonzero entry of each column of M (the *pivot* in that column) is 1.
- Let p_i be the row in which the i^{th} column has its pivot. Then $1 \le p_1 < p_2 < \cdots < p_k \le n$.
- Every entry below a pivot of M is 0, as is every entry to the right of a pivot.
- The remaining entries of M (i.e., other than the pivots and the 0s just described) can be anything whatsoever, depending on what W was in the first place.

For example, if $n = 4$ and $k = 2$, then M will have one of the following six forms:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & * \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & * \\ 0 & * \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} * & * \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} * & * \\ 1 & 0 \\ 0 & * \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} * & * \\ * & * \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.2}$$

Note that there is only one subspace W for which M ends up with the first form. At the other extreme, if the ground field $\mathbb{F}$ is infinite and you choose the entries of M randomly (for a suitable definition of "random" — for a precise formulation, consult your local probabilist), then you will almost always end up with a matrix M^* of the last form.

Definition 7.2. Let $0 \le k \le n$ and let $\mathbf{p} = \{p_1 < \cdots < p_k\} \in \binom{[n]}{k}$ (i.e., $p_1, \dots, p_k$ are distinct elements of $[n]$, ordered least to greatest). The *Schubert cell* $\Omega^\circ_{\mathbf{p}}$ is the set of all elements $W \in \mathrm{Gr}(k, \mathbb{F}^n)$ such that, for every i, the i^{th} column of $M(W)$ has its pivot in row p_i.

Theorem 7.3. (1) *Every $W \in \mathrm{Gr}(k, \mathbb{F}^n)$ belongs to exactly one Schubert cell; that is,* $\mathrm{Gr}(k, \mathbb{F}^n)$ *is the disjoint union of the subspaces* $\Omega^\circ_{\mathbf{p}}$.

(2) *For every $\mathbf{p} \in \binom{[n]}{k}$, there is a diffeomorphism*

$$\Omega^\circ_{\mathbf{p}} \xrightarrow{\sim} \mathbb{F}^{|\mathbf{p}|}$$

where $|\mathbf{p}| = (p_1 - 1) + (p_2 - 2) + \cdots + (p_k - k) = p_1 + p_2 + \cdots + p_k - \binom{k+1}{2}$.

(3) *Define a partial order on* $\binom{[n]}{k}$ *as follows: for* $\mathbf{p} = \{p_1 < \cdots < p_k\}$ *and* $\mathbf{q} = \{q_1 < \cdots < q_k\}$, *set* $\mathbf{p} \geq \mathbf{q}$ *if* $p_i \geq q_i$ *for every* i. *Then*

$$\mathbf{p} \geq \mathbf{q} \implies \overline{\Omega^\circ_{\mathbf{p}}} \supseteq \Omega^\circ_{\mathbf{q}}. \tag{7.3}$$

(4) *The poset* $\binom{[n]}{k}$ *is isomorphic to the interval* $Y_{k,n}$ *in Young's lattice.*

(5) $\mathrm{Gr}(k, \mathbb{F}^n)$ *is a compactification of the Schubert cell* $\Omega^\circ_{(n-k+1,n-k+2,\dots,n)}$, *which is diffeomorphic to* $\mathbb{F}^{k(n-k)}$. *In particular,* $\dim_{\mathbb{F}} \mathrm{Gr}(k, \mathbb{F}^n) = k(n-k)$.

The cell closures $\Omega_{\mathbf{p}} = \overline{\Omega^\circ_{\mathbf{p}}}$ are called **Schubert varieties**.

Proof. (1) is immediate from the definition.

For (2), the map $\Omega^\circ_{\mathbf{p}} \to \mathbb{F}^{|\mathbf{p}|}$ is given by reading off the $*$s in the reduced column-echelon form of $M(W)$. (For instance, let $n = 4$ and $k = 2$. Then the matrix representations in (7.2) give explicit diffeomorphisms of the Schubert cells of $\mathrm{Gr}(k, \mathbb{F}^n)$ to $\mathbb{C}^0$, $\mathbb{C}^1$, $\mathbb{C}^2$, $\mathbb{C}^2$, $\mathbb{C}^3$, $\mathbb{C}^4$ respectively.) The number of $*$s in the i-th column is $p_i - i$ ($p_i - 1$ entries above the pivot, minus $i - 1$ entries to the right of previous pivots), so the total number of $*$s is $|\mathbf{p}|$.

For (3): This is best illustrated by an example. Consider the second matrix in (7.2):

$$M = \begin{bmatrix} 1 & 0 \\ 0 & z \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

where I have replaced the entry labeled $*$ by a parameter z. Here's the trick: Multiply the second column of this matrix by the scalar $1/z$. Doing this doesn't change the column span, i.e.,

$$\operatorname{colspace} M = \operatorname{colspace} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1/z \\ 0 & 0 \end{bmatrix}.$$

Therefore, it makes sense to say that

$$\lim_{|z| \to \infty} \operatorname{colspace} M = \operatorname{colspace} \lim_{|z| \to \infty} M = \operatorname{colspace} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

which is the first matrix in (7.2). Therefore, the Schubert cell $\Omega^\circ_{1,2}$ is in the closure of the Schubert cell $\Omega^\circ_{1,3}$. In general, decrementing a single element of $\mathbf{p}$ corresponds to taking a limit of column spans in this way, so the covering relations in the poset $\binom{[n]}{k}$ give containment relations of the form (7.3).

For (4), the elements of $Y_{k,n}$ are partitions $\lambda = (\lambda_1, \dots, \lambda_k)$ such that $n - k \geq \lambda_1 > \cdots > \lambda_k \geq 0$. The desired poset isomorphism is $\mathbf{p} \mapsto (p_k - k, p_{k-1} - (k-1), \dots, p_1 - 1)$.

(5) now follows because $\mathbf{p} = (n-k+1, n-k+2, \dots, n)$ is the unique maximal element of $\binom{[n]}{k}$, and an easy calculation shows that $|\mathbf{p}| = k(n-k)$. $\square$

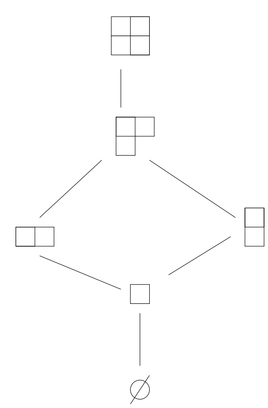
This theorem amounts to a description of $\mathrm{Gr}(k, \mathbb{F}^n)$ as a cell complex. (If you have not heard the term "cell complex" before, now you know what it means: a topological space that is the disjoint union of cells — that is, of copies of vector spaces — such that the closure of every cell is itself a union of cells.) Furthermore, the poset isomorphism with $Y_{n,k}$ says that for every i, the number of cells of $\mathrm{Gr}(k, \mathbb{F}^n)$ of dimension i is precisely the number of Ferrers diagrams with i blocks that fit inside k^{n-k} (recall that this means a $k \times (n-k)$ rectangle). Combinatorially, the best way to express this equality is this:

$$\sum_i \text{(number of Schubert cells of dimension } i\text{)}\ q^i \;=\; \sum_i \#\{\lambda \subseteq k^{n-k}\}\ q^i \;=\; \begin{bmatrix} n \\ k \end{bmatrix}_q .$$

(For those of you who know some algebraic topology: Suppose that $\mathbb{F} = \mathbb{C}$. Then $\mathrm{Gr}(k, \mathbb{F}^n)$ is a cell complex with no odd-dimensional cells (because, topologically, the dimension of cells is measured over $\mathbb{R}$). Therefore, in cellular homology, all the boundary maps are zero — because for each one, either the domain or the range is trivial — and so the homology groups are exactly the chain groups. So the Poincaré series of $\mathrm{Gr}(k, \mathbb{F}^n)$ is exactly the generating function for the dimensions of the cells. If $\mathbb{F} = \mathbb{R}$, then things are not nearly this easy — the boundary maps aren't necessarily all zero, and the homology can be more complicated.)

Example: If $k = 1$, then $\mathrm{Gr}(1, n)$ is the space of lines through the origin in $\mathbb{F}^n$; that is, projective space $\mathbb{F}P^{n-1}$. As a cell complex, this has one cell of every dimension; for instance, the projective plane consists of a 2-cell, the 1-cell and an 0-cell, i.e., a plane, a line and a point. In the standard geometric picture, the 1-cell and 0-cell together form the "line at infinity". Meanwhile, the interval $Y_{n,k}$ is a chain of rank $n-1$. Its rank-generating function is $1 + q + q^2 + \cdots + a^{n-1}$, which is the Poincaré polynomial of $\mathbb{R}P^{n-1}$. (For $\mathbb{F} = \mathbb{C}$, double the dimensions of all the cells, and substitute q^2 for q in the Poincaré polynomial.)

Example: If $n = 4$ and $k = 2$, then the interval in Young's lattice looks like this:



These correspond to the six matrix-types in (7.2). The rank-generating function is

$$\begin{bmatrix} 4 \\ 2 \end{bmatrix}_q \;=\; \frac{(1-q^4)(1-q^3)}{(1-q^2)(1-q)} \;=\; 1 + q + 2q^2 + q^3 + q^4.$$

Remark 7.4. What does all this have to do with enumerative geometry questions such as Problem 7.1? The answer (modulo technical details) is that the cohomology ring $H^*(X)$ encodes intersections of subvarieties[9] of X: for every subvariety $Z \subseteq \mathrm{Gr}(k, \mathbb{F}^n)$ of codimension i, there is a corresponding element $[Z] \in H^i(X)$ (the "cohomology class of Z") such that $[Z \cup Z'] = [Z] + [Z']$ and $[Z \cap Z'] = [Z][Z']$. These equalities hold only if Z and Z' are in general position with respect to each other (whatever that means), but the consequence is that Problem 7.1 reduces to a computation in $H^*(\mathrm{Gr}(k, \mathbb{F}^n))$: find the cohomology class $[Z]$ of the subvariety

$$Z = \{W \in \mathrm{Gr}(2, \mathbb{C}^4) \mid W \text{ meets some plane in } \mathbb{C}^4 \text{ nontrivially}\}$$

and compare $[Z]^4$ to the cohomology class $[\bullet]$ of a point. In fact, $[Z]^4 = 2[\bullet]$; this says that the answer to Problem 7.1 is (drum roll, please) **two**, which is hardly obvious! To carry out this calculation, one needs

[9] If you are more comfortable with differential geometry than algebraic geometry, feel free to think "submanifold" instead of "subvariety".

to calculate an explicit presentation of the ring $H^*(\mathrm{Gr}(k,\mathbb{F}^n))$ as a quotient of a polynomial ring (which requires the machinery of line bundles and Chern classes, but that's another story) and then figure out how to express the cohomology classes of Schubert cells with respect to that presentation. This is the theory of *Schubert polynomials.*

7.2. **Flag varieties.** There is a corresponding theory for the **flag variety**, which is the set $F\ell(n)$ of nested chains of vector spaces

$$F_\bullet = (0 = F_0 \subset F_1 \subset \cdots \subset F_n = \mathbb{F}^n)$$

or equivalently saturated chains in the (infinite) lattice $L_n(\mathbb{F})$. The flag variety is in fact a smooth manifold over $\mathbb{F}$ of dimension $\binom{n}{2}$. Like the Grassmannian, it has a decomposition into Schubert cells X_w°, which are indexed by permutations $w \in \mathfrak{S}_n$ rather than partitions,as we now explain.

For every flag $F_\bullet$, we can find a vector space basis $\{v_1,\dots,v_n\}$ for $\mathbb{F}^n$ such that $F_k = \mathrm{span}\{F_1,\dots,F_k\}$ for all k, and represent $F_\bullet$ by the invertible matrix $M \in G = GL(n,\mathbb{F})$ whose columns are $v_1,\dots,v_n$. OTOH, any ordered basis of the form

$$v_1' = b_{11}v_1, \quad v_2' = b_{12}v_1 + b_{22}v_2, \quad \dots, v_n' = b_{1n}v_1 + b_{2n}v_2 + \cdots + b_{nn}v_n,$$

where $b_{kk} \neq 0$ for all k, defines the same flag. That is, a flag is a coset of B in G, where B is the subgroup of invertible upper-triangular matrices (the *Borel subgroup*). Thus the flag variety can be (and often is) regarded as the quotient G/B. This immediately implies that it is an irreducible algebraic variety (as G is irreducible, and any image of an irreducible variety is irreducible). Moreover, it is smooth (e.g., because every point looks like every other point, and so either all points are smooth or all points are singular and the latter is impossible) and its dimension is $(n-1)+(n-2)+\cdots+0 = \binom{n}{2}$.

As in the case of the Grassmannian, there is a canonical representative for each coset of B, obtained by Gaussian elimination, and reading off its pivot entries gives a decomposition

$$F\ell(n) = \coprod_{w \in \mathfrak{S}_n} X_w^\circ.$$

Here the dimension of a Schubert cell X_w° is the number of **inversions** of w, i.e.,

$$\#\{(i,j):\ i<j \text{ and } w(i) > w(j)\}.$$

Recall that this is the rank function of the Bruhat and weak Bruhat orders on $\mathfrak{S}_n$. In fact, the (strong) Bruhat order is the cell-closure partial order (analogous to (7.3)). It follows that the Poincaré polynomial of $F\ell(n)$ is the rank-generating function of Bruhat order, namely

$$(1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}).$$

More strongly, it can be shown that the cohomology ring $H^*(F\ell(n);\mathbb{Z})$ is the quotient of $\mathbb{Z}[x_1,\dots,\times_n]$ by the ideal generated by symmetric function (coming soon).

The Schubert varieties in $F\ell(n)$ are

$$X_w = \overline{X_w^\circ} = \bigcup_{v \in \mathfrak{S}_n:\ v \leq w} X_v^\circ.$$

These are much-studied objects in combinatorics; for example, determining which Schubert varieties is singular turns out to to be a combinatorial question involving the theory of *pattern avoidance.*

Even more generally, instead of $\mathfrak{S}_n$, start with any finite Coxeter group G (roughly, a group generated by elements of order two — think of them as reflections). Then G has a combinatorially well-defined partial order also called the Bruhat order, and one can construct a G-analogue of the flag variety: that is, a smooth manifold whose structure as a cell complex is given by Bruhat order on G.

8. Group Representations

Definition 8.1. Let G be a group (typically finite) and let $V \cong \mathbb{F}^n$ be a finite-dimensional vector space over a field $\mathbb{F}$. A **representation of** G **on** V is a group homomorphism $\rho : G \to GL(V)$. That is, for each $g \in G$ there is an invertible $n \times n$ matrix $\rho(g)$, satisfying

$$\rho(g)\rho(h) = \rho(gh) \qquad \forall g, h \in G.$$

(That's matrix multiplication on the left side of the equation, and group multiplication in G on the right.) The number n is called the **dimension** (or **degree**) of the representation.

- ρ specifies an action of G on V that respects its vector space structure.
- We often abuse terminology by saying that ρ is a representation, or that V is a representation, or that the pair (ρ, V) is a representation.
- ρ is a **permutation representation** if $\rho(g)$ is a permutation matrix for all $g \in G$.
- ρ is **faithful** if it is injective as a group homomorphism.
- (For algebraists:) A representation of G is the same thing as a left module over the group algebra $\mathbb{F}G$.

Example 8.2. For any group G, the **trivial representation** is defined by $\rho_{\text{triv}}(g) = I_n$ (the $n \times n$ identity matrix).

Example 8.3 (The regular representation)**.** Let G be a finite group with n elements, and let $\mathbb{F}G$ be the vector space of formal $\mathbb{F}$-linear combinations of elements of G: that is,

$$\mathbb{F}G = \left\{ \sum_{h \in G} a_h h \mid a_h \in \mathbb{F} \right\}.$$

Then there is a representation ρ_{reg} of G on $\mathbb{F}G$, called the **regular representation**, defined by

$$g\left(\sum_{h \in G} a_h h\right) = \sum_{h \in G} a_h (gh).$$

That is, g permutes the standard basis vectors of $\mathbb{F}G$ according to the group multiplication law.

Example 8.4 (The defining representation of $\mathfrak{S}_n$)**.** Let $G = \mathfrak{S}_n$, the symmetric group on n elements. Then we can represent each permutation $\sigma \in G$ by the permutation matrix with 1's in the positions $(i, \sigma(i))$ for every $i \in [n]$, and 0's elsewhere. For instance, the permutation $4716253 \in \mathfrak{S}_7$ corresponds to the permutation matrix

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 8.5. Let $G = \mathbb{Z}/k\mathbb{Z}$ be the cyclic group of order k, and let ζ be a k^{th} root of unity (not necessarily primitive). Then G has a 1-dimensional representation given by $\rho(x) = \zeta^x$.

Example 8.6. Let G act on a finite set X. Then there is an associated permutation representation on $\mathbb{F}^X$, the vector space with basis X, given by

$$\rho(g)\left(\sum_{x \in X} a_x x\right) = \sum_{x \in X} a_x (g \cdot x).$$

For instance, the action of G on itself by left multiplication gives rise in this way to the regular representation, and the usual action of $\mathfrak{S}_n$ on an n-element set gives rise to the defining representation.

Example 8.7. Let $G = D_n$, the dihedral group of order $2n$, i.e., the group of symmetries of a regular n-gon, given in terms of generators and relations by

$$\langle s, r \ : \ s^2 = r^n = 1, \ srs = r^{-1} \rangle.$$

There are several associated representations of G.

First, we can regard s as a reflection and r as a rotation on $\mathbb{R}^2$:

$$\rho(s) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \rho(r) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

(where $\theta = 2\pi/n$). This is a faithful 2-dimensional representation.

Second, conside the permutation representations of G on vertices or on edges. These are faithful n-dimensional representations.

What about diameters? If n is odd, then we have another faithful n-dimensional representation. OTOH, if n is even, the action on diameters is $(2n)$-dimensional but not faithful — the map $r^{n/2}$, which rotates by $180°$, preserves all diameters.

Example 8.8. The symmetric group $\mathfrak{S}_n$ has a nontrivial 1-dimensional representation, the **sign representation**, given by

$$\rho_{\text{sign}}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{if } \sigma \text{ is odd.} \end{cases}$$

Note that $\rho_{\text{sign}}(g) = \det \rho_{\text{def}}(g)$, where ρ_{def} is the defining representation of $\mathfrak{S}_n$. In general, if ρ is any representation, then $\det \rho$ is a 1-dimensional representation.

Example 8.9. Let (ρ, V) and (ρ', V') be representations of G, where $V \cong \mathbb{F}^n$, $V' \cong \mathbb{F}^m$. The **direct sum** $\rho \oplus \rho' : G \to GL(V \oplus V')$ is defined by

$$(\rho \oplus \rho')(g)(v + v') = \rho(g)(v) + \rho'(g)(v')$$

for $v \in V$, $v' \in V'$. In terms of matrices, $(\rho \oplus \rho')(g)$ is a block-diagonal matrix

$$\left[\begin{array}{c|c} \rho(g) & 0 \\ \hline 0 & \rho'(g) \end{array}\right].$$

8.1. **Isomorphisms and Homomorphisms.** When two representations are the same? More generally, what is a map between representations?

Definition 8.10. Let (ρ, V) and (ρ', V') be representations of G. A linear transformation $\phi : V \to V'$ is **G-equivariant** if $g\phi = \phi g$. (Abusing notation as usual, we might write $\phi : \rho \to \rho'$.)

Equivalently, $g \cdot \phi(v) = \phi(g \cdot v)$ for all $g \in G$, $v \in V$. [Or, more precisely if less concisely: $\rho'(g) \cdot \phi(v) = \phi(\rho(g) \cdot v)$.]

Equivalently, the following diagram commutes for all $g \in G$:

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "g"] & V' \arrow[d, "g"] \\ V \arrow[r, "\phi"] & V' \end{tikzcd}$$

In the language of modules, G-equivariant transformations are just $\mathbb{F}G$-module homomorphisms.

Example 8.11. Let n be odd, and consider the dihedral group D_n acting on a regular n-gon. Label the vertices $1, \ldots, n$ in cyclic order. Label each edge the same as its opposite vertex. Then the permutation action on vertices is identical to that on edges. In other words, the diagram

$$\begin{tikzcd} \text{vertices} \arrow[r, "\text{opp}"] \arrow[d, "g"] & \text{edges} \arrow[d, "g"] \\ \text{vertices} \arrow[r, "\text{opp}"] & \text{edges} \end{tikzcd}$$

commutes for all $g \in D_n$.

Example 8.12. One way in which G-equivariant transformations occur is when an action "naturally" induces another action. For instance, consider the permutation action of $\mathfrak{S}_4$ on the vertices of K_4, which induces a 4-dimensional representation ρ_v. However, this action naturally determines an action on the six edges of K_4, which in turn induces a 6-dimensional representation ρ_e. This is to say that there is a G-equivariant transformation $\rho_v \to \rho_e$.

$$\begin{tikzcd} \text{vertex space} \arrow[r] \arrow[d, "\rho_v"] & \text{edge space} \arrow[d, "rho_e"] \\ \text{vertex space} \arrow[r] & \text{edge space} \end{tikzcd}$$

Definition 8.13. Two representations (ρ, V) and (ρ', V') of G are **isomorphic** if there is a G-equivariant map $\phi : V \to V'$ that is a vector space isomorphism.

Example 8.14. Let $\mathbb{F}$ be a field of characteristic $\neq 2$, and let $V = \mathbb{F}^2$, with standard basis $\{e_1, e_2\}$. Let $G = \mathfrak{S}_2 = \{12, 21\}$. The defining representation $\rho = \rho_{\text{def}}$ of G on V is given by

$$\rho(12) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \rho(21) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

On the other hand, the representation $\sigma = \rho_{\text{triv}} \oplus \rho_{\text{sign}}$ is given on V by

$$\sigma(12) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \sigma(21) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

These two representations are in fact isomorphic. Indeed, ρ acts trivially on $\operatorname{span}\{e_1 + e_2\}$ and acts by -1 on $\operatorname{span}\{e_1 - e_2\}$. Since these two vectors form a basis of V, one can check that

$$\phi = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

is an isomorphism $\rho \to \sigma$.

Our goal is to classify representations up to isomorphism. As we will see, we can do this without having to worry about every coordinate of every matrix $\rho(g)$ — all we really need to know is the *trace* of $\rho(g)$, known as the **character** of a representation. For instance, in this last example, we can detect the isomorphism $\rho \cong \sigma$ by observing that

$$\operatorname{tr}(\rho(12)) = \operatorname{tr}(\sigma(12)) = 2, \qquad \operatorname{tr}(\rho(21)) = \operatorname{tr}(\sigma(21)) = 0.$$

8.2. **Irreducibility, Indecomposability and Maschke's Theorem.**

Definition 8.15. Let (ρ, V) be a representation of G. A vector subspace $W \subset V$ is G**-invariant** if $\rho(g)W \subset W$ (equivalently, if W is a G-submodule of V). V is **irreducible** (or **simple**, or colloquially an "irrep") if it has no proper G-invariant subspace. A representation that can be decomposed into a direct sum of irreps is called **semisimple**.

For instance, any 1-dimensional representation is clearly irreducible.

Definition 8.16. The representation V is **decomposable** if there are G-invariant subspaces $W, W^\perp$ with $W \cap W^\perp = 0$ and $W + W^\perp = V$. Otherwise, V is **indecomposable**.

In general, we expect representations to have invariant subspaces, so being irreducible is a pretty strong condition. On the other hand, we would like to be able to decompose an arbitrary representation as a direct sum of irreps.

Clearly every representation can be written as the direct sum of indecomposable representations, and "irreducible" implies "indecomposable." But the converse is not true in general:

Example 8.17. Let $V = \{e_1, e_2\}$ be the standard basis for $\mathbb{F}^2$. Recall that the defining representation of $\mathfrak{S}_2 = \{12, 21\}$ is given by

$$\rho_{\text{def}}(12) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \rho_{\text{def}}(21) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and that

$$\rho_{\text{def}}(g)(e_1 + e_2) = \rho_{\text{triv}}(g)(e_1 + e_2), \qquad \rho_{\text{def}}(g)(e_1 - e_2) = \rho_{\text{sign}}(g)(e_1 - e_2).$$

Therefore, as we saw last time, the change of basis map

$$\phi = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}^{-1}$$

is a G-equivariant isomorphism between ρ_{def} and $\rho_{\text{triv}} \oplus \rho_{\text{sign}}$ — **unless $\mathbb{F}$ has characteristic 2**, in which case the matrix ϕ is not invertible and we cannot find a basis of G-eigenvectors. For instance, if $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ then W is the only G-invariant subspace of V, and consequently ρ_{def} is not semisimple.

Note that the space spanned by $e_1 - e_2$ is the orthogonal complement $W^\perp$ under the standard scalar product on $\mathbb{F}^2$. In general, the orthogonal complement of a G-invariant subspace ought to be G-invariant. As we will see in the proof of Maschke's Theorem, we don't need a scalar product per se: we can define "orthogonal complement" effectively just using the representation itself.

Fortunately, this kind of pathology does not happen in characteristic 0. Indeed, something stronger is true.

Theorem 8.18 (Maschke's Theorem)**.** *Let G be a finite group, and let $\mathbb{F}$ be a field whose characteristic does not divide $|G|$. Then every representation $\rho : G \to GL(V)$ is* completely reducible*, that is, every G-invariant subspace has an invariant complement.*

Proof. If ρ is an irreducible representation, then there is nothing to prove. Otherwise, let W be a G-invariant subspace, and let

$$\pi : V \to W$$

be any projection (i.e., a surjective $\mathbb{F}$-linear transformation that fixes the elements of W pointwise). (E.g., choose a basis for W, extend it to a basis for V, and tell π to fix all the basis elements in W and kill all the ones in $V \setminus W$.)

Note that we are *not* assuming that π is G-equivariant, merely that it is $\mathbb{F}$-linear. The trick is to construct a G-equivariant projection by "averaging π over the action of G". Specifically, define a map $\pi_G : V \to V$ by

$$\pi_G(v) = \frac{1}{|G|} \sum_{g \in G} g\pi(g^{-1}v). \tag{8.1}$$

Note first that $\pi_G(v) \in W$ for all $v \in V$, because $\pi(g^{-1}v) \in W$ and W is G-invariant. Second, if $w \in W$, then $g^{-1}w \in W$, so

$$\pi_G(w) = \frac{1}{|G|} \sum_{g \in G} gg^{-1}w = w.$$

Therefore π_G is also a projection $V \to W$.

Moreover, for $h \in G$, we have

$$\begin{aligned}\pi_G(hv) &= \frac{1}{|G|}\sum_{g\in G} g\pi(g^{-1}hv)\\ &= \frac{1}{|G|}\sum_{k\in G:\ hk=g}(hk)\pi((hk)^{-1}hv)\\ &= \frac{1}{|G|}h\sum_{k\in G}k\pi(k^{-1}v) \quad = \quad h\pi_G(v),\end{aligned}$$

which says that π_G is G-equivariant.

Now, define $W^\perp = \ker \pi_G$. Certainly $V \cong W \oplus W^\perp$ as vector spaces, and by G-equivariance, if $v \in W^\perp$ and $g \in G$, then $\pi_G(gv) = g\pi_G(v) = 0$, i.e., $gv \in W^\perp$. That is, $W^\perp$ is G-invariant. □

If the characteristic of $\mathbb{F}$ does divide $|G|$, then the problem is that the map $v \mapsto \sum_{g\in G} g\pi(g^{-1}v)$ will kill everything in W instead of preserving it.

Maschke's Theorem implies that a representation ρ is determined up to isomorphism by the multiplicity of each irreducible representation in ρ. Accordingly, to understand representations of G, we should first study irreps.

By the way, implicit in the proof is the following useful fact:

Proposition 8.19. *Any G-equivariant map has a G-equivariant kernel and G-equivariant image.*

8.3. **Characters.**

Definition 8.20. Let (ρ, V) be a representation of G over $\mathbb{F}$. Its *character* is the function $\chi_\rho : G \to \mathbb{F}$ given by

$$\chi_\rho(g) = \operatorname{tr}\rho(g).$$

Example 8.21. Some simple facts and some characters we've seen before:

(1) A one-dimensional representation is its own character.
(2) For any representation ρ, we have $\chi_\rho(1) = \dim \rho$, because $\rho(1)$ is the $n \times n$ identity matrix.
(3) The defining representation ρ_{def} of $\mathfrak{S}_n$ has character

$$\chi_{\text{def}}(\sigma) = \text{number of fixed points of } \sigma.$$

(4) The regular representation ρ_{reg} has character

$$\chi_{\text{reg}}(\sigma) = \begin{cases}|G| & \text{if } \sigma = 1_G\\ 0 & \text{otherwise.}\end{cases}$$

Example 8.22. Consider the two-dimensional representation ρ of the dihedral group $D_n = \langle r, s \mid r^n = s^2 = 0, srs = r^{-1}\rangle$ by rotations and reflections:

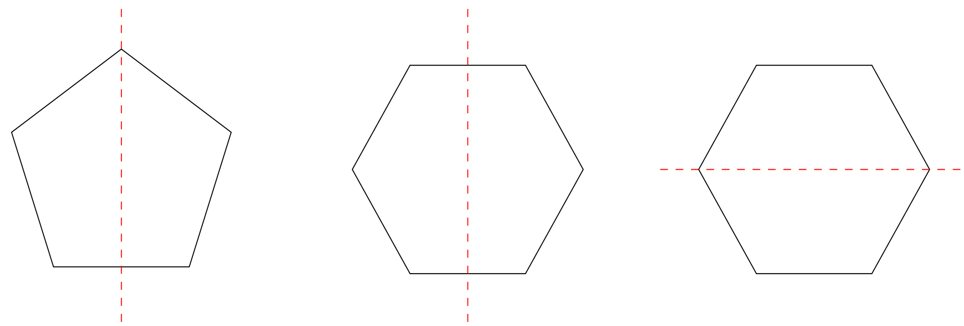
$$\rho(s) = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \qquad \rho(r) = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}$$

Its character is

$$\chi_\rho(r^j) = 2\cos i\theta \quad (0 \le j < n), \qquad \chi_\rho(sr^j) = 0 \quad (0 \le j < n).$$

On the other hand, if ρ' is the n-dimensional permutation representation on the vertices, then

$$\chi_{\rho'}(g) = \begin{cases} n & \text{if } g = 1, \\ 0 & \text{if } g \text{ is a nontrivial rotation,} \\ 1 & \text{if } n \text{ is odd and } g \text{ is a reflection,} \\ 0 & \text{if } n \text{ is even and } g \text{ is a reflection through two edges,} \\ 2 & \text{if } n \text{ is even and } g \text{ is a reflection through two vertices.} \end{cases}$$



Proposition 8.23. *Characters are* class functions*; that is, they are constant on conjugacy classes of G. Moreover, if $\rho \cong \rho'$, then $\chi_\rho = \chi_{\rho'}$.*

Proof. Recall from linear algebra that $\operatorname{tr}(ABA^{-1}) = \operatorname{tr}(B)$ in general. Therefore,

$$\operatorname{tr}\left(\rho(hgh^{-1})\right) = \operatorname{tr}\left(\rho(h)\rho(g)\rho(h^{-1})\right) = \operatorname{tr}\left(\rho(h)\rho(g)\rho(h)^{-1}\right) = \operatorname{tr}\rho(g).$$

For the second assertion, let $\phi : \rho \to \rho'$ be an isomorphism, i.e., $\phi \cdot \rho(g) = \rho'(g) \cdot \phi$ for all $g \in G$ (treating ϕ as a matrix in this notation). Since ϕ is invertible, we have therefore $\phi \cdot \rho(g) \cdot \phi^{-1} = \rho'(g)$. Now take traces. □

What we'd really like is the converse of this second assertion. In fact, much, much more is true. From now on, we consider only representations over $\mathbb{C}$.

Theorem 8.24 (Fundamental Theorem of Representaion Theory)**.** *Let G be any finite group.*

(1) *If $\chi_\rho = \chi_{\rho'}$, then $\rho \cong \rho'$. That is, a representation is determined up to isomorphism by its character.*
(2) *The characters of irreducible representations form a basis for the vector space $C\ell(G)$ of all class functions of G. Moreover, this basis is orthonormal with respect to the natural Hermitian inner product defined by*

$$\langle f, f' \rangle_G = \frac{1}{|G|} \sum_{g \in G} \overline{f(g)} f'(g).$$

(The bar denotes complex conjugate.)
(3) *As a consequence, the number of different irreducible representations of G equals the number of conjugacy classes.*
(4) *The regular representation ρ_{reg} satisfies*

$$\rho_{\text{reg}} \cong \bigoplus_{\text{irreps } \rho} \rho^{\oplus \dim \rho}$$

so in particular

$$|G| = \sum_{\text{irreps } \rho} (\dim \rho)^2.$$

We will prove this theorem bit by bit over the next several sections.

Example 8.25. The group $G = \mathfrak{S}_3$ has three conjugacy classes, determined by cycle shapes:

$$C_1 = \{1_G\}, \qquad C_2 = \{(12),\ (13),\ (23)\}, \qquad C_3 = \{(123),\ (132)\}.$$

We'll notate a character χ by the bracketed triple $[\chi(C_1), \chi(C_2), \chi(C_3)]$.

We know two irreducible 1-dimensional characters of $\mathfrak{S}_3$, namely the trivial character $\chi_{\mathrm{triv}} = [1,1,1]$ and the sign character $\chi_{\mathrm{sign}} = [1,-1,1]$.

Note that

$$\langle \chi_{\mathrm{triv}},\ \chi_{\mathrm{triv}} \rangle = 1, \qquad \langle \chi_{\mathrm{sign}},\ \chi_{\mathrm{sign}} \rangle = 1, \qquad \langle \chi_{\mathrm{triv}},\ \chi_{\mathrm{sign}} \rangle = 0.$$

Consider the defining representation. Its character is $\chi_{\mathrm{def}} = [3,1,0]$, and

$$\begin{aligned}
\langle \chi_{\mathrm{triv}},\ \chi_{\mathrm{def}} \rangle &= \frac{1}{6} \sum_{j=1}^{3} |C_j| \cdot \overline{\chi_{\mathrm{triv}}(C_j)} \cdot \chi_{\mathrm{def}}(C_j) \\
&= \frac{1}{6}\left(1 \cdot 1 \cdot 3 + 3 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0\right) \ = \ 1, \\
\langle \chi_{\mathrm{sign}},\ \chi_{\mathrm{def}} \rangle &= \frac{1}{6} \sum_{j=1}^{3} |C_j| \cdot \overline{\chi_{\mathrm{triv}}(C_j)} \cdot \chi_{\mathrm{def}}(C_j) \\
&= \frac{1}{6}\left(1 \cdot 1 \cdot 3 - 3 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0\right) \ = \ 0.
\end{aligned}$$

This tells us that ρ_{def} contains one copy of the trivial representation as a summand, and no copies of the sign representation. If we get rid of the trivial summand, the remaining two-dimensional representation ρ has character $\chi_\rho = \chi_{\mathrm{def}} - \chi_{\mathrm{triv}} = [2,0,-1]$.

Since

$$\langle \chi_\rho,\ \chi_\rho \rangle = \frac{1(2 \cdot 2) + 3(0 \cdot 0) + 2(-1 \cdot -1)}{6} = 1,$$

it follows that ρ is irreducible. So, up to isomorphism, $\mathfrak{S}_3$ has two distinct one-dimensional representations $\rho_{\mathrm{triv}}, \rho_{\mathrm{sign}}$ and one two-dimensional representation ρ. Note also that

$$\chi_{\mathrm{triv}} + \chi_{\mathrm{sign}} + 2\chi_\rho = [1,1,1] + [1,-1,1] + 2[2,0,-1] = [6,0,0] = \chi_{\mathrm{reg}}.$$

8.4. **New Characters from Old.** In order to investigate characters, we need to know how standard vector space (or, in fact, G-module) functors such as $\oplus$ and $\otimes$ affect the corresponding characters. Throughout, let (ρ, V), (ρ', V') be representations of G, with $V \cap V' = \emptyset$.

1. **Direct sum**. To construct a basis for $V \oplus V'$, we can take the union of a basis for V and a basis for V'. Equivalently, we can write the vectors in $V \oplus V'$ as column block vectors:

$$V \oplus V' = \left\{ \begin{bmatrix} v \\ v' \end{bmatrix} \ \mid v \in V,\ v' \in V' \right\}.$$

Accordingly, define $(\rho \oplus \rho', V \oplus V')$ by

$$(\rho \oplus \rho')(h) = \left[\begin{array}{c|c} \rho(h) & 0 \\ \hline 0 & \rho'(h) \end{array}\right].$$

From this it is clear that

$$\chi_{\rho \oplus \rho'}(h) = \chi_\rho(h) + \chi_{\rho'}(h). \tag{8.2}$$

2. **Duality**. Recall that the *dual space* V^* of V consists of all $\mathbb{F}$-linear transformations $\phi : V \to \mathbb{F}$. Given a representation (ρ, V), there is a natural action of G on V^* defined by

$$(h\phi)(v) = \phi(h^{-1}v)$$

for $h \in G$, $\phi \in V^*$, $v \in V$. (You need to define it this way in order for $h\phi$ to be a homomorphism — try it.) This is called the **dual representation** or **contragredient representation** and denoted ρ^*.

Proposition 8.26. *For every $h \in G$,*

$$\chi_{\rho^*}(h) = \overline{\chi_\rho(h)}. \tag{8.3}$$

where the bar denotes complex conjugate.

Proof. Choose a basis $\{v_1, \ldots, v_n\}$ of V consisting of eigenvectors of h (since we are working over $\mathbb{C}$); say $hv_i = \lambda_i v_i$.

In this basis, $\rho(h) = \operatorname{diag}(\lambda_i)$ (i.e., the diagonal matrix whose entries are the λ_i), and in the dual basis, $\rho^*(h) = \operatorname{diag}(\lambda_i^{-1})$.

On the other hand, some power of $\rho(h)$ is the identity matrix, so each λ_i must be a root of unity, so its inverse is just its complex conjugate. □

3. **Tensor product.** If $\{v_1, \ldots, v_n\}$, $\{v'_1, \ldots, v'_m\}$ are bases for V, V' respectively, then $V \otimes V'$ can be defined as the vector space with basis

$$\{v_i \otimes v'_j \mid 1 \le i \le n,\ 1 \le j \le m\}$$

equipped with a multilinear action of $\mathbb{F}$ (consult your favorite algebra textbook). In particular, $\dim V \otimes V' = (\dim V)(\dim V')$.

Accordingly, define a representation $(\rho \otimes \rho', V \otimes V')$ by

$$(\rho \otimes \rho')(h)(v \otimes v') = \rho(h)v \otimes v' + v \otimes \rho'(h)v'$$

or more concisely

$$h \cdot (v \otimes v') = (hv) \otimes v' + v \otimes (hv'),$$

extended bilinearly to all of $V \otimes V'$.

In terms of matrices, $(\rho \otimes \rho')(h)$ is represented by the block matrix

$$\begin{bmatrix} a_{11}B & a_{11}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{bmatrix}$$

where $\rho(h) = [a_{ij}]_{i,j=1\ldots n}$ and $\rho'(h) = B$. In particular,

$$\chi_{\rho\otimes\rho'}(h) = \chi_\rho(h)\chi_{\rho'}(h). \tag{8.4}$$

4.**Hom.** There are two kinds of Hom. Let V and W be representations of G. Then the vector space

$$\operatorname{Hom}_{\mathbb{C}}(V, W) = \{\mathbb{C}\text{-linear transformations } \phi : V \to W\}$$

itself admits a representation of G, defined by

$$(h \cdot \phi)(v) = h(\phi(h^{-1}v)) = \rho'(h)\Big(\phi(\rho(h^{-1})(v))\Big). \tag{8.5}$$

for $h \in G$, $\phi \in \operatorname{Hom}_{\mathbb{C}}(V, W)$, $v \in V$. That is, h sends ϕ to [the map $h \cdot \phi$ which acts on V as above]; It is not hard to verify that this is a genuine group action. Moreover, $\operatorname{Hom}_{\mathbb{C}}(V, W) \cong V^* \otimes W$ as vector spaces, so

$$\chi_{\operatorname{Hom}(\rho,\rho')}(h) = \overline{\chi_\rho(h)}\ \chi_{\rho'}(h). \tag{8.6}$$

The other kind of Hom is

$$\mathrm{Hom}_G(V,W) = \{G\text{-equivariant transformations } \phi : V \to W\}.$$

Evidently $\mathrm{Hom}_G(V,W) \subseteq \mathrm{Hom}_{\mathbb{C}}(V,W)$, but equality need not hold. For example, if V and W are the trivial and sign representations of $\mathfrak{S}_n$ $(n \geq 2)$, then $\mathrm{Hom}_{\mathbb{C}}(V,W) \cong \mathbb{C}$ but $\mathrm{Hom}_G(V,W) = 0$.

The two Homs are related as follows. In general, when G acts on a vector space V, the *subspace of G-invariants* is defined as

$$V^G = \{v \in V \mid hv = h \ \forall h \in G\}.$$

In our current setup, a linear map $\phi : V \to W$ is G-equivariant if and only if $h \cdot \phi = \phi$ for all $h \in G$, where the dot denotes the action of G on $\mathrm{Hom}_{\mathbb{C}}(V,W)$ (proof left to the reader). That is,

$$\mathrm{Hom}_G(V,W) = \mathrm{Hom}_{\mathbb{C}}(V,W)^G. \tag{8.7}$$

However, we still need a way of computing the character of $\mathrm{Hom}_G(\rho,\rho')$.

8.5. **The Inner Product.** Recall that a **class function** is a function $\chi : G \to \mathbb{C}$ that is constant on conjugacy classes of G. Define an inner product on the vector space $C\ell(G)$ of class functions by

$$\langle \chi,\, \psi \rangle_G = \frac{1}{|G|} \sum_{h \in G} \overline{\chi(h)} \psi(h).$$

Proposition 8.27. *Let (V,ρ) be a representation of G. Then*

$$\dim_{\mathbb{C}} V^G = \frac{1}{|G|} \sum_{h \in G} \chi_\rho(h) = \left\langle \chi_{\mathrm{triv}},\, \chi_\rho \right\rangle_G.$$

Proof. Define a linear map $\pi : V \to V$ by

$$\pi = \frac{1}{|G|} \sum_{h \in G} \rho(h).$$

In fact, $\pi(v) \in V^G$ for all $v \in V$, and if $v \in V^G$ then $\pi(v) = v$ (these assertions are not hard to verify). That is, π is a projection from $V \to V^G$, and if we choose a suitable basis for V, then it can be represented by the block matrix

$$\left[\begin{array}{c|c} I & 0 \\ \hline * & 0 \end{array}\right]$$

where the first and second column blocks (resp., row blocks) correspond to V^G and $(V^G)^\perp$ respectively. It is now evident that $\dim_{\mathbb{C}} V^G = \operatorname{tr} \pi$, giving the first equality. For the second equality, we know by Maschke's Theorem that V is semisimple, so we can decompose it as a direct sum of irreps. Then V^G is precisely the direct sum of the irreducible summands on which G acts trivially. $\square$

Example 8.28. Suppose that ρ is a permutation representation. Then V^G is the space of functions that are constant on the orbits. Therefore, the formula becomes

$$\#\text{ orbits } = \frac{1}{|G|} \sum_{h \in G} \#\text{ fixed points of } h$$

which is Burnside's Lemma from basic abstract algebra.

Proposition 8.29. *For any two representations ρ, ρ' of G, we have $\left\langle \chi_\rho,\, \chi_{\rho'} \right\rangle_G = \dim_{\mathbb{C}} \mathrm{Hom}_G(\rho,\rho')$.*

Proof.

$$\begin{aligned}
\langle \chi_\rho, \chi_{\rho'} \rangle_G &= \frac{1}{|G|} \sum_{h \in G} \overline{\chi_\rho(h)} \chi_{\rho'}(h) \\
&= \frac{1}{|G|} \sum_{h \in G} \chi_{\mathrm{Hom}(\rho,\rho')}(h) && \text{(by (8.6))} \\
&= \dim_{\mathbb{C}} \mathrm{Hom}(\rho,\rho')^G && \text{(by Proposition 8.27)} \\
&= \dim_{\mathbb{C}} \mathrm{Hom}_G(\rho,\rho') && \text{(by (8.7)).}
\end{aligned}$$

□

8.6. **Schur's Lemma and the Orthogonality Relations.**

Proposition 8.30 (Schur's Lemma). *Let G be a group, and let (ρ, V) and (ρ', V') be finite-dimensional representations of G over a field $\mathbb{F}$.*

(1) *If ρ and ρ' are irreducible, then every G-equivariant $\phi : V \to V'$ is either zero or an isomorphism.*
(2) *If in addition $\mathbb{F}$ is algebraically closed, then*

$$\mathrm{Hom}_G(V, V') \cong \begin{cases} \mathbb{F} & \text{if } \rho \cong \rho' \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the only G-equivariant maps from an G-irrep to itself are multiplication by a scalar.

Proof. For (1), recall from the proof of Maschke's Theorem[10] that $\ker\phi$ and $\operatorname{im}\phi$ are G-invariant subspaces. But since ρ, ρ' are simple, there are not many possibilities. Either $\ker\phi = 0$ and $\operatorname{im}\phi = W$, when ϕ is an isomorphism. Otherwise, $\ker\phi = V$ or $\operatorname{im}\phi = 0$, either of which implies that $\phi = 0$.

For (2), let $\phi \in \mathrm{Hom}_G(V, V')$. If $\rho \not\cong \rho'$ then $\phi = 0$ by (1) and we're done. Otherwise, we may as well assume that $V = V'$.

Since $\mathbb{F}$ is algebraically closed, ϕ has an eigenvalue λ. Then $\phi - \lambda I$ is G-equivariant and singular, hence zero by (1). So $\phi = \lambda I$. We've just shown that the only G-equivariant maps from V to itself are multiplication by λ. □

Theorem 8.31. *Let (ρ, V) and (ρ', V') be finite-dimensional representations of G over $\mathbb{C}$.*

(i) *If ρ and ρ' are irreducible, then*

$$\langle \chi_\rho, \chi_{\rho'} \rangle_G = \begin{cases} 1 & \text{if } \rho \cong \rho' \\ 0 & \text{otherwise.} \end{cases}$$

(ii) *If $\rho_1, \ldots, \rho_n$ are distinct irreducible representations and*

$$\rho = \bigoplus_{i=1}^n \underbrace{(\rho_i \oplus \cdot \oplus \rho_i)}_{m_i} = \bigoplus_{i=1}^n \rho_i^{\oplus m_i}$$

then

$$\langle \chi_\rho, \chi_{\rho_i} \rangle_G = m_i, \qquad \langle \chi_\rho, \chi_\rho \rangle_G = \sum_{i=1}^n m_i^2.$$

In particular, $\langle \chi_\rho, \chi_\rho \rangle_G = 1$ if and only if ρ is irreducible.

(iii) *If $\chi_\rho = \chi_{\rho'}$ then $\rho \cong \rho'$.*

---

[10] This part didn't require any assumption about the characteristic of $\mathbb{F}$.

(iv) *If $\rho_1, \dots, \rho_n$ is a complete list of irreducible representations of G, then*

$$\rho_{\text{reg}} \cong \bigoplus_{i=1}^{n} \rho_i^{\oplus \dim \rho_i}$$

and consequently

$$\sum_{i=1}^{n} (\dim \rho_i)^2 = |G|.$$

(v) *The irreducible characters (i.e., characters of irreducible representations) form an orthonormal basis for $C\ell(G)$. In particular, the number of irreducible characters equals the number of conjugacy classes of G.*

Proof of Theorem 8.31. Assertion (i) follows from part (2) of Schur's Lemma together with Proposition 8.29, and (ii) follows because the inner product is bilinear on direct sums. For (iii), Maschke's Theorem says that every complex representation ρ can be written as a direct sum of irreducibles. Their multiplicities determine ρ up to isomorphism, and can be recovered from χ_ρ by assertion (ii).

For (iv), recall that $\chi_{\text{reg}}(1_G) = |G|$ and $\chi_{\text{reg}}(g) = 0$ for $g \neq 1_G$. Therefore

$$\left\langle \chi_{\text{reg}},\, \rho_i \right\rangle_G = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_{\text{reg}}(g)} \rho_i(g) = \frac{1}{|G|} |G| \rho_i(1_G) = \dim \rho_i$$

so ρ_i appears in ρ_{reg} with multiplicity equal to its dimension.

That the irreducible characters are orthonormal (hence linearly independent in $C\ell(G)$) follows from Schur's Lemma together with assertion (3). The trickier part is to show that they in fact span $C\ell(G)$.

Let

$$Z = \left\{ \phi \in C\ell(G) \mid \left\langle \phi,\, \chi_\rho \right\rangle_G = 0 \text{ for every irreducible character } \rho \right\}.$$

That is, Z is the orthogonal complement of the span of the irreducible characters. We will show that in fact $Z = 0$.

Let $\phi \in Z$. For any representation (ρ, V), define a map $T_\rho = T_{\rho,\phi} : V \to V$ by

$$T_\rho = \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} \rho(g).$$

We will show that T_ρ is the zero map in disguise! First we show that it is G-equivariant. Indeed, for $h \in G$,

$$\begin{aligned}
T_\rho(hv) &= \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} (gh)(v) \\
&= \frac{h}{|G|} \sum_{g \in G} \overline{\phi(g)} (h^{-1} g h v) \\
&= \frac{h}{|G|} \sum_{k \in G} \overline{\phi(hkh^{-1})} (kv) && \text{(setting } k = h^{-1} g h,\ hkh^{-1} = g) \\
&= \frac{h}{|G|} \sum_{k \in G} \overline{\phi(k)} (kv) && \text{(because } \phi \in C\ell(G)) \\
&= h T_\rho(v).
\end{aligned}$$

Suppose now that ρ is irreducible. By Schur's Lemma, T_ρ is multiplication by a scalar (possibly zero). On the other hand, we assumed that ϕ is orthogonal to ρ, that is,

$$\begin{aligned} 0 \;=\; \langle \phi,\, \chi_\rho \rangle_G &= \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} \chi_\rho(g) \\ &= \operatorname{tr}\left(\frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} \rho(g) \right) \;=\; \operatorname{tr}(T_\rho). \end{aligned}$$

We've shown that T_ρ is multiplication by a scalar, and also that it has trace zero; therefore, $T_\rho = 0$ for every irreducible ρ. Also, T is additive on direct sums (that is, $T_{\rho \oplus \rho'} = T_\rho + T_{\rho'}$), so by Maschke's Theorem, $T_\rho = 0$ for *every* representation ρ. In particular, take $\rho = \rho_{\text{reg}}$: then

$$0 \;=\; T_{\rho_{\text{reg}}}(1_G) \;=\; \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} g.$$

This is an equation in the group algebra, and equating coefficients of g on either side implies that $\phi(g) = 0$ for every $g \in G$, as desired. □

Example 8.32. We calculate all the irreducible characters of $\mathfrak{S}_4$.

There are five conjugacy classes in $\mathfrak{S}_4$, corresponding to the cycle-shapes 1111, 211, 22, 31, and 4. The squares of their dimensions must add up to $|\mathfrak{S}_4| = 24$. The only list of five positive integers with that property is 1, 1, 2, 3, 3.

We start by writing down some characters that we know:

Cycle shape	1111	211	22	31	4
Size of conjugacy class	1	6	3	8	6
$\chi_1 = \chi_{\text{triv}}$	1	1	1	1	1
$\chi_2 = \chi_{\text{sign}}$	1	-1	1	1	-1
χ_{def}	4	2	0	1	0
χ_{reg}	24	0	0	0	0

Of course χ_{triv} and χ_{sign} are irreducible (since they are 1-dimensional). On the other hand, χ_{def} can't be irreducible because $\mathfrak{S}_4$ doesn't have a 4-dimensional irrep. Indeed,

$$\langle \chi_{\text{def}},\, \chi_{\text{def}} \rangle_G = 2$$

which means that ρ_{def} must be a direct sum of two distinct irreps. (If it were the direct sum of two copies of the unique 2-dimensional irrep, then $\langle \chi_{\text{def}},\, \chi_{\text{def}} \rangle_G$ would be 4, not 2, by (ii) of Theorem 8.31.) We calculate

$$\langle \chi_{\text{def}},\, \chi_{\text{triv}} \rangle_G = 1, \qquad \langle \chi_{\text{def}},\, \chi_{\text{sign}} \rangle_G = 0.$$

Therefore $\chi_3 = \chi_{\text{def}} - \chi_{\text{triv}}$ is an irreducible character.

Another 3-dimensional character is $\chi_4 = \chi_3 \otimes \chi_{\text{sign}}$. It is easy to check that $\langle \chi_4,\, \chi_4 \rangle_G = 1$, so χ_4 is irreducible.

The other irreducible character χ_5 has dimension 2. We can calculate it from the regular character and the other four irreducibles, because

$$\chi_{\text{reg}} \;=\; (\chi_1 + \chi_2) + 3(\chi_3 + \chi_4) + 2\chi_5$$

and so

$$\chi_5 = \frac{\chi_{\text{reg}} - \chi_1 - \chi_2 - 3\chi_3 - 3\chi_4}{2}.$$

and so the complete character table of $\mathfrak{S}_4$ is as follows.

Cycle shape	1111	211	22	31	4
Size of conjugacy class	1	6	3	8	6
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	3	1	-1	0	-1
χ_4	3	-1	-1	0	1
χ_5	2	0	2	-1	0

8.7. **One-Dimensional Characters.** Let G be a group and ρ a one-dimensional representation; that is, ρ is a group homomorphism $G \to \mathbb{C}^\times$. Note that $\chi_\rho = \rho$. Also, if ρ' is another one-dimensional representation, then

$$\rho(g)\rho'(g) = (\rho \otimes \rho')(g)$$

for all $g \in G$. Thus the group $Ch(G) = \mathrm{Hom}(G, \mathbb{C}^\times)$ of all one-dimensional characters forms a group under pointwise multiplication. The trivial character is the identity of $Ch(G)$, and the inverse of a character ρ is its dual $\rho^* = \bar{\rho}$.

Definition 8.33. The **commutator** of two elements $a, b \in G$ is the element $[a, b] = aba^{-1}b^{-1}$. The subgroup of G generated by all commutators is called the **commutator subgroup**, denoted $[G, G]$.

It is simple to check that $[G, G]$ is in fact a normal subgroup of G. Moreover, $\rho([a, b]) = 1$ for all $\rho \in Ch(G)$ and $a, b \in G$. Therefore, the one-dimensional characters of G are precisely those of the quotient $G^{ab} = G/[G, G]$, the *abelianization* of G. Accordingly, we would like to understand the characters of abelian groups.

Let G be an abelian group of finite order n. The conjugacy classes of G are all singleton sets (since $ghg^{-1} = h$ for all $g, h \in G$), so there are n distinct irreducible representations of G. On the other hand,

$$\sum_{\chi \text{ irreducible}} (\dim \chi)^2 = n$$

by Theorem 8.31 (iv), so in fact every irreducible character is 1-dimensional (and every representation of G is a direct sum of 1-dimensional representations).

Since a 1-dimensional representation equals its character, we just need to describe the homomorphisms $G \to \mathbb{C}^\times$.

The simplest case is that $G = \mathbb{Z}/n\mathbb{Z}$ is cyclic. Write G multiplicatively, and let g be a generator. Then each $\chi \in Ch(G)$ is determined by its value on g, which must be some n^{th} root of unity. There are n possibilities for χ, so all the irreducible characters of G arise in this way, and in fact form a group isomorphic to G.

Now we consider the general case. Every abelian group G can be written as

$$G \cong \prod_{i=1}^{r} \mathbb{Z}/n_i\mathbb{Z}.$$

Let g_i be a generator of the i^{th} factor, and let ζ_i be a primitive $(n_i)^{th}$ root of unity. Then each character χ is determined by the numbers $j_1, \ldots, j_r$, where $j_i \in \mathbb{Z}/n_i\mathbb{Z}$ and $\chi(g_i) = \zeta_i^{j_i}$. for all i. By now, it should be evident that

$$\mathrm{Hom}(G, \mathbb{C}^\times) \cong G,$$

an isomorphism known as *Pontrjagin duality*. More generally, for any group G we have

$$\mathrm{Hom}(G, \mathbb{C}^\times) \cong G^{ab}. \tag{8.8}$$

This is quite useful when computing irreducible characters, because it tells us right away about the one-dimensional characters of an arbitrary group.

Example 8.34. Consider the case $G = \mathfrak{S}_n$. Certainly $[\mathfrak{S}_n, \mathfrak{S}_n] \subseteq \mathfrak{A}_n$, and in fact equality holds. (This is trivial for $n \leq 2$. If $n \leq 3$, then the equation $(a\ b)(b\ c)(a\ b)(b\ c) = (a\ b\ c)$ in $\mathfrak{S}_n$ (multiplying left to right) shows that $[\mathfrak{S}_n, \mathfrak{S}_n]$ contains every 3-cycle, and it is not hard to show that the 3-cycles generate the full alternating group.) Therefore (8.8) gives

$$\operatorname{Hom}(\mathfrak{S}_n, \mathbb{C}^\times) \cong \mathfrak{S}_n/\mathfrak{A}_n \cong \mathbb{Z}/2\mathbb{Z}$$

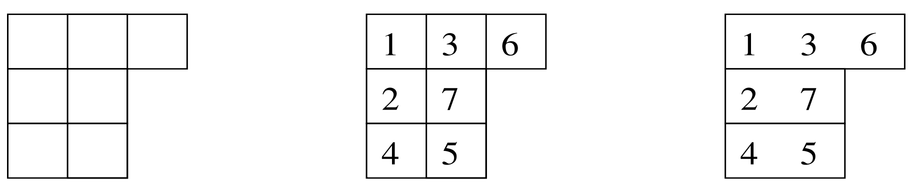
which says that χ_{triv} and χ_{sign} are the only one-dimensional characters of $\mathfrak{S}_n$.

8.8. **Characters of the Symmetric Group.** We worked out the irreducible characters of $\mathfrak{S}_4$ *ad hoc*. We'd like to have a way of calculating them in general.

Recall that a *partition* of n is a sequence $\lambda = (\lambda_1, \dots, \lambda_\ell)$ of weakly decreasing positive integers whose sum is n. We write $\lambda \vdash n$ to indicate that λ is a partition of n. The setof all partitions of n is $\operatorname{Par}(n)$, and the number of partitions of n is $p(n) = |\operatorname{Par}(n)|$.

For $\lambda \vdash n$, let C_λ be the conjugacy class in $\mathfrak{S}_n$ consisting of all permutations with cycle shape λ. Since the conjugacy classes are in bijection with $\operatorname{Par}(n)$, it makes sense to look for a set of representations indexed by partitions.

Definition 8.35. Let $\mu = (\mu_1, \dots, \mu_m) \vdash n$. The **Ferrers diagram** of shape μ is the top- and left-justified array of boxes with μ_i boxes in the i^{th} row. A **Young tableau**[11] **of shape** μ is a Ferrers diagram with the numbers $1, 2, \dots, n$ placed in the boxes, one number to a box. Two tableaux T, T' of shape μ are *row-equivalent*, written $T \sim T'$, if the numbers in each row of T are the same as the numbers in the corresponding row of T'. A **tabloid** of shape μ is an equivalence class of tableaux under row-equivalence. A tabloid can be represented as a tableau without vertical lines separating numbers in the same row. We write $\operatorname{sh}(T) = \mu$ to indicate that a tableau or tabloid T is of shape μ.



Ferrers diagram Young tableau Young tabloid

A Young tabloid can be regarded as a set partition $(T_1, \dots, T_m)$ of $[n]$, in which $|T_i| = \mu_i$. The order of the blocks T_i matters, but not the order of digits within each block. Thus the number of tabloids of shape μ is

$$\binom{n}{\mu} = \frac{n!}{\mu_1! \cdots \mu_m!}.$$

The symmetric group $\mathfrak{S}_n$ acts on tabloids by permuting the numbers. Accordingly, we have a permutation representation ρ_μ of $\mathfrak{S}_n$ on the vector space V^μ of all $\mathbb{C}$-linear combinations of tabloids of shape μ. I'll call this the **μ-tabloid representation** of $\mathfrak{S}_n$.

[11]The terminology surrounding tableaux is not consistent: some authors reserve the term "Young tableau" for a tableau in which the numbers increase downward and leftward. I'll call such a thing a "standard tableau". For the moment, we are not placing any restrictions on which numbers can go where.

Example 8.36. For $n = 3$, the characters of the tabloid representations ρ_μ are as follows.

(8.9)

		cycle shape λ		
	111	21	3	
	$\lvert C_\lambda \rvert$	1	3	2
	3	1	1	1
Tabloid shape μ	21	3	1	0
	111	6	0	0

Many familiar representations of $\mathfrak{S}_n$ can be expressed in this form.

- There is a unique tabloid of shape $\mu = (n)$: $T = \boxed{1\ 2\ \cdots\ n}$. Every permutation fixes T, so
$$\rho_{(n)} \cong \rho_{\text{triv}}.$$
- The tabloids of shape $\mu = (1, 1, \ldots, 1)$ are just the permutations of $[n]$. Therefore
$$\rho_{(1,1,\ldots,1)} \cong \rho_{\text{reg}}.$$
- A tabloid of shape $\mu = (n - 1, 1)$ is determined by its singleton part. So the representation ρ_μ is isomorphic to the action of $\mathfrak{S}_n$ on this part by permutation; that is
$$\rho_{(n-1,1)} \cong \rho_{\text{def}}.$$

For $n = 3$, the table in (8.9) is triangular, which implies immediately that the characters ρ_μ are linearly independent. In order to prove that this is the case for all n, we first need to define two orders on the set $\mathrm{Par}(n)$.

Definition 8.37. The *lexicographic order* on $\mathrm{Par}(n)$ is defined by $\lambda < \mu$ if $\lambda_k < \mu_k$ for the first k for which they differ. That is, for some $k > 0$,
$$\lambda_1 = \mu_1, \quad \lambda_2 = \mu_2, \quad \ldots, \lambda_{k-1} = \mu_{k-1}, \quad \lambda_k < \mu_k.$$

Note that this is a total order on $\mathrm{Par}(n)$. For instance, if $n = 5$, we have
$$(5) \ > \ (4,1) \ > \ (3,2) \ > \ (3,1,1) \ > \ (2,2,1) \ > \ (2,1,1,1) \ > \ (1,1,1,1,1).$$
(The lexicographically greater ones are short and wide; the lex-smaller ones are tall and skinny.)

Definition 8.38. The *dominance order* on $\mathrm{Par}(n)$ is defined as follows: $\lambda \trianglelefteq \mu$ if $\sum_{i=1}^k \lambda_i \leq \sum_{i=1}^k \mu_i$ for all k.

Note that $\lambda \lhd \mu$ if and only if there are set partitions L, M of $[n]$ such that L refines N. So dominance can be viewed as a "quotient order of Π_n".

Dominance is a partial order on $\mathrm{Par}(n)$: for example, $33, 411 \in \mathrm{Par}(6)$ are incomparable. (Dominance does happen to be a total order for $n \leq 5$.) Lexicographic order is a linear extension of dominance order: that is,

(8.10)
$$\lambda \lhd \mu \quad \Longrightarrow \quad \lambda < \mu.$$

Abbreviate χ_{ρ_μ} by χ_μ henceforth. Since the tabloid representations ρ_μ are permutation representations, we can calculate χ_μ by counting fixed points. That is,

(8.11)
$$\chi_\mu(C_\lambda) = \#\{\text{tabloids } T \mid \text{sh}(T) = \mu,\ w(T) = \lambda\}$$
for any $w \in C_\lambda$.

Proposition 8.39. *Let $\lambda, \mu \vdash n$. Then:*

(i) $\chi_\lambda(C_\lambda) \neq 0$.
(ii) $\chi_\mu(C_\lambda) \neq 0$ *if and only if $\lambda \trianglelefteq \mu$ (thus, only if $\lambda \leq \mu$ in lexicographic order).*

Proof. Let $w \in C_\lambda$. Take T to be any tabloid whose blocks are the cycles of w; then $wT = T$. For example, if $w = (1\ 3\ 6)(2\ 7)(4\ 5) \in \mathfrak{S}_7$, then T can be either of the following two tabloids:

$$\begin{array}{|ccc|}\hline 1 & 3 & 6 \\ \hline 2 & 7 & \multicolumn{1}{|c}{} \\ \cline{1-2} 4 & 5 & \multicolumn{1}{|c}{} \\ \cline{1-2}\end{array} \qquad \begin{array}{|ccc|}\hline 1 & 3 & 6 \\ \hline 4 & 5 & \multicolumn{1}{|c}{} \\ \cline{1-2} 2 & 7 & \multicolumn{1}{|c}{} \\ \cline{1-2}\end{array}$$

It follows from (8.11) that $\chi_\lambda(C_\lambda) \neq 0$.

For the second assertion, observe that $w \in \mathfrak{S}_n$ fixes a tabloid T of shape μ if and only if every cycle of w is contained in a row of P, which is possible if and only if $\lambda \trianglelefteq \mu$. $\square$

Corollary 8.40. *The characters $\{\chi_\mu \mid \mu \vdash n\}$ form a basis for $C\ell(G)$.*

Proof. Make the characters into a $p(n) \times p(n)$ matrix $X = [\chi_\mu(C_\lambda)]_{\mu,\lambda\vdash n}$ with rows and columns ordered by lex order on $\mathrm{Par}(n)$. By Proposition 8.39, X is triangular. Hence it is nonsingular. $\square$

We can transform the rows of the matrix X into a list of irreducible characters of $\mathfrak{S}_n$ by applying the Gram-Schmidt process with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathfrak{S}_n}$. Indeed, the triangularity of X means that we will be able to label the irreducible characters of $\mathfrak{S}_n$ as σ_ν, for $\nu \vdash n$, in such a way that

$$\begin{aligned} &\langle\sigma_\nu,\ \chi_\nu\rangle_G \neq 0, \\ &\left\langle\sigma_\nu,\ \chi_\mu\right\rangle_G = 0 \qquad \text{if } \nu < \mu. \end{aligned} \tag{8.12}$$

On the level of representations, this corresponds to decomposing the tabloid representation V^μ into its irreducible G-invariant subspaces Sp_λ (which are called *Specht modules*):

$$V^\mu = \bigoplus_\lambda (Sp_\lambda)^{\oplus K_{\lambda,\mu}}$$

for some nonnegative integers $K_{\lambda,\mu}$.

Example 8.41. Recall the table of characters (8.9) of the tabloid representations for $n = 3$. We will use this to produce the table of irreducible characters. For clarity of notation, I will write, e.g., χ_{311} instead of $\chi_{(3,1,1)}$.

First, $\chi_3 = [1,1,1] = \chi_{\mathrm{triv}}$ is irreducible and is the character σ_3 of the Specht module Sp_3.

Second, for the character χ_{21}, we observe that

$$\langle\chi_{21},\ \sigma_3\rangle_G = 1.$$

Applying Gram-Schmidt, we construct a character orthonormal to σ_3:

$$\sigma_{21} \ = \ \chi_{21} - \sigma_3 \ = \ [2,0,-1].$$

Notice that this character is irreducible.

Finally, for the character $\chi_{(111)}$, we have

$$\begin{aligned} \langle\chi_{111},\ \sigma_3\rangle_G &= 1, \\ \langle\chi_{111},\ \sigma_{21}\rangle_G &= 2. \end{aligned}$$

Accordingly, we apply Gram-Schmidt to obtain the character

$$\sigma_{111} \ = \ \chi_{111} - \sigma_3 - 2\sigma_{21} \ = \ [1,-1,1]$$

which is 1-dimensional, hence irreducible. In summary, the complete list of irreducible characters, labeled so as to satisfy (8.12), is as follows:

	λ 111	21	3	
σ_3	1	1	1	$= \chi_{\text{triv}}$
σ_{21}	2	0	-1	$= \chi_{\text{def}} - \chi_{\text{triv}}$
σ_{111}	1	-1	1	$= \chi_{\text{sign}}$

To summarize our calculation, we have shown that

$$\begin{bmatrix}\chi_\mu\end{bmatrix}_{\mu\vdash 3} = \begin{bmatrix}1&1&1\\3&1&0\\6&0&0\end{bmatrix} = \underbrace{\begin{bmatrix}1&0&0\\1&1&0\\1&2&1\end{bmatrix}}_{K}\begin{bmatrix}1&1&1\\2&0&-1\\1&-1&1\end{bmatrix} = \begin{bmatrix}K_{\lambda,\mu}\end{bmatrix}_{\lambda,\mu\vdash 3}\begin{bmatrix}\sigma_\lambda\end{bmatrix}_{\lambda\vdash 3}$$

that is,

$$\chi_\mu = \sum_\lambda K_{\lambda,\mu}\sigma_\lambda.$$

The numbers $K_{\lambda,\mu}$ are called the **Kostka numbers**. We will eventually find a combinatorial interpretation for them, from which it will also be easy to see that the matrix K is unitriangular.

8.9. Restricted and Induced Representations.

Definition 8.42. Let $H \subset G$ be finite groups, and let $\rho : G \to GL(V)$ be a representation of G. Then the restriction of ρ to H is a representation of G, denoted $\operatorname{Res}^G_H(\rho)$. (Alternate notation: $\rho\downarrow^G_H$.) Likewise, the restriction of $\chi = \chi_\rho$ to H is a character of H denoted by $\operatorname{Res}^G_H(\chi)$.

Notice that restricting a representation does not change its character. OTOH, whether or not a representation is irreducible can change upon restriction.

Example 8.43. Let C_λ denote the conjugacy class in $\mathfrak{S}_n$ of permutations of cycle-shape λ. Recall that $G = \mathfrak{S}_3$ has an irrep whose character $\psi = \chi_\rho$ is given by

$$\psi(C_{111}) = 2, \qquad \psi(C_{21}) = 0, \qquad \psi(C_3) = -1.$$

Let $H = \mathfrak{A}_3 \subseteq \mathfrak{S}_3$. This is an abelian group (isomorphic to $\mathbb{Z}/3\mathbb{Z}$), so the two-dimensional representation $\operatorname{Res}^G_H(\rho)$ is not irreducible. Specifically, let $\omega = e^{2\pi i/3}$ The table of irreducible characters of $\mathfrak{A}_3$ is as follows:

	1_G	$(1\,2\,3)$	$(1\,3\,2)$
χ_{triv}	1	1	1
χ_1	1	ω	ω^2
χ_2	1	ω^2	ω

Now it is evident that $\operatorname{Res}^G_H \psi = [2, -1, -1] = \chi_1 + \chi_2$. Note, by the way, that the conjugacy class $C_3 \subset \mathfrak{S}_3$ splits into two singleton conjugacy classes in $\mathfrak{A}_3$, a common phenomenon when working with restrictions.

Next, we construct a representation of G from a representation of a subgroup $H \subset G$.

Definition 8.44. Let $H \subset G$ be finite groups, and let $\rho : H \to GL(W)$ be a representation of H. Define the **induced representation** $\operatorname{Ind}^G_H(\rho)$ (alternate notation: $\rho\uparrow^G_H$) as follows. First, choose a set of left coset representatives for H in G, that is, a set $B = \{b_1, \dots, b_r\}$ such that $H = \bigsqcup_{j=1}^r b_j H$. Let $\mathbb{C}[G/H]$ be the $\mathbb{C}$-vector space with basis B, and let $V = \mathbb{C}[G/H] \otimes_{\mathbb{C}} W$. (Note that the notation G/H just means the set of left cosets — we are not assuming that H is a normal subgroup.)

Now let $g \in G$ act on $b_i \otimes w \in V$ as follows. Find the unique $b_i \in B$ and $h \in H$ such that $gb_i = b_j h$, i.e., $g = b_j h b_i^{-1}$. Then put

$$g \cdot (b_i \otimes w) = b_j h b_i^{-1} \cdot (b_i \otimes w) = b_j \otimes hw.$$

Extend this a representation of G on V by linearity.

Proposition 8.45. *$\mathrm{Ind}_H^G(\rho)$ is a representation of G that is independent of the choice of B. Moreover, for all $g \in G$,*

$$\chi_{\mathrm{Ind}_H^G(\rho)}(g) \;=\; \frac{1}{|H|} \sum_{k \in G:\; k^{-1}gk \in H} \chi_\rho(k^{-1}gk).$$

Proof. First, we verify that $\mathrm{Ind}_H^G(\rho)$ is a representation. Let $g, g' \in G$ and $b_i \otimes w \in V$. Then there is a unique $b_j \in B$ and $h \in H$ such that

$$gb_i = b_j h \tag{8.13}$$

and in turn there is a unique $b_\ell \in B$ and $h' \in H$ such that

$$g' b_j = b_\ell h'. \tag{8.14}$$

We need to verify that

$$g' \cdot (g \cdot (b_i \otimes w)) \;=\; (g'g) \cdot (b_i \otimes w). \tag{8.15}$$

Indeed,

$$\Big(g' \cdot (g \cdot (b_i \otimes w)) \;=\; g' \cdot (b_j \otimes hw)\Big) \;=\; b_\ell \otimes h'hw.$$

On the other hand, by (8.13) and (8.14), $gb_i = b_j h b_i^{-1}$ and $g' = b_\ell h' b_j^{-1}$, so

$$(g'g) \cdot (b_i \otimes w) \;=\; (b_\ell h' h b_i^{-1}) \cdot (b_i \otimes w) \;=\; b_\ell \otimes h'hw$$

as desired. Note by the way that

$$\dim \mathrm{Ind}_H^G \rho = \frac{|G|}{|H|} \dim \rho.$$

Now that we know that $\mathrm{Ind}_H^G(\rho)$ is a representation of G on V, we find its character on an arbitrary element $g \in G$. Regard $\mathrm{Ind}_H^G(\rho)(g)$ as a block matrix with r row and column blocks, each of size $\dim W$ and corresponding to the subspace of V of vectors of the form $b_i \otimes w$ for some fixed b_i. The block in position (i, j) is

- a copy of $\rho(h)$, if $gb_i = b_j h$ for some $h \in H$,
- zero otherwise.

Therefore,

$$\begin{aligned}
\chi_{\mathrm{Ind}_H^G(\rho)}(g) &= \mathrm{tr}\,(g:\; \mathbb{C}[G/H] \otimes_{\mathbb{C}} W \to \mathbb{C}[G/H] \otimes_{\mathbb{C}} W) \\
&= \sum_{i \in [r]:\; gb_i = b_i h \text{ for some } h \in H} \chi_\rho(h) \\
&= \sum_{i \in [r]:\; b_i^{-1} g b_i \in H} \chi_\rho(b_i^{-1} g b_i) \\
&= \frac{1}{|H|} \sum_{i \in [r]:\; b_i^{-1} g b_i \in H} \; \sum_{h \in H} \chi_\rho(h^{-1} b_i^{-1} g b_i h) \\
&= \frac{1}{|H|} \sum_{k \in G:\; k^{-1} g k \in H} \chi_\rho(k^{-1} g k).
\end{aligned}$$

Here we have put $k = b_i h$, which runs over all elements of G. The character of $\mathrm{Ind}_H^G(\rho)$ is independent of the choice of B; therefore, so is the representation itself. $\square$

Corollary 8.46. *Suppose H is a normal subgroup of G. Then*

$$\operatorname{Ind}_H^G \chi(g) = \begin{cases} \frac{|G|}{|H|}\chi(g) & \text{if } g \in H, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Since H is normal, $g \in H \iff$ some conjugate of g belongs to $H \iff$ every conjugate of g belongs to H. □

Corollary 8.47. *Let $H \subset G$ and let ρ be the trivial representation of H. Then*

$$\operatorname{Ind}_H^G \chi_{\text{triv}}(g) = \frac{\#\{k \in G:\ k^{-1}gk \in H\}}{|H|}. \tag{8.16}$$

Example 8.48. The character $\psi = \operatorname{Ind}_{\mathfrak{A}_3}^{\mathfrak{S}_3} \chi_{\text{triv}}$ is defined by $\psi(g) = 2$ for $g \in \mathfrak{A}_3$, $\psi(g) = 0$ for $g \notin \mathfrak{A}_3$. Thus $\psi = \chi_{\text{triv}} + \chi_{\text{sign}}$.

Example 8.49. Let $G = \mathfrak{S}_4$ and let H be the subgroup $\{\text{id},\ (1\ 2),\ (3\ 4),\ (1\ 2)(3\ 4)\}$ (note that this is not a normal subgroup). Let ρ be the trivial representation of G and χ its character. We can calculate $\psi = \operatorname{Ind}_H^G \chi$ using (8.16); letting C_λ denote the conjugacy class of permutations with cycle-shape λ we end up with

$$\psi(C_{1111}) = 6, \qquad \psi(C_{211}) = 2, \qquad \psi(C_{22}) = 2, \qquad \psi(C_{31}) = 0, \qquad \psi(C_4) = 0.$$

In the notation of Example 8.32, the decomposition into irreducible characters is $\chi_1 + \chi_2 + 2\chi_5$.

Restriction and inducing are related by the following useful formula.

Theorem 8.50 (Frobenius Reciprocity). $\left\langle \operatorname{Ind}_H^G \chi,\ \psi \right\rangle_G = \left\langle \chi,\ \operatorname{Res}_H^G \psi \right\rangle_H.$

Proof.

$$\begin{aligned}
\left\langle \operatorname{Ind}_H^G \chi,\ \psi \right\rangle_G &= \frac{1}{|G|}\sum_{g\in G} \overline{\operatorname{Ind}_H^G \chi(g)}\psi(g) \\
&= \frac{1}{|G|}\sum_{g\in G} \frac{1}{|H|} \sum_{k\in G:\ k^{-1}gk\in H} \overline{\chi(k^{-1}gk)}\psi(g) && \text{(by Prop. 8.45)} \\
&= \frac{1}{|G||H|}\sum_{h\in H}\sum_{k\in G}\sum_{g\in G:\ k^{-1}gk=h} \overline{\chi(h)}\psi(k^{-1}gk) \\
&= \frac{1}{|G||H|}\sum_{h\in H}\sum_{k\in G}\overline{\chi(h)}\psi(h) && \text{(i.e., } g = khk^{-1}) \\
&= \frac{1}{|H|}\sum_{h\in H}\overline{\chi(h)}\psi(h) \ = \ \left\langle \chi,\ \operatorname{Res}_H^G \psi \right\rangle_H.
\end{aligned}$$

□

Example 8.51. Sometimes, Frobenius reciprocity suffices to calculate the isomorphism type of an induced representation. Let ψ, χ_1 and χ_2 be as in Example 8.43. We would like to compute $\operatorname{Ind}_H^G \chi_1$. By Frobenius reciprocity

$$\left\langle \operatorname{Ind}_H^G \chi_1,\ \psi \right\rangle_G = \left\langle \chi_1,\ \operatorname{Res}_H^G \psi \right\rangle_H = 1.$$

But ψ is irreducible. Therefore, it must be the case that $\operatorname{Ind}_H^G \chi_1 = \psi$, and the corresponding representations are isomorphic. The same is true if we replace χ_1 with χ_2.

9. Symmetric Functions

Definition 9.1. Let R be a commutative ring, typically $\mathbb{Q}$ or $\mathbb{Z}$. A **symmetric function** is (temporarily) a polynomial in $R[x_1, \dots, x_n]$ that is invariant under permuting the variables.

For example, if $n = 3$, then up to scalar multiplication, the only symmetric function of degree 1 in x_1, x_2, x_3 is $x_1 + x_2 + x_3$.

In degree 2, here are two:

$$x_1^2 + x_2^2 + x_3^2, \qquad x_1x_2 + x_1x_3 + x_2x_3.$$

Every other symmetric function that is homogeneous of degree 2 is a R-linear combination of these two, because the coefficients of x_1^2 and x_1x_2 determine the coefficients of all other monomials. Note that the set of all degree-2 symmetric functions forms a vector space.

In degree 3, the following three polynomials form a basis for the space of symmetric functions:

$$\begin{aligned}
&x_1^3 + x_2^3 + x_3^3,\\
&x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2,\\
&x_1x_2x_3.
\end{aligned}$$

Each member of this basis is a sum of the monomials in a single orbit under the action of $\mathfrak{S}_3$. Accordingly, we call them **monomial symmetric functions**, and index each by the partition whose parts are the exponents of one of its monomials. That is,

$$\begin{aligned}
m_3(x_1, x_2, x_3) &= x_1^3 + x_2^3 + x_3^3,\\
m_{21}(x_1, x_2, x_3) &= x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2,\\
m_{111}(x_1, x_2, x_3) &= x_1x_2x_3.
\end{aligned}$$

In general, for $\lambda = (\lambda_1, \dots, \lambda_\ell)$, we define

$$m_\lambda(x_1, \dots, x_n) = \sum_{\{a_1,\dots,a_\ell\} \subset [n]} x_{a_1}^{\lambda_1} x_{a_2}^{\lambda_2} \cdots x_{a_\ell}^{\lambda_\ell}.$$

But unfortunately, this is zero if $\ell > n$. So we need more variables! In fact, we will in general work with an *infinite*[12] set of variables $\{x_1, x_2, \dots\}$.

Definition 9.2. Let $\lambda \vdash n$. The **monomial symmetric function** m_λ is the power series

$$m_\lambda = \sum_{\{a_1,\dots,a_\ell\} \subset \mathbb{P}} x_{a_1}^{\lambda_1} x_{a_2}^{\lambda_2} \cdots x_{a_\ell}^{\lambda_\ell}.$$

That is, m_λ is the sum of all monomials whose exponents are the parts of λ. Another way to write this is

$$m_\lambda = \sum_{\text{rearrangements } \alpha \text{ of } \lambda} x^\alpha$$

where x^α is shorthand for $x_1^{\alpha_1} x_2^{\alpha_2} \cdots$. Here we are regarding λ as a countably infinite sequence in which all but finitely many terms are 0.

We then define

$$\begin{aligned}
\Lambda_d = \Lambda_{R,d} &= \{\text{degree-}d \text{ symmetric functions with coefficients in } R\},\\
\Lambda = \Lambda_R &= \bigoplus_{d \geq 0} \Lambda_d.
\end{aligned}$$

[12] This (understandably) bothers some people. In practice, we rarely have to worry about more than finitely many variables when carrying out calculations.

Each Λ_d is a finite-dimensional vector space, with basis $\{m_\lambda \mid \lambda \vdash d\}$. $\dim_{\mathbb{C}} \Lambda_d = p(d)$ (the number of partitions of d), and the dimension does not change even if we zero out all but d variables, so for many purposes it is permissible (and less intimidating) to regard Λ_d as the space of degree-d symmetric functions in d variables.

Moreover, Λ is a graded ring. In fact, let $\mathfrak{S}_\infty$ be the group whose members are the permutations of $\{x_1, x_2, \dots\}$ with only finitely many non-fixed points; that is,

$$\mathfrak{S}_\infty = \bigcup_{n=1}^{\infty} \mathfrak{S}_n.$$

Then

$$\Lambda = R[[x_1, x_2, \dots,]]^{\mathfrak{S}_\infty}.$$

Where is all this going? The punchline is that we will eventually construct an isomorphism

$$\Lambda \xrightarrow{F} \bigoplus_{n \geq 0} C\ell(\mathfrak{S}_n)$$

called the **Frobenius characteristic**. Thus will allow us to translate symmetric function identities into statements about representations and characters of $\mathfrak{S}_n$, and vice versa. Many of these statements are best stated in terms of bases for Λ other than the monomial symmetric functions, so we now consder several important families.

9.1. **Elementary symmetric functions.** For $k \in \mathbb{N}$ we define

$$e_k \quad = \quad \sum_{\substack{S \subset \mathbb{N} \\ |S|=k}} \prod_{s \in S} x_s \quad = \quad \sum_{0 < i_1 < i_2 < \cdots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k} \quad = \quad m_{11\cdots 1}$$

where there are k $1's$ in the last expression. (In particular $e_0 = 1$.) We then define

$$e_\lambda = e_{\lambda_1} \cdots e_{\lambda_\ell}.$$

For example, $e_{11} = (x_1 + x_2 + x_3 + \cdots)^2 = (x_1^2 + x_2^2 + \cdots) + 2(x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4 + \cdots) = m_2 + 2m_{11}$. In degree 3, we have

$$\begin{aligned}
e_3 &= \sum_{i<j<k} x_i x_j x_k && = m_{111}, \\
e_{21} &= (x_1 + x_2 + x_3 + \cdots)(x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4 + \cdots) && = m_{21} + 3m_{111}, \\
e_{111} &= (x_1 + x_2 + x_3 + \cdots)^3 && = m_3 + 3m_{21} + 6m_{111},
\end{aligned}$$

We can conveniently express all the e's together as a generating function. Observe that

$$E(t) := \prod_{i \geq 1} (1 + tx_i) \;=\; \sum_{k \geq 0} t^k e_k \tag{9.1}$$

by expanding $E(t)$ as a power series in t whose coefficients are power series in $\{x_i\}$. Note that there are no issues of convergence: we are working in the ring of formal power series $R[[t, x_1, x_2, \dots]]$.

9.2. **Complete homogeneous symmetric functions.** For $k \in \mathbb{N}$, we define h_k to be the sum of *all* monomials of degree k:

$$h_k = \sum_{0<i_1\leq i_2\leq\cdots\leq i_k} x_{i_1}x_{i_2}\cdots x_{i_k} = \sum_{\lambda\vdash k} m_\lambda.$$

We then define $h_\lambda = h_{\lambda_1}\cdots h_{\lambda_\ell}$.

For example, $h_{11} = e_{11}$ and $h_2 = m_{11} + m_2$. In degree 3, we have

$$\begin{aligned} h_{111} &= m_1^3 = (x_1+x_2+x_3+\cdots)^3 = 6m_{111}+3m_{21}+m_3, \\ h_{21} &= h_1h_2 = e_1(m_{11}+m_2) = e_1(e_{11}-e_2) = e_{111}-e_{21} = m_3+2m_{21}+3m_{111}, \\ h_3 &= m_{111}+m_{21}+m_3. \end{aligned}$$

The analogue of (9.2) for the homogeneous symmetric functions is

$$H(t) := \prod_{i\geq 1}\frac{1}{1-tx_i} = \sum_{k\geq 0} t^k h_k \tag{9.2}$$

(because each factor in the infinite product is a geometric series $1 + tx_i + t^2x_i^2 + \cdots$, so when we expand and collect like powers of t, the coefficient of t^k will be the sum of all possible ways to build a monomial of degree k). It is immediate from the algebra that

$$H(t)E(-t) = 1$$

as formal power series. Extracting the coefficients of positive powers of t gives the **Jacobi-Trudi relations**:

$$\sum_{k=0}^{n}(-1)^k e_k h_{n-k} = 0 \tag{9.3}$$

for all $n > 0$, where we put $e_0 = h_0 = 1$. Explicitly,

$$h_1 - e_1 = 0, \quad h_2 - e_1h_1 + e_2 = 0, \quad h_3 - e_1h_2 + e_2h_1 - e_3 = 0, \quad \ldots$$

These equations can be used (iteratively) to solve for the h_k as polynomials in the e_k:

$$\begin{aligned} h_1 &= e_1, \\ h_2 &= e_1h_1 - e_2 = e_1^2 - e_2, \\ h_3 &= e_1h_2 - e_2h_1 + e_3 = e_1(e_1^2 - e_2) - e_2e_1 + e_3, \end{aligned} \tag{9.4}$$

etc. The roles of e and h can be reversed in all of this. Indeed, in many situations, the elementary and homogeneous symmetric functions behave dually.

9.3. **Power-sum symmetric functions.** Define

$$\begin{aligned} p_k &= x_1^k + x_2^k + \cdots = m_k, \\ p_\lambda &= p_{\lambda_1}\cdots p_{\lambda_\ell}. \end{aligned}$$

For example, in degree 2,

$$\begin{aligned} p_2 &= m_2, \\ p_{11} &= (x_1+x_2+\cdots)^2 = m_2 + 2m_{11}. \end{aligned}$$

While $\{p_2, p_{11}\}$ is a $\mathbb{Q}$-vector space basis for $\Lambda_\mathbb{Q}$, it is not a $\mathbb{Z}$-module basis for $\Lambda_\mathbb{Z}$. To put this in a more elementary way, not every symmetric function with integer coefficients can be expressed as an integer combination of the power-sums; for example, $m_{11} = (p_{11} - p_2)/2$.

9.4. **Schur functions.** The definition of these power series is very different from the preceding ones, and it looks quite weird at first. However, the Schur functions turn out to be essential in the study of symmetric functions.

Definition 9.3. A **column-strict tableau** T of shape λ, or **λ-CST** for short, is a labeling of the boxes of a Ferrers diagram with integers (not necessarily distinct) that is

- weakly increasing across every row; and
- strictly increasing down every column.

The partition λ is called the **shape** of T, and the set of all column-strict tableaux of shape λ is denoted $CST(\lambda)$. The **content** of a CST is the sequence $\alpha = (\alpha_1, \alpha_2, \dots)$, where α_i is the number of boxes labelled i, and the **weight** of T is the monomial $x^T = x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots$. For example:

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 3 \\ \hline 2 & 3 & \multicolumn{1}{c}{} \\ \cline{1-2}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 4 & 8 & \multicolumn{1}{c}{} \\ \cline{1-2}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 1 & 4 & \multicolumn{1}{c}{} \\ \cline{1-2}\end{array}$$

$$x_1^2\, x_2\, x_3^2 \qquad\qquad x_1^3 x_4 x_8 \qquad\qquad \text{Not a CST}$$

Definition 9.4. The **Schur function** corresponding to a partition λ is

$$s_\lambda = \sum_{T \in CST(\lambda)} x^T.$$

It is far from obvious that s_λ is symmetric, but in fact it is. We will prove this shortly.

Example 9.5. Suppose that $\lambda = (n)$ is the partition with one part, so that the corresponding Ferrers diagram has a single row. Each multiset of n positive integers (with repeats allowed) corresponds to exactly one CST, in which the numbers occur left to right in increasing order. Therefore

$$s_{(n)} = h_n = \sum_{\lambda \vdash n} m_\lambda. \tag{9.5}$$

At the other extreme, suppose that $\lambda = (1, 1, \dots, 1)$ is the partition with n singleton parts, so that the corresponding Ferrers diagram has a single column. To construct a CST of this shape, we need n distinct labels, which can be arbitrary. Therefore

$$s_{(1,1,\dots,1)} = e_n = m_{(1,1,\dots,1)}. \tag{9.6}$$

Let $\lambda = (2, 1)$. We will express s_λ as a sum of monomial symmetric functions. No tableau in $CST(\lambda)$ can have three equal entries, so the coefficient of m_3 is zero.

For weight $x_a x_b x_c$ with $a < b < c$, there are two possibilities, shown below.

$$\begin{array}{|c|c|}\hline a & b \\ \hline c & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline a & c \\ \hline b & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

Therefore, the coefficient of m_{111} is 1.

Finally, for every $a \neq b \in \mathbb{N}$, there is one tableau of shape λ and weight $x_a^2 x_b$ — either the one on the left if $a < b$, or the one on the right if $a > b$.

a	b
b	

b	b
a	

Therefore, $s_{(2,1)} \;=\; 2m_{111} + m_{21}$.

Proposition 9.6. *s_λ is a symmetric function for all λ.*

Proof. First, observe that the number

(9.7) $$c(\lambda, \alpha) = |\{T \in CST(\lambda) \mid x^T = x^\alpha\}|$$

depends only on the ordered sequence of nonzero exponents[13] in α. For instance, for any $\lambda \vdash 8$, there are the same number of λ-CST's with weights

$$x_1^1 x_2^2 x_3^4 x_9^1 \quad \text{and} \quad x_1^1 x_2^2 x_7^4 x_9^1$$

because there is an obvious bijection between them given by changing all 3's to 7's or vice versa.

To complete the proof that s_λ is symmetric, it suffices to show that swapping the powers of adjacent variables does not change $c(\lambda, \alpha)$. That will imply that s_λ is invariant under every transposition $(k\ k+1)$, and these transpositions generate the group $\mathfrak{S}_\infty$.

We will prove this by a bijection, which is easiest to show by example. Let $\lambda = (9, 7, 4, 3, 2)$. We would like to show that there are the same number of λ-CST's with weights

$$x_1^3 x_2^2 x_3^3 x_4^3 \boldsymbol{x_5^4 x_6^7} x_7^3 \quad \text{and} \quad x_1^3 x_2^2 x_3^3 x_4^3 \boldsymbol{x_5^7 x_6^4} x_7^3.$$

Let T be the following λ-CST:

1	1	1	2	3	5	6	6	6
2	3	4	5	6	7	7		
3	4	5	6					
4	6	6						
5	7							

Observe that the occurrences of 5 and of 6 each form "snakes" from southwest to northeast.

1	1	1	2	3	5	6	6	6
2	3	4	5	6	7	7		
3	4	5	6					
4	6	6						
5	7							

[13] This is precisely the statement that s_λ is a **quasisymmetric** function.

To construct a new tableau in which the numbers of 5's and of 6's are switched, we ignore all the columns containing both a 5 and a 6, and then group together all the other strings of 5's and 6's in the same row.

1	1	1	2	3	5	6	6	6
2	3	4	5	6	7	7		
3	4	5	6					
4	6	6						
5	7							

Then, we swap the numbers of 5's and 6's in each of those contiguous blocks.

1	1	1	2	3	5	5	5	6
2	3	4	5	5	7	7		
3	4	5	6					
4	5	6						
6	7							

Observe that this construction is an involution (because the ignored columns do not change).

This construction allows us to swap the exponents on x_k and x_{k+1} for any k, concluding the proof. $\square$

Theorem 9.7. *Let R be a commutative ring and let $n \geq 1$.*

- *The sets $\{m_\lambda \mid \lambda \vdash n\}$ $\{e_\lambda \mid \lambda \vdash n\}$ $\{h_\lambda \mid \lambda \vdash n\}$ and $\{s_\lambda \mid \lambda \vdash n\}$ are all bases for $\Lambda_{R,n}$ as a free R-module.*
- *If R is a field, then $\{p_\lambda \mid \lambda \vdash n\}$ is a basis for $\Lambda_{R,n}$ as a vector space.*
- *The sets $\{e_1, e_2, \dots\}$ and $\{h_1, h_2, \dots\}$ generate Λ as a polynomial algebra over R. (The first of these assertions is sometimes (misleadingly) called the "fundamental theorem of symmetric functions".)*

Proof. It is more or less obvious that the m_λ are a $\mathbb{Z}$-basis.

By the definition of Schur functions, we have for every λ

$$s_\lambda = \sum_{\lambda \vdash n} K_{\lambda\mu} m_\mu$$

where $K_{\lambda\mu}$ is the number of column-strict tableaux T with shape λ and content μ. The $K_{\lambda\mu}$ are called **Kostka numbers**.

First, suppose that $\lambda = \mu$. Then there is exactly one possibility for T: fill the i^{th} row full of i's. Therefore

$$\forall \lambda \vdash n : \quad K_{\lambda\lambda} = 1. \tag{9.8}$$

Second, observe that

- every 1 in T must appear in the 1st row;
- every 2 in T must appear in the 1st or 2nd row;
- ...
- every i in T must appear in one of the first i rows;
- ...

and therefore

$$\mu_1 \leq \lambda_1, \qquad \mu_1+\mu_2 \leq \lambda_1+\lambda_2, \qquad \dots, \qquad \mu_1+\cdots+\mu_i \leq \lambda_1+\cdots+\lambda_i, \qquad ,\dots$$

That is

$$\forall \lambda,\mu \vdash n: \quad K_{\lambda\mu} > 0 \implies \lambda \trianglerighteq \mu \tag{9.9}$$

where $\trianglerighteq$ means "dominates" (see Definition 8.38). But that means that the matrix $[K_{\lambda\mu}]_{\lambda,\mu\vdash n}$ is unitriangular (provided that we order rows and columns by some linear extension of dominance, such as lex order). Therefore, it is invertible over R, and in particular the Schur functions are a vector space basis for $\Lambda_{\mathbb{Q}}$ and a free module basis for $\Lambda_{\mathbb{Z}}$.

A similar unitriangularity argument can be used to show that the elementary symmetric functions are a $\mathbb{Z}$-basis. Expand the e_λ's in the monomial basis as $e_\lambda = \sum_{\mu\vdash n} a_{\lambda\mu} m_\mu$. Then $a_{\lambda\tilde{\lambda}} = 1$ for all λ, and $M_{\lambda\mu} > 0$ only if $\tilde{\lambda} \trianglerighteq \mu$ (where $\tilde{\lambda}$ means the conjugate partition). To see this, it suffices to consider the coefficient of $x^\mu = x_1^{\mu_1} x_2^{\mu_2} \cdots$ in $e_{\lambda_1} \cdots e_{\lambda_\ell}$. If $\tilde{\lambda} = \mu$, then there is exactly one way to make this monomial: choose the term $x_1 x_2 \cdots x_{\lambda_i}$ from e_{λ_i}. Meanwhile, the total degree of $x_1,\dots,x_i$ in every monomial in e_λ is at most $\tilde{\lambda}_1 + \cdots + \tilde{\lambda}_i$, so if this number is $< \mu_1 + \cdots + \mu_i$ then the monomial x^μ does not occur at all in e_λ. (More generally, the numbers $a_{\lambda\mu}$ have a nice combinatorial interpretation; see [14, §7.4].)

Meanwhile, the Jacobi-Trudi relations between the h's and e's given by (9.4) says that one family is a basis iff the other is.

The proof that the p_λ form a $\mathbb{Q}$-basis is analogous to that for the e's. Write $p_\lambda = \sum_\mu b_{\lambda\mu} m_\mu$. Consider the monomials x^α that show up in $p_\lambda = \prod_{i=1}^{\ell} (\sum x_i^{\lambda_i})$. For each variable x_i, the power α_i with which it shows up must be a sum of some of the parts of λ; it follows that $b_{\lambda\mu} = 0$ unless λ refines μ. Meanwhile, if a_i is the number of occurrences of i in λ, then $b_{\lambda\lambda} = a_1! a_2! \cdots$. It follows that the matrix $[b_{\lambda\mu}]$ is triangular, with nonzero but non-unit elements on its main diagonal. $\square$

Here is another way to see that the h's are a R-basis. Define a ring endomorphism $\omega : \Lambda \to \Lambda$ by $\omega(e_i) = h_i$ for all i, so that $\omega(e_\lambda) = h_\lambda$. This is well-defined since the elementary symmetric functions are algebraically independent (recall that $\Lambda \cong R[e_1, e_2, \dots]$).

Proposition 9.8. *$\omega(\omega(f)) = f$ for all $f \in \Lambda$. In particular, the map ω is a ring automorphism.*

Proof. Recall the Jacobi-Trudi relations (9.3). Applying ω, we find that

$$\begin{aligned} 0 &= \sum_{k=0}^{n} (-1)^{n-k} \omega(e_k)\omega(h_{n-k}) \\ &= \sum_{k=0}^{n} (-1)^{n-k} h_k \omega(h_{n-k}) \\ &= \sum_{k=0}^{n} (-1)^{k} h_{n-k} \omega(h_k) \\ &= (-1)^n \sum_{k=0}^{n} (-1)^{n-k} h_{n-k} \omega(h_k) \end{aligned}$$

and comparing this last expression with the original Jacobi-Trudi relations gives $\omega(h_k) = e_k$. $\square$

Corollary 9.9. *$\{h_\lambda\}$ is a graded $\mathbb{Z}$-basis for Λ. Moreover, $\Lambda_R \cong R[h_1, h_2, \dots]$.*

9.5. **The Cauchy kernel and the Hall inner product.** Our next step in studying Λ will be to define an inner product structure on it. These will come from considering the **Cauchy kernel** and the **dual Cauchy kernel**, which are the formal power series

$$\Omega = \prod_{i,j \geq 1} (1 - x_i y_j)^{-1}, \qquad \Omega^* = \mho = \prod_{i,j \geq 1} (1 + x_i y_j).$$

These series are symmetric with respect to each of the variable sets $\mathbf{x} = \{x_1, x_2, \dots\}$ and $\mathbf{y} = \{y_1, y_2, \dots\}$. As we'll see, the Cauchy kernel can be expanded in many different ways in terms of symmetric functions in the variable sets $\mathbf{x}$ and $\mathbf{y}$.

For a partition $\lambda \vdash n$, let r_i be the number of i's in λ, and define[14]

$$z_\lambda = 1^{r_1} r_1! \, 2^{r_2} r_2! \cdots, \qquad \varepsilon_\lambda = (-1)^{r_2 + r_4 + \cdots}. \tag{9.10}$$

For example, if $\lambda = (3, 3, 2, 1, 1, 1)$ then $z_\lambda = 1^3 3! \, 2^1 1! \, 3^2 2! = 216$. The notation "$z_\lambda$" comes from the fact that this is the size of the centralizer of a permutation $\sigma \in \mathfrak{S}_n$ with cycle-shape λ (that is, the group of permutations that commute with σ). Meanwhile, ε_λ is just the sign of a permutation with cycle-shape λ.

Proposition 9.10. *We have*

$$\prod_{i,j \geq 1} (1 - x_i y_j)^{-1} = \sum_\lambda h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y}) \ = \ \sum_\lambda \frac{p_\lambda(\mathbf{x}) p_\lambda(\mathbf{y})}{z_\lambda}, \tag{9.11}$$

$$\prod_{i,j \geq 1} (1 + x_i y_j) = \sum_\lambda e_\lambda(\mathbf{x}) m_\lambda(\mathbf{y}) \ = \ \sum_\lambda \varepsilon_\lambda \frac{p_\lambda(\mathbf{x}) p_\lambda(\mathbf{y})}{z_\lambda}, \tag{9.12}$$

where the sums run over all partitions λ.

[14] Stanley uses m_i where I am using r_i. I want to avoid conflict with the notation for monomial symmetric functions.

Proof. For the first identity in (9.11),

$$\begin{aligned}
\prod_{i,j\geq 1}(1-x_iy_j)^{-1} &= \prod_{j\geq 1}\left(\prod_{i\geq 1}(1-x_it)^{-1}\ \Big|_{t=y_j}\right)\\
&= \prod_{j\geq 1}\left(\sum_{k\geq 0}h_k(\mathbf{x})t^k\ \Big|_{t=y_j}\right) = \prod_{j\geq 1}\sum_{k\geq 0}h_k(\mathbf{x})\ y_j^k \qquad (9.13)\\
&= \sum_\lambda h_\lambda(\mathbf{x})m_\lambda(\mathbf{y})
\end{aligned}$$

(since the coefficient on the monomial $y_1^{\lambda_1}y_2^{\lambda_2}\cdots$ in (9.13) is $h_{\lambda_1}h_{\lambda_2}\cdots$).

For the second identity in (9.11), recall the power series expansions

$$\log(1+q) = q - \frac{q^2}{2} + \frac{q^3}{3} - \cdots = \sum_{n\geq 1}(-1)^{n+1}\frac{q^n}{n}, \qquad \exp(q) = \sum_{n\geq 0}\frac{q^n}{n!}.$$

These are formal power series that obey the rules you would expect; for instance, $\log(\prod_i q_i) = \sum_i(\log q_i)$ and $\exp\log(q) = q$. In particular:

$$\begin{aligned}
\log\Omega = \log\prod_{i,j\geq 1}(1-x_iy_j)^{-1} &= -\log\prod_{i,j\geq 1}(1-x_iy_j) = -\sum_{i,j\geq 1}\log(1-x_iy_j)\\
&= \sum_{i,j\geq 1}\sum_{n\geq 1}\frac{x_i^ny_j^n}{n} = \sum_{n\geq 1}\frac{1}{n}\sum_{i,j\geq 1}x_i^ny_j^n\\
&= \sum_{n\geq 1}\frac{p_n(\mathbf{x})p_n(\mathbf{y})}{n}
\end{aligned}$$

and now exponentiating both sides and applying the power series expansion for exp, we get

$$\begin{aligned}
\Omega &= \exp\left(\sum_{n\geq 1}\frac{p_n(\mathbf{x})p_n(\mathbf{y})}{n}\right) = \sum_{k\geq 0}\frac{1}{k!}\left(\sum_{n\geq 1}\frac{p_n(\mathbf{x})p_n(\mathbf{y})}{n}\right)^k\\
&= \sum_{k\geq 0}\frac{1}{k!}\left[\sum_{\lambda:\,\ell(\lambda)=k}\binom{k}{r_1!\,r_2!\,\dots}\left(\frac{p_1(\mathbf{x})p_1(\mathbf{y})}{1}\right)^{r_1(\lambda)}\left(\frac{p_2(\mathbf{x})p_2(\mathbf{y})}{2}\right)^{r_2(\lambda)}\cdots\right]\\
&= \sum_\lambda\frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}.
\end{aligned}$$

□

The proofs of the identities for the dual Cauchy kernel are analogous, and left to the reader.

Corollary 9.11. *We have*

$$h_n = \sum_{\lambda\vdash n}\frac{p_\lambda}{z_\lambda}\ ; \qquad (9.14)$$

$$e_n = \sum_{\lambda\vdash n}\varepsilon_\lambda\frac{p_\lambda}{z_\lambda}\ ; \quad \textit{and} \qquad (9.15)$$

$$\omega(p_\lambda) = \varepsilon_\lambda p_\lambda. \qquad (9.16)$$

Proof. For (9.14), we start with the identity of (9.11):

$$\sum_\lambda h_\lambda(\mathbf{x})m_\lambda(\mathbf{y}) = \sum_\lambda\frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}.$$

Set $y_1 = t$, and $y_k = 0$ for all $k > 1$. This kills all terms on the left side for which λ has more than one part, so we get

$$\sum_{\lambda=(n)} h_n(\mathbf{x})t^n \;=\; \sum_\lambda \frac{p_\lambda(\mathbf{x})t^{|\lambda|}}{z_\lambda}$$

and extracting the coefficient of t^n gives (9.14).

Starting with (9.12) and doing the same thing yields (9.15).

In order to obtain (9.16), let ω act on symmetric functions in $\mathbf{x}$ while fixing those in $\mathbf{y}$. Using (9.11) and (9.12), we obtain

$$\begin{aligned}\sum_\lambda \frac{\boxed{p_\lambda(\mathbf{x})}\; p_\lambda(\mathbf{y})}{z_\lambda} \;&=\; \sum_\lambda h_\lambda(\mathbf{x})m_\lambda(\mathbf{y}) \;=\; \omega\left(\sum_\lambda e_\lambda(\mathbf{x})m_\lambda(\mathbf{y})\right) \;=\; \omega\left(\sum_\lambda \varepsilon_\lambda \frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}\right)\\ &=\; \sum_\lambda \frac{\boxed{\varepsilon_\lambda\omega(p_\lambda(\mathbf{x}))}\; p_\lambda(\mathbf{y})}{z_\lambda}\end{aligned}$$

and equating coefficients of $p_\lambda(\mathbf{y})/z_\lambda$, as shown, yields the desired result. $\square$

Definition 9.12. The **Hall inner product** $\langle \cdot, \cdot \rangle$ on $\Lambda_\mathbb{Q}$ is defined by declaring $\{h_\lambda\}$ and $\{m_\mu\}$ to be dual bases: $\langle h_\lambda, m_\mu \rangle = \delta_{\lambda\mu}$.

Two bases $\{u_\lambda\}$, $\{v_\lambda\}$ for Λ are dual under the Hall inner product (i.e., $\langle u_\lambda, v_\mu \rangle = \delta_{\lambda\mu}$) if and only if

$$\prod_{i,j\geq 1} \frac{1}{1 - x_i y_j} = \sum_\lambda u_\lambda v_\lambda.$$

(Stanley: "The proof is a straightforward exercise in linear algebra and can be omitted without significant loss of understanding.") In particular, $\{p_\lambda\}$ and $\{p_\lambda/z_\lambda\}$ are dual bases, so $\{p_\lambda/\sqrt{z_\lambda}\}$ is an orthonormal basis for $\Lambda_\mathbb{R}$. (This, by the way, shows that $\langle \cdot, \cdot \rangle$ is a genuine inner product in the sense of being a *nondegenerate* bilinear form.)

More generally, if $\Omega = \sum_{\lambda,\mu} q_{\lambda\mu} u_\lambda(\mathbf{x}) v_\mu(\mathbf{y})$, then $\langle u_\lambda, v_\mu$ is the (λ, μ) entry of the matrix $[q_{\lambda\mu}]^{-1}$ (I think).

The involution ω is an isometry, i.e., $\langle a, b \rangle = \langle \omega(a), \omega(b) \rangle$. The easiest way to see this is in terms of the power-sum basis: by (9.16), we have

$$\langle \omega p_\lambda, \omega p_\mu \rangle = \langle \epsilon_\lambda p_\lambda, \epsilon_\lambda p_\mu \rangle = \epsilon_\lambda^2 \langle p_\lambda, p_\mu \rangle = \langle p_\lambda, p_\mu \rangle$$

because $\epsilon_\lambda \in \{1, -1\}$ for all λ. (Note that we don't even need the fact that $\{p_\lambda/\sqrt{z_\lambda}\}$ is an orthonormal basis for Λ.)

In fact, the Schur functions are an orthonormal basis for $\Lambda_\mathbb{Z}$. We will prove this by showing that

$$\Omega \;=\; \prod_{i,j\geq 1} \frac{1}{1 - x_i y_j} \;=\; \sum_\lambda s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}). \tag{9.17}$$

The proof requires a marvelous bijection called the **RSK correspondence** (for Robinson, Schensted and Knuth).

9.6. **The RSK Correspondence.**

Definition 9.13. Let $\lambda \vdash n$. A *standard [Young] tableau of shape* λ is a filling of the Ferrers diagram of λ with the numbers $1, 2, \ldots, n$ that is increasing left-to-right and top-to-bottom. We write $SYT(\lambda)$ for the set of all standard tableaux of shape λ, and set $f^\lambda = |SYT(\lambda)|$.

For example, if $\lambda = (3, 3)$, then $f^\lambda = 5$; the members of $SYT(\lambda)$ are as follows:

$$\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4&6\\\hline\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5&6\\\hline\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&2&5\\\hline 3&4&6\\\hline\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&5&6\\\hline\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5&6\\\hline\end{array}$$

Each Young tableau of shape λ corresponds to a saturated chain in the interval $[\emptyset, \lambda]$ of Young's lattice, namely

$$\emptyset = \lambda_{(0)} \subset \lambda_{(1)} \subset \cdots \subset \lambda_{(n)} = \lambda$$

where $\lambda_{(k)}$ denotes the subtableau consisting only of the boxes filled with the numbers $1, \ldots, k$. This correspondence between Young tableaux and saturated chains in $[\emptyset, \lambda]$ is a bijection, and is of fundamental importance.

The *RSK correspondence* (for Robinson-Schensted-Knuth) constructs, for every permutation $w \in \mathfrak{S}_n$, a pair $RSK(w) = (P, Q)$ of standard tableaux of the same shape $\lambda \vdash n$, using the following *row-insertion* operation defined as follows.

Definition 9.14. Let T be a column-strict tableau and let $x \in \mathbb{N}$. The *row-insertion* $T \leftarrow x$ is defined as follows:

- If $T = \emptyset$, then $T \leftarrow x = \boxed{x}$
- If $x \geq u$ for all entries u in the top row of T, then append x to the end of the top row.
- Otherwise, find the rightmost entry u such that $x < u$. Replace u with x, and then insert u into the subtableau consisting of the second and succeeding rows. (For short, "x bumps u.")
- Repeat until the bumping stops.

Got that? Now, for $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$, let P be the tableau $((\emptyset \leftarrow w_1) \leftarrow w_2) \leftarrow \cdots \leftarrow w_n \in \mathfrak{S}_n$. Let Q be the standard tableau of the same shape as P that records which box appears at which step of the insertion. The tableaux P and Q are respectively called the **insertion tableau** and the **recording tableau**, and the map $w \mapsto (P, Q)$ is the **RSK correspondence**.

Example 9.15. Let $w = 57214836 \in \mathfrak{S}_8$. We start with a pair (P, Q) of empty tableaux.

Step 1: Row-insert $w_1 = 5$ into P. We do this in the obvious way. Since it's the first cell added, we add a cell containing 1 to Q.

$$\mathbf{P} = \begin{array}{|c|}\hline 5\\\hline\end{array}\qquad\qquad \mathbf{Q} = \begin{array}{|c|}\hline 1\\\hline\end{array} \tag{9.18a}$$

Step 2: Row-insert $w_2 = 7$ into P. Since $5 < 7$, we can do this by appending the new cell to the top row, and adding a cell labeled 2 to Q to record where we've put the new cell in P.

$$\mathbf{P} = \begin{array}{|c|c|}\hline 5&7\\\hline\end{array}\qquad\qquad \mathbf{Q} = \begin{array}{|c|c|}\hline 1&2\\\hline\end{array} \tag{9.18b}$$

Step 3: Row-insert $w_3 = 2$ into P. This is a bit trickier. We can't just append a 2 to the first row of P, because the result would not be a standard tableau. The 2 has to go in the top left cell, but that already contains a 5. Therefore, the 2 "bumps" the 5 out of the first row into a new second row. Again, we record the location of the new cell by adding a cell labeled 3 to Q.

$$\begin{array}{|c|c|}\hline 2&7\\\hline 5\\\cline{1-1}\end{array}\qquad\qquad \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array} \tag{9.18c}$$

Step 4: Row-insert $w_4 = 1$ into P. This time, the new 1 bumps the 2 out of the first row. The 2 has to go into the second row, but again we can't simply append it to the right. Instead, the 2 bumps the 5 out of the second row into the (new) third row.

$$\begin{array}{|c|c|}\hline 1 & 7 \\ \hline 2 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array} \qquad\qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \tag{9.18d}$$

Step 5: Row-insert $w_5 = 4$ into P. The 4 bumps the 7 out of the first row. The 7, however, can comfortably fit at the end of the second row, without any more bumping.

$$\begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 7 \\ \hline 5 \\ \cline{1-1}\end{array} \qquad\qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 \\ \cline{1-1}\end{array} \tag{9.18e}$$

Step 6: Row-insert $w_6 = 8$ into P. The 8 just goes at the end of the first row.

$$\begin{array}{|c|c|c|}\hline 1 & 4 & 8 \\ \hline 2 & 7 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array} \qquad\qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \tag{9.18f}$$

Step 7: Row-insert $w_7 = 3$ into P. 3 bumps 4, and then 4 bumps 7.

$$\begin{array}{|c|c|c|}\hline 1 & 3 & 8 \\ \hline 2 & 4 \\ \cline{1-2} 5 & 7 \\ \cline{1-2}\end{array} \qquad\qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 \\ \cline{1-2} 4 & 7 \\ \cline{1-2}\end{array} \tag{9.18g}$$

Step 8: Row-insert $w_8 = 6$ into P. 6 bumps 8 into the second row.

$$\begin{array}{|c|c|c|}\hline 1 & 3 & 6 \\ \hline 2 & 4 & 8 \\ \hline 5 & 7 \\ \cline{1-2}\end{array} \qquad\qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 & 8 \\ \hline 4 & 7 \\ \cline{1-2}\end{array} \tag{9.18h}$$

A crucial feature of the RSK correspondence is that it can be reversed. That is, given a pair (P, Q), we can recover the permutation that gave rise to it.

Example 9.16. Suppose that we were given the pair of tableaux in (9.18h). What was the previous step? To get the previous Q, we just delete the 8. As for P, the last cell added must be the one containing 8.

This is in the second row, so somebody must have bumped 8 out of the first row. That somebody must be the largest number less than 8, namely 6. So 6 must have been the number inserted at this stage, and the previous pair of tableaux must have been those in (9.18g).

Example 9.17. Suppose P is the standard tableau with 18 boxes shown on the left.

1	2	5	8	10	18
3	4	11	12		
6	7	13			
9	15	17			
14	16				

1	2	5	8	10	18
3	4	11	12		
6	7	13			
9	15	17			
14	16				

Suppose in addition that we know that the cell labeled 16 was the last one added (because the corresponding cell in Q contains an 18). Then the "bumping path" must be as shown on the right. (That is, the 16 was bumped by the 15, which was bumped by the 13, and so on.) To find the previous tableau in the algorithm, we push every number in the bumping path up and toss out the top one.

1	2	5	8	12	18
3	4	11	13		
6	7	15			
9	16	17			
14					

That is, we must have gotten the original tableau by row-inserting 10 into the tableau just shown.

Iterating this "de-insertion" allows us to recover w from the pair (P, Q). We have proved the following fact:

Theorem 9.18. *The RSK correspondence is a bijection*

$$\mathfrak{S}_n \xrightarrow{RSK} \bigcup_{\lambda \vdash n} SYT(\lambda) \times SYT(\lambda).$$

Corollary 9.19. $\sum_{\lambda \vdash n} (f^\lambda)^2 = n!$.

To confirm this for $n = 3$, here are all the SYT's with 3 boxes:

1	2	3

1	2
3	

1	3
2	

1
2
3

Note that $f^{(3)} = f^{(1,1,1)} = 1$ and $f^{(2,1)} = 2$, and $1^2 + 1^2 + 2^2 = 6 = 3!$. This calculation ought to look familiar.

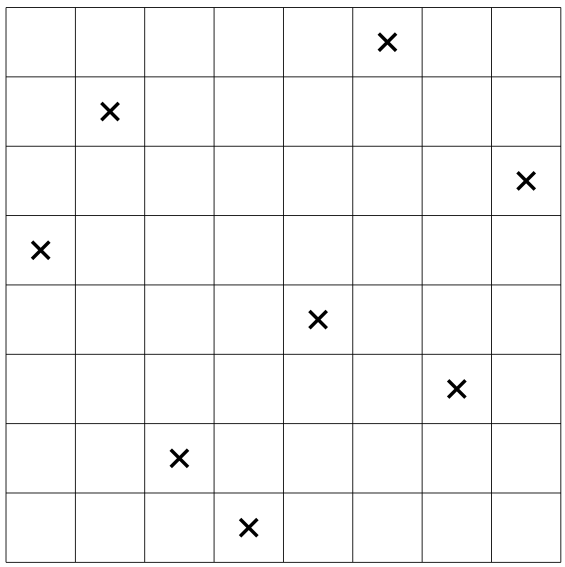
Another neat fact about the RSK correspondence is this:

Proposition 9.20. *Let $w \in \mathfrak{S}_n$. If $RSK(w) = (P, Q)$, then $RSK(w^{-1}) = (Q, P)$. In particular, the number of involutions in $\mathfrak{S}_n$ is $\sum_{\lambda \vdash n} f^\lambda$.*

This is hard to see from the standard RSK algorithm, where it looks like P and Q play inherently different roles. In fact, they are more symmetric than they look. There is an alternate description of RSK [14, §7.13] from which the symmetry is more apparent.

9.7. **An alternate version of RSK.** Fix $w \in \mathfrak{S}_n$. Start by drawing an $n \times n$ grid, numbering columns west to east and rows south to north. For each i, place an X in the i-th column and w_i-th row. We are now going to label each of the $(n+1) \times (n+1)$ intersections of the grid lines with a partition, such that the partitions either stay the same or get bigger as we move north and east. We start by labeling each intersection on the west and south sides with the empty partition $\emptyset$.

For instance, if $w = 57214836$, the grid is as follows.



For each box whose SW, SE and NW corners have been labeled λ, μ, ν respectively, label the NE corner ρ according to the following rules:

Rule 1: If $\lambda = \mu = \nu$ and the box doesn't contain an X, then set $\rho = \lambda$.

Rule 2: If $\lambda \subsetneq \mu = \nu$ and the box doesn't contain an X, then it must be the case that $\mu_i = \lambda_i + 1$ for some i. Define ρ by incrementing μ_{i+1}.

Rule 3: If $\mu \neq \nu$, then set $\rho = \mu \vee \nu$ (where that $\vee$ means the join in Young's lattice: i.e., take the componentwise maximum of the elements of μ and ν).

Rule X: If there is an X in the box, then it must be the case that $\lambda = \mu = \nu$. Define ρ by incrementing λ_1.

Rule 6: There is no Rule 6.

Note that the underlined assertions need to be proved; this can be done by induction.

Example 9.21. Let $n = 8$ and $w = 57214836$. In Example 9.15, we found that $RSK(w) = (P, Q)$, where P, Q are as follows:

1	3	6
2	4	8
5	7	

1	2	6
3	5	8
4	7	

The following extremely impressive figure shows what happens when we run the alternate RSK algorithm on w. The partitions λ are shown in red. The numbers in parentheses indicate which rules were used.

0	1 (3)	2 (3)	21 (3)	211 (3)	221 (3)	321 ×	322 (3)	332 (2)
0	1 (3)	2 ×	21 (3)	211 (3)	221 (2)	221 (3)	222 (2)	322 (3)
0	1 (3)	1 (1)	11 (3)	111 (3)	211 (3)	211 (1)	221 (3)	321 ×
0	1 ×	1 (3)	11 (2)	111 (2)	211 (3)	211 (3)	221 (3)	221 (3)
0	0 (1)	0 (1)	1 (3)	11 (3)	21 ×	21 (3)	22 (2)	22 (3)
0	0 (1)	0 (1)	1 (3)	11 (3)	11 (1)	11 (1)	21 ×	21 (3)
0	0 (1)	0 (1)	1 ×	11 (2)	11 (3)	11 (3)	11 (3)	11 (3)
0	0 (1)	0 (1)	0 (1)	1 ×	1 (3)	1 (3)	1 (3)	1 (3)
0	0	0	0	0	0	0	0	0

Observe that:

- Rule 1 is used exactly in those squares that have no X either due west or due south.

- For all squares s, $|\rho|$ is the number of X's in the rectangle whose northeast corner is s. In particular, the easternmost partition $\lambda_{(k)}$ in the k^{th} row, and the northernmost partition $\mu_{(k)}$ in the k^{th} column, both have size k.

- Therefore, the sequences

$$\emptyset = \lambda_{(0)} \subset \lambda_{(1)} \subset \cdots \subset \lambda_{(n)},$$
$$\emptyset = \mu_{(0)} \subset \mu_{(1)} \subset \cdots \subset \mu_{(n)}$$

correspond to SYT's of the same shape (in this case 332). These are precisely the tableaux P and Q of the RSK correspondence.

9.8. **Generalized RSK and Schur Functions.** The RSK correspondence can be extended to obtain more general tableaux. Think of a permutation in two-line notation, i.e.,

$$57214836 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 7 & 2 & 1 & 4 & 8 & 3 & 6 \end{pmatrix}.$$

More generally, we can allow "generalized permutations", i.e., things of the form

$$w = \binom{\mathbf{q}}{\mathbf{p}} = \begin{pmatrix} q_1 & q_2 & \cdots & q_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \tag{9.19}$$

where $\mathbf{q} = (q_1, \ldots, q_n), \mathbf{p} = (p_1, \ldots, p_n) \in [n]^n$, and the ordered pairs $(q_1, p_1), \ldots, (q_n, p_n)$ are in lexicographic order, but repeats are allowed.

Example 9.22. Consider the generalized permutation

$$w = \begin{pmatrix} 1 & 1 & 2 & 4 & 4 & 4 & 5 & 5 & 5 \\ 2 & 4 & 1 & 1 & 3 & 3 & 2 & 2 & 4 \end{pmatrix}.$$

We can row-insert the elements of the bottom row into a tableau P while recording the elements of the top row in a tableau Q:

$$\mathbf{P} = \begin{array}{|c|}\hline 2 \\ \hline\end{array} \qquad\qquad \mathbf{Q} = \begin{array}{|c|}\hline 1 \\ \hline\end{array}$$

$$\begin{array}{|c|c|}\hline 2 & 4 \\ \hline\end{array} \qquad\qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline\end{array}$$

$$\begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad\qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array}$$

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 4 \\ \hline\end{array} \qquad\qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 4 \\ \hline\end{array}$$

etc.

The tableaux P, Q we get in this way will always have the same shape be weakly increasing eastward and strictly increasing southward—that is, they will be *column-strict tableaux*, precisely the things for which the Schur functions are generating functions. Moreover, the generalized permutation w can be recovered from the pair (P, Q) as follows. For each n, there is at most one occurrence of n in each column of Q (since Q is column-strict), and a little thought should convince you that those n's show up from left to right as Q is constructed. Therefore, the rightmost instance of the largest entry in Q indicates the last box added to P, which can then be "unbumped" to recover the previous P and thus the last column of w. Iterating this process allows us to recover w.

Accordingly, we have a bijection

$$\left\{ \text{generalized permutations } \binom{\mathbf{q}}{\mathbf{p}} \text{ of length } n \right\} \xrightarrow{RSK} \{(P, Q) \mid P, Q \in CST(\lambda),\ \lambda \vdash n\}. \tag{9.20}$$

in which the weights of the tableaux P, Q are $x^P = x_{p_1} \cdots x_{p_n}$, $x^Q = x_{q_1} \cdots x_{q_n}$.

On the other hand, a generalized permutation $w = \binom{\mathbf{q}}{\mathbf{p}}$ as in (9.19) can be specified by an infinite matrix $M = [m_{ij}]_{i,j \in \mathbb{N}}$ in which m_{ij} is the number of occurrences of (q_i, p_i) in w. For example, the generalized permutation w of Example 9.22 corresponds to the integer matrix

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 2 & 0 & 0 & \cdots \\ 0 & 2 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

For each such integer matrix M, let x_i and y_j be the sum of the entries in the i^{th} row and the j^{th} column, respectively. Now here comes a key observation: The generating function for matrices by these weights is precisely the Cauchy kernel. That is,

$$\sum_{M=[m_{ij}]\in\mathbb{N}^{n\times n}} \prod_{i,j} (x_i y_j)^{m_{ij}} \;=\; \prod_{i,j\geq 1} \sum_{m_{ij}=0}^{\infty} (x_i y_j)^{m_{ij}} \;=\; \prod_{i,j\geq 1} \frac{1}{1-x_i y_j} \;=\; \Omega.$$

On the other hand,

$$\begin{aligned} \Omega = \prod_{i,j\geq 1} \frac{1}{1-x_i y_j} &= \sum_{M=[m_{ij}]\in\mathbb{N}^{n\times n}} \prod_{i,j} (x_i y_j)^{m_{ij}} = \sum_{n\in\mathbb{N}} \sum_{w=\binom{\mathbf{q}}{\mathbf{p}}} x_{a_1}\cdots x_{a_n}\, y_{b_1}\cdots y_{b_n} \\ &= \sum_{\lambda} \sum_{P,Q\,\in\, CST(\lambda)} x^P y^Q \qquad\qquad \text{(by RSK)} \\ &= \sum_{\lambda} \left(\sum_{P\in CST(\lambda)} x^P \right) \left(\sum_{Q\in CST(\lambda)} y^Q \right) \;=\; \sum_{\lambda} s_\lambda(x) s_\lambda(y). \end{aligned}$$

We have proven:

Theorem 9.23. $\Omega = \sum_\lambda s_\lambda(x)s_\lambda(y)$. *Consequently, the Schur functions form an orthonormal $\mathbb{Z}$-basis for Λ under the Hall inner product.*

9.9. **The Frobenius Characteristic.** Let R be a ring. Denote by $C\ell_R(\mathfrak{S}_n)$ the vector space of R-valued class functions on the symmetric group $\mathfrak{S}_n$. If no R is specified, we assume $R = \mathbb{C}$. Define

$$C\ell(\mathfrak{S}) \;=\; \bigoplus_{n\geq 0} C\ell(\mathfrak{S}_n).$$

We make $C\ell(\mathfrak{S})$ into a graded ring as follows. Let $n_1, n_2 \in \mathbb{N}$ and $n = n_1 + n_2$. For $f_1 \in C\ell(\mathfrak{S}_{n_1})$ and $f_2 \in C\ell(\mathfrak{S}_{n_2})$, we can define a function $f_1 \otimes f_2 \in C\ell(\mathfrak{S}_{n_1,n_2}) := C\ell(\mathfrak{S}_{n_1} \times \mathfrak{S}_{n_2})$ by

$$(f_1 \otimes f_2)(w_1, w_2) = f_1(w_1) f_2(w_2).$$

There is a natural inclusion of groups $\mathfrak{S}_{n_1} \times \mathfrak{S}_{n_2} \hookrightarrow \mathfrak{S}_n$: let $\mathfrak{S}_{n_1}$ act on $\{1, 2, \ldots, n_1\}$ and let $\mathfrak{S}_{n_2}$ act on $\{n_1+1, n_1+2, \ldots, n_1+n_2\}$. Thus we can define $f_1 \cdot f_2 \in C\ell(\mathfrak{S}_n)$ by means of the induced "character", applying formula (8.45):

$$f_1 \cdot f_2 \;=\; \operatorname{Ind}_{\mathfrak{S}_{n_1,n_2}}^{\mathfrak{S}_n} (f_1 \otimes f_2) \;=\; \frac{1}{n_1!\, n_2!} \sum_{\substack{g\in\mathfrak{S}_n:\\ g^{-1}wg\in\mathfrak{S}_{n_1,n_2}}} (f_1 \otimes f_2)(g^{-1}wg).$$

This product makes $C\ell(\mathfrak{S})$ into a commutative graded $\mathbb{C}$-algebra. (We omit the proof; one has to check properties like associativity.)

For a partition $\lambda \vdash n$, let 1_λ be the indicator function on the conjugacy class $C_\lambda \subset \mathfrak{S}_n$, and let

$$\mathfrak{S}_\lambda \;=\; \mathfrak{S}_{\{1,\,\dots,\,\lambda_1\}} \;\times\; \mathfrak{S}_{\{\lambda_1+1,\,\dots,\,\lambda_1+\lambda_2\}} \;\times\; \cdots \;\times\; \mathfrak{S}_{\{n-\lambda_\ell+1,\,\dots,\,n\}} \;\subset\; \mathfrak{S}_n.$$

For $w \in \mathfrak{S}_n$, denote by $\lambda(w)$ the cycle-shape of w, expressed as a partition.

Definition 9.24. The **Frobenius characteristic** is the map

$$\mathbf{ch} : \; C\ell_{\mathbb{C}}(\mathfrak{S}) \to \Lambda_{\mathbb{C}}$$

defined on $f \in C\ell(\mathfrak{S}_n)$ by

$$\mathbf{ch}(f) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \overline{f(w)} p_{\lambda(w)}.$$

An equivalent definition is as follows. Define a function $\psi : \mathfrak{S}_n \to \Lambda^n$ by

$$\psi(w) = p_{\lambda(w)}. \tag{9.21}$$

This is in fact a class function, albeit with values in Λ instead of the usual $\mathbb{C}$. Nonetheless, we can write

$$\mathbf{ch}(f) = \langle f, \psi \rangle_{\mathfrak{S}_n}$$

and it is often convenient to work with this formula for the Frobenius characteristic.

Theorem 9.25. *The Frobenius characteristic* $\mathbf{ch} : C\ell(\mathfrak{S}) \to \Lambda$ *has the following properties:*

(i) $\mathbf{ch}(1_\lambda) = p_\lambda / z_\lambda$.
(ii) $\mathbf{ch}$ *is an isometry, i.e., it preserves inner products:*

$$\langle f, g \rangle_{\mathfrak{S}_n} \;=\; \langle \mathbf{ch}(f), \mathbf{ch}(g) \rangle_\Lambda.$$

(iii) $\mathbf{ch}$ *is a ring isomorphism.*
(iv) $\mathbf{ch}$ *restricts to an isomorphism* $C\ell_{\mathbb{Z}}(\mathfrak{S}) \to \Lambda_{\mathbb{Z}}$.
(v) $\mathbf{ch}(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \chi_{\mathrm{triv}}) = h_\lambda$.
(vi) $\mathbf{ch}(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \chi_{\mathrm{sign}}) = e_\lambda$.
(vii) *The irreducible characters of* $\mathfrak{S}_n$ *are* $\{\mathbf{ch}^{-1}(s_\lambda) \mid \lambda \vdash n\}$.
(viii) *For all characters* χ, *we have* $\mathbf{ch}(\chi \otimes \chi_{\mathrm{sign}}) = \omega(\mathbf{ch}(\chi))$.

Proof. Recall (9.10) that $|C_\lambda| = n!/z_\lambda$, where $z_\lambda = 1^{r_1} r_1!\, 2^{r_2} r_2! \dots$, where r_i is the number of occurrences of i in λ. Therefore

$$\mathbf{ch}(1_\lambda) \;=\; \frac{1}{n!} \sum_{w \in C_\lambda} p_\lambda \;=\; p_\lambda / z_\lambda.$$

It follows that $\mathbf{ch}$ is (at least) a graded $\mathbb{C}$-vector space isomorphism, since $\{1_\lambda\}$ and $\{p_\lambda/z_\lambda\}$ are graded $\mathbb{C}$-bases for $C\ell(\mathfrak{S})$ and Λ respectively).

We now show that $\mathbf{ch}$ is an isometry. It suffices to check it on these bases. Let $\lambda, \mu \vdash n$; then

$$\langle 1_\lambda,\, 1_\mu \rangle_{\mathfrak{S}_n} \;=\; \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \overline{1_\lambda(w)} 1_\mu(w) \;=\; \frac{1}{n!} |C_\lambda| \delta_{\lambda\mu} \;=\; \delta_{\lambda\mu}/z_\lambda,$$

$$\left\langle \frac{p_\lambda}{z_\lambda},\, \frac{p_\mu}{z_\mu} \right\rangle_\Lambda \;=\; \frac{1}{\sqrt{z_\lambda z_\mu}} \left\langle \frac{p_\lambda}{\sqrt{z_\lambda}},\, \frac{p_\mu}{\sqrt{z_\mu}} \right\rangle_\Lambda \;=\; \frac{1}{\sqrt{z_\lambda z_\mu}} \delta_{\lambda\mu} \;=\; \delta_{\lambda\mu}/z_\lambda.$$

Next we check that $\mathbf{ch}$ is a ring homomorphism (hence an isomorphism). Let $f \in \mathfrak{S}_j$, $g \in \mathfrak{S}_k$, and $n = j + k$. Then

$$\mathbf{ch}(f \cdot g) \;=\; \left\langle \operatorname{Ind}_{\mathfrak{S}_j \times \mathfrak{S}_k}^{\mathfrak{S}_n} (f \otimes g),\; \psi \right\rangle_{\mathfrak{S}_n}$$

(where ψ is as in (9.21))

$$= \left\langle f \otimes g,\ \mathrm{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_j \times \mathfrak{S}_k} \psi \right\rangle_{\mathfrak{S}_j \times \mathfrak{S}_k}$$

(by Frobenius reciprocity)

$$\begin{aligned}
&= \frac{1}{j!\,k!} \sum_{(w,x)\in \mathfrak{S}_j \times \mathfrak{S}_k} \overline{f \otimes g(w,x)}\ p_{\lambda(w,x)} \\
&= \left(\frac{1}{j!} \sum_{w\in\mathfrak{S}_j} \overline{f(w)}\ p_{\lambda(w)} \right) \left(\frac{1}{k!} \sum_{x\in\mathfrak{S}_k} \overline{g(x)}\ p_{\lambda(x)} \right) \\
&= \mathbf{ch}(f)\ \mathbf{ch}(g).
\end{aligned}$$

Next, observe that

$$\begin{aligned}
\mathbf{ch}(\chi_{\mathrm{triv}}(\mathfrak{S}_n)) &= \langle \chi_{\mathrm{triv}}, \psi \rangle_{\mathfrak{S}_n} \\
&= \frac{1}{n!} \sum_{w\in\mathfrak{S}_n} p_{\lambda(w)} \\
&= \sum_{\lambda\vdash n} \frac{p_\lambda}{z_\lambda} \ = \ h_n, \\
\mathbf{ch}(\chi_{\mathrm{sign}}(\mathfrak{S}_n)) &= \langle \chi_{\mathrm{sign}}, \psi \rangle_{\mathfrak{S}_n} \\
&= \frac{1}{n!} \sum_{w\in\mathfrak{S}_n} \varepsilon_{\lambda(w)} p_{\lambda(w)} \\
&= \sum_{\lambda\vdash n} \varepsilon_\lambda \frac{p_\lambda}{z_\lambda} \ = \ e_n,
\end{aligned}$$

and since $\mathbf{ch}$ is a ring homomorphism, we obtain

$$h_\lambda = \prod_{i=1}^{\ell} h_{\lambda_i} = \prod_{i=1}^{\ell} \mathbf{ch}(\chi_{\mathrm{triv}}(\mathfrak{S}_n)) = \mathbf{ch}\left(\prod_{i=1}^{\ell} \chi_{\mathrm{triv}}(\mathfrak{S}_n) \right) = \mathbf{ch}(\mathrm{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_\lambda} \chi_{\mathrm{triv}})$$

and likewise $e_\lambda = \mathbf{ch}(\mathrm{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_\lambda} \chi_{\mathrm{sign}})$. □

The Frobenius characteristic allows us to translate back and forth between representations (equivalently, characters) of symmetric groups, and symmetric functions; in particular, it reveals that the Schur functions, which seem much less natural than the m's, e's, h's or p's, are in some ways the most important basis for Λ. It is natural to ask how to multiply them. That is, suppose that μ, ν are partitions with $|\mu| = q$, $|\nu| = r$. Then the product $s_\mu s_\nu$ is a symmetric function that is homogeneous of degree $n = q + r$, so it has a unique expansion as a linear combination of Schur functions:

$$s_\mu s_\nu = \sum_\lambda c^\lambda_{\mu,\nu} s_\lambda \tag{9.22}$$

with $c^\lambda_{\mu,\nu} \in \mathbb{Z}$. These numbers $c^\lambda_{\mu,\nu} \in \mathbb{Z}$ (i.e., the structure coefficients for Λ, regarded as an algebra in the Schur functions) are called the *Littlewood-Richardson coefficients*. (Note that they must be integers, because $s_\mu s_\nu$ is certainly a $\mathbb{Z}$-linear combination of the monomial symmetric functions, and the Schurs have the same $\mathbb{Z}$-linear span.) Equivalently, we can define the $c^\lambda_{\mu,\nu}$ in terms of the Hall inner product:

$$c^\lambda_{\mu,\nu} = \langle s_\mu s_\nu,\ s_\lambda \rangle_\Lambda.$$

Via the Frobenius characteristic, we can interpret the $c^\lambda_{\mu,\nu}$ in terms of representations of the symmetric group:

$$c^\lambda_{\mu,\nu} \ = \ \left\langle \mathrm{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_q \times \mathfrak{S}_r}(\chi^\mu \otimes \chi^\nu),\ \chi^\lambda \right\rangle_{\mathfrak{S}_n} \ = \ \left\langle \chi^\mu \otimes \chi^\nu,\ \mathrm{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_q \times \mathfrak{S}_r}(\chi^\lambda) \right\rangle_{\mathfrak{S}_q \times \mathfrak{S}_r}$$

where the second equality comes from Frobenius reciprocity.

A number of natural questions about representations of $\mathfrak{S}_n$ can now be translated into tableau combinatorics.

(i) Find a combinatorial interpretation for the coefficients $c^{\lambda}_{\mu,\nu}$. (Answer: The Littlewood-Rchardson Rule, in all its various guises.)
(ii) Find a combinatorial interpretation for the value of the irreducible characters χ_λ on the conjugacy class C_μ. (Answer: The Murnaghan-Nakayama Rule.)
(iii) Determine the dimension f^λ of the irreducible character χ_λ. (Answer: The Hook-Length Formula.)

9.10. **Equivalent Versions of RSK: Knuth equivalence and jeu de taquin.**

Definition 9.26. Let $\mathbf{b}, \mathbf{b}'$ be finite ordered lists of integers (or "words in the alphabet $\mathbb{N}$"). We say that $\mathbf{b}, \mathbf{b}'$ are *Knuth equivalent*, written $\mathbf{b} \underset{K}{\sim} \mathbf{b}'$, if one can be obtained from the other by a sequence of transpositions as follows:

1. If $x \leq y < z$, then $\cdots xzy \cdots \underset{K}{\sim} \cdots zxy \cdots$.
2. If $x < y \leq z$, then $\cdots yxz \cdots \underset{K}{\sim} \cdots yzx \cdots$.

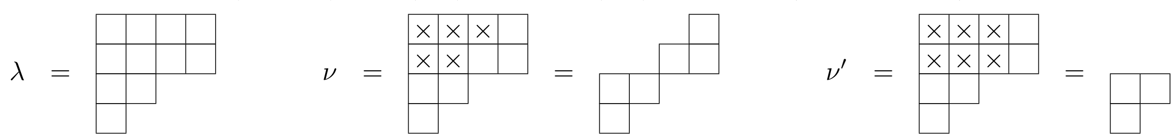
(Here the notation $\cdots xzy \cdots$ means a word that contains the letters x, z, y consecutively.)

For example, $2122\underline{131}2 \underset{K}{\sim} 2122\underline{311}2$ by Rule 1, and $212\underline{231}12 \underset{K}{\sim} 212\underline{213}12$ by Rule 2 (applied in reverse).

This definition looks completely unmotivated at first, but hold that thought!

Definition 9.27. Let λ, μ be partitions with $\mu \subseteq \lambda$. The *skew (Ferrers) shape* λ/μ is defined by removing from λ the boxes in μ.

For example, if $\lambda = (4,4,2,1)$, $\mu = (3,2)$, and $\mu' = (3,3)$, then $\nu = \lambda/\mu$ and $\nu' = \lambda/\mu'$ are as follows:



(where the $\times$'s mean "delete this box"). Note that there is no requirement that a skew shape be connected.

Definition 9.28. Let $\nu = \lambda/\mu$ be a skew shape. A *column-strict (skew) tableau of shape ν* is a filling of the boxes of ν with positive integers such that each row is weakly increasing eastward and each column is strictly increasing southward. (Note that if $\mu = \emptyset$, this is just a CST; see Definition 9.3.)

Here are a couple of examples. Again, there is no requirement that a skew tableau be connected (as in T' below).

We now define an equivalence relation on column-strict skew tableaux, called *jeu de taquin*[15]. The rule is as follows:

[15] French for "sliding game", roughly; it refers to the 15-square puzzle with sliding tiles, invented and popularized by Sam Loyd, that used to come standard on every Macintosh in about 1985.

That is, for each inner corner of T — that is, an empty cell that has numbers to the south and east, say x and y — then we can either slide x north into the empty cell (if $x \leq y$) or slide y west into the empty cell (if $x > y$). It is not hard to see that any such slide (hence, any sequence of slides) preserves the property of column-strictness.

For example, the following is a sequence of jeu de taquin moves. The bullets $\bullet$ denote the inner corner that is being slid into.

$$
\begin{array}{l}
\begin{ytableau} \none & \bullet & 1 & 4 \\ \none & 1 & 2 \\ 2 & 3 & 4 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} \none & 1 & 1 & 4 \\ \none & \bullet & 2 \\ 2 & 3 & 4 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} \none & 1 & 1 & 4 \\ \none & 2 & \bullet \\ 2 & 3 & 4 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} \none & 1 & 1 & 4 \\ \bullet & 2 & 4 \\ 2 & 3 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} \none & 1 & 1 & 4 \\ 2 & 2 & 4 \\ \bullet & 3 \end{ytableau}
\\[2ex]
\;\rightarrow\;
\begin{ytableau} \bullet & 1 & 1 & 4 \\ 2 & 2 & 4 \\ 3 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} 1 & \bullet & 1 & 4 \\ 2 & 2 & 4 \\ 3 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} 1 & 1 & \bullet & 4 \\ 2 & 2 & 4 \\ 3 \end{ytableau}
\;\rightarrow\;
\begin{ytableau} 1 & 1 & 4 & 4 \\ 2 & 2 \\ 3 \end{ytableau}
\end{array}
\tag{9.23}
$$

If two skew tableaux T, T' can be obtained from each other by such slides (or by their reverses), we say that they are *jeu de taquin equivalent*, denoted $T \underset{J}{\sim} T'$. Note that any skew column-strict tableau T is jeu de taquin equivalent to an ordinary CST (called the *rectification* of T); see, e.g., the example (9.23) above. In fact, the rectification is unique; the order in which we choose inner corners does not matter.

Definition 9.29. Let T be a column-strict skew tableau. The *row-reading word* of T, denoted $\operatorname{row}(T)$, is obtained by reading the rows left to right, bottom to top.

For example, the reading words of the skew tableaux in (9.23) are

$$2341214, \quad 2342114, \quad 2342114, \quad 2324114, \quad 3224114, \quad 3224114, \quad 3224114, \quad 3224114, \quad 3221144.$$

If T is an ordinary (not skew) tableau, then it is determined by its row-reading word, since the "line breaks" occur exactly at the strict decreases of $\operatorname{row}(T)$. For skew tableaux, this is not the case. Note that some of the slides in (9.23) do not change the row reading word; as a simpler example, the following skew tableaux both have reading word 122:

$$
\begin{ytableau} 1 & 2 & 2 \end{ytableau}
\qquad\qquad
\begin{ytableau} \none & \none & 2 \\ 1 & 2 \end{ytableau}
\qquad\qquad
\begin{ytableau} \none & 2 & 2 \\ 1 \end{ytableau}
$$

On the other hand, it's not hard to se that rectifying the second or third tableau will yield the first; therefore, they are all jeu de taquin equivalent.

For a word $\mathbf{b}$ on the alphabet $\mathbb{N}$, let $P(\mathbf{b})$ denote its insertion tableau under the RSK algorithm. (That is, construct a generalized permutation $\binom{\mathbf{q}}{\mathbf{b}}$ in which $\mathbf{q}$ is any word; run RSK; and remember only the tableau P, so that the choice of $\mathbf{q}$ does not matter.)

Theorem 9.30. *(Knuth–Schützenberger) For two words $\mathbf{b}, \mathbf{b}'$, the following are equivalent:*

(i) $P(\mathbf{b}) = P(\mathbf{b}')$.
(ii) $\mathbf{b} \underset{K}{\sim} \mathbf{b}'$.
(iii) $T \underset{J}{\sim} T'$*, for any (or all) column-strict skew tableaux T, T' with row-reading words $\mathbf{b}, \mathbf{b}'$ respectively.*

This is sometimes referred to (e.g., in Fulton's book) as the equivalence of "bumping" (the RSK algorithm as presented in Section 9.6) and "sliding" (jeu de taquin).

9.11. **Skew Tableaux and the Littlewood-Richardson Rule.** Let $\nu = \lambda/\mu$ be a skew shape, and let $CST(\lambda/\mu)$ denote the set of all column-strict skew tableaux of shape λ/μ. It is natural to define the *skew Schur function*

$$s_{\lambda/\mu}(x_1, x_2, \dots) = \sum_{T \in CST(\lambda/\mu)} x_T.$$

For example, suppose that $\lambda = (2,2)$ and $\mu = (1)$, so that

$$\nu = \begin{array}{cc} & \boxed{} \\ \boxed{} & \boxed{} \end{array}.$$

What are the possibilities for $T \in CST(\nu)$? Clearly the entries cannot all be equal. If $a < b < c$, then there are two ways to fill ν with a, b, c (left, below). If $a < b$, then there is one way to fill ν with two a's and one b (center), and one way to fill ν with one a and two b's (right).

$$\begin{array}{|c|c|} \hline & a \\ \hline b & c \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline & b \\ \hline a & c \\ \hline \end{array} \qquad\qquad \begin{array}{|c|c|} \hline & a \\ \hline a & b \\ \hline \end{array} \qquad\qquad \begin{array}{|c|c|} \hline & a \\ \hline b & b \\ \hline \end{array}$$

Therefore, $s_\nu = 2m_{111} + m_{21}$ (these are monomial symmetric functions). In fact, skew Schur functions are always symmetric. This is not obvious, but is not too hard to prove. (Like ordinary Schur functions, it is fairly easy to see that they are *quasisymmetric*.) Therefore, we can write

$$s_{\lambda/\mu} = \sum_\nu \tilde{c}_{\lambda/\mu,\nu} s_\nu$$

where $\tilde{c}_{\lambda/\mu,\nu} \in \mathbb{Z}$ for all λ, μ, ν. The punchline is that the tildes are unnecessary: these numbers are in fact the Littlewood-Richardson coefficients $c^\lambda_{\mu,\nu}$ of equation (9.22). Better yet, they are symmetric in μ and ν.

Proposition 9.31. *Let* $\mathbf{x} = \{x_1, x_2, \dots\}$, $\mathbf{y} = \{y_1, y_2, \dots\}$ *be two countably infinite sets of variables. Think of them as an alphabet with* $1 < 2 < \cdots < 1' < 2' < \cdots$. *Then*

$$s_\lambda(\mathbf{x}, \mathbf{y}) = \sum_{\mu \subseteq \lambda} s_\mu(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}).$$

Proof. Every $T \in CST(\lambda)$ labeled with $1, 2, \dots, 1', 2', \dots$ consists of a CST of shape μ filled with $1, 2, \dots$ (for some $\mu \subseteq \lambda$) together with a CST of shape λ/μ filled with $1', 2', \dots$. $\square$

Theorem 9.32. *For all partitions* λ, μ, ν, *we have*

$$\tilde{c}_{\lambda/\mu,\nu} = c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}.$$

Equivalently,

$$\langle s_\mu s_\nu,\ s_\lambda \rangle_\Lambda \ = \ \langle s_\nu,\ s_{\lambda/\mu} \rangle_\Lambda.$$

Proof. We need three countably infinite sets of variables $\mathbf{x}, \mathbf{y}, \mathbf{z}$ for this. Consider the "double Cauchy kernel"

$$\Omega(\mathbf{x}, \mathbf{z})\Omega(\mathbf{y}, \mathbf{z}) \ = \ \prod_{i,j} (1 - x_i z_j)^{-1} \prod_{i,j} (1 - y_i z_j)^{-1}.$$

On the one hand, expanding both factors in terms of Schur functions and then applying the definition of the Littlewood-Richardson coefficients to the $\mathbf{z}$ terms gives

$$\begin{aligned} \Omega(\mathbf{x}, \mathbf{z})\Omega(\mathbf{y}, \mathbf{z}) \ &= \ \left(\sum_\mu s_\mu(\mathbf{x}) s_\mu(\mathbf{z})\right)\left(\sum_\nu s_\nu(\mathbf{y}) s_\nu(\mathbf{z})\right) \ = \ \sum_{\mu,\nu} s_\mu(\mathbf{x}) s_\nu(\mathbf{y}) s_\mu(\mathbf{z}) s_\nu(\mathbf{z}) \\ &= \ \sum_{\mu,\nu} s_\mu(\mathbf{x}) s_\nu(\mathbf{y}) \sum_\lambda c^\lambda_{\mu,\nu} s_\lambda(\mathbf{z}). \end{aligned} \tag{9.24}$$

On the other hand, we also have (formally setting $s_{\lambda/\mu} = 0$ if $\mu \not\subseteq \lambda$)

$$
\begin{aligned}
\Omega(\mathbf{x},\mathbf{z})\Omega(\mathbf{y},\mathbf{z}) &= \sum_\lambda s_\lambda(\mathbf{x}\,\mathbf{y})s_\lambda(\mathbf{z}) \;=\; \sum_\lambda \sum_{\mu\subseteq\lambda} s_\mu(\mathbf{x})s_{\lambda/\mu}(\mathbf{y})s_\lambda(\mathbf{z}) \\
&= \sum_\lambda \sum_\mu s_\mu(\mathbf{x})s_\lambda(\mathbf{z}) \sum_\nu \tilde{c}_{\lambda/\mu,\nu} s_\nu(\mathbf{y}) \\
&= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y}) \sum_\lambda s_\lambda(\mathbf{z})\tilde{c}_{\lambda/\mu,\nu}. \qquad (9.25)
\end{aligned}
$$

(The first equality is perhaps clearer in reverse; think about how to express the right-hand side as an infinite product over the variable sets $\mathbf{x} \cup \mathbf{y}$ and $\mathbf{z}$. The second equality uses Proposition 9.31.) Now the theorem follows from the equality of (9.24) and (9.25). ☐

There are a lot of combinatorial interpretations of the Littlewood-Richardson numbers. Here is one:

Theorem 9.33 (Littlewood-Richardson Rule)**.** *$c^\lambda_{\mu,\nu}$ equals the number of column-strict tableaux T of shape λ/μ, and content ν such that the reverse of* row(T) *is a* ballot sequence *(or* Yamanouchi word*, or* lattice permutation*): that is, each initial sequence of it contains at least as many 1's as 2's, at least as many 2's as 3's, et cetera.*

Important special cases are the *Pieri rules*, which describe how to multiply by the Schur function corresponding to a single row or column (i.e., by an h or an e.)

Theorem 9.34 (Pieri Rules)**.** *Let (k) denote the partition with a single row of length k, and let (1^k) denote the partition with a single column of length k. Then:*

$$s_\lambda s_{(k)} s_\lambda h_k = \sum_\mu s_\mu$$

where mu ranges over all partitions obtained from λ by adding k boxes, no more than one in each column; and

$$s_\lambda s_{(1^k)} = s_\lambda e_k \sum_\mu s_\mu$$

where mu ranges over all partitions obtained from λ by adding k boxes, no more than one in each row.

Another important, even more special case is

$$s_\lambda s_1 = \sum_\mu s_\mu$$

where μ ranges over all partitions obtained from λ by adding a single box. Via the Frobenius characteristic, this gives a "branching rule" for how the restriction of an irreducible character of $\mathfrak{S}_n$ splits into a sum of irreducibles when restricted:

$$\mathrm{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_{n-1}}(\chi^\lambda) = \oplus_\mu \chi^\mu$$

where now μ ranges over all partitions obtained from λ by *deleting* a single box.

9.12. **The Murnaghan-Nakayama Rule.** We know from Theorem 9.25 that the irreducible characters of $\mathfrak{S}_n$ are $\chi^\lambda = \mathbf{ch}^{-1}(s_\lambda)$ for $\lambda \vdash n$. The Murnaghan-Nakayama Rule gives a formula for the value of the character χ^λ on the conjugacy class C_μ in terms of *rim-hook tableaux*. Here is an example of a rim-hook tableau of shape $\lambda = (5,4,3,3,1)$ and content $\mu = (6,3,3,2,1,1)$:

1	1	1	4	4
1	2	3	3	
1	2	3		
1	2	6		
5				

Note that the columns and row are weakly increasing, and for each i, the set $H_i(T)$ of cells containing an i is contiguous.

Theorem 9.35 (Murnaghan-Nakayama Rule (1937))**.**

$$\chi^\lambda(C_\mu) = \sum_{\substack{\text{rim-hook tableaux } T\\ \text{of shape } \lambda \text{ and content } \mu}} \prod_{i=1}^{n} (-1)^{1+\mathrm{ht}(\mathrm{H_i(T)})}.$$

For example, the heights of $H_1, \dots, H_6$ in the rim-hook tableau above are 4, 3, 2, 1, 1, 1. There are an even number of even heights, so this rim-hook tableau contributes 1 to $\chi\lambda(C_\mu)$.

An important special case is when $\mu = (1, 1, \dots, 1)$, i.e., since then $\chi^\lambda(C_\mu) = \chi^\lambda(1_{\mathfrak{S}_n})$ i.e., the dimension of the irreducible representation S^λ of $\mathfrak{S}_n$ indexed by λ. On the other hand, a rim-hook tableau of content μ is just a standard tableau. So the Murnaghan-Nakayama Rule implies the following:

Corollary 9.36. $\dim S^\lambda = f^\lambda$.

This begs the question of how to calculate f^λ (which you may have been wondering anyway). There is a beautiful formula for f^λ called the *hook-length formula*; we will first need another elegeant piece of symmetric-function combinatorics.

9.13. **The Jacobi-Trudi Determinant Definition of Schur Functions.** There is a formula for the Schur function s_λ as a determinant of a matrix whose entries are h_n's or e_n's, with a stunning proof due to the ideas of Lindström, Gessel, and Viennot. This exposition follows closely that of [12, §4.5]. Define $h_0 = e_0 = 1$ and $h_k = e_k = 0$ for $k < 0$.

Theorem 9.37. *For any* $\lambda = (\lambda_1 \dots, \lambda_\ell)$ *we have*

$$s_\lambda = \left| h_{\lambda_i - i + j} \right|_{i,j=1\dots,\ell} \tag{9.26}$$

and

$$s_{\tilde\lambda} = \left| e_{\lambda_i - i + j} \right|_{i,j=1\dots,\ell}. \tag{9.27}$$

For example,

$$s_{311} = \begin{vmatrix} h_3 & h_4 & h_5 \\ h_0 & h_1 & h_2 \\ h_{-1} & h_0 & h_1 \end{vmatrix} = \begin{vmatrix} h_3 & h_4 & h_5 \\ 1 & h_1 & h_2 \\ 0 & 1 & h_1 \end{vmatrix} = h_{311} + h_5 - h_{41} - h_{32}.$$

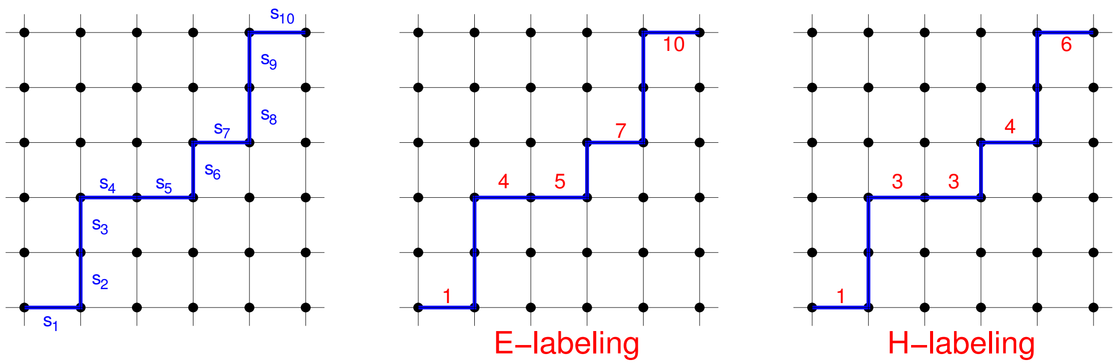
We will just prove the h case

Proof. **Step 1: For $n \in \mathbb{N}$, express h_n, e_n as generating functions for *lattice paths*.**

Consider walks that start at a point $(a, b) \in \mathbb{N}^2$ and move north or east one step at a time. For every path that we consider, the number of eastward steps must be finite, but the number of northward steps is allowed to be infinite (so that the "ending point" can be of the form (x, ∞) for some $x \in \mathbb{N}$).

Let $s_1, s_2, \ldots,$ be the steps of P. We assign labels to the eastward steps in the two following ways.

E-labeling: $L(s_i) = i$.
H-labeling: $\hat{L}(s_i) = 1+$(number of northward steps preceding s_i). (Alternately, if the path starts on the x-axis, this is $1 +$ the y-coordinate of the step.)



The northward steps don't get labels. The set of labels of all eastward steps of P gives rise to monomials

$$x^P = \prod_i x_{L(s_i)}, \qquad \hat{x}^P = \prod_i x_{\hat{L}(s_i)}$$

in countably infinitely many variables $x_1, x_2, \ldots$. Note that the path P can be recovered from either of these two monomials. Moreover, the monomial $x_{L(s_i)}$ is always square-free, while $x_{\hat{L}(s_i)}$ can be any monomial. Therefore

$$e_n = \sum_P x^P, \qquad h_n = \sum_P \hat{x}^P$$

where both sums run over all lattice paths from some fixed starting point (a, b) to $(a + n, \infty)$.

Step 2: For a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$, express h_λ, e_λ as generating functions for families of lattice paths.

Let there be given sets $U = \{u_1, \ldots, u_\ell\}$, $V = \{v_1, \ldots, v_\ell\}$ of lattice points in $\mathbb{Z} \times (\mathbb{Z} \cup \{\infty\})$. A *$U,V$-lattice path family* is a tuple $\mathbf{P} = \{\pi, P_1, \ldots, P_\ell\}$, where $\pi \in \mathfrak{S}_\ell$ and each P_i is a path from u_i to $v_{\pi(i)}$. Define

$$x^{\mathbf{P}} = \prod_{i=1}^{\ell} x^{P_i}, \qquad \hat{x}^{\mathbf{P}} = \prod_{i=1}^{\ell} \hat{x}^{P_i}, \qquad (-1)^{\mathbf{P}} = \varepsilon(\pi)$$

where ε denotes the sign of π.

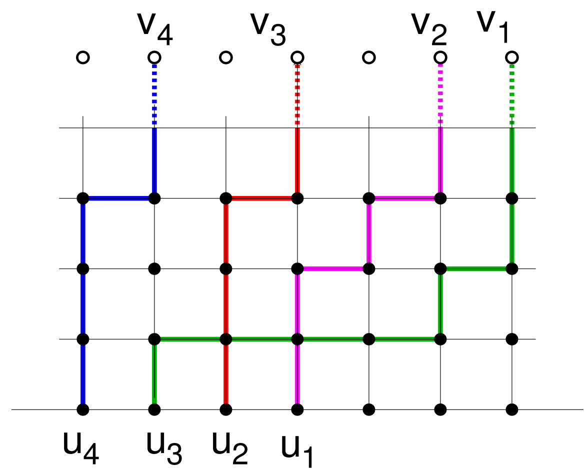
For a partition λ of length ℓ, a **λ-path family** is a (U, V)-path family, where U, V are defined by

$$u_i = (1 - i, 0), \qquad v_i = (\lambda_i - i + 1, \infty).$$

for $1 \leq i \leq \ell$. For instance, if $\lambda = (3, 3, 2, 1)$ then

$$\begin{array}{llll} u_1 = (0,0), & u_2 = (-1,0), & u_3 = (-2,0), & u_4 = (-3,0), \\ v_1 = (3,\infty), & v_2 = (2,\infty), & v_3 = (0,\infty), & v_4 = (-2,\infty) \end{array}$$

and the following figure is an example of a path family corresponding to λ. Note that $\pi = (2\ 3)$ and so $(-1)^{\mathbf{P}} = -1$.



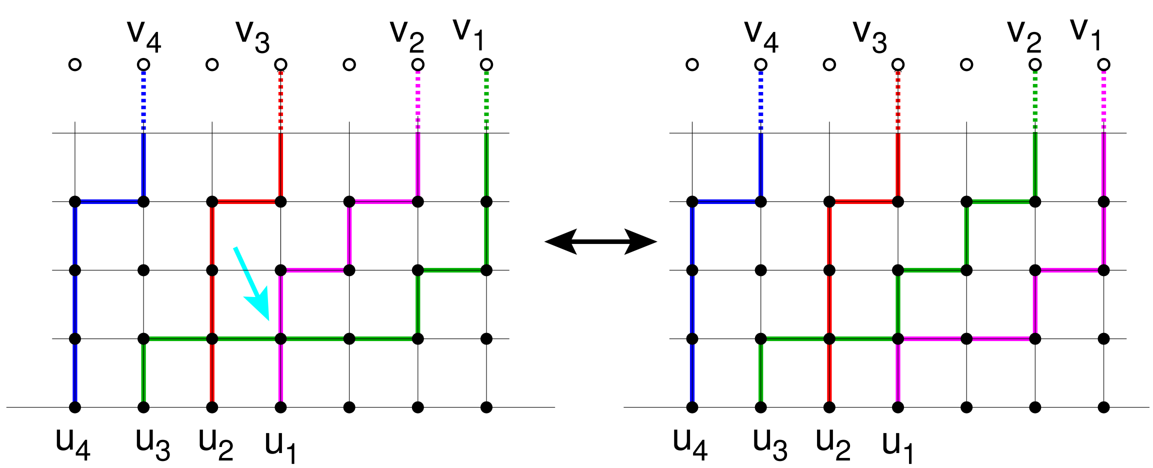
Expanding the determinant on the right-hand side of (9.26), we see that

$$
\begin{aligned}
|h_{\lambda_i-i+j}|_{i,j=1\dots,\ell} &= \sum_{\pi\in\mathfrak{S}_\ell} \varepsilon(\pi) \prod_{i=1}^{\ell} h_{\lambda_i-i+\pi(i)} \\
&= \sum_{\pi\in\mathfrak{S}_\ell} \varepsilon(\pi) \sum_{\mathbf{P}=(\pi,P_1,\dots,P_\ell)} \hat{x}^{P_1}\cdots\hat{x}^{P_\ell} \\
&= \sum_{\mathbf{P}} (-1)^{\mathbf{P}} \hat{x}^{\mathbf{P}}. \qquad (9.28)
\end{aligned}
$$

Call a path family *good* if no two of its paths meet in a common vertex.

Step 3: Show that all the terms cancel out except for the good families.

Suppose that two of the paths in $\mathbf{P}$ meet at a common vertex. Define a sign-reversing, weight-preserving involution $\mathbf{P} \mapsto \mathbf{P}^\sharp$ on non-good λ-path families by interchanging two partial paths to the northeast of an first intersection point. The picture looks like this:



We have to have a canonical way of choosing the intersection point so that this interchange gives an involution. For example, choose the southeasternmost point contained in two or more paths, and choose the two paths through it with largest indices. One then checks that

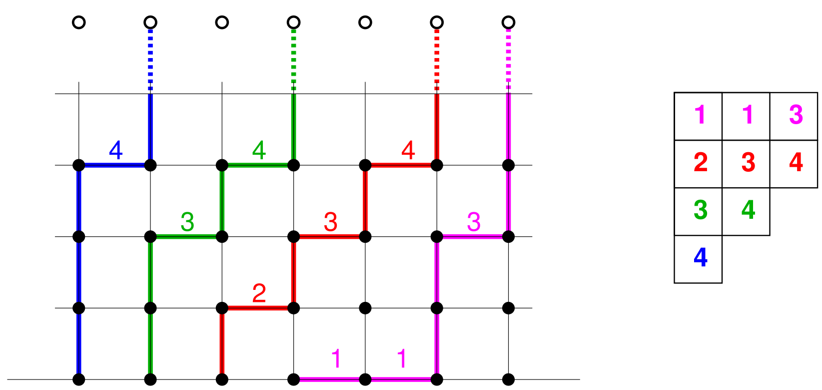
- this operation is an involution on non-good path families;
- $x^{\mathbf{P}} = x^{\mathbf{P}^\sharp}$; and
- $(-1)^{\mathbf{P}} = -(-1)^{\mathbf{P}^\sharp}$.

Therefore by (9.28) we have

$$|h_{\lambda_i - i + j}|_{i,j=1\dots,\ell} = \sum_{\mathbf{P}} (-1)^{\mathbf{P}} \hat{x}^{\mathbf{P}} = \sum_{\mathbf{P}\text{ good}} \hat{x}^{\mathbf{P}}. \tag{9.29}$$

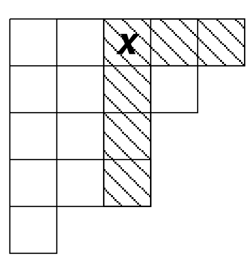
Step 4: Enumerate weights of good path families.

Now, for each good path family, label each path using the H-labeling. Clearly the labels weakly increase as we move north along each path. Moreover, for each j and $i < i'$, the j^{th} step of the path P_i is strictly southeast of the j^{th} step of $P_{i'}$. What this means is that we can construct a column-strict tableau of shape λ by reading off the labels of each path. For example:



Therefore, (9.29) implies that $|h_{\lambda_i - i + j}|_{i,j=1\dots,\ell} = s_\lambda$ as desired. $\square$

9.14. **The Hook-Length Formula.** We now return to the problem of calculating f^λ, the number of standard tableaux of shape λ. As usual, let $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell > 0) \vdash n$. For each cell x in row i and column j of the Ferrers diagram of λ, let $h(x) = h(i,j)$ denote its *hook length*: the number of cells due east of, due south of, or equal to x. In the following example, $h(x) = 6$.



To be precise, if $\tilde{\lambda}$ is the conjugate partition to λ, then

$$h(i,j) = \lambda_i - (i-1) + \tilde{\lambda}_j - (j-1) - 1 = \lambda_i + \tilde{\lambda}_j - i - j + 1. \tag{9.30}$$

In particular,

$$h(i,1) = \lambda_i + \tilde{\lambda}_1 - i = \lambda_i + \ell - i. \tag{9.31}$$

Theorem 9.38 (Hook Formula of Frame, Robinson, and Thrall (1954))**.** *Let $\lambda \vdash n$. Then*

$$f^\lambda \;=\; \frac{n!}{\displaystyle\prod_{x\in\lambda} h(x)}.$$

Example 9.39. For $\lambda = (5,4,3,3,1) \vdash 16$ as above, the tableau of hook lengths is

$$\mathrm{Hook}(\lambda) = \begin{array}{|c|c|c|c|c|} \hline 9 & 7 & 6 & 3 & 1 \\ \hline 7 & 5 & 4 & 1 \\ \cline{1-4} 5 & 3 & 2 \\ \cline{1-3} 4 & 2 & 1 \\ \cline{1-3} 1 \\ \cline{1-1} \end{array}$$

so $f^\lambda = 14!/(9 \cdot 7^2 \cdot 6 \cdot 5^2 \cdot 4^2 \cdot 3^2 \cdot 2^2 \cdot 1^4) = 2288$. As another example, if $\lambda = (n,n) \vdash 2n$, the hook lengths are $n+1,\ n,\ n-1,\ \ldots,\ 2$ (in the top row) and $n,\ n-1,\ n-2,\ \ldots,\ 1$ (in the bottom row). Therefore $f^\lambda = \frac{(2n)!}{(n+1)!\ n!} = \frac{1}{n+1}\binom{2n}{n}$, the n^{th} Catalan number (as we already know).

The following method of proof appears in [1]. We need a couple of lemmas about hook numbers.

Lemma 9.40. *Let $\lambda \vdash n$, $\ell = \ell(\lambda)$, and $1 \le i \le \ell$. Then the sequence*

$$\underbrace{h(i,1), \quad \ldots, \quad h(i,\lambda_i)}_{A}, \quad \underbrace{h(i,1) - h(i+1,1), \quad \ldots, \quad h(i,1) - h(\ell,1)}_{B}. \tag{9.32}$$

is a permutation of $\{1,\ldots,h(i,1)\} = \{1,\ldots,\lambda_i + \ell - i\}$.

For example, if $\lambda = (5,4,3,3,1) \vdash 16$ is as above, then the sequences are as follows:

$$\begin{array}{llllll} i=1: & 9,\ 7,\ 6,\ 3,\ 1, & 9-7,\ 9-5,\ 9-4,\ 9-1 & = & 9,\ 7,\ 6,\ 3,\ 1,\ 2,\ 4,\ 5,\ 8 \\ i=2: & 7,\ 5,\ 4,\ 1, & 7-5,\ 7-4,\ 7-1 & = & 7,\ 5,\ 4,\ 1,\ 2,\ 3,\ 6 \\ i=3: & 5,\ 3,\ 2, & 5-4,\ 5-1 & = & 5,\ 3,\ 2,\ 1,\ 4 \\ i=4: & 4,\ 2,\ 1, & 4-1 & = & 4,\ 2,\ 1,\ 3 \\ i=5: & 1 & & = & 1. \end{array}$$

Proof of Lemma 9.40. From the definition, it is immediate that A and B are respectively strictly decreasing and strictly increasing sequences of positive integers $\le h(i,1)$. Moreover, the total length is $\lambda_i + r - i$, which is precisely $h(i,1)$. Therefore, it is sufficient to prove that no element of A equals any element of B, i.e.,

$$\begin{aligned} z = h(i,j) - (h(i,1) - h(k-1)) &= (\lambda_i + \tilde{\lambda}_j - i - j + 1) - (\lambda_i + \tilde{\lambda}_1 - i - 1 + 1) + (\lambda_k + \tilde{\lambda}_1 - k - 1 + 1) \\ &= \lambda_i + \tilde{\lambda}_j - i - j + 1 - \lambda_i - \tilde{\lambda}_1 + i + \lambda_k + \tilde{\lambda}_1 - k \\ &= \tilde{\lambda}_j - j + 1 + \lambda_k - k \\ &= (\tilde{\lambda}_j - k) + (\lambda_k - j) + 1 \end{aligned}$$

is nonzero. (Here we have used (9.30) in the first line.) Indeed, either λ has a box in the position (j,k) or it doesn't. If it does, then $\lambda_k \ge j$ and $\tilde{\lambda}_j \ge k$, so $z > 0$. If it doesn't, then $\lambda_k \le j-1$ and $\tilde{\lambda}_j \le k-1$, so $z < 0$. $\square$

Remark 9.41. Here is a pictorial way to see Lemma 9.40. Make a new tableau $U = U(\lambda)$ of shape λ whose $(1, j)$ entry is $h(1, j)$, and whose (i, j) entry for $i > 1$ is $h(1, j) - h(i, j)$. For $\lambda = (5, 4, 3, 3, 1)$, this tableau is

$$\begin{array}{ccccc} 9 & 7 & 6 & 3 & 1 \\ 2 & 2 & 2 & 2 & \\ 4 & 4 & 4 & & \\ 5 & 5 & 5 & & \\ 8 & & & & \end{array}.$$

Lemma 9.40 says that every hook whose corner is in the first row has distinct entries in U. Observe in the example that the outer corners (boxes that can be deleted) are in the positions (i, j) such that $U(1, j) = U(i, 1)+1$, and the inner corners (boxes that can be added) are the ones with $U(1, j) = U(i, 1)-1$. Adding or deleting a box switches the entry in its row label with the entry above it in the top row:

$$\lambda = (5,4,3,3,1) \qquad \mathrm{Hook}(\lambda) = \begin{array}{ccccc} 9 & 7 & 6 & 3 & 1 \\ 7 & 5 & 4 & 1 & \\ 5 & 3 & 2 & & \\ 4 & 2 & 1 & & \\ 1 & & & & \end{array} \qquad U(\lambda) = \begin{array}{ccccc} 9 & 7 & 6 & 3 & 1 \\ 2 & 2 & 2 & 2 & \\ 4 & 4 & 4 & & \\ 5 & 5 & 5 & & \\ 8 & & & & \end{array}$$

$$\lambda = (5,4,4,3,1) \qquad \mathrm{Hook}(\lambda) = \begin{array}{ccccc} 9 & 7 & 6 & 4 & 1 \\ 7 & 5 & 4 & 2 & \\ 6 & 4 & 3 & 1 & \\ 4 & 2 & 1 & & \\ 1 & & & & \end{array} \qquad U(\lambda) = \begin{array}{ccccc} 9 & 7 & 6 & 4 & 1 \\ 2 & 2 & 2 & 2 & \\ 3 & 3 & 3 & 3 & \\ 5 & 5 & 5 & & \\ 8 & & & & \end{array}$$

This observation can probably be made into a proof of Lemma 9.40 All this can probably be proved by induction on the number of boxes.

Corollary 9.42. *For each i we have*

$$\prod_{j=1}^{\lambda_i} h(i,j) \prod_{k=i+1}^{r} (h(i,1) - h(k,1)) \;=\; h(i,1)!$$

or equivalently

$$\prod_{j=1}^{\lambda_i} h(i,j) \;=\; \frac{h(i,1)!}{\prod_{k=i+1}^{r} h(i,1) - h(k,1)}$$

and therefore

$$\prod_{x\in\lambda} h(x) \;=\; \prod_{i=1}^{r}\prod_{j=1}^{\lambda_i} h(i,j) \;=\; \frac{\prod_{i=1}^{r} h(i,1)!}{\prod_{i=1}^{r}\prod_{k=i+1}^{r} (h(i,1) - h(k,1))} \;=\; \frac{\prod_{i=1}^{r} h(i,1)!}{\prod_{1\le i<j\le r} (h(i,1) - h(j,1))}. \tag{9.33}$$

Proof of the Hook-Length Formula. A standard tableau is just a column-strict tableau with entries $1, \ldots, n$, each occurring once. Therefore, if we expand $s_\lambda = \sum_{T\in CST(\lambda)} x^T$ in the basis of monomial symmetric functions, then the coefficient of $m_{11\cdots1}$ will equal f^λ. Denote by $\Phi(F)$ the coefficient of $m_{11\cdots1}$ in a degree-n symmetric function F; note that Φ is a $\mathbb{C}$-linear map $\Lambda_n \to \mathbb{C}$.

$$\Phi(h_{(\nu_1,\ldots,\nu_r)}) = \frac{n!}{\nu_1!\cdots\nu_r!} \tag{9.34}$$

since the right-hand side counts the number of ways to factor a squarefree monomial in $|\nu|$ variables as the product of an ordered list of r squarefree monomials of degrees $\nu_1, \ldots \nu_r$.

For $1 \le i \le \ell$, let

$$\mu_i = h(i,1) = \lambda_i + \ell - i.$$

Now, the Jacobi-Trudi identity (9.26) gives

$$\begin{aligned}
f^\lambda = \Phi(s_\lambda) &= \Phi\Big(\det[h_{\lambda_i - i + j}]\Big) \\
&= \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) \Phi(h_{\lambda_1 + \pi(1) - 1,\ \dots,\ \lambda_\ell + \pi(\ell) - \ell}) \\
&= \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) \frac{n!}{(\lambda_1 + \pi(1) - 1)! \cdots (\lambda_\ell + \pi(\ell) - \ell)!} && \text{(by (9.34))} \\
&= n!\ \det\left[\frac{1}{(\lambda_i + j - i)!}\right] && \text{(this is the tricky part)} \\
&= \frac{n!}{(\lambda_1 + \ell - 1)! \cdots (\lambda_\ell)}\ \det\left[\frac{(\lambda_i + \ell - i)!}{(\lambda_i + j - i)!}\right] && \text{(this is just multiplying by 1} \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!}\ \det\left[\frac{\mu_i!}{(\mu_i - \ell + j)!}\right] \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!} \begin{vmatrix} \mu_1(\mu_1 - 1)\cdots(\mu_1 - \ell + 1) & \mu_1(\mu_1 - 1)\cdots(\mu_1 - \ell + 2) & \cdots & \mu_1 & 1 \\ \mu_2(\mu_2 - 1)\cdots(\mu_2 - \ell + 1) & \mu_2(\mu_2 - 1)\cdots(\mu_2 - \ell + 2) & \cdots & \mu_2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \mu_\ell(\mu_\ell - 1)\cdots(\mu_\ell - \ell + 1) & \mu_\ell(\mu_\ell - 1)\cdots(\mu_\ell - \ell + 2) & \cdots & \mu_\ell & 1 \end{vmatrix}
\end{aligned}$$

If we regard this last determinant as a polynomial in indeterminates $\mu_1, \dots, \mu_\ell$, we see that it has degree $\binom{\ell}{2}$ and is divisible by $\mu_i - \mu_j$ for every $i \neq j$ (since setting $\mu_i = \mu_j$ makes two rows equal, hence makes the determinant zero). Therefore, it must actually equal $\prod_{1 \le i < j \le \ell}(\mu_i - \mu_j)$ (this is known as the *Vandermonde determinant*). Therefore,

$$\begin{aligned}
f^\lambda &= \frac{n!}{\mu_1! \cdots \mu_\ell!} \prod_{1 \le i < j \le \ell} (\mu_i - \mu_j) \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!} \prod_{1 \le i < j \le \ell} (h(i,1) - h(j,1)) \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!} \ \frac{\prod_{i=1}^{\ell} h(i,1)!}{\prod_{x \in \lambda} h(x)} && \text{(by (9.33))} \\
&= \frac{n!}{\prod_{x \in \lambda} h(x)}.
\end{aligned}$$

$\square$

9.15. **Quasisymmetric functions and Hopf algebras.**

Definition 9.43. A quasisymmetric function is a formal power series $F \in \mathbb{C}[[x_1, x_2, \dots]]$ with the following property: if $i_1 < \cdots < i_r$ and $j_1 < \cdots < j_r$ are two sets of indices in strictly increasing order and $\alpha_1, \dots, \alpha_r \in \mathbb{N}$, then
$$[x_{i_1}^{\alpha_1} \cdots x_{i_r}^{\alpha_r}]F = [x_{j_1}^{\alpha_1} \cdots x_{j_r}^{\alpha_r}]F$$
where $[\mu]F$ denotes the coefficient of μ in F.

Symmetric functions are automatically quasisymmetric, but not vice versa. For example,
$$\sum_{i<j} x_i^2 x_j$$
is quasisymmetric but not symmetric (in fact, it is not preserved by *any* permutation of the variables). On the other hand, the set of quasisymmetric functions forms a graded ring $QSym \subset \mathscr{C}[[x_1, \dots]]$. We now describe a vector space basis for $QSym$.

A *composition* α is a sequence $(\alpha_1, \dots, \alpha_r)$ of positive integers, called its *parts*. Unlike a partition, we do not require that the parts be in weakly decreasing order. If $\alpha_1 + \cdots + \alpha_r = n$, we write $\alpha \models n$; the set of all compositions of n will be denoted $\mathrm{Comp}(n)$. Sorting the parts of a composition in decreasing order produces a partition of n, denoted by $\lambda(\alpha)$.

Compositions are much easier to count than partitions. Consider the set of partial sums
$$S(\alpha) = \{\alpha_1,\ \alpha_1 + \alpha_2,\ \dots,\ \alpha_1 + \cdots + \alpha_{r-1}\}.$$
The map $\alpha \mapsto S(\alpha)$ is a bijection from compositions of n to subsets of $[n-1]$; in particular, $|\mathrm{Comp}(n)| = 2^{n-1}$. We can define a partial order on $\mathrm{Comp}(n)$ via S by setting $\alpha \preceq \beta$ if $S(\alpha) \subseteq S(\beta)$; this is called *refinement*. The covering relations are merging two adjacent parts into one part.

The *monomial quasisymmetric function* of a composition $\alpha = (\alpha_1, \dots, \alpha_r) \models n$ is the power series
$$M_\alpha \;=\; \sum_{i_1 < \cdots < i_r} x_{i_1}^{\alpha_1} \cdots x_{i_r}^{\alpha_r} \;\in\; \mathbb{Z}[[x_1, x_2, \dots]]_n.$$
(For example, the quasisymmetric function $\sum_{i<j} x_i^2 x_j$ mentioned above is M_{21}.) This is the sum of all the monomials whose coefficient is constrained by the definition of quasisymmetry to be the same as that of any one of them. Therefore, the set $\{M_\alpha\}$ is a graded basis for $QSym$.

Example 9.44. Let $\mathcal{M}$ be a matroid on ground set E of size n. Consider weight functions $f : E \to \mathbb{N}$; one of the definitions of a matroid (see the problem set) is that a smallest-weight basis of $\mathcal{M}$ can be chosen via the following greedy algorithm (list E in weakly increasing order by weight $e_1, \dots, e_n$; initialize $B = \emptyset$; for $i = 1, \dots, n$, if $B \cup \{e_i\}$ is independent, then replace B with $B \cup \{e_i\}$). The *Billera-Jia-Reiner invariant* of $\mathcal{M}$ is the formal power series
$$W(\mathcal{M}) = \sum_f x_{f(1)} x_{f(2)} \cdots x_{f(n)}$$
where the sum runs over all weight functions f for which there is a *unique* smallest-weight basis. The correctness of the greedy algorithm implies that $W(\mathcal{M})$ is quasisymmetric.

For example, let $E = \{e_1, e_2, e_3\}$ and $\mathcal{M} = U_2(3)$. The bases are e_1e_2, e_1e_3, e_2e_3. Then E has a unique smallest-weight basis iff f has a unique maximum; it doesn't matter if the two smaller weights are equal or not. If the weights are all distinct then they can be assigned to E in $3! = 6$ ways; if the two smaller weights are equal then there are three choices for the heaviest element of E. Thus
$$W(U_2(3)) = \sum_{i<j<k} 6x_i x_j x_k + \sum_{i<j} 3x_i x_j^2 = 6M_{111} + 3M_{12}.$$

Questions: How are $W(\mathcal{M})$ and $W(\mathcal{M}^*)$ related?

Now for Hopf algebras. First, here is an algebraic description.

What is a $\mathbb{C}$-algebra? It is a $\mathbb{C}$-vector space A equipped with a ring structure. Its multiplication can be thought of as a $\mathbb{C}$-linear map

$$\mu : A \otimes A \to A$$

that is associative, i.e., $\mu(\mu(a,b),c) = \mu(a,\mu(b,c))$. Associativity can be expressed as the commutativity of the diagram

$$\begin{tikzcd}
A \otimes A \otimes A \arrow[r, "\mu\otimes I"] \arrow[d, "I\otimes\mu"'] & A \otimes A \arrow[d, "\mu"] \\
A \otimes A \arrow[r, "\mu"'] & A
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
a \otimes b \otimes c \arrow[r, mapsto] \arrow[d, mapsto] & ab \otimes c \arrow[d, mapsto] \\
a \otimes bc \arrow[r, mapsto] & abc
\end{tikzcd}$$

where I denotes the identity map. (Diagrams like this rely on the reader to interpret notation such as $\mu \otimes I$ as the only thing it could be possibly be; in this case, "apply μ to the first two tensor factors and tensor what you get with [I applied to] the third tensor factor".)

What then is a $\mathbb{C}$-*coalgebra*? It is a $\mathbb{C}$-vector space Z equipped with a $\mathbb{C}$-linear *comultiplication* map

$$\Delta : Z \to Z \otimes Z$$

that is *coassociative*, a condition defined by reversing the arrows in the previous diagram:

$$\begin{tikzcd}
Z & Z \otimes Z \arrow[l, "\Delta"'] \\
Z \otimes Z \arrow[u, "\Delta"] & Z \otimes Z \otimes Z \arrow[l, "\Delta\otimes I"] \arrow[u, "I\otimes\Delta"']
\end{tikzcd}$$

Just as an algebra has a unit, a coalgebra has a *counit*. To say what this is, let us diagrammatize the defining property of the multiplicative unit 1_A in an algebra A: it is the image of $1_{\mathbb{C}}$ under a map $u : \mathbb{C} \to A$ such that the diagram

$$\begin{tikzcd}
& A \arrow[dl] \arrow[dr] & \\
\mathbb{C} \otimes A \arrow[dr, "u\otimes I"'] & & A \otimes \mathbb{C} \arrow[dl, "I\otimes u"] \\
& A \otimes A \arrow[uu, "\mu"'] &
\end{tikzcd}$$

commutes. (Here I is the identity map, and the top diagonal maps take $a \in A$ to $1 \otimes A$ and $a \otimes 1$ respectively.) Thus a counit of a coalgebra is a map $\varepsilon : Z \to \mathbb{C}$ such that the diagram

$$\begin{tikzcd}
& A \arrow[dl, leftrightarrow] \arrow[dr, leftrightarrow] \arrow[dd, "\Delta"'] & \\
\mathbb{C} \otimes A & & A \otimes \mathbb{C} \\
& A \otimes A \arrow[ul, "\varepsilon\otimes I"] \arrow[ur, "I\otimes\varepsilon"'] &
\end{tikzcd}$$

A *bialgebra* is a vector space that has both a multiplication and a comultiplication, and such that multiplication is a coalgebra morphism and comultiplication is an algebra morphism. Both of these conditions are expressible by commutativity of the diagram

$$
\begin{tikzcd}
A \otimes A \arrow[r, "\Delta\otimes\Delta"] \arrow[d, "\mu"'] & A \otimes A \otimes A \otimes A \arrow[d, "\mu\otimes\mu"] \\
A \arrow[r, "\Delta"'] & A \otimes A
\end{tikzcd}
$$

Comultiplication takes some getting used to. In combinatorial settings, one should generally think of multiplication as putting two objects together, and comultiplication as taking an object apart into two subobjects. A unit is a trivial object (putting it together with another object has no effect), and the counit is the linear functional that picks off the coefficient of the unit.

Example 9.45 (The graph algebra)**.** For $n \geq 0$, let $\mathcal{G}_n$ be the set of formal $\mathbb{C}$-linear combinations of unlabeled simple graphs on n vertices (or if you prefer, of isomorphism classes $[G]$ of simple graphs G), and let $\mathcal{G} = \bigoplus_{n\geq 0} \mathcal{G}_n$. Thus $\mathcal{G}$ is a graded vector space, which we make into a $\mathbb{C}$-algebra by defining $\mu([G] \otimes [H]) = [G \sqcup H]$, where $\sqcup$ denotes disjoint union. The unit is the unique graph K_0 with no vertices (or, technically, the map $u : \mathbb{C} \to \mathcal{G}_0$ sending $c \in \mathbb{C}$ to $c[K_0]$). Comultiplication in $\mathcal{G}$ can be defined by

$$\Delta[G] = \sum_{A,B} [G|_A] \otimes [G|_B]$$

which can be checked to be a coassociative algebra morphism, making $\mathcal{G}$ into a bialgebra. This comultiplication is in fact *cocommutative*[16]. Let f be the "switching map" that sends $a \otimes b$ to $b \otimes a$; then commutativity and cocommutativity of multiplication and comultiplication on a bialgebra B are expressed by the diagrams

$$
\begin{tikzcd}
B \otimes B \arrow[rr, "f"] \arrow[dr, "\mu"'] & & B \otimes B \arrow[dl, "\mu"] \\
& B &
\end{tikzcd}
\qquad
\begin{tikzcd}
B \otimes B \arrow[rr, "f"] & & B \otimes B \\
& B \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] &
\end{tikzcd}
$$

So cocommutativity means that $\Delta(G)$ is symmetric under switching; for the graph algebra this is clear because A and B are interchangeable in the definition.

Example 9.46 (Rota's Hopf algebra of posets)**.** For $n \geq 0$, let $\mathcal{P}_n$ be the vector space of formal $\mathbb{C}$-linear combinations of isomorphism classes $[P]$ of finite graded posets P of rank n. Thus $\mathcal{P}_0$ and $\mathcal{P}_1$ are one-dimensional (generated by $\bullet$ and $\begin{smallmatrix}\bullet\\ |\\ \bullet\end{smallmatrix}$), but $\dim \mathcal{P}_n = \infty$ for $n \geq 2$. We make $\mathcal{P} = \bigoplus_n \mathcal{P}_n$ into a graded $\mathbb{C}$-algebra by defining $\mu([P] \otimes [Q]) = [P \times Q]$, where $\times$ denotes Cartesian product; thus $u(1) = \bullet$. Comultiplication is defined by

$$\Delta[P] = \sum_{x\in P} [\hat{0}, x] \otimes [x, \hat{1}].$$

[16]There are those who call this "mmutative".

Coassociativity is checked by the following calculation:

$$\begin{aligned}
\Delta \otimes I(\Delta(P)) &= \Delta \otimes I\left(\sum_{x \in P} [\hat{0}, x] \otimes [x, \hat{1}]\right) \\
&= \sum_{x \in P} \Delta([\hat{0}, x]) \otimes [x, \hat{1}] \\
&= \sum_{x \in P} \left(\sum_{y \in [\hat{0}, x]} [\hat{0}, y] \otimes [y, x]\right) \otimes [x, \hat{1}] \\
&= \sum_{x \leq y \in P} [\hat{0}, y] \otimes [y, x] \otimes [x, \hat{1}] \\
&= \sum_{y \in P} [\hat{0}, y] \otimes \left(\sum_{x \in [y, \hat{1}]} [y, x] \otimes [x, \hat{1}]\right) \\
&= \sum_{y \in P} [\hat{0}, y] \otimes \Delta([y, \hat{1}]) \ = \ I \otimes \Delta(\Delta(P)).
\end{aligned}$$

This Hopf algebra is commutative, but not cocommutative; there's no reason for the switching map to fix $\Delta(P)$ unless P is self-dual.

The ring Λ of symmetric functions is a coalgebra in the following way. We regard Λ as a subring of the ring of formal power series $\mathbb{C}[[\mathbf{x}]] = \mathbb{C}[[x_1, x_2, \dots]]$ First, the counit is just the map that takes a formal power series to its constant term. To figure out the coproduct, we then make a "Hilbert Hotel substitution":

$$x_1, x_2, x_3, x_4, \dots \ \mapsto \ x_1, y_1, x_2, y_2$$

to obtain a power series in $\mathbb{C}[[\mathbf{x}, \mathbf{y}]] = \mathbb{C}[[\mathbf{x}]] \otimes \mathbb{C}[[\mathbf{y}]]$. This is symmetric in each of the variable sets $\mathbf{x}$ and $\mathbf{y}$, i.e.,

$$\Lambda(\mathbf{x}, \mathbf{y}) \subseteq \Lambda(\mathbf{x}) \otimes \Lambda(\mathbf{y}).$$

So every symmetric function $F(\mathbf{x}, \mathbf{y})$ can be written in the form $\sum F_1(\mathbf{x}) F_2(\mathbf{y})$; we set $\Delta(F) = \sum F_1 \otimes F_2$.

For example, clearly $\Delta(c) = c = c \otimes 1 = 1 \otimes c$ for any scalar c.

$$\begin{aligned}
h_1(\mathbf{x}, \mathbf{y}) &= \sum_{i \geq 1} x_i + \sum_{i \geq 1} y_i \\
&= \left(\sum_{i \geq 1} x_i\right) \cdot 1 + 1 \cdot \left(\sum_{i \geq 1} y_i\right) \\
&= h_1(\mathbf{x}) \cdot 1 + 1 \cdot h_1(\mathbf{y}) \\
\therefore \quad \Delta(h_1) &= h_1 \otimes 1 + 1 \otimes h_1
\end{aligned}$$

$$\begin{aligned}
h_2(\mathbf{x}, \mathbf{y}) &= h_2(\mathbf{x}) + h_1(\mathbf{x}) h_1(\mathbf{y}) + h_2(\mathbf{y}) \\
\Delta(h_2) &= h_2 \otimes 1 + h_1 \otimes h_1 + 1 \otimes h_2
\end{aligned}$$

and more generally

$$\Delta(h_k) = \sum_{j=0}^{k} h_j \otimes h_{k-j}, \qquad \Delta(e_k) = \sum_{j=0}^{k} e_j \otimes e_{k-j}.$$

We can finally define a Hopf algebra!

Definition 9.47. A Hopf algebra is a bialgebra $\mathcal{H}$ with a **antipode** $S : \mathcal{H} \to \mathcal{H}$, which satisfies the commutative diagram

(9.35)
$$\begin{tikzcd} \mathcal{H} \arrow[r, "\Delta"] \arrow[drr, "\varepsilon"'] & \mathcal{H}\otimes\mathcal{H} \arrow[rr, "S\otimes I"] & & \mathcal{H}\otimes\mathcal{H} \arrow[r, "\mu"] & \mathcal{H} \\ & & \mathbb{C} \arrow[urr, "u"'] & & \end{tikzcd}$$

In other words, to calculate the antipode of something, comultiply it to get $\Delta g = \sum g_1 \otimes g_2$. Now hit every first tensor factor with S and then multiply it out again to obtain $\sum S(g_1)\cdot g_2$. If you started with the unit then this should be 1, while if you started with any other homogeneous object then you get 0. This enables calculating the antipode recursively. For example, in QSym:

$$\begin{aligned} \mu(S\otimes I(\Delta 1)) &= \mu(S\otimes I(1\otimes 1)) = \mu(S(1)\otimes 1) = S(1)\\ u(\varepsilon(1)) &= 1\\ S(1) &= 1 \end{aligned}$$

$$\begin{aligned} \mu((S\otimes I)(\Delta h_1)) &= \mu((S\otimes I)(h_1\otimes 1 + 1\otimes h_1)) = \mu(Sh_1\otimes 1 + S(1)\otimes h_1) = Sh_1 + h_1\\ u(\varepsilon(h_1)) &= 0\\ Sh_1 &= -h_1\\ &\dots \end{aligned}$$

Proposition 9.48. *Let B be a bialgebra that is graded and* connected, *i.e., the 0th graded piece has dimension 1 as a vector space. Then the commutative diagram* (**??**) *defines a unique antipode $S : B \to B$, and thus B can be made into a Hopf algebra in a unique way.*

Combinatorics features lots of graded connected bialgebras (such as all those we have seen so far), so this proposition gives us a Hopf algebra structure "for free".

In general the antipode is not very nice, but for symmetric functions it is. Our calculation of $\Delta(h_k)$ says that

$$\sum_{j=0}^{k} S(h_j)h_{k-j} = \begin{cases} 1 & \text{if } k=0,\\ 0 & \text{if } k>0 \end{cases}$$

and comparing with the Jacobi-Trudi relations says that $S(h_k) = (-1)^k e_k$, i.e., $S = (-1)^k\omega$.

Exercises (to be expanded):

(i) Let $E(M)$ denote the ground set of a matroid M, and call $|E(M)|$ the *order* of M. Let $\mathcal{M}_n$ be the vector space of formal $\mathbb{C}$-linear combinations of isomorphism classes $[M]$ of matroids M of order n. Let $\mathcal{M} = \bigoplus_{n\geq 0}\mathcal{M}_n$. Define a graded multiplication on $\mathcal{M}$ by $[M][M'] = [M\oplus M']$ and a graded comultiplication by

$$\Delta[M] = \sum_{A\subseteq E(M)} [M|_A]\otimes M/A]$$

where $M|_A$ and M/A denote restriction and contraction respectively. Check that these maps make $\mathcal{M}$ into a graded bialgebra, and therefore into a Hopf algebra by Proposition 9.48.

(ii) Prove that the Billera-Jia-Reiner invariant defines a Hopf algebra morphism $\mathcal{M} \to QSym$. (First I'd need to tell you how to comultiply in QSym....)

References

1. Martin Aigner, *A course in enumeration*, Graduate Texts in Mathematics, vol. 238, Springer, Berlin, 2007.
2. Egbert Brieskorn, *Sur les groupes de tresses [d'après V. I. Arnol′d]*, Séminaire Bourbaki, 24ème année (1971/1972), Exp. No. 401, Springer, Berlin, 1973, pp. 21–44. Lecture Notes in Math., Vol. 317.
3. Winfried Bruns and Jürgen Herzog, *Cohen-Macaulay rings*, Cambridge Studies in Advanced Mathematics, vol. 39, Cambridge University Press, Cambridge, 1993. MR 1251956 (95h:13020)
4. William Fulton, *Young tableaux*, London Mathematical Society Student Texts, vol. 35, Cambridge University Press, Cambridge, 1997.
5. Curtis Greene, *Some partitions associated with a partially ordered set*, J. Combinatorial Theory Ser. A **20** (1976), no. 1, 69–79.
6. Curtis Greene and Daniel J. Kleitman, *The structure of Sperner k-families*, J. Combinatorial Theory Ser. A **20** (1976), no. 1, 41–68.
7. Branko Grünbaum, *Convex polytopes*, second ed., Graduate Texts in Mathematics, vol. 221, Springer-Verlag, New York, 2003, Prepared and with a preface by Volker Kaibel, Victor Klee and Günter M. Ziegler.
8. Ezra Miller and Bernd Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics, vol. 227, Springer-Verlag, New York, 2005. MR 2110098 (2006d:13001)
9. Peter Orlik and Louis Solomon, *Combinatorics and topology of complements of hyperplanes*, Invent. Math. **56** (1980), no. 2, 167–189.
10. James G. Oxley, *Matroid theory*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1992.
11. G. L. Rybnikov, *On the fundamental group of the complement of a complex hyperplane arrangement*, Funktsional. Anal. i Prilozhen. **45** (2011), no. 2, 71–85.
12. Bruce E. Sagan, *The symmetric group*, second ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001.
13. Richard P. Stanley, *Combinatorics and commutative algebra*, second ed., Progress in Mathematics, vol. 41, Birkhäuser Boston Inc., Boston, MA, 1996. MR 1453579 (98h:05001)
14. ———, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999, With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. MR 1676282 (2000k:05026)
15. ———, *Enumerative combinatorics. Volume 1*, second ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012. MR 2868112