Lecture Notes on Algebraic Combinatorics

Jeremy L. Martin

jlmartin@ku.edu

April 29, 2015

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 `jlmartin@ku.edu` so I can fix it. Thanks to those who have suggested additions and pointed out errors, including but not limited to: Kevin Adams, Nitin Aggarwal, Lucas Chaffee, Ken Duna, Josh Fenton, Logan Godkin, Bennet Goeckner, Alex Lazar, Nick Packauskas, Billy Sanders, and Tony Se. Thanks to Marge Bayer for contributing the material on Ehrhart theory (§**??**).

1. The Fundamentals: Posets, Simplicial Complexes, and Polytopes

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 say that x is **covered** by y, written $x \lessdot y$, if $x < y$ and there exists no z such that $x < z < y$. Two posets P, Q are **isomorphic** if there is a bijection $\phi : P \to Q$ that is order-preserving; that is, $x \leq y$ in P iff $\phi(x) \leq \phi(y)$ in Q.

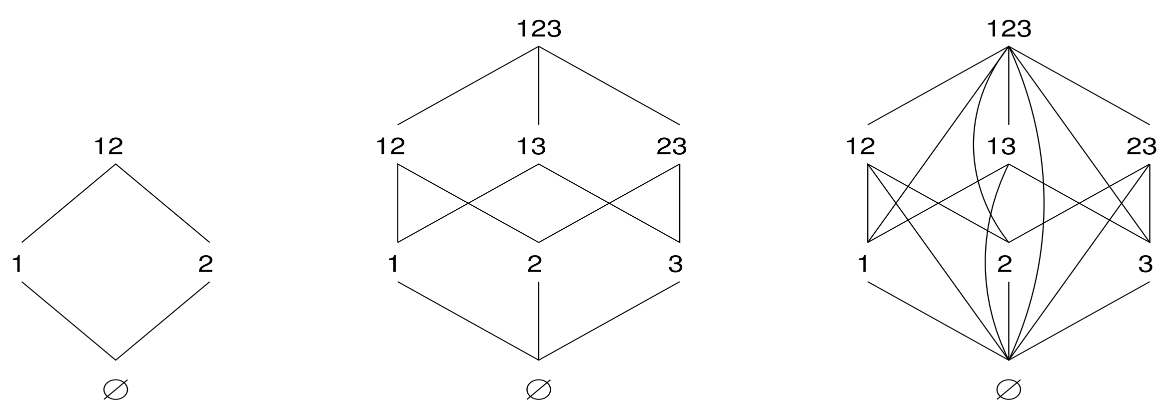
We'll usually assume that P is finite.

Definition 1.2. 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,

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

We'll have a lot more to say about lattices in Section 2.

Example 1.3 (**Boolean algebras**). Let $[n] = \{1, 2, \ldots, n\}$ (a standard piece of notation in combinatorics) and let $2^{[n]}$ be the power set of $[n]$. We can partially order $2^{[n]}$ by writing $S \leq T$ if $S \subseteq T$. A poset isomorphic to $2^{[n]}$ is called a **Boolean algebra of rank** n.



Note that $2^{[n]}$ is a lattice, with meet and join given by intersection and union respectively.

The first two pictures are **Hasse diagrams**: graphs whose vertices are the elements of the poset and whose edges represent the **covering relations**, which are enough to generate all the relations in the poset by transitivity. (As you can see on the right, including *all* the relations would make the diagram unnecessarily complicated.) By convention, bigger elements in P are at the top of the picture

The Boolean algebra $2^{[n]}$ 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 $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ respectively (or if necessary $\hat{\mathbf{0}}_P$ and $\hat{\mathbf{1}}_P$).

Definition 1.4. A poset that has both a $\hat{\mathbf{0}}$ and a $\hat{\mathbf{1}}$ is called **bounded**.[1] An element that covers $\hat{\mathbf{0}}$ is called an **atom**, and an element that is covered by $\hat{\mathbf{1}}$ is called a **coatom**. (For example, the atoms in $2^{[n]}$ are the singleton subsets of $[n]$.)

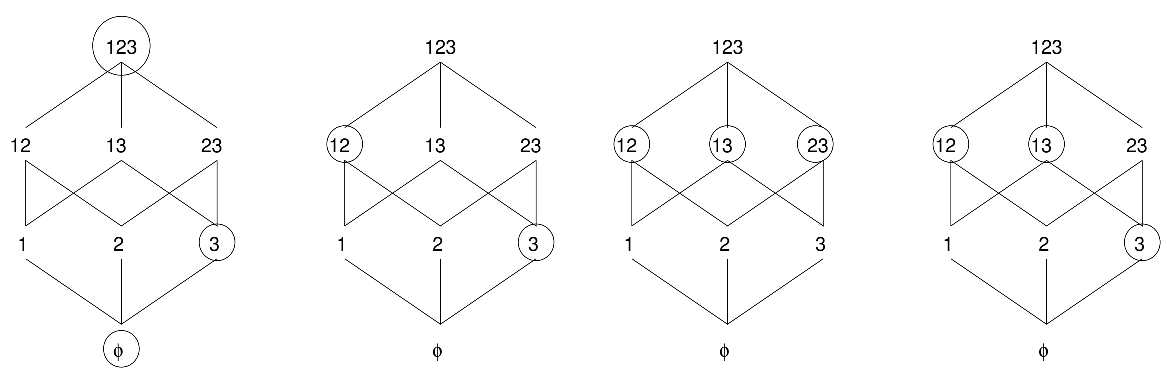
We can make a poset P bounded: define a new poset $\hat{P}$ by adjoining new elements $\hat{\mathbf{0}}, \hat{\mathbf{1}}$ such that $\hat{\mathbf{0}} < x < \hat{\mathbf{1}}$ for every $x \in P$. Meanwhile, sometimes we have a bounded poset and want to delete the bottom and top elements.

Definition 1.5. 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 $x = y$. For example, every nonempty interval $[A, B] \subseteq 2^{[n]}$ is itself a Boolean algebra of rank $|B| - |A|$.

Definition 1.6. A subset $C \subset P$ is called a **chain** if its elements are pairwise comparable. Thus every chain is of the form $C = \{x_0, \ldots, x_n\}$, where $x_0 < \cdots < x_n$. An **antichain** is a subset of P in which no two of its elements are comparable.[2]



chain antichain antichain neither

1.2. **Ranked Posets.** One of the many nice properties of $2^{[n]}$ 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 1.7. 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

[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".

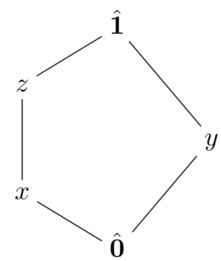
chain. A poset P is **ranked** if for every $x \in P$, all saturated chains with top element x have the same length; this number is called the **rank** of x and denoted $r(x)$. It follows that

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

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.
(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 $\hat{\mathbf{1}} \gtrdot y$, but $r(\hat{\mathbf{1}}) = 3$ and $r(y) = 1$.



Definition 1.8. 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,

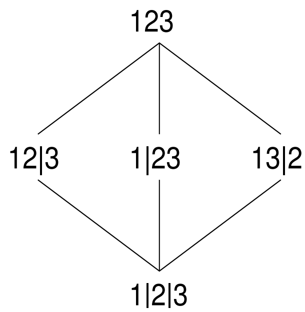
$$F_{2^{[n]}}(q) = \sum_{S \subset [n]} q^{|S|} = (1+q)^n.$$

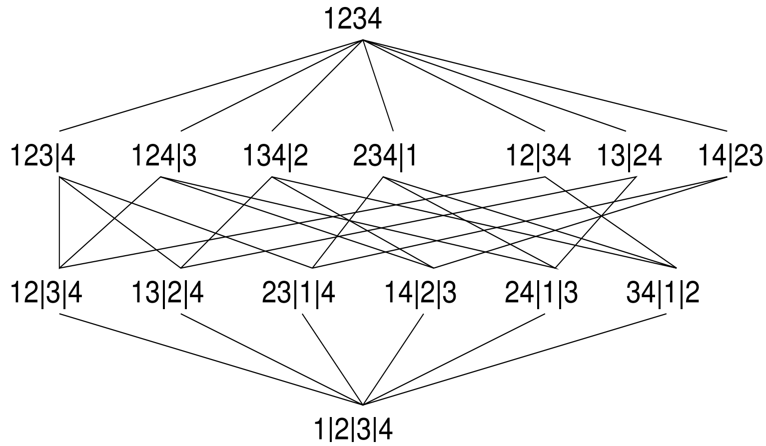
The expansion of this polynomial is palindromic, because the coefficients are a row of Pascal's Triangle. That is, $2^{[n]}$ is **rank-symmetric**. In fact, much more is true. For any poset P, we can define the **dual poset** P^* by reversing all the order relations, or equivalently turning the Hasse diagram upside down. It's not hard to prove (easy exercise) that the Boolean algebra is **self-dual**, i.e., $2^{[n]} \cong (2^{[n]})^*$, from which it immediately follows that it is rank-symmetric.

Example 1.9 (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.



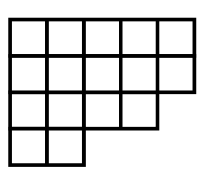


- The covering relations are of the form "merge two blocks into one".
- Π_n is graded, with $\hat{\mathbf{0}} = 1|2|\cdots|n$ and $\hat{\mathbf{1}} = 12\cdots n$. The rank function is $r(S) = n - |S|$.
- The coefficients of the rank-generating function of Π_n are the Stirling numbers of the second kind: $S(n,k)$ = 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 $F_4(q) = 1 + 6q + 7q^2 + q^3$.
- Π_n is a lattice. The meet of two partitions is their "coarsest common refinement": x, y belong to the same block of $S \wedge T$ if and only if they belong to the same block of S and to the same block of T. The join is the transitive closure of the union of the equivalence relations corresponding to S and T.

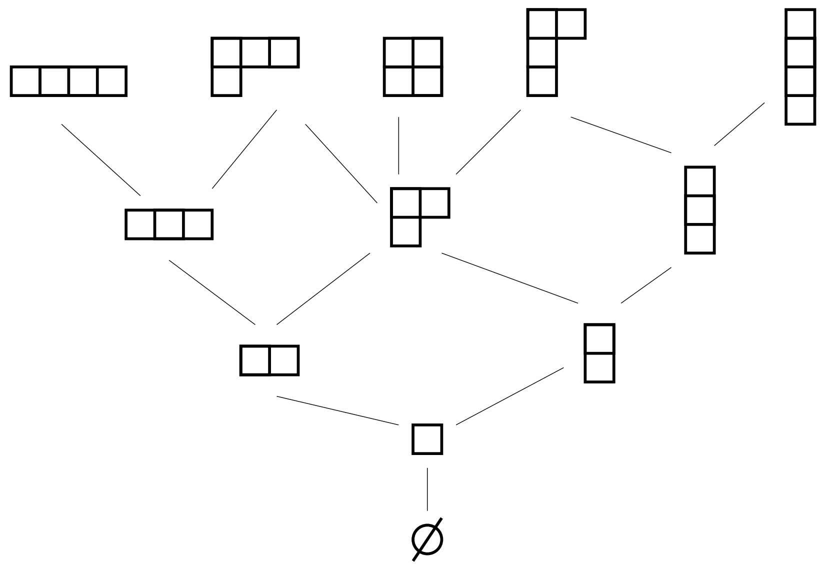
Example 1.10 (**Young's lattice**)**.** A *partition* is a sequence $\lambda = (\lambda_1, \dots, \lambda_\ell)$ of weakly decreasing positive integers: i.e., $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. If $n = \lambda_1 + \cdots + \lambda_\ell$, we write $\lambda \vdash n$ and/or $n = |\lambda|$. 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 $i = 1, 2, \dots$. Then Y is a ranked lattice, with rank function $r(\lambda) = |\lambda|$. Join and meet are given by component-wise max and min — we'll shortly see another description of the lattice operations.

This is an infinite poset, but the number of partitions at any given rank is finite. So in particular Y is *locally finite* (if X is any adjective, then "poset P is locally X" means "every interval in P is X"). Moreover, the rank-generating function
$$\sum_{\lambda} q^{|\lambda|} = \sum_{n \geq 0} \sum_{\lambda \vdash n} q^n$$
is a well-defined formal power series, and it 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:



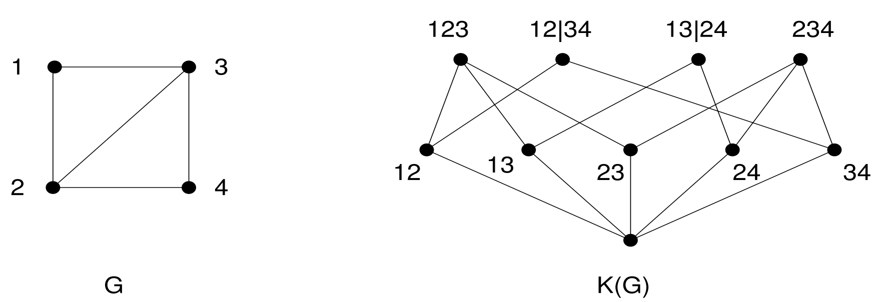
Then $\lambda \geq \mu$ if and only the Ferrers diagram of λ contains that of μ. The bottom part of the Hasse diagram of Y looks like this:



In terms of Ferrers diagrams, join and meet are simply union and intersection respectively.

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.



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 $\hat{\mathbf{1}}$. Notice that $\Pi_n = K(K_n)$ (the complete graph on n vertices).

1.3. **Simplicial Complexes.** The canonical references for this material are [Sta96], [BH93, Ch. 5]. See also [MS05] (for the combinatorics and algebra) and [Hat02] (for the topology).

Definition 1.12. Let V be a finite set of *vertices*. An **(abstract) simplicial complex** Δ on V is a nonempty family of subsets of V with the property that if $\sigma \in \Delta$ and $\tau \subseteq \sigma$, then $\tau \in \Delta$. Equivalently, Δ is an order ideal in the Boolean algebra 2^V. The elements of Δ are called its **faces** or **simplices**.

The **dimension** of a face σ is $\dim\sigma = |\sigma| - 1$. A face of dimension k is a k**-face** or k**-simplex**. The **dimension** of a non-void simplicial complex Δ is $\dim\Delta = \max\{\dim\sigma \mid \sigma \in \Delta\}$. A complex is **pure** if all its facets have the same dimension.

Throughout this section, I use the (fairly) standard convention

$$\dim\Delta = d - 1.$$

This looks a little funny, but the idea is that d is then the maximum number of vertices in a face. (It is a constant source of confusion that the dimension of a simplex is one less than its cardinality...)

The simplest simplicial complexes are the *void complex* $\Delta = \emptyset$ (which is often excluded from consideration) and the *irrelevant complex* $\Delta = \{\emptyset\}$. It is sometimes required that every singleton subset of V is a face (since if $v \in V$ and $\{v\} \notin \Delta$, then $v \notin \sigma$ for all $\sigma \in \Delta$, so you might as well replace V with $V \setminus \{v\}$). A simplicial complex with a single facet is called a **simplex**.

Since Δ is an order ideal, it is determined by its maximal elements, which are called **facets**. Thus we may define the simplicial complex **generated** by a list of faces $\sigma_1, \dots, \sigma_r$, namely

$$\langle\sigma_1, \dots, \sigma_r\rangle = \bigcup_{i=1}^{r} 2^{\sigma_i}.$$

The set of facets of a complex is the unique minimal set of generators for it.

Simplicial complexes are combinatorial models for topological spaces. The vertices $V = [n]$ should be regarded as points $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$, and a simplex $\sigma = \{v_1, \dots, v_r\}$ should be thought of as (homeomorphic to) the convex hull of the corresponding points:

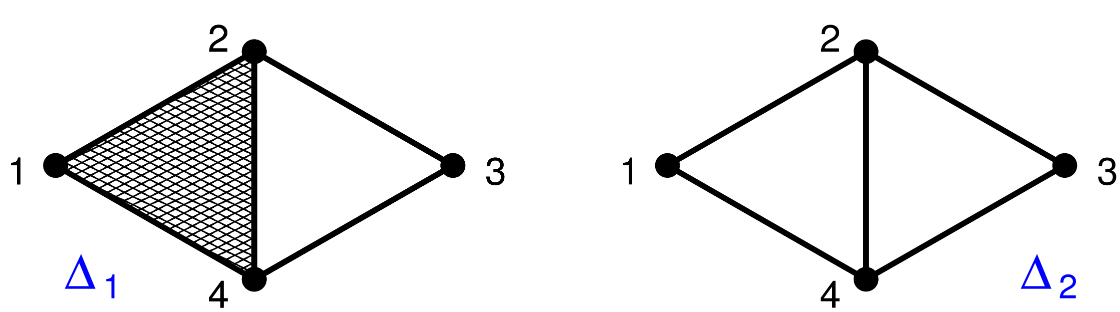
$$|\sigma| = \operatorname{conv}\{\mathbf{e}_{v_1}, \dots, \mathbf{e}_{v_r}\} = \{c_1\mathbf{e}_{v_1} + \cdots + c_{v_r}\mathbf{e}_r \mid \ 0 \le c_i \le 1 \ (\forall i), \quad c_1 + \cdots + c_n = 1\}.$$

For example, faces of sizes 1, 2, and 3 correspond respectively to line segments, triangles, and tetrahedra. Taking $\{\mathbf{e}_i\}$ to be the standard basis of $\mathbb{R}^n$ gives the **standard geometric realization** $|\Delta|$ of Δ:

$$|\Delta| = \bigcup_{\sigma\in\Delta} \operatorname{conv}\{\mathbf{e}_i \mid i \in \sigma\}.$$

It is usually possible to realize Δ geometrically in a space of much smaller dimension. For example, every graph can be realized in $\mathbb{R}^3$, and planar graphs can be realized in $\mathbb{R}^2$. It is common to draw geometric pictures of simplicial complexes, just as we draw pictures of graphs. We sometimes use the notation $|\Delta|$ to denote any old geometric realization (i.e., any topological space homeomorphic to the standard geometric realization), and it is also common to ignore the distinction between Δ and $|\Delta|$ and to treat Δ itself as a topological space.

For instance, here are pictorial representations of the simplicial complexes $\Delta_1 = \langle 124, 23, 24, 34\rangle$ and $\Delta_2 = \langle 12, 14, 23, 24, 34\rangle$.



The filled-in triangle indicates that 124 is a face of Δ_1, but not of Δ_2. Note that Δ_2 is the subcomplex of Δ_1 consisting of all faces of dimensions ≤ 1 — that is, it is the **1-skeleton** of Δ_1.

Definition 1.13. Let Δ be a simplicial complex of dimension $d-1$. The ***f*-vector of** Δ is $(f_{-1}, f_0, f_1, \ldots, f_{d-1})$, where $f_i = f_i(\Delta)$ is the number of faces of dimension i. The term f_{-1} is often omitted, because $f_{-1} = 1$ unless Δ is the void complex. The ***f*-polynomial** is the generating function for the nonnegative f-numbers (essentially the rank-generating function of Δ as a poset):

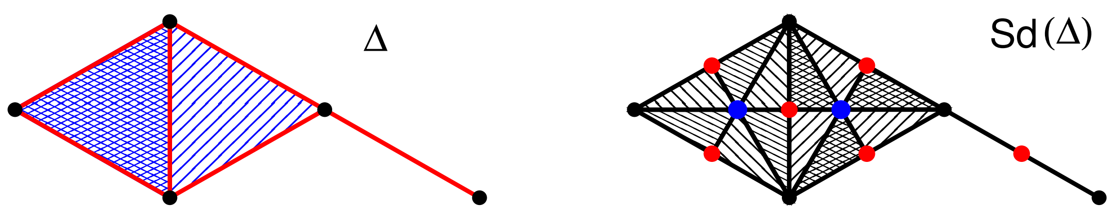
$$f(\Delta, t) = f_0 + f_1 t + f_2 t^2 + \cdots + f_d t^d.$$

For instance, if Δ_1, Δ_2 are the simplicial complexes pictured above, then

$$f(\Delta_1, t) = (4, 5, 1) \quad \text{and} \quad f(\Delta_2, t) = (4, 5).$$

Example 1.14. Let P be a finite poset and let $\Delta(P)$ be the set of chains in P. Since every subset of a chain is a chain, $\Delta(P)$ is a simplicial complex; it is called the *order complex* of P. The minimal nonfaces of $\Delta(P)$ are precisely the pairs of incomparable elements of P; in particular every minimal nonface has size two, which is to say that $\Delta(P)$ is a *flag complex*. Note that $\Delta(P)$ is pure if and only if P is ranked.

What if P itself is the set of faces of a simplicial complex? In this case $\Delta(P)$ is the *barycentric subdivision*. To construct the barycentric subdivision $\mathrm{Sd}(\Delta)$ of a simplicial complex Δ, construct a new vertex in the middle of each face and start connecting them — this is best illustrated by a picture.



Each vertex of $\mathrm{Sd}(\Delta)$ corresponds to a face of Δ; in this figure the colors encode dimension. Note in particular that the barycentric subdivision of a simplicial complex is homeomorphic to it.

There is a fundamental connection, the *Stanley-Reisner correspondence*, between simplicial complexes and commutative algebra. Let $R = \Bbbk[x_1, \ldots, x_n]$ be the ring of polynomials in n variables over your favorite field $\Bbbk$. Define the **support** of a monomial $\mu \in R$ as

$$\operatorname{supp} \mu = \{i : \ x_i \text{ divides } \mu\}$$

Definition 1.15. Let Δ be a simplicial complex on vertex set $[n]$. Its **Stanley-Reisner ideal** in R is

$$I_\Delta = \langle x_{i_1} \cdots x_{i_r} \mid \{i_1, \ldots, i_r\} \notin \Delta \rangle \, .$$

The **Stanley-Reisner ring** or **face ring** is $\Bbbk[\Delta] := R/I_\Delta$.

Since Δ is a simplicial complex, the monomials in I_Δ are *exactly* those whose support is not a face of Δ. Therefore, the monomials supported on a face of Δ are a natural vector space basis for the graded ring $\Bbbk[\Delta]$.

Its Hilbert series can be calculated by counting these monomials, using basic generatingfunctionology:

$$\begin{aligned}
\operatorname{Hilb}(\Bbbk[\Delta],q) \overset{\text{def}}{\equiv} \sum_{i\geq 0} q^i \dim_{\Bbbk}(\Bbbk[\Delta])_i &= \sum_{\sigma\in\Delta}\ \sum_{\mu:\ \operatorname{supp}\mu=\sigma} q^{|\mu|}\\
&= \sum_{\sigma\in\Delta}\left(\frac{q}{1-q}\right)^{|\sigma|}\\
&= \sum_{i=0}^{d} f_{i-1}\left(\frac{q}{1-q}\right)^{i} = \frac{\sum_{i=0}^{d} f_{i-1}q^i(1-q)^{d-i}}{(1-q)^d} = \frac{\sum_{i=0}^{d} h_i q^i}{(1-q)^d}
\end{aligned}$$

The numerator of this rational expression is a polynomial in q, called the **h-polynomial** of Δ, and its list of coefficients $(h_0, h_1, \dots)$ is called the **h-vector** of Δ. Applying the binomial theorem yields a formula for the h-numbers in terms of the f-numbers:

$$\begin{aligned}
\sum_{i=0}^{d} f_{i-1}q^i(1-q)^{d-i} &= \sum_{i=0}^{d} f_{i-1}q^i \sum_{j=0}^{d-i}\binom{d-i}{j}(-1)^j q^j\\
&= \sum_{i=0}^{d}\sum_{j=0}^{d-i}\binom{d-i}{j}(-1)^j q^{i+j} f_{i-1}
\end{aligned}$$

and now extracting the q^k coefficient (i.e., the summand in the second sum with $j = k - i$) yields

$$h_k = \sum_{i=0}^{k}\binom{d-i}{k-i}(-1)^{k-i} f_{i-1}. \tag{1.2}$$

(Note that the upper limit of summation might as well be k instead of d, since the binomial coefficient in the summand vanishes for $i > k$.) These equations can be solved to give the f's in terms of the h's.

$$f_{i-1} = \sum_{k=0}^{i}\binom{d-k}{i-k} h_k. \tag{1.3}$$

So the f-vector and h-vector contain equivalent information about a complex. On the level of generating functions, the conversions look like this [BH93, p. 213]:

$$\sum_i h_i q^i = \sum_i f_{i-1} q^i (1-q)^{d-i}, \tag{1.4}$$

$$\sum_i f_i q^i = \sum_i h_i q^{i-1}(1+q)^{d-i}. \tag{1.5}$$

The equalities (1.2) and (1.3) can be obtained by applying the binomial theorem to the right-hand sides of (1.4) and (1.5) and equating coefficients. Note that it is most convenient simply to sum over all $i \in \mathbb{Z}$.

Two useful special cases are as follows: (1.2) gives

$$h_d = \sum_{i=0}^{d}\binom{d-i}{d-i}(-1)^{d-i} f_{i-1} = (-1)^{d-1}\tilde{\chi}(\Delta)$$

and (1.3) gives

$$f_{d-1} = \sum_{k=0}^{d} h_k.$$

For certain complexes, the h-numbers themselves have a direct combinatorial interpretation. The last formula suggests that they should enumerate facets of a pure complex in some way.

Definition 1.16. A pure simplicial complex Δ of dimension $d-1$ is **shellable** if its facets can be ordered $F_1, \dots, F_n$ such that any of the following conditions are satisfied:

(1) For every $i \in [n]$, the set $\langle F_i \rangle \setminus \langle F_1, \dots, F_{i-1} \rangle$ has a unique minimal element, usually denoted R_i.
(2) Δ decomposes as a disjoint union of intervals $[R_1, F_1] \cup \cdots \cup [R_n, F_n]$.
(3) For every $i > 1$, the complex $\langle F_i \rangle \cap \langle F_1, \dots, F_{i-1} \rangle$ is pure of dimension $d - 2$.

The proof of equivalence is left as an exercise.

Proposition 1.17. *Let Δ be shellable of dimension $d-1$, with h-vector $(h_0, \dots, h_d)$. Then*
$$h_j = \#\{F_i \mid \#R_i = j\}.$$
Moreover, if $h_j(\Delta) = 0$ for some j, then $h_k(\Delta) = 0$ for all $k > j$.

The proof is left as an exercise. One consequence is that the h-vector of a shellable complex is strictly nonnegative, since its coefficients count something. This statement is emphatically not true about the Hilbert series of arbitrary graded rings, or even arbitrary Stanley-Reisner rings!

If a simplicial complex is shellable, then its Stanley-Reisner ring is *Cohen-Macaulay* (CM). This is an important and subtle algebraic condition that can be expressed algebraically in terms of depth or local cohomology (topics beyond the scope of these notes) or in terms of simplicial homology (coming shortly). Shellability is the most common combinatorial technique for proving that a ring is CM. The constraints on the h-vectors of CM complexes are the same as those on shellable complexes, although it is an open problem to give a general combinatorial interpretation of the h-vector of a CM complex.

Since simplicial complexes are (models of) topological spaces, the tools of algebraic topology play a crucial role in studying their combinatorics. For a full treatment, see Chapter 2 of Hatcher [Hat02], but here are the basic tools.

Fix a ring R (typically $\mathbb{Z}$ or a field). Let Δ be a simplicial complex on vertex set $[n]$. The k^{th} **simplicial chain group** of Δ over R, denoted $C_k(\Delta, R)$ or simply $C_k(\Delta)$, is the free R-module with basis elements $[\sigma]$ indexed by the k-simplices in Δ. The **(simplicial) boundary map** $\partial_k : C_k(\Delta) \to C_{k-1}(\Delta)$ is defined as follows: if $\sigma = \{v_0, \dots, v_k\}$ is a k-face, with $1 \leq v_0 < \cdots < v_k \leq n$, then
$$\partial_k[\sigma] = \sum_{i=0}^{n} (-1)^i [v_0, \dots, \widehat{v_i}, \dots, v_k]$$
where the hat denotes removal. The map is then extended R-linearly to all of $C_k(\Delta)$.

Recall that σ itself is represented by a k-dimensional ball; the k-chain $\partial[\sigma]$ should be thought of as the $k-1$-sphere that is its boundary, expressed as a sum of $(k-1)$-simplices with consistent orientations (as represented by the signs). Often it is convenient to abbreviate ∂_k by ∂, since either the subscript is clear from context or else we want to say something about all boundary maps at once.

The fundamental fact about boundary maps is that
$$\partial^2 = 0$$
(that is, $\partial_k \circ \partial_{k+1}$ for all k). This is equivalent to saying that $\ker \partial_k \supseteq \operatorname{im} \partial_{k+1}$ for all k (if you haven't seen this before, convince yourself it's true). That is, the string of R-modules
$$0 \to C_{\dim \Delta}(\Delta) \to \cdots \to C_{k+1}(\Delta) \xrightarrow{\partial_{k+1}} C_k(\Delta) \xrightarrow{\partial_k} C_{k-1}(\Delta) \to \cdots \to C_{-1}(\Delta) \to 0$$
is an **algebraic chain complex**. In particular, we can define the **reduced**[4] **simplicial homology groups**
$$\tilde{H}_k(\Delta) = \tilde{H}_k(\Delta; R) = \ker \partial_k / \operatorname{im} \partial_{k+1}.$$
It turns out that these groups depend only on the homotopy type of $|\Delta|$ (although this takes quite a bit of effort to prove), so for instance they do not depend on the choice of labeling of vertices and are invariant

[4]The unreduced homology groups $H_k(\Delta; R)$ are defined by deleting $C_{-1}(\Delta)$ from the simplicial chain complex; this results in an extra summand of R in the dimension-0 term and has no effect elsewhere

under operations like barycentric subdivision (and have many other good properties). A complex all of whose homology groups vanish is called **R-acyclic**; for example, if $|\Delta|$ is contractible then Δ is R-acyclic for every R. If $\Delta \cong \mathbb{S}^d$ (i.e., $|\Delta|$ is a d-dimensional sphere), then

$$\tilde{H}_k(\Delta; R) \cong \begin{cases} R & \text{if } k = d, \\ 0 & \text{if } k < d. \end{cases} \tag{1.6}$$

The Cohen-Macaulay condition can be expressed homologically. Define the **link** of a face $\sigma \in \Delta$ by

$$\operatorname{lk}_\Delta(\sigma) = \{\tau \in \Delta \mid \tau \cap \sigma = \emptyset,\ \tau \cup \sigma \in \Delta\}.$$

The link of σ is a subcomplex of Δ, and can be thought of as "what you see if you stand at σ and look outward." For example, if $|\Delta|$ is a manifold of dimension d, then the link of any k-face is a simplicial $(d-k-1)$-sphere, so in particular every link has zero homology in non-top dimension. The Cohen-Macaulay condition is a relaxation of this condition: specifically, Reisner's theorem says that Δ is Cohen-Macaulay over R iff it is pure (so that $\dim \operatorname{lk}_\Delta(\sigma) = \dim \Delta - \dim \sigma - 1$ for all σ) and for every face σ, one has

$$\tilde{H}_k(\operatorname{lk}_\Delta(\sigma); R) = 0 \qquad \forall k < \dim \Delta - \dim \sigma - 1.$$

In other words, every link has the homology type of a **wedge of spheres** of the appropriate dimension. (The wedge sum of a collection of spaces is obtained by identifying a point of each; for example, the wedge of n circles looks like a flower with n petals. Reduced homology is additive on wedge sums, so by (1.6) the wedge sum of n copies of $\mathbb{S}^d$ has reduced homology R^n in dimension d and 0 in other dimensions.)

A Cohen-Macaulay complex Δ is **Gorenstein** (over R) if in addition $\tilde{H}_{\dim \Delta - \dim \sigma - 1}(\operatorname{lk}_\Delta(\sigma); R) \cong R$ for all σ. That is, every link has the homology type of a sphere. This is very close to being a manifold (I don't know offhand of a Gorenstein complex that is not a manifold, although I'm sure examples exist).

1.4. **Polytopes.** One reference for this material is chapter 2 of Schrijver's notes [Sch13].

Polytopes are familiar objects: cubes, pyramids, Platonic solids, etc. Polytopes can be described either by their vertices or by their facets (maximal faces); our first task is to show that the two descriptions are in fact equivalent.

First, a couple of key terms. A subset $S \subset \mathbb{R}^n$ is **convex** if, for any two points in S, the line segment joining them is also a subset of S. The smallest convex set containing a given set T is called its **convex hull**, denoted $\operatorname{conv}(T)$. Explicitly, one can show (exercise; not too hard) that

$$\operatorname{conv}(\mathbf{x}_1, \dots, \mathbf{x}_r) = \left\{ c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r \mid 0 \le c_i \le 1 \text{ for all } i \text{ and } \sum_{i=1}^r c_i = 1 \right\}. \tag{1.7}$$

These points are called **convex linear combination** of the $\mathbf{x}_i$. A related definition is the **affine hull** of a point set:

$$\operatorname{aff}(\mathbf{x}_1, \dots, \mathbf{x}_r) = \left\{ c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r \mid \sum_{i=1}^r c_i = 1 \right\}. \tag{1.8}$$

This is the smallest affine space containing all $\mathbf{x}_i$. ("Affine space" means "translate of a vector subspace of $\mathbb{R}^n$".)

Clearly $\operatorname{conv}(T) \subset \operatorname{aff}(T)$. For example, if $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are three non-collinear points in $\mathbb{R}^3$, then $\operatorname{conv}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is the triangle having those points as vertices, while $\operatorname{aff}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is the unique plane containing all three points.

Definition 1.18. A **polyhedron** is a nonempty intersection of finitely many closed half-spaces in $\mathbb{R}^n$, or equivalently the solution space of finitely many linear inequalities $a_{i1}x_1 + \cdots + a_{in}x_n \ge b_i$. The equations are often written as a single matrix equation $A\mathbf{x} \ge \mathbf{b}$, where $A = [a_{ij}] \in \mathbb{R}^{m \times n}$, $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, and $\mathbf{b} = (b_1, \dots, b_m) \in \mathbb{R}^m$. A bounded polyhedron is called a **polytope**.

Definition 1.19. A **polytope** is the convex hull of a finite set of points.

Theorem 1.20. *These two definitions of "polytope" are equivalent.*

Before sketching the proof, let's connect the geometry and the algebra. A **face** of a polyhedron $P \subset \mathbb{R}^n$ is a subset of P that maximizes some linear functional. This seems strange but is actually quite natural when you get used to it. Let's say we have a polytope sitting in $\mathbb{R}^2$. What point or set of points is highest? In other words, what points maximize the linear functional $(x, y, z) \mapsto z$? The answer could be a single vertex, or an edge, or a polygonal face. Of course, there's nothing special about the z-direction. If you pick any direction (as expressed by a vector $\mathbf{v}$), then the extreme points of P in that direction are the maxima on P of the linear functional $\mathbf{x} \mapsto \mathbf{x} \cdot \mathbf{v}$, and they correspond to what you would want to call a "face". If you pick a linear functional "at random", then with probability 1 the face it determines will be a vertex of P — higher-dimensional faces correspond to more special directions. (What is the dimension of a face? It is just the dimension of its affine span.)

With that in mind...

Sketch of proof of Theorem 1.20. For the details, see Schrijver, §2.2 (or better yet, fill them in yourself).

First, let P be a bounded intersection of finitely many half-spaces, i.e., $P = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq b\}$, where $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^{m\times 1}$. <u>Show</u> that a point $\mathbf{z} \in P$ is the unique maximum of some linear functional iff the submatrix

$$A_{\mathbf{z}} := \text{rows } a_i \text{ for which } a_i \cdot \mathbf{z} = b_i$$

has full rank n. Therefore the vertices are all of the form $A_R^{-1} b_R$, where R is a row basis of A and A_R, b_R denote restrictions. Not every point of this form necessarily lies in P, but this argument does show that there are only finitely many vertices (specifically, at most $\binom{m}{n}$). <u>Show</u> that P is their convex hull.

Second, let $P = \operatorname{conv}(\mathbf{x}_1, \dots, \mathbf{x}_r) \subset \mathbb{R}^n$. Assume without loss of generality that $\operatorname{aff}(P) = \mathbb{R}^n$ (otherwise, replace $\mathbb{R}^n$ with the affine hull) and that the origin is in the interior of P (translating if necessary). Define

$$P^* = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} \leq 1 \ \ \forall \mathbf{x} \in P\}. \tag{1.9}$$

This is called the **(polar) dual**. <u>Show</u> that in fact

$$P^* = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x}_i \cdot \mathbf{y} \leq 1 \ \ \forall i \in [r]\}$$

which means that P^* is an intersection of finitely many half-spaces. So, by the first part of the theorem, $P^* = \operatorname{conv}(\mathbf{y}_1, \dots, \mathbf{y}_s)$ for some $\mathbf{y}_1, \dots, \mathbf{y}_s$. Now, <u>show</u> that

$$P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y}_j \leq 1 \ \ \forall j \in [s]\}$$

which expresses P as an intersection of finitely many half-spaces. $\square$

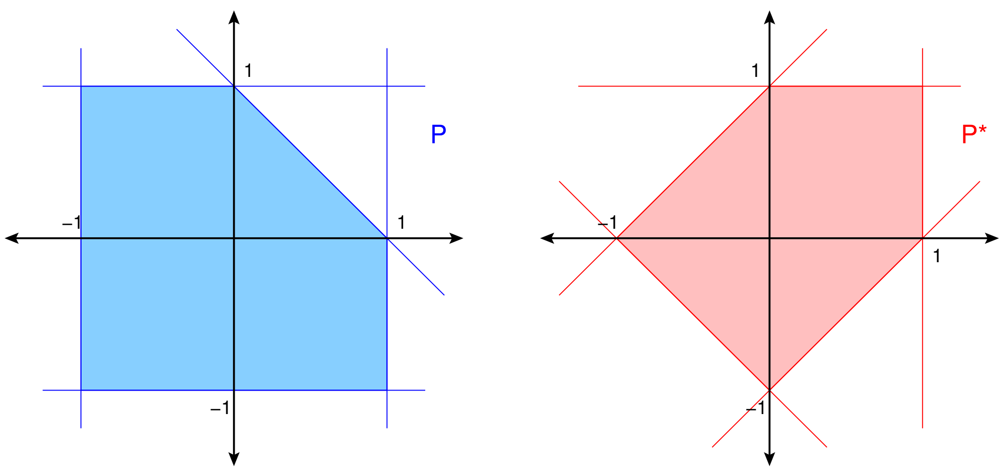
Example 1.21. Let P be the polytope shown on the left below. It can be expressed as an intersection of hyperplanes (i.e., solution set to a system of linear inequalities) as follows:

$$P = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

If we number the rows $R_0, \ldots, R_4$, every pair of rows other than $\{R_0, R_2\}$ and $\{R_1, R_3\}$ is of full rank. The points corresponding to the other eight pairs of rows are:

Rows	Point		Constraint on P^*
$0,1$	$(-1,-1)$	vertex	$-x-y \leq 1$
$0,3$	$(-1,1)$	vertex	$-x+y \leq 1$
$0,4$	$(-1,2)$	not in P	
$1,2$	$(1,-1)$	vertex	$x-y \leq 1$
$1,4$	$(2,-1)$	not in P	
$2,3$	$(1,1)$	not in P	
$2,4$	$(1,0)$	vertex	$x \leq 1$
$3,4$	$(0,1)$	vertex	$y \leq 1$

Thus the vertices of P correspond to the bounding hyperplanes (i.e., lines) of P^*, and vice versa.



Some more key terms. Throughout, let P be an n-dimensional polytope in $\mathbb{R}^n$.

- A **facet** of P is a face of codimension 1 (that is, dimension $n-1$). In this case there is a unique (up to scaling) linear functional that is maximized on F, given by the normal vector to it (pointing outward from the rest of P). Faces of codimension 2 are called **ridges** and faces of codimension 3 are sometimes called **peaks**.
- A **supporting hyperplane** of P is a hyperplane that meets P in a nonempty face.
- A polytope is **simplicial** if every face is a simplex. For example, every 2-dimensional polytope is simplicial, but of the Platonic solids in $\mathbb{R}^3$, only the tetrahedron, octahedron and icosahedron are simplicial — the cube and dodecahedron are not. The boundary of a simplicial polytope is thus a simplicial $(n-1)$-sphere.
- A polytope is **simple** if every vertex belongs to exactly n faces.
- The **face poset** of P is the poset of all faces, ordered by inclusion. This is clearly ranked by dimension; one can decide whether to include the empty face as $\hat{\mathbf{0}}$ and all of P as $\hat{\mathbf{1}}$. For instance, the face poset of a simplex is a Boolean algebra. Two polytopes are *combinatorially isomorphic* if their face posets are isomorphic.

Proposition 1.22. *A polytope P is simple if and only if its dual P^* is simplicial. Also, the face poset of P is the dual of the face poset of P^*.*

One of the big questions about polytopes is to classify their possible f-vectors and, more generally, the structure of their face posets. Here is a result of paramount importance in this area.

Theorem 1.23. *Let Δ be the boundary sphere of a convex simplicial polytope $P \subset \mathbb{R}^n$. Then Δ is shellable, and its h-vector is a palindrome, i.e., $h_i = h_{d-i}$ for all i.*

Palindromicity of the h-vector is known as the *Dehn-Sommerville relations.* They were first proved early in the 20th century, but the following proof, due to Bruggesser and Mani [BM71], is undoubtedly the one in the Book.

Sketch of proof. Let $\mathcal{H}$ be the collection of hyperplanes spanned by facets of P. Let ℓ be a line that passes through the interior of P and meets each hyperplane in $\mathcal{H}$ in a distinct point. (Note that almost any line will do.) Imagine walking along this line, starting just outside P so that only one facet is visible. Call that facet F_1. As you continue to walk, more and more facets become visible. Label the facets $F_2, \dots, F_m$ in the order in which they become visible (equivalently, order them in the order in which the line ℓ meets their affine spans). When you get to infinity, come back the other way (so that all of a sudden "invisible" and "visible" switch meanings) and continue to label the facets $F_{m+1}, \dots, F_n$ in the order in which they disappear.

In fact, this is a shelling order, because for $1 \leq j \leq m$, $\partial F_j \cap \langle F_1, \dots, F_{j-1}\rangle$ consists of the maximal subfaces of F_j that were visible before crossing $\operatorname{aff}(F_j)$ but after crossing $\operatorname{aff}(F_{j-1})$, while for $m+1 \leq j \leq n$,

In fact, this is a shelling order (called a *line shelling*), because

$$\partial F_j \cap \langle F_1, \dots, F_{j-1}\rangle = \begin{cases} \langle \text{maximal subfaces of } F_j \text{ that become visible upon crossing } \operatorname{aff}(F_j)\rangle & \text{for } 1 \leq j \leq m, \\ \langle \text{maximal subfaces of } F_j \text{ that become invisible upon crossing } \operatorname{aff}(F_j)\rangle & \text{for } m+1 \leq j \leq n. \end{cases}$$

Moreover, each facet F contributes to $h_i(P)$, where

$$i = i(F) = \#\{G < F \mid F, G \text{ share a ridge}\}.$$

On the other hand, the reversal of $<$ is another instance of this construction, hence is also a shelling order. Since each facet shares a ridge with exactly n other facets, the previous formula says that if facet F contributes to h_i with respect to the first shelling order then it contributes to h_{n-i} in its reversal. Since the h-vector is an invariant of P, it follows that $h_i = h_{n-i}$ for all i. □

The Dehn-Sommerville relations are a basic tool in classifying h-vectors, and therefore f-vectors, of simplicial polytopes. Since $h_0 = 1$ for shellable complexes, it follows immediately that the only possible h-vectors for simplicial polytopes in $\mathbb{R}^2$ and $\mathbb{R}^3$ are $(1, k, 1)$ and $(1, k, k, 1)$, respectively (where k is a positive integer).

1.5. **Exercises.**

Exercise 1.1. A *directed acyclic graph* or DAG, is a pair $G = (V, E)$, where V is a finite set of *vertices*; E is a finite set of *edges*, each of which is an ordered pair of distinct vertices; and E contains no directed cycles, i.e., no subsets of the form $\{(v_1, v_2),\ (v_2, v_3),\ \dots,\ (v_{n-1}, v_n),\ (v_n, v_1)\}$ for any $v_1, \dots, v_n \in V$.

(a) Let P be a poset with order relation $<$. Let $E = \{(v, w) \mid v, w \in P,\ v < w\}$. Prove that the pair (P, E) is a DAG.
(b) Let $G = (V, E)$ be a DAG. Define a relation $<$ on V by setting $v < w$ iff there is some directed path from v to w in G, i.e., iff E has a subset of the form $\{(v_1, v_2),\ (v_2, v_3),\ \dots,\ (v_{n-1}, v_n)\}$ with $v = v_1$ and $w = v_n$. Prove that this relation makes V into a poset.

(This problem is purely a technical exercise, but it does show that posets and DAGs are essentially the same thing.)

Exercise 1.2. Let n be a positive integer. Let D_n be the set of all positive-integer divisors of n (including n itself), partially ordered by divisibility.

(a) Prove that D_n is a ranked poset, and describe the rank function.

(b) For which values of n is D_n (i) a chain; (ii) a Boolean algebra? For which values of n, m is it the case that $D_n \cong D_m$?
(c) Prove that D_n is a *distributive lattice*, i.e., a lattice such that $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ for all $x, y, z \in D_n$. Describe its meet and join operations and its join-irreducible elements.
(d) Prove that D_n is *self-dual*, i.e., there is a bijection $f : D_n \to D_n$ such that $f(x) \leq f(y)$ if and only if $x \geq y$.

Exercise 1.3. Let G be a graph with connected components $G_1, \dots, G_r$. Describe the clique poset of G in terms of the clique posets of $G_1, \dots, G_r$.

Exercise 1.4. Let Δ be a simplicial complex on vertex set V, and let $v_0 \notin V$. The **cone over** Δ is the simplicial complex $C\Delta$ generated by all faces $\sigma \cup \{v_0\}$ for $\sigma \in \Delta$.

(a) Determine the f- and h-vectors of $C\Delta$ in terms of those of Δ.
(b) Show that Δ is shellable if and only if $C\Delta$ is shellable.

Exercise 1.5. Let Δ be a graph (that is, a 1-dimensional simplicial complex) with c components, v vertices, and e edges. Determine the isomorphism types of the simplicial homology groups $\tilde{H}_0(\Delta; R)$ and $\tilde{H}_1(\Delta; R)$ for any coefficient ring R.

Exercise 1.6. Construct two simplicial complexes with the same f-vector such that one is shellable and one isn't.

Exercise 1.7. Prove that conditions (1), (2) and (3) in the definition of shellability (Defn. 1.16) are equivalent.

Exercise 1.8. Prove Proposition 1.17.

Exercise 1.9. Prove that the link operation commutes with union and intersection of complexes. That is, if X, Y are simplicial complexes that are subcomplexes of a larger complex $X \cup Y$, and $\sigma \in X \cup Y$, then prove that

$$\mathrm{lk}_{X \cup Y}(\sigma) = \mathrm{lk}_X(\sigma) \cup \mathrm{lk}_Y(\sigma) \qquad \text{and} \qquad \mathrm{lk}_{X \cap Y}(\sigma) = \mathrm{lk}_X(\sigma) \cap \mathrm{lk}_Y(\sigma).$$

Exercise 1.10. (Requires some experience with homological algebra.) Prove that shellable simplicial complexes are Cohen-Macaulay. (Hint: First do the previous problem. Then use a Mayer-Vietoris sequence.)

Exercise 1.11. (Requires less experience with homological algebra than you might think.) The **reduced Euler characteristic** of a simplicial complex Δ is the alternating sum of its f-numbers

$$\tilde{\chi}(\Delta) = -1 + f(\Delta, -1) = -f_{-1} + f_0 - f_1 + f_2 \pm \cdots$$

(Topologists usually work with the unreduced Euler characteristic $\chi(\Delta) = f(\Delta, -1)$, which corresponds to ignoring the empty face.) Prove the *Euler-Poincaré formula*

$$\tilde{\chi}(\Delta) = \sum_{k \geq -1} (-1)^k \dim_{\Bbbk} \tilde{H}_k(\Delta; \Bbbk)$$

(the choice of ground field $\Bbbk$ is immaterial).

Exercise 1.12. Confirm that the convex hull of a finite point set is the set of convex linear combinations of it.

Exercise 1.13. Fill in the details in the proof of Theorem 1.20.

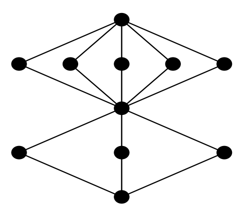
Definition 2.1. A poset L is a **lattice** if every pair $x, y \in L$ has a unique **meet** $x \wedge y$ and **join** $x \vee y$. That is,

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

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 $\hat{\mathbf{0}} = \wedge L$ and $\hat{\mathbf{1}} = \vee L$). For convenience, we define $\wedge \emptyset = \hat{\mathbf{1}}$ and $\vee \emptyset = \hat{\mathbf{0}}$.

Example 2.2. The Boolean algebra $2^{[n]}$ is a lattice, with $S \wedge T = S \cap T$ and $S \vee T = S \cup T$.

Example 2.3. The *complete graded poset* $P(a_1, \ldots, a_n)$ has $r(\hat{\mathbf{1}}) = n + 1$ and $a_i > 0$ elements at rank i for every $i > 0$, with every possible order relation (i.e., $r(x) > r(y) \implies x > y$).

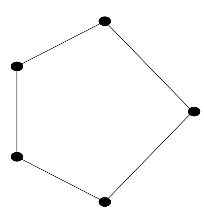


P(3,1,5)

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

Example 2.4. 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 $\hat{\mathbf{1}}$ element. In the case that $G = K_n$, the clique poset is a lattice, namely the partition lattice Π_n.

Example 2.5. Lattices don't have to be ranked. For example, the poset N_5 shown below is a perfectly good lattice.



Proposition 2.6 (**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 2.7. *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. Similarly, every bounded join semi-lattice is a lattice.*

Proof. Let P be a bounded meet-semilattice. 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 $\hat{\mathbf{1}} \in B$. 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. The proof of the second assertion is dual to this proof. □

Definition 2.8. 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

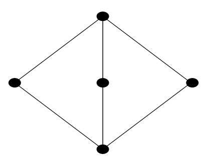
$$x \wedge_{L'} y = x \wedge_L y \quad \text{and} \quad x \vee_{L'} y = x \vee_L y.$$

A sublattice $L' \subset L$ does not have to have the same $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ 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 $\hat{\mathbf{0}}_{L'} = x$ and $\hat{\mathbf{1}}_{L'} = z$.)

Example 2.9 (**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, $L_2(2) \cong M_5$.



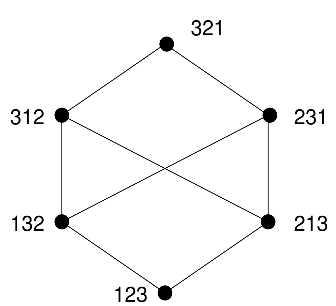
Note that $L_n(q)$ is self-dual, under the anti-automorphism $W \to W^{\perp}$. (An *anti-automorphism* is an isomorphism $P \to P^*$.)

Example 2.10 (**Bruhat order and weak Bruhat order**)**.** Let $\mathfrak{S}_n$ be the set of permutations of $[n]$ (i.e., the symmetric group).[5] 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:

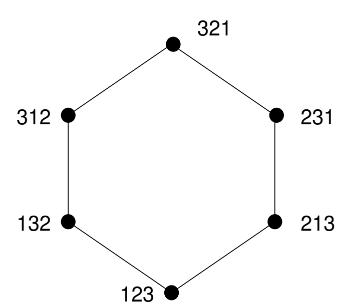
(1) $\sigma \lessdot \sigma'$ if σ' can be obtained by swapping σ_i with σ_{i+1}, where $\sigma_i < \sigma_{i+1}$. For example,
$$4718\underline{26}35 \lessdot 4718\underline{62}35 \quad \text{and} \quad 4\underline{71}82635 \gtrdot 4\underline{17}82635.$$

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

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

[5]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.

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**:

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

2.1. **Distributive Lattices.**

Definition 2.11. A lattice L is **distributive** if the following two equivalent conditions hold:

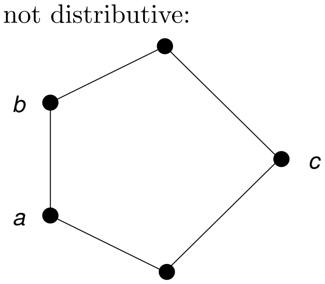
$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \qquad \forall x, y, z \in L, \tag{2.1a}$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \qquad \forall x, y, z \in L. \tag{2.1b}$$

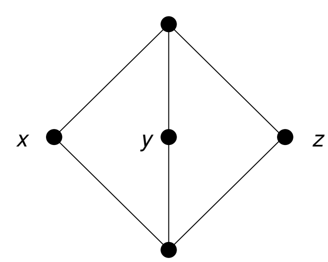
Proving that the two conditions (2.1a) and (2.1b) are equivalent is not too hard, but is not trivial (it's a homework problem). Note that replacing the equalities with $\geq$ and $\leq$ respectively gives statements that are true for all lattices.

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

(1) The Boolean algebra $2^{[n]}$ 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

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

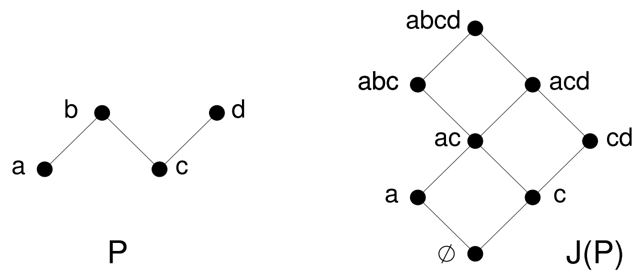
$$(a \wedge b) \vee (a \wedge c) = a \qquad\qquad (x \wedge z) \vee (y \wedge z) = \hat{\mathbf{0}}$$

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 a sublattice of a Boolean lattice (because, remember, meet and join are given by intersection and union on Ferrers diagrams).
(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 2.12. 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

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



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.

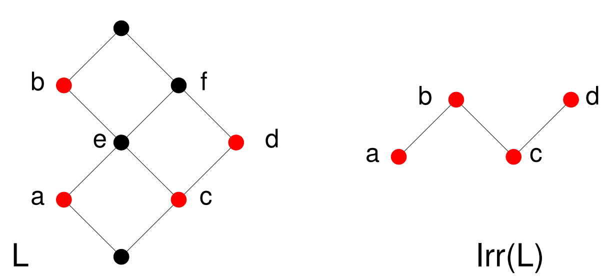
Proposition 2.13. *The operations $A \vee B = A \cup B$ and $A \wedge B = A \cap B$ make $J(P)$ into a distributive lattice, partially ordered by set containment.*

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 2.14. 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 $\hat{\mathbf{0}}$), every element $x \in L$ can be written as the join of join-irreducibles (but not necessarily uniquely; e.g., M_5). Given such an expression, we can throw out every join-irreducible strictly less than another, giving an expression for x as the join of an antichain worth of join-irreducibles; such an expression is called *irredundant*.

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.



Theorem 2.15 (**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 2.16. *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 2.16 says when applied to the divisor lattice D_n.)

Proposition 2.17. *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

$$x = p_1 \vee \cdots \vee p_n = q_1 \vee \cdots \vee q_m \tag{2.2}$$

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

By Lemma 2.16, $p_1 \leq q_j$ for some j. Again by Lemma 2.16, $q_j \leq p_i$ for some i. If $i \neq 1$, then $p_1 \lneq p_i$, 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 (2.2) 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

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

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 2.18. *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 $2^{[n]}$ are the singleton sets $\{i\}$ for $i \in [n]$; these form an antichain.

Corollary 2.19. *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 $x \vee \bar{x} = \hat{\mathbf{1}}$ and $\bar{x} \wedge y = \hat{\mathbf{0}}$.*
(6) *L is* relatively complemented. *That is, for every interval $[x, z] \subseteq L$ 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 $[\hat{\mathbf{0}}, z]$. Then Then

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

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 $\bar{x} \wedge x = x \neq \hat{\mathbf{0}}$, 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.)

2.2. **Modular Lattices.**

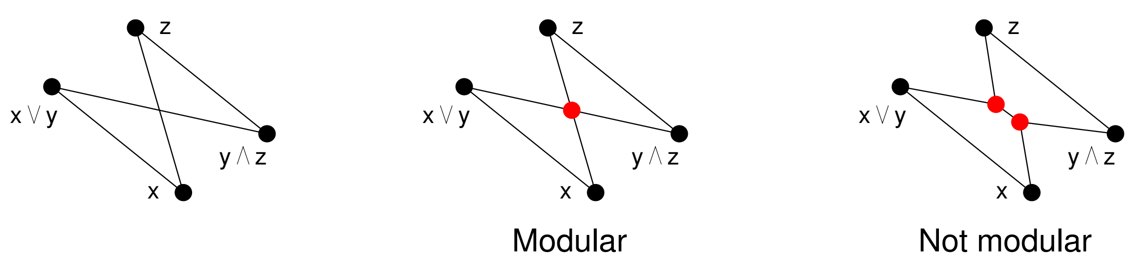
Definition 2.20. A lattice L is **modular** if for every $x, y, z \in L$ with $x \leq z$, the *modular equation* holds:

$$x \vee (y \wedge z) = (x \vee y) \wedge z. \tag{2.3}$$

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.



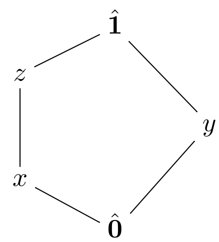
Note that for all lattices, if $x \leq z$, then $x \vee (y \wedge z) \leq (x \vee y) \wedge z$. Modularity says that, in fact, equality holds.

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
$$x \vee (y \wedge z) \;=\; (x \vee y) \wedge (x \vee z) \;=\; (x \vee y) \wedge z,$$
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 immediate, because the modular equation is invariant under dualization.

(3) The nonranked lattice N_5 is not modular:



$$x \leq z,$$
$$x \vee (y \wedge z) \;=\; x \vee \hat{\mathbf{0}} \;=\; x,$$
$$(x \vee y) \wedge z \;=\; \hat{\mathbf{1}} \wedge z \;=\; z.$$

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).

Proposition 2.21. *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) $\Longrightarrow$ (2): If $y \wedge z \leq x \leq z$, then the modular equation (2.3) becomes $x = (x \vee y) \wedge z$.

(2) $\Longrightarrow$ (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) $\Longleftrightarrow$ (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). $\square$

The following example may explain the term "modular lattice."

Corollary 2.22. *Let R be a (not necessarily commutative) ring and M a (left) R-submodule. Then the poset $L(M)$ of (left) R-submodules of M, ordered by inclusion, is a modular lattice with operations $A \vee B = A + B$, $A \wedge B = A \cap B$ (although it may not be a finite poset)*

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. 2.21. $\square$

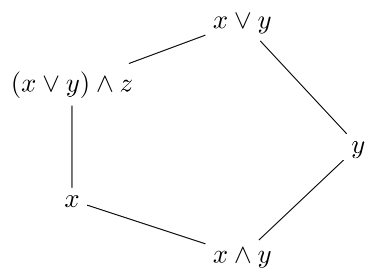
Examples include the lattice of subspaces of a vector space (e.g., $L_n(q)$, which we have seen previously) and the lattice of subgroups of an abelian group (i.e., of a $\mathbb{Z}$-module). In contrast, the lattice of subgroups of a nonabelian group need not be modular. For example, let G be the symmetric group $\mathfrak{S}_4$ and consider the subgroups $X = \langle (1\ 2\ 3) \rangle$, $Y = \langle (1\ 2) \rangle$ (using cycle notation), $Z = \mathfrak{A}_4$ (the alternating group). Then $(XY) \cap Z = Z$ but $X(Y \cap Z) = Z$.

Theorem 2.23. *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 $\Longrightarrow$ 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).

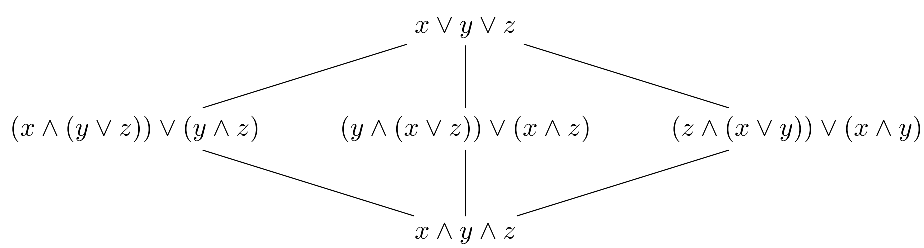
Suppose that L is not distributive. If it isn't modular then it contains an N_5, so there's nothing to prove. If it is modular, then choose x, y, z such that

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

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;
(3) $x \vee y = x \vee z = y \vee z$, and likewise for meets.

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



$\square$

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.

2.3. **Semimodular Lattices.**

Definition 2.24. A lattice L is **(upper) semimodular** if for all $x, y \in L$,

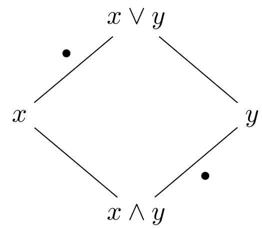
$$x \wedge y \lessdot y \implies x \lessdot x \vee y. \tag{2.4}$$

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

The implication (2.4) is trivially true 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 joining 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 2.25. *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 2.21 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. □

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

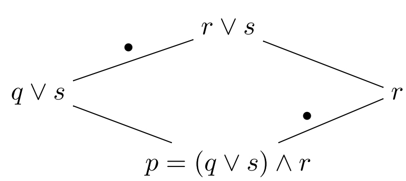
$$r(x \vee y) + r(x \wedge y) \spadesuit r(x) + r(y)$$

where $\spadesuit$ stands for $\leq$, $\geq$ or $=$ according as L is USM, LSM, or modular, respectively.

Lemma 2.26. *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 2.27. *Let L be a lattice. Then L is semimodular if and only if it is ranked, with rank function r satisfying the semimodular inequality*

$$r(x \vee y) + r(x \wedge y) \leq r(x) + r(y) \qquad \forall x, y \in L. \tag{2.5}$$

Proof. ($\Longleftarrow$) Suppose that L is a ranked lattice with rank function r satisfying (2.5). 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 (2.5)

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

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

($\Longrightarrow$) First, observe that if L is semimodular, then

$$x \wedge y \lessdot x, y \implies x, y \lessdot x \vee y. \tag{2.6}$$

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.

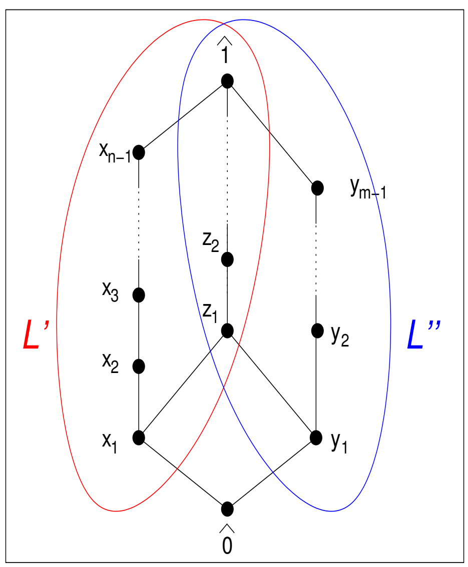
Base case: If $c(L) = 0$ or $c(L) = 1$, then this is trivial.

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 (2.5).

First, we show that L is ranked.

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

Let $L' = [x_1, \hat{\mathbf{1}}]$ and $L'' = [y_1, \hat{\mathbf{1}}]$. 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 (2.6), z_1 covers both x_1 and y_1. Let $z_1, z_2, \ldots, \hat{\mathbf{1}}$ be a saturated chain in L (thus, in $L' \cap L''$).



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

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

Second, we show that the rank function r of L satisfies (2.5). 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

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

[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, $c(2^{[n]}) = n$, not $n + 1$.

By Lemma 2.26, 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

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

which implies the semimodular inequality (2.5). □

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 (2.5).

Theorem 2.28. *L is modular if and only if it is ranked, with a rank function r satisfying*

$$r(x \vee y) + r(x \wedge y) = r(x) + r(y) \qquad \forall x, y \in L. \tag{2.7}$$

Proof. If L is modular, then it is both upper and lower semimodular, so the conclusion follows by Theorem 2.27. On the other hand, suppose that L is a lattice whose rank function r satisfies (2.7). 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,

$$\begin{aligned} r(x \vee (y \wedge z)) &= r(x) + r(y \wedge z) - r(x \wedge y \wedge z) \\ &= r(x) + r(y \wedge z) - r(x \wedge y) \\ &= r(x) + r(y) + r(z) - r(x \vee z) - r(x \wedge y) \end{aligned}$$

and

$$\begin{aligned} r((x \vee y) \wedge z) &= r(x \vee y) + r(z) - r(x \vee y \vee z) \\ &= r(x \vee y) + r(z) - r(y \vee z) \\ &= r(x) + r(y) - r(x \wedge y) + r(z) - r(y \vee z). \end{aligned}$$

□

2.4. **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

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

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 $r(A) = \dim \mathbb{F}A$. (Exercise: Check that r satisfies the submodular inequality.) It is *not* in general modular; e.g., see Example 2.31 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 $\mathbb{F}\langle v_1, \ldots, v_k \rangle = \mathbb{F}v_1 + \cdots + \mathbb{F}v_k$. This motivates the following definition:

Definition 2.29. A lattice L is **geometric** if it is (upper) semimodular and atomic. If $L \cong L(E)$ for some set of vectors E, we say that E is a **(linear) representation** of L.

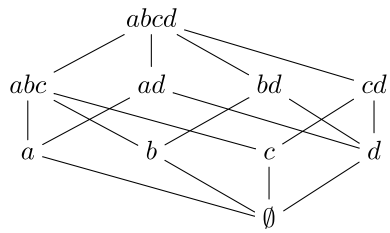
For example, the set $E = \{(0,1), (1,0), (1,1)\} \subseteq \mathbb{F}_2^2$ is a linear representation of the geometric lattice M_5. (For that matter, so is any set of three nonzero vectors in a two-dimensional space over any field, provided none is a scalar multiple of another.)

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. If $L \cong L^{\text{aff}}(E)$, we could say that E is a **(affine) representation** of L.

Example 2.30. Let E consist of four distinct points a, b, c, d, where a, b, c are collinear but no other set of three points is. Then $L^{\text{aff}}(E)$ is the lattice shown below (which happens to be modular).



Example 2.31. 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 2.32. *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 is 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 2.26, and $u' \wedge y = x$ (this takes a little more work; the proof is left as exercise). 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 $x = \hat{\mathbf{0}}$ 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 $[\hat{\mathbf{0}}, x]$. 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 $x_n = \hat{\mathbf{0}}$, and $x = a_1 \vee \cdots \vee a_n$. $\square$

2.5. **Exercises.**

Exercise 2.1. Prove that the two formulations of distributivity of a lattice L are equivalent, i.e.,

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

Exercise 2.2. Fill in the verification that $u' \wedge y = x$ in the first part of the proof of Proposition 2.32.

Exercise 2.3. Let S be a set and let $\mathcal{A}$ be a family of subsets of S such that $S = \bigcup_{A \in \mathcal{A}} A$. Let $U(\mathcal{A})$ be the set of unions of A_i's, i.e.,

$$U(\mathcal{A}) = \left\{ \bigcup_{A \in \mathcal{A}'} A \mid \mathcal{A}' \subseteq \mathcal{A} \right\}.$$

(a) Prove that $U(\mathcal{A})$ is a lattice. Under what conditions on $\mathcal{A}$ is $U(\mathcal{A})$ atomic?
(b) Let L be any lattice. Construct a set family $\mathcal{A}$ such that $U(\mathcal{A}) \cong L$.

Exercise 2.4. Let $L_n(q)$ be the poset of subspaces of an n-dimensional vector space over the finite field $\mathbb{F}_q$ (so $L_n(q)$ is a modular lattice by Corollary 2.22).

(a) Calculate its rank-generating function. (Giving a closed formula for the number of elements at each rank is good; giving a closed form for $\sum_{V \in L_n(q)} x^{r(V)}$ is even better.)
(b) Count the maximal chains in $L_n(q)$.

Exercise 2.5. Let Y be Young's lattice (which we know is distributive).

(a) Describe the join-irreducible elements of Young's lattice Y. If $\lambda = \mu_1 \vee \cdots \vee \mu_k$ is an irredundant decomposition into join-irreducibles, then what quantity does k correspond to in the Ferrers diagram of λ?
(b) Count the maximal chains in the interval $[\emptyset, \lambda] \subset Y$ if the Ferrers diagram of λ is a $2 \times n$ rectangle.
(c) Ditto if λ is a hook shape (i.e., $\lambda = (n+1, 1, 1, \ldots, 1)$, with a total of m copies of 1).

Exercise 2.6. Prove that the rank-generating function of Bruhat order on $\mathfrak{S}_n$ is

$$\sum_{\sigma \in \mathfrak{S}_n} q^{r(\sigma)} = \prod_{i=1}^{n} \frac{1-q^i}{1-q}$$

where $r(\sigma) = \#\{\{i,j\} \mid i < j \text{ and } \sigma_i > \sigma_j\}$. (Hint: Induct on n, and use one-line notation for permutations, not cycle notation.)

Exercise 2.7. Fill in the details in the proof of Birkhoff's theorem by showing the following facts.

(a) For a finite distributive lattice L, show that the map $\phi : L \to J(\mathrm{Irr}(L))$ given by

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

is indeed a lattice isomorphism.
(b) For a finite poset P, show that an order ideal in P is join-irreducible in $J(P)$ if and only if it is principal (i.e., generated by a single element).

Exercise 2.8. Prove that the partition lattice Π_n is geometric by finding an explicit linear representation E. (Hint: What are the atoms of Π_n? Start by finding vectors corresponding to them. Also, there is an "optimal" construction that works over any field.)

Exercise 2.9. Verify that the lattice Π_4 is not modular.

3. 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.

3.1. **Closure operators.** In what follows, we will frequently be adding or deleting single elements to or from a set. Accordingly, we abbreviate $A \cup e = A \cup \{e\}$ and $A \setminus e = A \setminus \{e\}$.

Definition 3.1. 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 $\bar{A} \subseteq \bar{B}$. A set A is called **closed** if $\bar{A} = A$.

As a consequence,

$$\overline{\bar{A} \cap \bar{B}} = \bar{A} \cap \bar{B} \qquad \forall A, B \subseteq E, \tag{3.1}$$

because $\overline{\bar{A} \cap \bar{B}} \subset \bar{\bar{A}} = \bar{A}$. In particular, intersections of closed sets are closed.

Definition 3.2. A closure operator on E is a **matroid closure operator** if it satisfies the *exchange axiom*:

$$\text{if } e \notin \bar{A} \text{ but } e \in \overline{A \cup f}, \text{ then } f \in \overline{A \cup e}. \tag{3.2}$$

Definition 3.3. 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 3.4. *Vector matroids.* Let V be a vector space over a field $\mathbb{F}$, and let $E \subseteq V$ be a finite set. Then

$$A \mapsto \bar{A} := \mathbb{F}A \cap E$$

is a matroid closure operator on E. It is easy to check the conditions for a closure operator. To check condition (3.2), 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 (3.2). 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 $\mathbb{F}$).

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. Note that a vector matroid is simple if none of the vectors is zero (so that $\bar{\emptyset} = \emptyset$) and if no vector is a scalar multiple of another (so that all singleton sets are closed).

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

Theorem 3.5. *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. **First,** let M be a simple matroid on E.

Step 1a: Show that $L(M)$ is an atomic lattice. The intersection of flats is a flat by (3.1), so the operation $A \wedge B = A \cap B$ makes $L(M)$ into a meet-semilattice. It's bounded (with $\hat{\mathbf{0}} = \bar{\emptyset}$ and $\hat{\mathbf{1}} = E$), so it's a lattice by Proposition 2.7. 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.

<u>Claim</u>: 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 (3.2), which implies $F \vee \{a\} \subseteq F \vee \{b\} \subseteq G$. So the $\subseteq$ above is actually an equality.

Moreover, 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 Claim, 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.)

Second, let L be a geometric lattice with atoms E.

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 (3.2).

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. ∎

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:

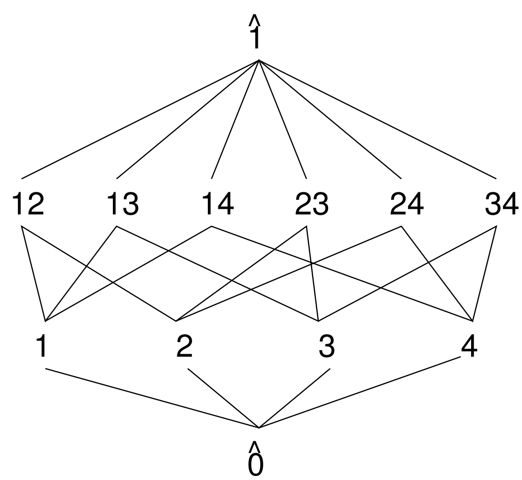
Definition 3.6. A *matroid rank function* on E is a function $r : 2^E \to \mathbb{N}$ satisfying the following conditions for all $A, B \subseteq E$:

(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 3.7. 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 3.6). 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 $2^{[n]}$ 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 2.31).



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.

Definition 3.8. 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 $A \subseteq E$.

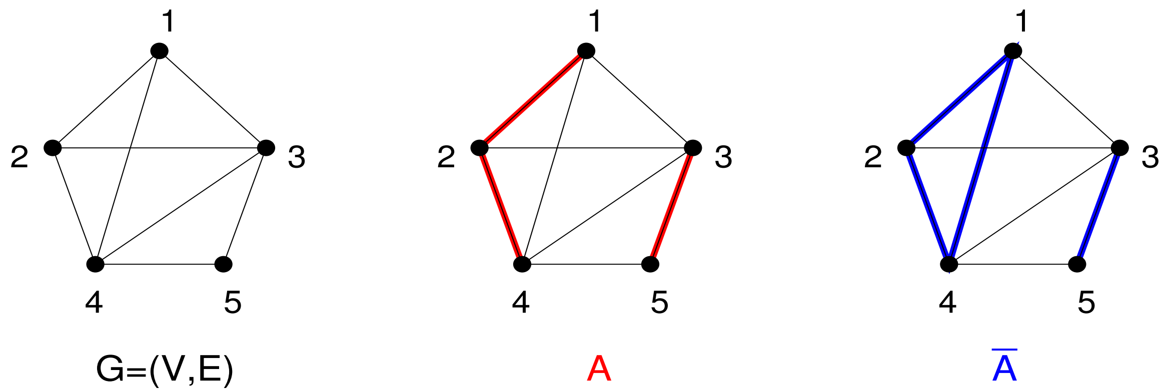
In general, every equivalent definition of "matroid" (and there are several more coming) will induce a corresponding equivalent notion of "isomorphic".

3.3. **Graphic Matroids.** Let G be a finite graph with vertices V and edges E. For convenience, we will write $e = xy$ to mean "e is an edge with endpoints x, y". This notation does not 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 3.9. 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

$$\bar{A} = \{e = xy \in E \mid x, y \text{ belong to the same component of } G|_A\}. \tag{3.3}$$

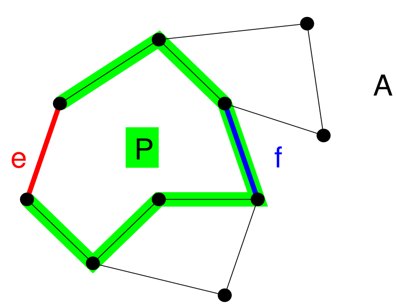
Equivalently, an edge $e = xy$ belongs to $\bar{A}$ if there is a path between x and y consisting of edges in A (for short, an *A-path*). For example, in the following graph, $14 \in \bar{A}$ because $\{12, 24\} \subset A$.



Proposition 3.10. *The operator $A \mapsto \bar{A}$ defined by (3.3) 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 $M(G)$. (I haven't yet said what a basis is — see the next section.)

Theorem 3.11. *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 3.12. 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 3.13. 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 3.14. 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.

3.4. **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 3.15. A **(matroid) independence system** $\mathscr{I}$ is a family of subsets of E such that

$$\emptyset \in \mathscr{I}; \tag{3.4a}$$

$$\text{if } I \in \mathscr{I} \text{ and } I' \subseteq I, \text{ then } I' \in \mathscr{I}; \quad \text{and} \tag{3.4b}$$

$$\text{if } I, J \in \mathscr{I} \text{ and } |I| < |J|, \text{ then there exists } x \in J \setminus I \text{ such that } I \cup x \in \mathscr{I}. \tag{3.4c}$$

Of course, conditions (3.4a) and (3.4b) say that $\mathscr{I}$ is an 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 (3.4a) and (3.4b) are clear. For condition (3.4c), 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 3.16. *Let E be a finite set.*

(1) If r is a matroid rank function on E, then

$$\mathscr{I} = \{A \subset E : \ r(A) = |A|\} \tag{3.5a}$$

is an independence system.

(2) If $\mathscr{I}$ is an independence system on E, then

$$r(A) = \max\{|I| : \ I \subseteq A, \ I \in \mathscr{I}\} \tag{3.5b}$$

is a matroid rank function.

(3) These constructions are mutual inverses.

Proof. Let r be a matroid rank function on E and define $\mathscr{I}$ as in (3.5a). First, $r(I) \leq |I|$ for all $I \subseteq E$, so (3.4a) follows immediately. Second, suppose $I \in \mathscr{I}$ and $I' \subseteq I$; say $I' = \{x_1, \ldots, x_k\}$ and $I = \{x_1, \ldots, x_n\}$. Consider the "flag" (nested family of subsets)

$$\emptyset \subsetneq \{x_1\} \subsetneq \{x_1, x_2\} \subsetneq \cdots \subsetneq I' \subsetneq \cdots \subsetneq I.$$

The rank starts at 0 and increases at most 1 each time by submodularity. But since $r(I) = |I|$, it must increase by exactly 1 each time. In particular $r(I') = l = |I'|$ and so $I' \in \mathscr{I}$, proving (3.4b).

To prove (3.4c), let $I, J \in \mathscr{I}$ with $|I| < |J|$ and let $J \setminus I = \{x_1, \dots, x_n\}$. If $n = 1$ then $J = I \cup \{x_1\}$ and there is nothing to show. Now suppose that $n \geq$ and $r(I \cup \{x_k\}) = r(I)$ for every $k \in [n]$. By submodularity,

$$\begin{aligned} r(I \cup \{x_1, x_2\}) &\leq r(I \cup x_1) + r(I \cup x_2) - r(I) &&= r(I), \\ r(I \cup \{x_1, x_2, x_3\}) &\leq r(I \cup \{x_1, x_2\}) + r(I \cup x_3) - r(I) &&= r(I), \\ &\cdots \\ r(I \cup \{x_1, x_2, \dots, x_n\}) &\leq r(I \cup \{x_1, \dots, x_{n-1}\}) + r(I \cup x_n) - r(I) &&= r(I), \end{aligned}$$

and equality must hold throughout. But then $r(I \cup J) = r(I) < r(J)$, which is a contradiction.

Now suppose that $\mathscr{I}$ is an independence system on E, and define a function $r : 2^E \to \mathbb{Z}$ as in (3.5b). It is immediate from the definition that $r(A) \leq |A|$ and that $A \subseteq B$ implies $r(A) \leq r(B)$ for all $A, B \in \mathscr{I}$.

Let $A, B \subset E$ and let I be a maximal independent subset of $A \cap B$. Extend I to a maximal independent subset $J \subseteq A \cup B$, which we can do by the following algorithm:

```
Let J' be any maximal independent subset of A ∪ B
Initialize J := I
while |J| < r(A ∪ B):
     Find x ∈ J' \ J such that J ∪ {x} is independent
     Set J := J ∪ {x}
```

Note that $J \setminus I \subseteq A \triangle B$ (where $\triangle$ means symmetric difference), since I was chosen to be a *maximal* independent subset of $A \cap B$. Therefore, by inclusion-exclusion,

$$\begin{aligned} r(A \cup B) = |J| &= |J \cap A| + |J \cap B| - |J \cap A \cap B| \\ &= |J \cap A| + |J \cap B| - |I| \\ &= |J \cap A| + |J \cap B| - r(|A \cap B) \\ &\leq r(A) + r(B) - r(|A \cap B) \end{aligned}$$

since $J \cap A$ and $J \cap B$ are certainly independent subsets of A, B respectively (but not necessarily maximal ones). So we have proved that r is submodular. $\square$

Remark 3.17. The algorithm in the proof implies another useful fact (which makes perfect sense in the context of vector matroids): if $P \subseteq Q$ are subsets of the ground set of a matroid, then any basis of P can be extended to a basis of Q.

If $M = M(G)$ is a graphic matroid, the associated independence system $\mathscr{I}$ 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 (3.4c) 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 3.18. *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 $\Delta|_F = \{\sigma \in \Delta \mid \sigma \subseteq F\}$ is shellable.
(3) For every $F \subseteq E$, the induced subcomplex $\Delta|_F = \{\sigma \in \Delta \mid \sigma \subseteq F\}$ is Cohen-Macaulay.
(4) For every $F \subseteq E$, the induced subcomplex $\Delta|_F$ is pure.

Proof. The implications (2) $\implies$ (3) $\implies$ (4) are consequences of the material in Section 1.3 (the first is a homework problem and the second is easy).

(4) $\implies$ (1): Suppose I, J are independent sets with $|I| < |J|$. Then the induced subcomplex $\Delta|_{I\cup J}$ is pure, which means that I is not a maximal face of it. Therefore there is some $x \in (I \cup J) \setminus I = J \setminus I$ such that $I \cup x \in \Delta$, establishing the donation condition.

(1) $\implies$ (4): Let $F \subseteq E$. If I is a non-maximum face of $\Delta|_F$, then we can pick J to be a maximum element of it, and then donation says that there is some $x \in J$ such that $I \cup \{x\}$ is a face of Δ, hence of $\Delta|_F$.

(4) $\implies$ (2): More interesting; left as an exercise. $\square$

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

Definition 3.19. A **(matroid) basis system on** E is a family $\mathscr{B} \subseteq 2^E$ such that for all $B, B' \in \mathscr{B}$,

$$|B| = |B'|; \quad \text{and} \tag{3.6a}$$

$$\text{for all } e \in B \setminus B',\ \text{there exists } e' \in B' \setminus B \text{ such that } B \setminus e \cup e' \in \mathscr{B}. \tag{3.6b}$$

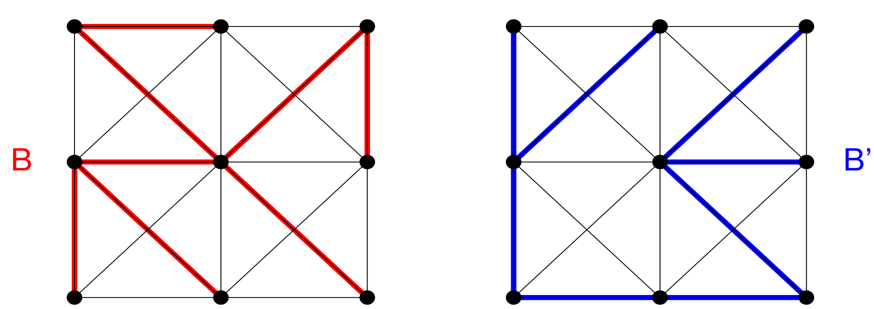
Given (3.6a), the condition (3.6b) can be replaced with

$$\text{for all } e \in B \setminus B',\ \text{there exists } e' \in B' \setminus B \text{ such that } B' \setminus e' \cup e \in \mathscr{B}, \tag{3.6c}$$

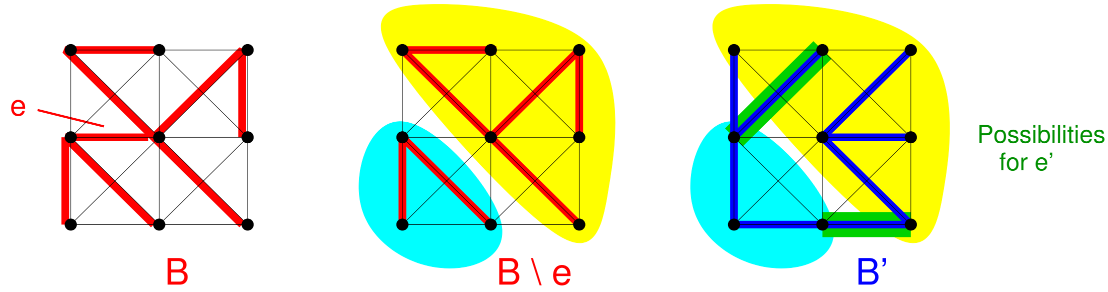
although this requires some proof (left as an exercise).

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 (3.6b).

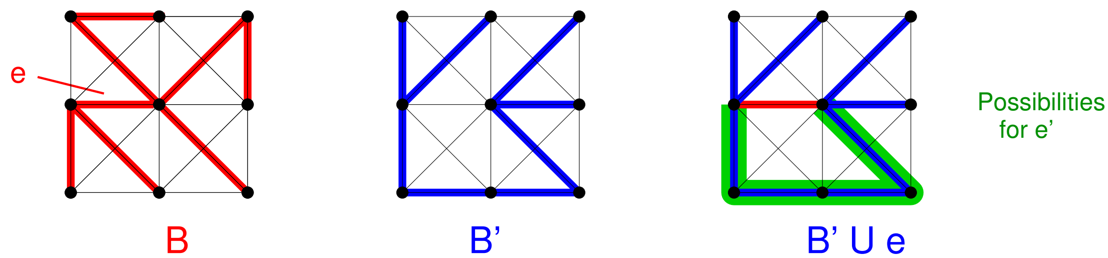
If G is a graph, then the bases of $M(G)$ are its **spanning forests**, i.e., its maximal acyclic edge sets. If G is connected (which, as we will see, we may as well assume when studying graphic matroids) then the bases of $M(G)$ are its spanning trees.



Here is the graph-theoretic interpretation of (3.6b). Let G be a connected graph, let B, B' be spanning trees, and let $e \in B \setminus B'$. Then $B \setminus e$ has exactly 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 (3.6c), 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 3.20. *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 3.20 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 3.21. A **(matroid) circuit system on** E is a family $\mathscr{C} \subseteq 2^E$ such that, for all $C, C' \in \mathscr{C}$,

$$C \not\subsetneq C'; \quad \text{and} \tag{3.7a}$$

$$\text{for all } e \in C \cap C',\ C \cup C' \setminus e \text{ contains an element of } \mathscr{C}. \tag{3.7b}$$

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 a (non-loop) edge $e = xy$, then $C \setminus e$ and $C' \setminus e$ are paths between x and y, so concatenating them forms a closed path. This path is not necessarily itself a cycle, but must contain some cycle.

Proposition 3.22. *Let E be a finite set.*

(1) If $\mathscr{I}$ is an independence system on E, then $\{C \notin \mathscr{I} \mid C' \in \mathscr{I}\ \forall C' \subsetneq C\}$ 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.

In other words, the circuits are the minimal nonfaces of the independence complex (hence they correspond to the generators of the Stanley-Reisner ideal; see Defn. 1.15). The proof is left as an 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 3.15). 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 $A \in \Delta$; the maximum will certainly be achieved on a facet. A naive approach to find a maximal-weight A, which may or may not work for a given Δ and w, is the following "greedy" algorithm (known as **Kruskal's 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 3.23. *Δ is a matroid independence system if and only if Kruskal's 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. This interpretation can be useful in algebraic combinatorics; see Example 9.47 below.

Summary of Matroid Axiomatizations

- *Geometric lattice*: a lattice that is atomic and semimodular. Only simple 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 $w : E \to \mathbb{R}_{\geq 0}$.

3.5. **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 3.24. 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.

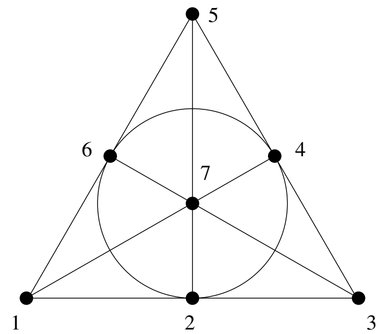
For example:

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

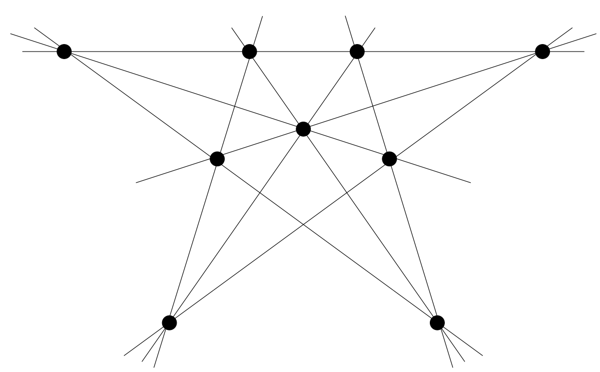


This point configuration can't be represented over $\mathbb{R}$ — if you try to draw seven non-collinear points such that the six triples 123, 345, 156, 147, 257, 367 are each collinear, then 246 will not be collinear (and in fact this is true over any field of characteristic $\neq 2$) — but it can be 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 is an affine representation of a rank-3 matroid over $\mathbb{R}$, but the matroid is not representable over $\mathbb{Q}$ [Grü03, 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 $\mathbb{F}_2$.)

Recall that a *minor* of a matrix is the determinant of some square submatrix of M. A matrix is called **totally unimodular** if every minor is either 0, 1, or -1.

Theorem 3.26. *A matroid M is regular if and only if it can be represented by the columns of a totally unimodular matrix.*

One direction is easy: if M has a unimodular representation then the coefficients can be interpreted as lying in any field, and the linear dependence of a set of columns does not depend on the choice of field (because $-1 \neq 0$ and $1 \neq 0$ in every field). The reverse direction is harder (see [Oxl92, chapter 6]), and the proof is omitted. In fact, something more is true: M is regular if and only if it is binary (representable over $\mathbb{F}_2$) and representable over at least one field of characteristic $\neq 2$.

Example 3.27. The matrix

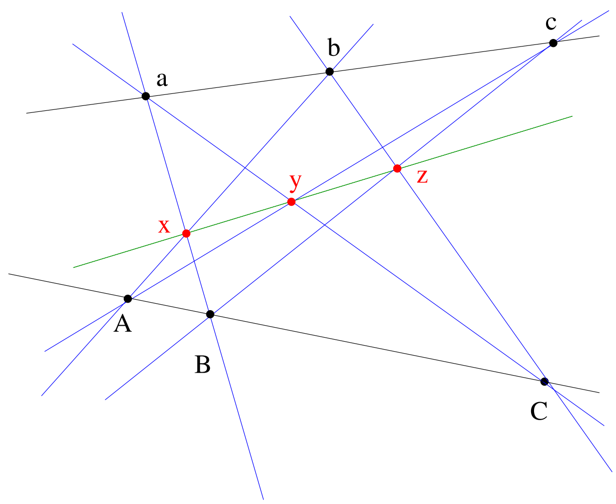
$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

represents $U_2(4)$ over any field of characteristic $\neq 2$, but the last two columns are dependent (in fact equal) in characteristic 2.

There exist matroids that are not representable over *any* field. A standard example is the **non-Pappus matroid**, which has a ground set of size 9 (see Example 3.28).

Example 3.28. Pappus' Theorem from Euclidean geometry says that if a, b, c, A, B, C are distinct points in $\mathbb{R}^2$ such that a, b, c and A, B, C are collinear, then x, y, z are collinear, where

$$x = \overline{aB} \cap \overline{Ab}, \qquad y = \overline{aC} \cap \overline{Ac}, \qquad z = \overline{bC} \cap \overline{Bc}.$$



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

$$\emptyset, \qquad a, b, c, a, b, c, x, y, z, \qquad abc,\ ABC,\ aBx,\ Abx,\ aCy,\ Acy,\ bCz,\ Bcz,\ xyz, \qquad E.$$

It turns out that deleting xyz from this list produces the family of closed sets of a matroid NP. Since Pappus' theorem can be proven using analytic geometry, and the equations that say that x, y, z are collinear are valid over any field (i.e., involve only ± 1 coefficients), it follows that NP 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 [Oxl92, p. 511].

3.6. **Operations on Matroids.** There are several ways to construct new matroids from old ones. We'll begin with a boring but useful one (direct sum) and then move on to the more exciting constructions of duality and deletion/contraction.

3.6.1. *Direct Sum.*

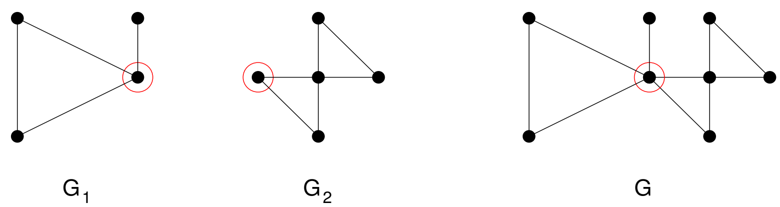
Definition 3.29. Let M_1, M_2 be matroids on disjoint sets E_1, E_2, with basis systems $\mathscr{B}_1, \mathscr{B}_2$. The **direct sum** $M_1 \oplus M_2$ is the matroid on $E_1 \cup E_2$ with basis system

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

I will omit the routine proof that $\mathscr{B}$ 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, then $M(G_1) \oplus M(G_2) \cong M(G_1 + G_2)$, where $+$ denotes disjoint union. Actually, 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 3.30. A matroid that cannot be written nontrivially as a direct sum of two smaller matroids is called **connected** or **indecomposable**.[7]

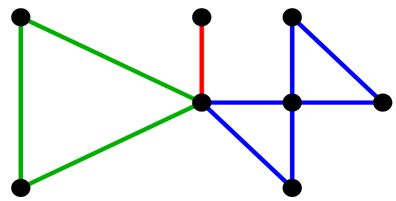
Proposition 3.31. *Let $G = (V, E)$ be a loopless graph. Then $M(G)$ is indecomposable if and only if G is **2-connected** — i.e., not only is it connected, but so is every subgraph obtained by deleting a single vertex.*

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.

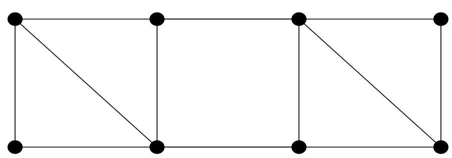
[7]The first term is more common among matroid theorists, but I prefer "indecomposable" to avoid potential confusion with the graph-theoretic meaning of "connected".

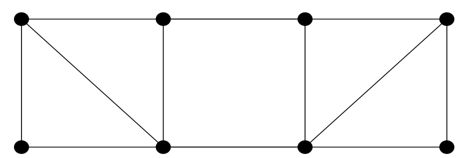


We'll prove the other direction later (maybe).

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 $2^{[n]}$ (and, for that matter, so is the lattice of flats).

In light of Proposition 3.31, 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.





3.6.2. *Duality.*

Definition 3.32. 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 (3.6a) is clearly invariant under complementation, and complementation swaps (3.6b) and (3.6c) — 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 $S = \{v_1, \dots, v_n\} \subset \mathbb{F}^r$, 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

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

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 $\operatorname{rank}(Y) = \dim \operatorname{nullspace}(X) = n - r$, 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 3.33. 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^*.

Remark 3.34. Invertible row operations on a matrix $X \in \mathbb{F}^{r\times n}$ (i.e., action on the left by $GL_r(\mathscr{F})$) do not change the matroid represented by its columns, since they correspond to changing the basis of the ambient space of the column vectors. Therefore, one can often assume without loss of generality that a matroid representation X is in *reduced* row-echelon form, i.e., that

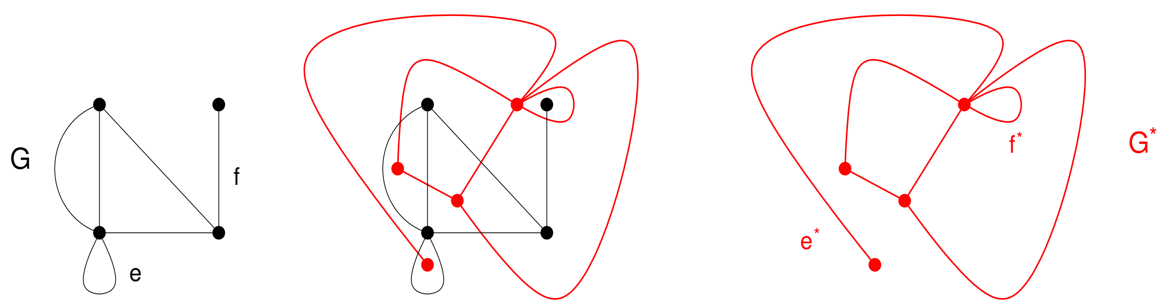
$$X = [I_r \mid A]$$

where I_r is the $r\times r$ identity matrix and A is arbitrary. It is easy to check that nullspace $X = (\text{rowspace}\, X^*)^T$, where

$$X^* = [A^T \mid -I_{n-r}],$$

(this is a standard recipe), so the columns of X^* are a representation for the dual matroid M^*. This construction also shows that representability over a particular field is unchanged by dualization.

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:

- $A \subseteq E$ is acyclic if and only if $E^* \setminus A^*$ is connected.
- $A \subseteq E$ 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).

If G is not planar then in fact $M(G)^*$ is not a graphic matroid (although it is certainly regular).

Definition 3.35. 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** (probably not standard terminology, but natural and useful).

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.

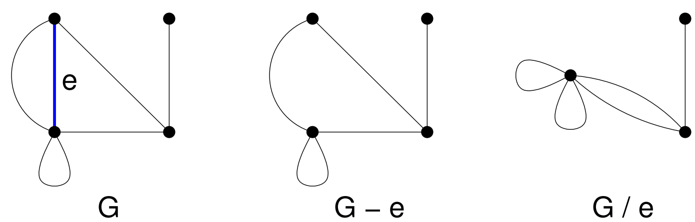
A **cocircuit** of M is by definition a circuit of the dual matroid M^*. Set-theoretically, a cocircuit is a minimal set not contained in any basis of M^*, so it is a minimal set that meets every basis of M. For a connected graph G, the cocircuits of the graphic matroid $M(G)$ are the *bonds* of G: the 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 3.21).

3.6.3. *Deletion and Contraction.*

Definition 3.36. 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 $M \setminus e$ (or $M - e$) 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 shrink it down so that its two endpoints merge into one.



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 representing a matroid M, and let $e \in E$. Then $M \setminus e$ 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:
$$(M \setminus e)^* \cong M^*/e \quad \text{and} \quad (M/e)^* \cong M^* \setminus e. \tag{3.8}$$
(Exercise: Prove this.)

Example 3.37. If M is the uniform matroid $U_k(n)$, then $M \setminus e \cong U_k(n-1)$ 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 3.36.

Proposition 3.38. *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*
$$b(M) = \begin{cases} b(M \setminus e) & \text{if } e \text{ is a loop;} \\ b(M/e) & \text{if } e \text{ is a coloop;} \\ b(M \setminus e) + b(M/e) & \text{otherwise.} \end{cases}$$

Example 3.39. If $M \cong U_k(n)$, then $b(M) = \binom{n}{k}$, and the recurrence of Proposition 3.38 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

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

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

Recall that the reduced Euler characteristic of Δ is

$$\tilde{\chi}(\Delta) = \sum_{\sigma\in\Delta} (-1)^{\dim\sigma} = \sum_{\sigma\in\Delta} (-1)^{|\sigma|-1}. \tag{3.9}$$

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

$$\begin{aligned}\tilde{\chi}(\Delta) &= \sum_{\sigma\in\Delta:\ e\notin\sigma} (-1)^{\dim\sigma} + \sum_{\sigma\in\Delta:\ e\in\sigma} (-1)^{\dim\sigma}\\ &= \sum_{\sigma\in\operatorname{del}_\Delta(e)} (-1)^{\dim\sigma} + \sum_{\tau\in\operatorname{lk}_\Delta(e)} (-1)^{1+\dim\tau}\\ &= \tilde{\chi}(\operatorname{del}_\Delta(e)) - \tilde{\chi}(\operatorname{lk}_\Delta(e)). \end{aligned}\tag{3.10}$$

These observations are the tip of an iceberg that we will explore in Section 4.

3.7. Exercises.

Exercise 3.1. Let $\mathbb{F}$ be a field, let $n \in \mathbb{N}$, and let S be a finite subset of the vector space $\mathbb{F}^n$. Recall from Section 2.4 the definitions of the lattices $L(S)$ and $L^{\text{aff}}(S)$. For $s = (s_1, \dots, s_n) \in S$, let $\hat{s} = (1, s_1, \dots, s_n) \in \mathbb{F}^{n+1}$, and let $\hat{S} = \{\hat{s} \mid s \in S\}$. Prove that $L(\hat{S}) = L^{\text{aff}}(S)$. (This is useful because it saves us a dimension — e.g., many geometric lattices of rank 3 can be represented conveniently as affine point configurations in $\mathbb{R}^2$.)

Exercise 3.2. Determine, with proof, all pairs of integers $k \le n$ such that there exists a graph G with $M(G) \cong U_k(n)$. (Here $U_k(n)$ denotes the uniform matroid of rank k on n elements; see Example 3.7.)

Exercise 3.3. Prove the equivalence of the two forms of the basis exchange condition (3.6b) and (3.6c). (Hint: Examine $|B \setminus B'|$.)

Exercise 3.4. (Proposed by Kevin Adams.) Let B, B' be bases of a matroid M. Prove that there exists a bijection $\phi : B \setminus B' \to B \setminus B$ such that $B \setminus e \cup \phi(e)$ is a basis of M for every $e \in B \setminus B'$.

Exercise 3.5. Prove Proposition 3.22, which describes the equivalence between matroid independence systems and matroid circuit systems.

Exercise 3.6. Complete the proof of Proposition 3.18 by showing that if Δ is a simplicial complex in ground set V in which every induced subcomplex is pure, then Δ is shellable. To do this, pick a vertex v. Use induction to show that the two complexes

$$\begin{aligned}\Delta_1 &= \langle \text{facets } F \in \Delta \mid v \notin F\rangle\,,\\ \Delta_2 &= \langle \text{facets } F \in \Delta \mid v \in F\rangle\end{aligned}$$

are both shellable, then concatenate the shelling orders to produce a shelling order on Δ. Derive a relationship among the h-polynomials of Δ, Δ_1, and Δ_2.

Exercise 3.7. Let M be a matroid on ground set E with rank function $r : 2^E \to \mathbb{N}$. Prove that the rank function r^* of the dual matroid M^* is given by $r^*(A) = r(E \setminus A) + |A| - r(E)$ for all $A \subseteq E$.

Exercise 3.8. Let E be a finite set and let Δ be an abstract simplicial complex on E. Let $w : E \to \mathbb{R}_{\ge 0}$ be any function; think of $w(A)$ as the "weight" of A. For $A \subseteq E$, define $w(A) = \sum_{e\in A} w(e)$. Consider the problem of maximizing $w(A)$ over all maximal[2] elements $A \in \Delta$ (also known as *facets* of Δ). A naive approach to try to produce such a set A is the following *greedy algorithm*:

[2] Recall that "maximal" means "not contained in any other element of Δ", which is logically weaker than "of largest possible cardinality".

(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.

This algorithm may or may not work for a given simplicial complex Δ and weight function w.

(a) Construct a simplicial complex and a weight function for which this algorithm does not produce a facet of maximal weight. (Hint: The smallest example has $|E| = 3$.)
(b) Prove that the following two conditions are equivalent:
- The greedy algorithm produces a facet of maximal weight for *every* weight function w.
- Δ is a matroid independence system.

Note: Do not use Proposition 3.18, which makes the problem too easy. Instead, work directly with the definition of a matroid independence system.

Exercise 3.9. Prove equation 3.8.

Exercise 3.10. Let X and Y be disjoint sets of vertices, and let B be an X, Y-bipartite graph: that is, every edge of B has one endpoint in each of X and Y. For $V = \{x_1, \ldots, x_n\} \subset X$, a *transversal* of V is a set $W = \{y_1, \ldots, y_n\} \subset Y$ such that $x_i y_i$ is an edge of B. (The set of all edges $x_i y_i$ is called a *matching.*) Let $\mathscr{I}$ be the family of all subsets of X that have a transversal; it is immediate that $\mathscr{I}$ is a simplicial complex.

Prove that $\mathscr{I}$ is in fact a matroid independence system by verifying that the donation condition holds. (Suggestion: Write down an example or two of a pair of independent sets I, J with $|I| < |J|$, and use the corresponding matchings to find a systematic way of choosing a vertex that J can donate to I.) These matroids are called *transversal matroids*; along with linear and graphic matroids, they are the other "classical" examples of matroids in combinatorics.

Exercise 3.11. (Requires a bit of abstract algebra.) Let n be a positive integer, and let ζ be a primitive n^{th} root of unity. The *cyclotomic matroid* Y_n is represented over $\mathbb{Q}$ by the numbers $1, \zeta, \zeta^2, \ldots, \zeta^{n-1}$, regarded as elements of the cyclotomic field extension $\mathbb{Q}(\zeta)$. Thus, the rank of Y_n is the dimension of $\mathbb{Q}(\zeta)$ as a $\mathbb{Q}$-vector space, which is given by the Euler phi function. Prove the following:

(a) if n is prime, then $Y_n \cong U_{n-1}(n)$.
(b) if m is the square-free part of n (i.e., the product of all the distinct primes dividing n — e.g., the square-free part of $56 = 2^3 \cdot 7$ is $2 \cdot 7 = 14$) then Y_n is the direct sum of n/m copies of Y_m.
(c) if $n = pq$, where p, q are distinct primes, then $Y_n \cong M(K_{p,q})^*$ — that is, the dual of the graphic matroid of the complete bipartite graph $K_{p,q}$.

4. The Tutte Polynomial

Throughout this section, let M be a matroid on ground set E with rank function r, and let $n = |E|$.

For $A \subseteq E$, we define

$$\operatorname{corank} A = r(E) - r(A),$$
$$\operatorname{nullity} A = |A| - r(A).$$

Corank and nullity measure how far A is from being spanning and independent, respectively. That is, the corank is the minimum number of elements needed to adjoin to A to produce a spanning set (i.e., to intersect all cocircuits), while the nullity is the minimum number of elements needed to delete from A to produce an independent set (i.e., to break all circuits).

Definition 4.1. The **Tutte polynomial** of M is

$$T_M = T_M(x, y) \ := \ \sum_{A \subseteq E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)}. \tag{4.1}$$

Example 4.2. If $E = \emptyset$ then $T_M(x, y) = 1$. Mildly less trivially, if every element is a coloop, then $r(A) = |A|$ for all A, so

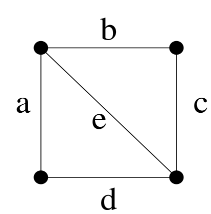
$$T_M = \sum_{A \subseteq E} (x-1)^{n-|A|} = (x-1+1)^n = x^n$$

by the binomial theorem. If every element is a loop, then the rank function is identically zero and we get

$$T_M \sum_{A \subseteq E} (y-1)^{|A|} = y^n.$$

Example 4.3. ($U_1(2)$, $U_1(3)$, $U_2(3)$)

Example 4.4. Let G be the graph below (known as the "diamond"):



The formula (4.1) gives

A	$\lvert A\rvert$	$r(A)$	$3-r(A)$	$\lvert A\rvert - r(A)$	contribution to (4.1)		
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 set	5	3	0	2	$1(x-1)^0(y-1)^2$	$=$	$y^2 - 2y + 1$
Total					$x^3 + 2x^2 + x + 2xy + y^2 + y$		

Many invariants of M an be obtained by specializing the variables x, y appropriately. Some easy ones:

(1) $T_M(2, 2) = \sum_{A \subseteq E} 1 = w^{|E|}$.

(2) Consider $T_M(1,1)$. This kills off all summands whose corank is nonzero (i.e., all non-spanning sets) and whose nullity is nonzero (i.e., all non-independent sets). What's left are the bases, ash each of which contributes a summand of $0^0 = 1$. So $T_M(1,1) = b(M)$, the number of bases. We previously observed that this quantity satisfies a deletion/contraction recurrence; this will show up again soon.
(3) Similarly, $T_M(1,2)$ and $T_M(2,1)$ count respectively the number of spanning sets and the number of independent sets.
(4) A little more generally, we can enumerate independent and spanning sets by their cardinality:

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

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

(5)

$$\begin{aligned} T_M(0,1) &= \sum_{A\subseteq E} (-1)^{r(E)-r(A)} 0^{|A|-r(A)} = \sum_{A\subseteq E \text{ independent}} (-1)^{r(E)-r(A)} \\ &= (-1)^{r(E)} \sum_{A\in\mathscr{I}(M)} (-1)^{|A|} \\ &= (-1)^{r(E)-1} \tilde{\chi}(\mathscr{I}(M)). \end{aligned}$$

The fundamental theorem about the Tutte polynomial is that it satisfies a deletion/contraction recurrence. In a sense it is the most general such invariant — we will give a "recipe theorem" that expresses any deletion/contraction invariant as a Tutte polynomial specialization (more or less).

Theorem 4.5. *The Tutte polynomial satisfies (and can be computed by) the following* ***Tutte recurrence****:*

(T1) *If $E = \emptyset$, then $T_M = 1$.*
(T2a) *If $e \in E$ is a loop, then $T_M = yT_{M\setminus e}$.*
(T2b) *If $e \in E$ is a coloop, then $T_M = xT_{M/e}$.*
(T3) *If $e \in E$ is ordinary, then $T_M = T_{M\setminus e} + T_{M/e}$.*

We can use this recurrence to compute the Tutte polynomial, by picking one element at a time to delete and contract. The miracle is that *it doesn't matter what order we choose on the elements of E* — all orders will give the same final result! (In the case that M is a uniform matroid, then it *is* clear at this point that T_M is well-defined by the Tutte recurrence, because, up to isomorphism, $M \setminus e$ and M/e are independent of the choices of $e \in E$.)

Before proving the theorem, here are some examples.

Example 4.6. Suppose that $M \cong U_n(n)$, that is, every element is a coloop. By induction, $T_M(x,y) = x^n$. Dually, if $M \cong U_0(n)$ (every element is a loop), then $T_M(x,y) = y^n$.

Example 4.7. Let $M \cong U_1(2)$ (the graphic matroid of the "digon", two vertices joined by two parallel edges). Let $e \in E$; then

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

Example 4.8. 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 the Tutte recurrence 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,

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

The Tutte recurrence says we can represent a calculation of T_M by a binary tree in which moving down corresponds to deleting or contracting:

$$\begin{array}{ccccccc} & & & M & & & \\ & M \setminus e & & & & M \setminus f & \\ M \setminus e \setminus f & & M \setminus e/f & & M/e \setminus f & & M/e/f \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

Example 4.9. Consider the diamond of Example 4.4. 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

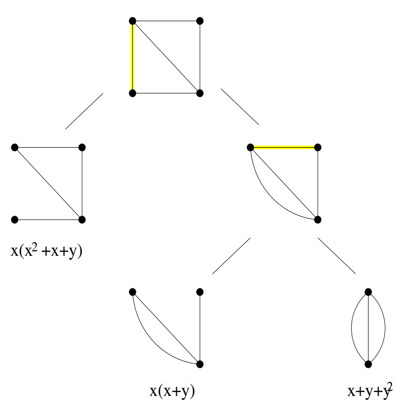
$$T(G \setminus a) \;=\; x(x^2 + x + y)$$

by Example 4.8. Next, apply the Tutte recurrence to the edge b in G/a. The graph $G/a \setminus b$ has a coloop c, contracting which produces a digon. Meanwhile, $M(G/a/b) \cong U_1(3)$. Therefore

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

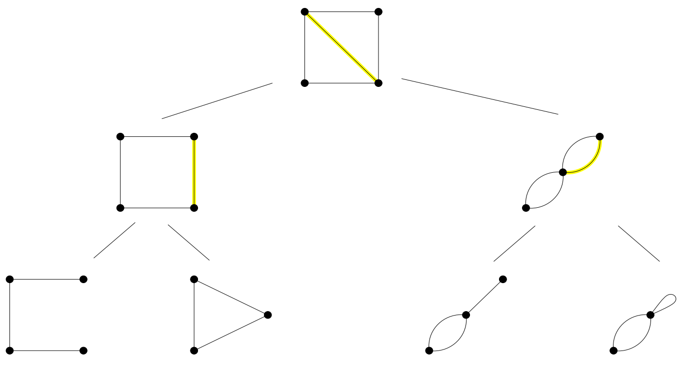
Putting it all together, we get

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



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

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



Proof of Theorem 4.5. 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 $r'(A) = r_{M\setminus e}(A) = r_M(A)$.
(2) If e is not a loop, then $r''(A) = r_{M/e}(A) = r_M(A \cup e) - 1$.

To save space, set $X = x - 1$, $Y = y - 1$. We already know that if $E = \emptyset$, then $T_M = 1$.

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

$$\begin{aligned}
T_M &= \sum_{A\subseteq E} X^{r(E)-r(A)}Y^{|A|-r(A)} \\
&= \sum_{A\subseteq E:\ e\not\in A} X^{r(E)-r(A)}Y^{|A|-r(A)} + \sum_{B\subseteq E:\ e\in B} X^{r(E)-r(B)}Y^{|A|-r(B)} \\
&= \sum_{A\subseteq E\setminus e} X^{r'(E\setminus e)-r'(A)}Y^{|A|-r'(A)} + \sum_{A\subseteq E\setminus e} X^{r'(E\setminus e)-r'(A)}Y^{|A|+1-r'(A)} \\
&= (1+Y)\sum_{A\subseteq E\setminus e} X^{r'(E\setminus e)-r'(A)}Y^{|A|-r'(A)} \quad = \quad yT_{M\setminus e}.
\end{aligned}$$

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

$$\begin{aligned}
T_M &= \sum_{A\subseteq E} X^{r(E)-r(A)}Y^{|A|-r(A)} \\
&= \sum_{e\not\in A\subseteq E} X^{r(E)-r(A)}Y^{|A|-r(A)} + \sum_{e\in B\subseteq E} X^{r(E)-r(B)}Y^{|B|-r(B)} \\
&= \sum_{A\subseteq E\setminus e} X^{(r''(E\setminus e)+1)-r''(A)}Y^{|A|-r''(A)} + \sum_{A\subseteq E\setminus e} X^{(r''(E\setminus e)+1)-(r''(A)+1)}Y^{|A|+1-(r''(A)+1)} \\
&= \sum_{A\subseteq E\setminus e} X^{r''(E\setminus e)+1-r''(A)}Y^{|A|-r''(A)} + \sum_{A\subseteq E\setminus e} X^{r''(E\setminus e)-r''(A)}Y^{|A|-r''(A)} \\
&= (X+1)\sum_{A\subseteq E\setminus e} X^{r''(E\setminus e)-r''(A)}Y^{|A|-r''(A)} \quad = \quad xT_{M/e}.
\end{aligned}$$

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

$$\begin{aligned}
T_M &= \sum_{A\subseteq E} X^{r(E)-r(A)}Y^{|A|-r(A)}\\
&= \sum_{A\subseteq E\setminus e} \left[X^{r(E)-r(A)}Y^{|A|-r(A)}\right] + \left[X^{r(E)-r(A\cup e)}Y^{|A\cup e|-r(A\cup e)}\right]\\
&= \sum_{A\subseteq E\setminus e} \left[X^{r'(E\setminus e)-r'(A)}Y^{|A|-r'(A)}\right] + \left[X^{(r''(E)+1)-(r''(A)+1)}Y^{|A|+1-(r''(A)-1)}\right]\\
&= \sum_{A\subseteq E\setminus e} X^{r'(E\setminus e)-r'(A)}Y^{|A|-r'(A)} + \sum_{A\subseteq E\setminus e} X^{r''(E\setminus e)-r''(A)}Y^{|A|-r''(A)}\\
&= T_{M\setminus e} + T_{M/e}.
\end{aligned}$$

□

Some easy and useful observations:

(1) The Tutte polynomial is multiplicative on direct sums, i.e., $T_{M_1\oplus M_2} = T_{M_1}T_{M_2}$. This is probably easier to see from the rank-nullity generating function than from the recurrence.
(2) Duality interchanges x and y, i.e.,

$$T_M(x,y) = T_{M^*}(y,x). \tag{4.2}$$

This can be deduced either from the Tutte recurrence (since duality interchanges deletion and contraction; see (3.8)) or from the corank-nullity generating function, by expressing r^* in terms of r.
(3) The Tutte recurrence implies that every coefficient of T_M is a *nonnegative* integer, a property which is not obvious from the closed formula (4.1).

4.1. **Recipes.** The results in this section describe how to "reverse-engineer" a general deletion/contraction recurrence for a graph or matroid isomorphism invariant to express it in terms of a specialization of the Tutte polynomial.

Theorem 4.10 (Tutte Recipe Theorem for Matroids)**.** *Let $u(M)$ be a matroid isomorphism invariant that satisfies a recurrence of the form*

$$u(M) = \begin{cases} 1 & \text{if } E=\emptyset,\\ Xu(M/e) & \text{if } e\in E \text{ is a coloop,}\\ Yu(M\setminus e) & \text{if } e\in E \text{ is a loop,}\\ au(M/e)+bu(M\setminus e) & \text{if } e\in E \text{ is ordinary}\end{cases}$$

where E denotes the ground set of M and X,Y,a,b are either indeterminates or numbers, with $a,b\neq 0$. Then

$$u(M) = a^{r(M)}b^{n(M)}T_M(X/a,Y/b).$$

Proof. Denote by $r(M)$ and $n(M)$ the rank and nullity of M, respectively. Note that

$$r(M) = r(M\setminus e) = r(M/e)+1 \qquad\text{and}\qquad n(M) = n(M\setminus e)+1 = n(M/e)$$

whenever deletion and contraction are well-defined. Define a new matroid invariant

$$\tilde{u}(M) = a^{-r(M)}b^{-n(M)}u(M)$$

and rewrite the recurrence in terms of $\tilde{u}$, abbreviating $r=r(M)$ and $n=n(M)$, to obtain

$$a^rb^n\tilde{u}(M^E) = \begin{cases} 1 & \text{if } E=\emptyset,\\ Xa^{r-1}b^n\tilde{u}(M/e) & \text{if } e\in E \text{ is a coloop,}\\ Ya^rb^{n-1}\tilde{u}(M\setminus e) & \text{if } e\in E \text{ is a loop,}\\ a^rb^nu(M/e)+a^rb^n\tilde{u}(M\setminus e) & \text{if } e\in E \text{ is ordinary.}\end{cases}$$

Setting $X = xa$ and $Y = yb$, we see that $\tilde{u}(M) = T_M(x, y) = T_M(X/a, Y/b)$ by Theorem 4.5, and rewriting in terms of $u(M)$ gives the desired formula. □

Bollobás [Bol98, p.340] gives an analogous result for graphs:

Theorem 4.11 (Tutte Recipe Theorem for Graphs)**.** *Let $u(G)$ be a graph isomorphism invariant that satisfies a recurrence of the form*

$$u(G) = \begin{cases} a^{|V|} & \text{if } E = \emptyset, \\ Xu(G \setminus e) & \text{if } e \in E \text{ is a coloop,} \\ Yu(G \setminus e) & \text{if } e \in E \text{ is a loop,} \\ bu(G \setminus e) + cu(G/e) & \text{if } e \in E \text{ is ordinary,} \end{cases}$$

where $G = (V, E)$ and X, Y, a, b, c are either indeterminates or numbers (with $b, c \neq 0$). Then

$$u(G) = a^{k(G)} b^{n(G)} c^{r(G)} T_G(aX/c, Y/b).$$

We omit the proof, which is similar to that of the previous result. A couple of minor complications are that many deletion/contraction graph invariants involve the numbers of vertices or components, which cannot be deduced from the matroid of a graph, and that deletion of a cut-edge makes sense in the context of graphs (no, that's not a misprint in the second case!). The invariant U is described by Bollobás as "the universal form of the Tutte polynomial."

4.2. **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

$$T_M(x, y) \ = \sum_{B \in \mathscr{B}(M)} x^{i(B)} y^{e(B)}. \tag{4.3}$$

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 $e_1 < \cdots < e_n$ on the ground set E, so they are not really invariants of B itself. However, another miracle occurs: the right-hand side of (4.3) does not depend on this choice of total order.

Definition 4.12. Let M be a matroid on E with basis system $\mathscr{B}$ and let $B \in \mathscr{B}$. For $e \in B$, the **fundamental cocircuit** of e with respect to B, denoted $C^*(e, B)$, is the unique cocircuit in $(E \setminus B) \cup e$. That is,

$$C^*(e, B) = \{e' \mid B \setminus e \cup e' \in \mathscr{B}\}.$$

Dually, for $e \notin B$, then the **fundamental circuit** of e with respect to B, denoted $C(e, B)$, is the unique circuit in $B \cup e$. That is,

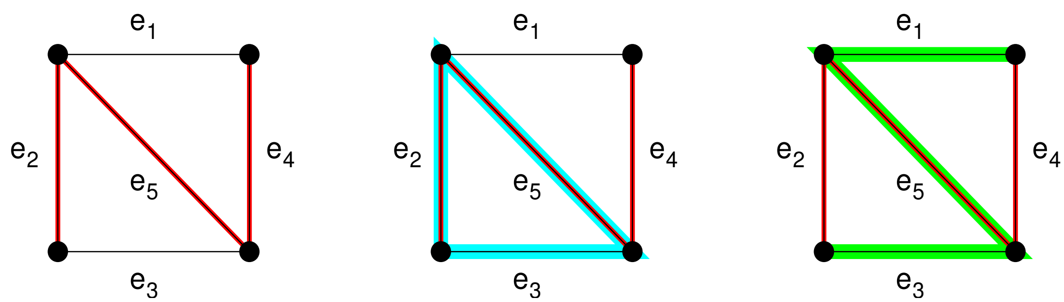
$$C(e, B) = \{e' \mid B \setminus e' \cup e \in \mathscr{B}\}.$$

Definition 4.13. Let M be a matroid on a totally ordered vertex set $E = \{e_1 < \cdots < e_n\}$, and let B be a basis of M. An element $e \in B$ is **internally active** with respect to B if e is the minimal element of $C^*(e, B)$. An element $e \notin B$ is **externally active** with respect to B if e is the minimal element of $C(e, B)$. We set

$$i(B) = \#\{e \in B \mid e \text{ is internally active with respect to } B\},$$
$$e(B) = \#\{e \in E \setminus B \mid e \text{ is externally active with respect to } B\}.$$

Note that these numbers depend on the choice of ordering of E.

Here is 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

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

Theorem 4.14. *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* (4.3) *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 boils down to showing that the generating function on the right-hand side of (4.3) satisfies the Tutte 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 generating function (4.3) is divisible by y (resp. x).

4.3. **The Chromatic Polynomial.**

Definition 4.15. 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.

Some important special cases:

- 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, the graph $G = \overline{K_n}$ with n vertices and no edges has chromatic polynomial k^n, since every coloring is proper.

- If T 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 $\hat{\mathbf{1}}$ 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 $p_T(k) = k(k-1)^{n-1}$.
- 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 4.16. *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

$$
\begin{aligned}
p_G(k) &= k^n && \text{if } E = \emptyset; && (4.4)\\
p_G(k) &= 0 && \text{if } G \text{ has a loop}; && (4.5)\\
p_G(k) &= (k-1)p_{G/e}(k) && \text{if } e \text{ is a coloop}; && (4.6)\\
p_G(k) &= p_{G\setminus e}(k) - p_{G/e}(k) && \text{otherwise.} && (4.7)
\end{aligned}
$$

We already know (4.4) and (4.5). Suppose $e = xy$ is not a loop. Let f be a proper k-coloring of $G \setminus e$. 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 (4.7) 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 $G \setminus e$, 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 (4.6).

The graph $G \setminus e$ 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 induction

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

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 an arbitrary (i.e., not necessarily graphic) matroid. Stay tuned!

4.4. **Acyclic Orientations.** An **orientation** $\mathcal{O}$ 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 $(x_0, x_1, \dots, x_{n-1})$ 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 $a(G) = |A(G)|$.

For example:

(1) If G has a loop then $a(G) = 0$.
(2) If G has no loops, then every total order on the vertices gives rise to an acyclic orientation: orient each edge from smaller to larger vertex. Of course, different total orders can produce the same a.o.
(3) If G has no edges than $a(G) = 1$. Otherwise, $a(G)$ is even, since reversing all edges is a fixed-point free involution on $A(G)$.
(4) Removing parallel copies of an edge does not change $a(G)$, since all parallel copies would have to be oriented in the same direction to avoid any 2-cycles.
(5) If G is a forest then every orientation is acyclic, so $a(G) = 2^{|E(G)}$.
(6) If $G = K_n$ then the acyclic orientations are in bijection with the total orderings, so $a(G) = n!$.
(7) If $G = C_n$ (the cycle of graph of length n) then it has 2^n orientations, of which exactly two are not acyclic, so $a(C_n) = 2^n - 2$.

Colorings and orientations are intimately connected. Given a proper coloring $f : V(G) \to [k]$, one can naturally define an acyclic orientation by directing each edge from the smaller to the larger color. (So #2 in the above list is a special case of this.) The connection between them is the prototypical example of what is called *combinatorial reciprocity.*

Theorem 4.17 (Stanley's Acyclic Orientation Theorem)**.** *Let $G = (V, E)$ be a graph with $|V| = n$. A **(compatible) k-pair** is a pair $(\mathcal{O}, f)$, where $\mathcal{O}$ is an acyclic orientation of G and $f : V \to [k]$ is a coloring such that for every directed edge $x \to y$ in D we have $f(x) \le f(y)$. Let $C(G, k)$ be the number of compatible k-pairs of G. Then*

$$(-1)^n p_G(-k) = |C(G, k)|. \tag{4.8}$$

Note that combining this equality with Theorem 4.16 gives

$$|C(G, k)| = (-1)^n(-1)^{n-c}(-k)^c\, T_G(1 - (-k), 0) = k^c T_G(1 + k, 0). \tag{4.9}$$

Proof. If G has no edges then $|C(G, k)| = k^n = (-1)^n(-k)^n = (-1)^n p_G(-k)$, confirming (4.8).

If G has a loop then it has no acyclic orientations, hence no k-pairs for any k, so both sides of (4.8) are zero.

Let $e = xy$ be an edge of G that is not a loop. Denote the left-hand side of (4.8) by $\bar{p}_G(k)$ Then

$$\begin{aligned}\bar{p}_G(k) = (-1)^n p_G(-k) &= (-1)^n (p_{G\setminus e}(k) - p_{G/e}(k))\\ &= (-1)^n((-1)^n \bar{p}_{G\setminus e}(-k) - (-1)^{n-1}\bar{p}_{G/e}(-k))\\ &= \bar{p}_{G\setminus e}(-k) + \bar{p}_{G/e}(-k)\end{aligned}$$

so all we have to show is that $|C(G, k)|$ satisfies the same recurrence. Write

$$C(G,k) \;=\; \underbrace{\{(\mathcal{O}, f) :\; f(x) \ne f(y)\}}_{C'} \cup \underbrace{\{(\mathcal{O}, f) :\; f(x) = f(y)\}}_{C''}.$$

Observe that it always possible to extend a k-pair of $G \setminus e$ to G by choosing an appropriate orientation for e. (If neither choice worked, then the orientation on $G \setminus e$ must contain a directed path from each of x, y to the other, and then it wouldn't be acyclic.) Call a pair *reversible* if either choice of orientation for e works, and *irreversible* otherwise, and write $C(G \setminus e, k) = C_{\mathrm{rev}} \cup C_{\mathrm{irr}}$.

For $(\mathcal{O}, f) \in C'$, we can simply delete e to obtain a k-pair of $G \setminus e$. The k-pairs we obtain are precisely the irreversible ones.

On the other hand, for $(\mathcal{O}, f) \in C''$, deleting e produces a reversible k-pair of $G \setminus e$, and contracting e produces a compatible k-pair of G/e (whose coloring is well-defined because $f(x) = f(y)$). Since e could,

in fact, be oriented in two ways, we obtain every reversible element of $C(G \setminus e, k)$, and every element of $C(G/e, k)$, twice. That is,
$$|C''| = 2|C_{\text{rev}}| = 2|C(G/e,k)|$$
but then
$$|C''| = |C_{\text{rev}}| + |C(G/e,k)|$$
and so
$$\begin{aligned} |C(G,k)| = |C'| + |C''| &= |C_{\text{irr}} + |C_{\text{rev}}| + |C(G/e,k)| \\ &= |C(G \setminus e,k)| + |C(G/e,k)| \end{aligned}$$
as desired. □

In particular, if $k = 1$ then there is only one choice for f and every acyclic orientation is compatible with it, which produces the following striking corollary:

Theorem 4.18. *The number of acyclic orientations of G is $|p_G(-1)| = T_G(2,0)$.*

Combinatorial reciprocity can be viewed geometrically. For more detail, look ahead to Section 6.5 and/or see a source such as Beck and Robins [BR07], but here is a brief taste.

Let G be a simple graph on n vertices. The **graphic arrangement** $\mathcal{A}_G$ is the union of all hyperplanes in $\mathbb{R}^n$ defined by the equations $x_i = x_j$ where ij is an edge of G. The complement $\mathbb{R}^n \setminus \mathcal{A}_G$ consists of finitely many disjoint open polyhedra (the "regions" of the arrangement), each of which is defined by a set of inequalities, including either $x_i < x_j$ or $x_i > x_j$ for each edge. Thus each region naturally gives rise to an orientation of G, and it is not hard to see that the regions are in fact in bijection with the acyclic orientations. Meanwhile, a k-coloring of G can be regarded as an integer point in the cube $[1,k]^n \subset \mathbb{R}^n$, and a proper coloring corresponds to a point that does not lie on any hyperplane in $\mathcal{A}_G$. In this setting, Stanley's theorem is an instance of something more general called *Ehrhart reciprocity* (which I will add notes on at some point).

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

Definition 4.19. A **linear code** $\mathscr{C}$ is a subspace of $(\mathbb{F}_q)^n$, 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., $\dim \mathscr{C}^\perp = n - \dim \mathscr{C}$. (Note, though, that a subspace and its orthogonal complement can intersect nontrivially. A space can even be its own orthogonal complement, such as $\{00, 11\} \subset \mathbb{F}_2^2$. This does not happen over $\mathbb{R}$, where the inner product is not only nondegenerate but also positive-definite, but "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 4.20 (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*

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

The proof is again a deletion-contraction argument and is omitted. As an example, if $\mathscr{C} = \{000, 101, 011, 110\} \subseteq \mathbb{F}_2^3$ as above, then the matroid M is $U_2(3)$. Its Tutte polynomial is $x^2 + x + y$, and Greene's theorem gives

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

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 $T_M(x, y) = T_{M^\perp}(y, x)$ by (4.2), so setting $s = (1-t)/(1+(q-1)t)$ (so $t = (1-s)/(1+(q-1)s)$; isn't that convenient?) gives

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

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.

4.6. **Exercises.**

Exercise 4.1. An orientation of a graph is called **totally cyclic** if every edge belongs to a directed cycle. Prove that the number of totally cyclic orientations of G is $T_G(0, 2)$.

Exercise 4.2. Let $G = (V, E)$ be a connected graph and A be an abelian group, written additively. An A-**flow** on G is a function $f : E \to A$ such that for every vertex v,

$$\sum_{e \in E(v)} f(e) = 0$$

where $E(v)$ means the set of edges incident to v. Prove that the number of everywhere nonzero A-flows is $(-1)^{|E|-|V|-1} T_G(0, 1 - |A|)$.

Exercise 4.3. Let $G = (V, E)$ be a graph with n vertices and c components. For a vertex coloring $f : V \to \mathbb{P}$, let $i(f)$ denote the number of "improper" edges, i.e., whose endpoints are assigned the same color. *Crapo's coboundary polynomial* of G is

$$\bar{\chi}_G(q; t) \;=\; q^{-c} \sum_{f : V \to [q]} t^{i(f)}.$$

This is evidently a stronger invariant than the chromatic polynomial of G, which can be obtained as $q\bar{\chi}_G(q, 0)$. In fact, the coboundary polynomial provides the same information as the Tutte polynomial. Prove that

$$\bar{\chi}_G(q; t) \;=\; (t-1)^{n-c} T_G\left(\frac{q+t-1}{t-1}, t\right)$$

by finding a deletion/contraction recurrence for the coboundary polynomial.

Exercise 4.4. Let M be a matroid on E. The **reliability polynomial** $R_M(p)$ of a matroid is the probability that the rank of M stays the same when each edge is independently deleted with probability p. In other words, suppose we have a lot of i.i.d. random variables X_e, one for each $e \in E$, each of which is zero with probability p and one with probability $1-p$. Let $A = \{e \in E \mid X_e = 1\}$. Then $R_M(p)$ is the probability that $r(A) = r(E)$. Obtain $R_G(p)$ as an evaluation of the Tutte polynomial.

Exercise 4.5. Express the h-vector of a matroid complex in terms of the Tutte polynomial of the underlying matroid. (Hint: First figure out a deletion/contraction recurrence for the h-vector, using Exercise 3.6.)

Exercise 4.6. Merino's theorem on critical configurations of the chip-firing game.

Exercise 4.7. Prove Theorem 4.14.

Exercise 4.8. Prove Theorem 4.20.

Much, much more about the Tutte polynomial can be found in [BO92], the MR review of which begins, "The reviewer, having once worked on that polynomial himself, is awed by this exposition of its present importance in combinatorial theory." (The reviewer was one W.T. Tutte.)

5. Poset Algebra

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

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

In this section, we will always assume that P is **locally finite**, i.e., every interval is finite.

Definition 5.1. The **incidence algebra** $I(P)$ is the set of functions $f : \mathrm{Int}(P) \to \mathbb{C}$ ("incidence functions"), made into a $\mathbb{C}$-vector space with pointwise addition, subtraction and scalar multiplication, and equipped with the *convolution product*

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

Here we abbreviate $f([x, y])$ by $f(x, y)$, and it is often convenient to set $f(x, y) = 0$ if $x \not\leq y$. Note that the assumption of local finiteness is both necessary and sufficient for convolution to be well-defined.

Proposition 5.2. *Convolution is associative (although it is not in general commutative).*

Proof. This is a straight-up calculation:

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

□

The multiplicative identity of $I(P)$ is just the Kronecker delta function, regarded as an incidence function:

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

Therefore, we sometimes write 1 for δ.

Proposition 5.3. *An incidence 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. In that case, the inverse is given by the formula*

$$f^{-1}(x, y) = \begin{cases} f(x, x)^{-1} & \text{if } x = y, \\ -f(y, y)^{-1} \sum_{z:\ x \leq z < y} g(x, z)f(z, y) & \text{if } x < y. \end{cases}$$

Proof. Let g be a left convolution inverse of f. In particular, $f(x, x) = g(x, x)^{-1}$ for all x, so it is necessary that $f(x, x) \neq 0$ for all x. On the other hand, if $x < y$, then

$$(g * f)(x, y) = \sum_{z:\ z \in [x,y]} g(x, z)f(z, y) = \delta(x, y) = 0$$

and solving for $g(x, y)$ gives the formula above, which is well-defined provided that $f(y, y) \neq 0$, so the condition is also sufficient. Similarly, the same condition jus necessary and sufficient for f to have a right convolution inverse h, and if Moreover, if $g * f = \delta = f * h$ then $g = g * f * h = h$ by associativity, so the left and right inverses coincide. □

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, namely 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

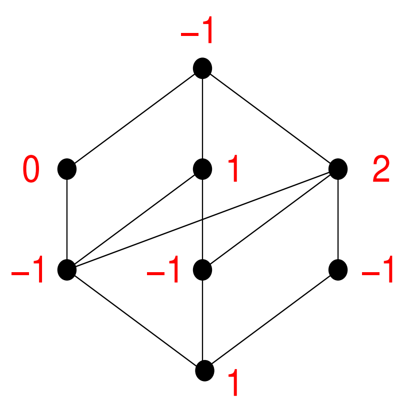
$$\eta^k(x,y) = \Big|\{x_1, \dots, x_{k-1} : \ x < x_1 < x_2 < \cdots < x_{k-1} < y\}\Big|,$$

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 $n \gg 0$.

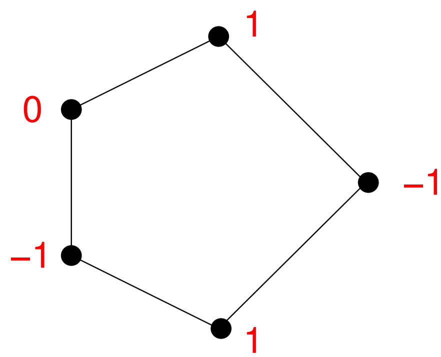
5.2. **The Möbius function.** The **Möbius function** μ_P of a poset P is defined as the convolution inverse of its zeta function: $\mu_P = \zeta_P^{-1}$. Proposition 5.3 provides a recursive formula for μ:

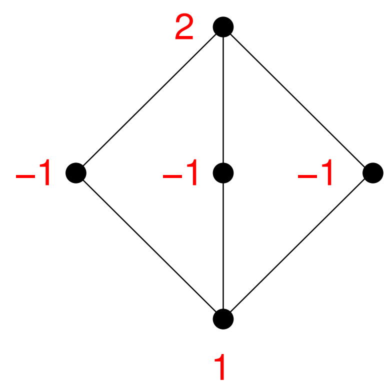
$$\begin{cases} \mu(x,x) = 1, \\ \mu(x,y) = -\sum_{z:\ x \leq z < y} \mu(x,z) & \text{if } x < y. \end{cases}$$

Example 5.4. In the diagram of the following poset P, the red numbers indicate $\mu_P(\hat{\mathbf{0}}, x)$.



Example 5.5. In these diagrams of the posets M_5 and N_5, the red numbers indicate $\mu_P(\hat{\mathbf{0}}, x)$.





Example 5.6. *The Möbius function of a boolean algebra.* Let $2^{[n]}$ be the boolean algebra of rank n and let $A \subseteq [n]$. Then $\mu(\hat{\mathbf{0}}, A) = (-1)^{|A|}$. To prove this, induct on $|A|$. The case $|A| = 0$ is clear. For $|A| > 0$, we have

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

More generally, if $B \subseteq A$, then $\mu(B, A) = (-1)^{|B \setminus A|}$, because every interval of $2^{[n]}$ 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

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

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

Example 5.7. *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

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

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.

The Möbius function is useful in many ways. It can be used to formulate a more general version of inclusion-exclusion called *Möbius inversion.* It behaves nicely under poset operations such as product, and has geometric and topological applications. Even just the single number $\mu_P(\hat{\mathbf{0}},\hat{\mathbf{1}})$ tells you a lot about a bounded poset P. Here is an application to counting chains.

Theorem 5.8 (Philip Hall's Theorem)**.** [Sta12, Prop. 3.8.5] *Let P be a chain-finite, bounded poset, and let*

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

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

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

Proof. Recall that $c_k=\eta^k(\hat{\mathbf{0}},\hat{\mathbf{1}})=(\zeta-\delta)^k(\hat{\mathbf{0}},\hat{\mathbf{1}})$. 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

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

Example 5.9. For the poset P of Example 5.4, we have $c_0=0$, $c_1=1$, $c_2=6$, $c_3=6$, and $c_k=0$ for $k>3$. So $c_0-c_1+c_2-c_3=-1=\mu_P(\hat{\mathbf{0}},\hat{\mathbf{1}})$.

This alternating sum looks like an Euler characteristic, and in fact it is. Define the **order complex** $\Delta(P)$ of a poset 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.) Then $c_k(P) = f_k(\Delta(P))$, so $\mu_{\hat{P}}(\hat{\mathbf{0}}, \hat{\mathbf{1}}) = \tilde{\chi}(\Delta(P))$, the reduced Euler characteristic (see Exercise 1.11).

5.3. Möbius inversion.

Theorem 5.10 (**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*

$$g(x) = \sum_{y\colon y \leq x} f(y) \qquad \forall x \in P \qquad \iff \qquad f(x) = \sum_{y\colon y \leq x} \mu(y,x) g(y) \qquad \forall x \in P, \tag{5.1a}$$

$$g(x) = \sum_{y\colon y \geq x} f(y) \qquad \forall x \in P \qquad \iff \qquad f(x) = \sum_{y\colon y \geq x} \mu(x,y) g(y) \qquad \forall x \in P. \tag{5.1b}$$

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

$$(f \cdot \alpha)(x) = \sum_{y\colon y \leq x} \alpha(y,x) f(y),$$
$$(\alpha \cdot f)(x) = \sum_{y\colon y \geq x} \alpha(x,y) f(y).$$

for $\alpha \in I(P)$. In terms of these actions, formulas (5.1a) and (5.1b) are respectively just the "trivial" observations

$$g = f \cdot \zeta \qquad \iff \qquad f = g \cdot \mu, \tag{5.2a}$$
$$g = \zeta \cdot f \qquad \iff \qquad f = \mu \cdot g. \tag{5.2b}$$

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 $2^{[n]}$, 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 5.11. 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

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

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 $I(2^{[n]})$, this is just $g = \zeta \cdot f$. Thus $f = \mu \cdot g$, or in terms of the Möbius inversion formula (5.1b),

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

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 5.12. 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

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

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 $g(W) = \sum_{U \supseteq W} f(U)$, 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 $0 = \hat{\mathbf{0}}$, we obtain

$$f(\hat{\mathbf{0}}) \;=\; \sum_{U \in L} \mu_L(\hat{\mathbf{0}}, U) x^{n - r(U)} \;=\; \#\{\text{1-1 linear maps } V \to X\} \;=\; (x-1)(x-q)(x-q^2)\cdots(x-q^{n-1}). \quad (5.3)$$

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 (5.3) 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

$$\mu_{L_n(q)}(\hat{\mathbf{0}}, \hat{\mathbf{1}}) = (-1)^n q^{\binom{n}{2}}$$

as predicted in Example 5.7.

5.4. **The Characteristic Polynomial.**

Definition 5.13. Let P be a finite graded poset with rank function r, and suppose that $r(\hat{\mathbf{1}}) = n$. The *characteristic polynomial* of P is defined as

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

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

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

Proof. Let $A \mapsto \bar{A}$ be the matroid closure operator of 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 $f(K) = \mu_L(\hat{\mathbf{0}},K)$. 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, (5.1a)) we have

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

On the other hand, $g(K)$ is the binomial expansion of $(1-1)^{|\bar{K}|}$, so the only nonzero summand is $J = \hat{\mathbf{0}}$. So the previous equation gives $f(K) = \mu(\hat{\mathbf{0}},K)$ as desired. $\square$

Example 5.15. Let G be a simple graph with n vertices and c components so that its graphic matroid $M = M(G)$ has rank $n-c$. Let $L = L(G)$ 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 5.14, we see that

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

Another major application of the characteristic polynomial is in the theory of hyperplane arrangements, which we'll get to later.

5.5. **Möbius Functions of Lattices.**

Theorem 5.16. *The Möbius function of a geometric lattice weakly alternates in sign. I.e., if L is a geometric lattice and r is its rank function, then $(-1)^{r(x)}\mu_L(\hat{\mathbf{0}}, x) \geq 0$ 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 5.14 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 $T_M \in \mathbb{N}[x, y]$. Meanwhile, every interval $[\hat{\mathbf{0}}, z] \subset L$ is a geometric lattice, so the sign of $\mu(\hat{\mathbf{0}}, z)$ is the same as that of $(-1)^{r(z)}$ (or zero). $\square$

In fact this theorem holds for any semimodular lattice — it is not necessary to assume that L is atomic. This can be proven algebraically using tools we're about to develop (Stanley, Prop. 3.10.1) or combinatorially, by interpreting $(-1)^{r(M)}\mu(L)$ 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 5.17. 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, $\hat{\mathbf{1}}$ is the multiplicative unit of $A(L)$.)

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 $L = 2^{[n]}$ 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 5.18. *The set*

$$\{\varepsilon_x = \sum_{y \leq x} \mu(y, x) y \mid x \in L\}$$

is a $\mathbb{C}$-vector space basis for $A(L)$, with $\varepsilon_x \varepsilon_y = \delta_{xy}\varepsilon_x$. In particular, $A(L) \cong \mathbb{C}^{|L|}$ as rings.

Proof. By Möbius inversion,

$$x = \sum_{y \leq x} \varepsilon_y. \tag{5.4}$$

which shows that the ε_x form a vector space basis. Let $\mathbb{C}_x$ be a copy of $\mathbb{C}$ with unit 1_x, so that we can identify $\mathbb{C}^{|L|}$ with $\prod_{x \in L} \mathbb{C}_x$. This is the direct product of rings, with multiplication $1_x 1_y = \delta_{xy} 1_x$. We claim that the $\mathbb{C}$-linear map $\phi : A(L) \to \mathbb{C}^{|L|}$ given by

$$\phi(\varepsilon_x) = 1_x$$

is a ring isomorphism. It is certainly a vector space isomorphism, and by (5.4) we have

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

Computations in the Möbius algebra will show that we can compute $\mu(x,y)$ more easily by summing over a cleverly chosen *subset* of $[x,y]$. Of course we know that $\mu(P) = -\sum_{x\neq\hat{\mathbf{1}}} -\mu(\hat{\mathbf{0}},x) = 0$ for any poset P, but this leads to a recursive computation that can be quite inefficient. The special structure of a lattice L leads to much more streamlined expressions for $\mu(L)$. The first result in this vein is known as *Weisner's theorem.*

Proposition 5.19. *Let L be a finite lattice with at least two elements. Then for every $a \in L \setminus \{\hat{\mathbf{1}}\}$ we have*

$$\sum_{\substack{x\in L:\\ x\wedge a=\hat{\mathbf{0}}}} \mu(x,\hat{\mathbf{1}}) = 0$$

and in particular,

$$\mu(L) \;=\; \mu_L(\hat{\mathbf{0}},\hat{\mathbf{1}}) \;=\; -\sum_{\substack{x\in L\setminus\{\hat{\mathbf{0}}\}:\\ x\wedge a=\hat{\mathbf{0}}}} \mu(x,\hat{\mathbf{1}}). \tag{5.5}$$

Proof. We work in $A(L)$ and calculate $a\varepsilon_{\hat{\mathbf{1}}}$ in two ways. On the one hand

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

On the other hand

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

Now taking the coefficient of $\hat{\mathbf{0}}$ on both sides gives the desired equation. $\square$

Example 5.20 (**The Möbius function of the partition lattice** Π_n**).** Let $a = 1|23\cdots n \in \Pi_n$. Then the partitions x that show up in the sum of (5.5) are just the atoms whose non-singleton block is $\{1,i\}$ for some $i>1$. For each such x, the interval $[x,\hat{\mathbf{1}}]\subset\Pi_n$ is isomorphic to Π_{n-1}, so (5.5) 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 5.21 (**The 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 $[X,\hat{\mathbf{1}}]\subset L$ is isomorphic to $L_{n-1}(q)$. Therefore

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

and by induction

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

A drawback of Weisner's theorem is that it is still recursive; the right-hand side of (5.5) involves other values of the Möbius function. This is not a problem for integer-indexed families of lattices $\{L_n\}$ such that every rank-k element $x\in L_n$ has $[\hat{\mathbf{0}},x]\cong L_k$ (as we have just seen), but this is too much to hope for in general. The next result, Rota's *crosscut theorem*, gives a non-recursive way of computing the Möbius function.

Definition 5.22. Let L be a lattice. An **upper crosscut** of L is a set $X\subset L\setminus\{\hat{\mathbf{1}}\}$ such that if $y\in L\setminus X\setminus\{\hat{\mathbf{1}}\}$, then $y<x$ for some $x\in X$. A **lower crosscut** of L is a set $X\subset L\setminus\{\hat{\mathbf{0}}\}$ such that if $y\in L\setminus X\setminus\{\hat{\mathbf{0}}\}$, then $y>x$ for some $x\in X$.

It would be simpler to define an upper (resp., lower) crosscut as a set that contains all coatoms (resp., atoms), but in practice the formulation in the previous definition is typically a convenient way to show that a particular set is a crosscut.

Theorem 5.23 (Rota's crosscut theorem)**.** *Let L be a finite lattice and let X be an upper crosscut. Then*

$$\mu(L) = \sum_{Y \subseteq X:\ \bigwedge Y = \hat{\mathbf{0}}} (-1)^{|Y|}. \tag{5.6a}$$

Dually, if X is a lower crosscut, then

$$\mu(L) = \sum_{Y \subseteq X:\ \bigvee Y = \hat{\mathbf{1}}} (-1)^{|Y|}. \tag{5.6b}$$

Proof. The two statements are dual, so it suffices to prove (5.6a). For any $x \in L$, we have the following simple equation in the Möbius algebra of L:

$$\hat{\mathbf{1}} - x \ = \ \sum_{y \in L} \varepsilon_y \ - \ \sum_{y \leq x} \varepsilon_y \ = \ \sum_{y \not\leq x} \varepsilon_y.$$

Therefore, for any $X \subseteq L$, we have

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

where $Y = \{y \in L \mid y \not\leq x \text{ for all } x \in X\}$. (Expand the sum and recall that $\varepsilon_y \varepsilon_{y'} = \delta_{yy'} \varepsilon_y$.) But if X is an upper crosscut, then $Y = \{\hat{\mathbf{1}}\}$, and this last equation becomes

$$\prod_{x \in X} (\hat{\mathbf{1}} - x) \ = \ \varepsilon_{\hat{\mathbf{1}}} \ = \ \sum_{y \in L} \mu(y, \hat{\mathbf{1}}) y. \tag{5.7}$$

On the other hand, a direct binomial expansion gives

$$\prod_{x \in X} (\hat{\mathbf{1}} - x) \ = \ \sum_{A \subseteq X} (-1)^{|A|} \bigwedge A. \tag{5.8}$$

Now equating the coefficients of $\hat{\mathbf{0}}$ on the right-hand sides of (5.7) and (5.8) yields (5.6a). The proof of (5.6b) is similar. $\square$

Corollary 5.24. *Let L be a lattice in which $\hat{\mathbf{1}}$ is not a join of atoms (for example, a distributive lattice that is not Boolean). Then $\mu(L) = 0$.*

Another consequence is the promised strengthening of Theorem 5.16.

Corollary 5.25. *The Möbius function of any semimodular lattice L weakly alternates in sign. That is, $(-1)^{r(x)} \mu(\hat{\mathbf{0}}, x) \geq 0$ for all $x \in L$.*

Proof. Let L' be the join-sublattice of L generated by atoms. Then L' is geometric so $(-1)^{r(x)} \mu(x) \geq 0$ for $x \in L$. On the other hand, if $x \in L \setminus L'$, then applying Corollary 5.24 to $[\hat{\mathbf{0}}, x]$ gives $\mu(x) = 0$. $\square$

The crosscut theorem will be useful in studying hyperplane arrangements. Another topological application is the following result due to J. Folkman (1966), whose proof (omitted) uses the crosscut theorem.

Theorem 5.26. *Let L be a geometric lattice of rank r, and let $P = L \setminus \{\hat{\mathbf{0}}, \hat{\mathbf{1}}\}$. 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 $r-2$.

5.6. **Exercises.**

Exercise 5.1. Let P be a chain-finite poset. The *kappa function* of P is the element of the incidence algebra $I(P)$ defined by $\kappa(x,y)=1$ if $x \lessdot y$, $\kappa(x,y)=0$ otherwise.

(a) Give a condition on the convolution powers of κ that is equivalent to P being ranked.
(b) Give combinatorial interpretations of $\kappa * \zeta$ and $\zeta * \kappa$.

Exercise 5.2. Let Π_n be the lattice of set partitions of $[n]$. Recall that the order relation on Π_n is given as follows: if $\pi,\sigma \in \Pi_n$, then $\pi \leq \sigma$ if every block of π is contained in some block of σ (for short, "π refines σ"). In this problem, you're going to calculate the number $\mu_n := \mu_{\Pi_n}(\hat{\mathbf{0}},\hat{\mathbf{1}})$.

(a) Calculate μ_n by brute force for $n=1,2,3,4$. Make a conjecture about the value of μ_n in general.
(b) Define a function $f:\Pi_n \to \mathbb{Q}[x]$ as follows: if X is a finite set of cardinality x, then
$$f(\pi) = \#\big\{h:[n]\to X \;\;\big|\;\; h(s)=h(s') \iff s,s' \text{ belong to the same block of } \pi\big\}.$$
For example, if $\pi = \hat{\mathbf{1}} = \{\{1,2,\dots,n\}\}$ is the one-block partition, then $f(\pi)$ counts the constant functions from $[n]$ to X, so $f(\pi)=x$. Find a formula for $f(\pi)$ in general.
(c) Let $g(\pi)=\sum_{\sigma\geq\pi} f(\sigma)$. Prove that $g(\pi)=x^{|\pi|}$ for all $\pi\in\Pi_n$. (Hint: What kinds of functions are counted by the sum?)
(d) Apply Möbius inversion and an appropriate substitution for x to calculate μ_n.

Exercise 5.3. This problem is about how far Proposition 5.18 can be extended. Suppose that R is a commutative $\mathbb{C}$-algebra of finite dimension n as a $\mathbb{C}$-vector space, and that $x_1,\dots,x_n \in R$ are linearly independent idempotents (i.e., $x_i^2 = x_i$ for all i), but we do not know how to write each x_ix_j as a linear combination $\sum_k c_{ij}^k a_k$. (The scalars c_{ij}^k are the *structure constants* of the algebra.) Does it necessarily 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 $e_ie_j=0$ for $i\neq j$? If so, there should be an explicit construction of the e_i's through linear algebra. If not, there should be some condition on the c_{ij}^k's that says when the rings are isomorphic. (I don't know the answer to this question; someone ought to supply it.)

Exercise 5.4. The q**-binomial coefficient** is the rational function
$$\begin{bmatrix} n \\ k \end{bmatrix}_q \;=\; \frac{(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})}{(q^k-1)(q^k-q)\cdots(q^k-q^{k-1})}.$$

(a) Check that setting $q=1$ (after canceling out common terms), or equivalently applying $\lim_{q\to 1}$, recovers the ordinary binomial coefficient $\binom{n}{k}$.
(b) Prove the *q-Pascal identities*:
$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^k \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \qquad \text{and} \qquad \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k}\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q.$$
Deduce that $\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]_q$ is actually a polynomial in q (not merely a rational function).
(c) (Stanley, EC1, 2nd ed., 3.119) Prove the *q-binomial theorem*:
$$\prod_{k=0}^{n-1}(x-q^k) \;=\; \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q (-1)^k q^{\binom{k}{2}} x^{n-k}.$$
(Hint: Let $V=\mathbb{F}_q^n$ and let X be a vector space over $\mathbb{F}_q$ with x elements. Count the number of one-to-one linear transformations $V\to X$ in two ways.)

Exercise 5.5. (Stanley, EC1, 3.129) Here is a cute application of combinatorics to elementary number theory. Let P be a finite poset, and let μ be the Möbius function of $\hat{P} = P \cup \{\hat{\mathbf{0}},\hat{\mathbf{1}}\}$. Suppose that P has a fixed-point-free automorphism $\sigma: P\to P$ of prime order p; that is, $\sigma(x)\neq x$ and $\sigma^p(x)=x$ for all $x\in P$. Prove that $\mu_{\hat{P}}(\hat{\mathbf{0}},\hat{\mathbf{1}}) \cong -1 \pmod{p}$. What does this say in the case that $\hat{P} = \Pi_p$?

6. Hyperplane Arrangements

A canonical source for the combinatorial theory of hyperplane arrangements is Stanley's book chapter [Sta07].

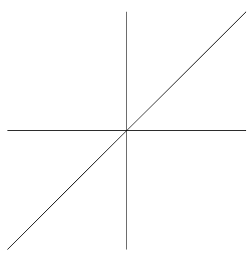
6.1. Basic definitions.

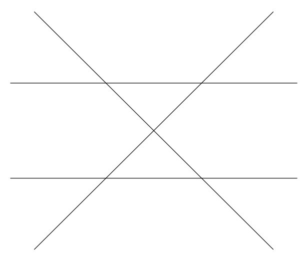
Definition 6.1. Let $\mathbb{F}$ be a field, typically either $\mathbb{R}$ or $\mathbb{C}$, and let $n \geq 1$. A **linear hyperplane** in $\mathbb{F}^n$ is a vector subspace of codimension 1. An **affine hyperplane** is a translate of a linear hyperplane. A **hyperplane arrangement** $\mathcal{A}$ is a finite collection of (distinct) hyperplanes. The number n is called the **dimension** of $\mathcal{A}$, and the space $\mathbb{F}^n$ is its **ambient space**.

We often abuse notation and use $\mathcal{A}$ for the set $\bigcup_{H\in\mathcal{A}} H$. In particular, we are frequently interested in the complement $\mathbb{F}^n \setminus \mathcal{A} := \mathbb{F}^n \setminus \bigcup_{H\in\mathcal{A}} H$.

Throughout, the symbol $\mathbf{v}_H$ means a nonzero normal vector to H. Such a normal is determined up to nonzero scalar multiples.

Example 6.2. The left-hand arrangement $\mathcal{A}_1 \subset \mathbb{R}^2$ 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 $\ell_1, \ell_2, \ell_3, \ell_4$ given by the equations $x = y$, $x = -y$, $y = 1$, $y = -1$ respectively





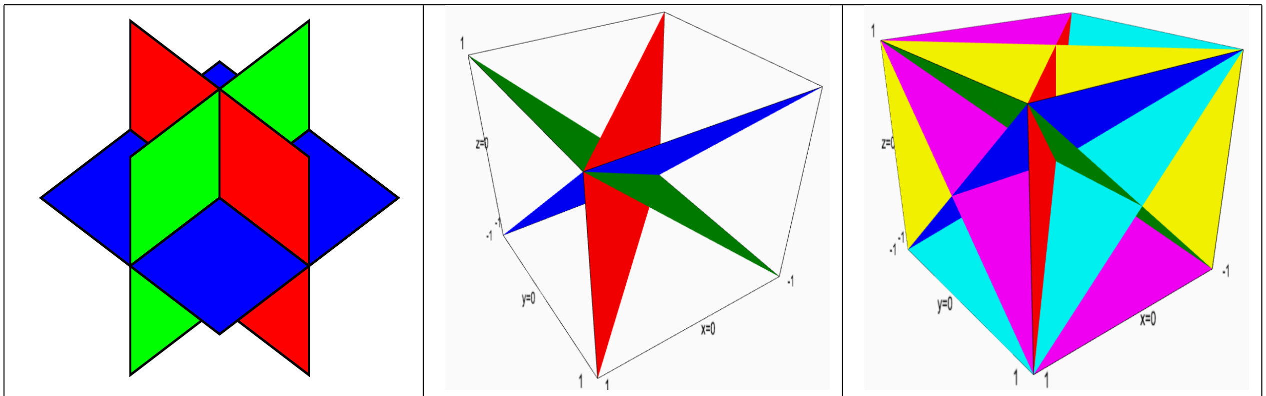
Each hyperplane is the zero set of some linear form, so their union is the zero set of the product of those s linear forms. We can specify an arrangement concisely by that product, called the **defining polynomial** of $\mathcal{A}$ (which is well-defined up to a scalar multiple). For example, the defining polynomials of $\mathcal{A}_1$ and $\mathcal{A}_2$ are $xy(x - y)$ and $xy(y - 1)(y + 1)$ respectively.

Example 6.3. The *Boolean arrangement* Bool_n consists of the coordinate hyperplanes in n-space. Its defining polynomial is $x_1x_2 \dots x_n$.

The *braid arrangement* Br_n consists of the $\binom{n}{2}$ hyperplanes $x_i = x_j$ in n-space. (Its defining polynomial $\prod_{1\leq i<j\leq n}(x_i - x_j)$ is known as the **Vandermonde determinant**.)

If $G = (V, E)$ is a simple graph on vertex set $V = [n]$, then the corresponding *graphic arrangement* $\mathcal{A}_G$ is the subarrangement of Br_n consisting of those hyperplanes $x_i = x_j$ for which $ij \in E$. Thus Br_n itself is the graphic arrangement of the complete graph K_n.

Here are some pictures. Note that every hyperplane in Br_n contains the line $x_1 = x_2 = \cdots = x_n$, so projecting $\mathbb{R}^4$ along that line allows us to picture Br_4 as an arrangement $\mathrm{ess}(\mathrm{Br}_4)$ in $\mathbb{R}^3$ ("ess" means *essentialization*, to be explained soon).

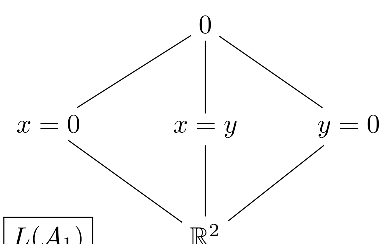


Bool_3 | Br_3 | $\mathrm{ess}(\mathrm{Br}_4)$

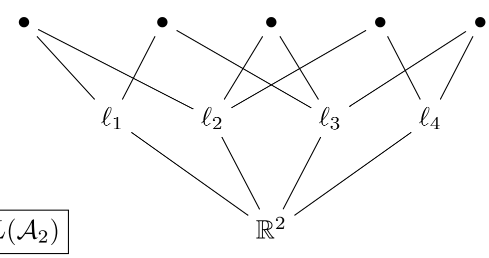
The second two figures were produced using the computer algebra system `Sage` [S+14].

Definition 6.4. 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.

For example, the intersection posets of the 2-dimensional arrangements $\mathcal{A}_1$ and $\mathcal{A}_2$ of Example 6.2 are as follows.



$L(\mathcal{A}_1)$



$L(\mathcal{A}_2)$

The poset $L(\mathcal{A})$ is the fundamental combinatorial invariant of $\mathcal{A}$. Some easy observations:

(1) If $T : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear transformation, then $L(T(\mathcal{A})) \cong L(\mathcal{A})$, where $T(\mathcal{A}) = \{T(H) \mid H \in \mathcal{A}\}$. In fact, the intersection poset is invariant under any *affine* transformation. (The group of affine transformations is generated by the invertible linear transformations together with translations.)

(2) The poset $L(\mathcal{A})$ is a meet-semilattice, with meet given by

$$\left(\bigcap_{H \in \mathcal{B}} H\right) \wedge \left(\bigcap_{H \in \mathcal{C}} H\right) = \bigcap_{H \in \mathcal{B} \cap \mathcal{C}} H$$

for all $\mathcal{B}, \mathcal{C} \subseteq \mathcal{A}$. Its $\hat{\mathbf{0}}$ element is the empty intersection, which by convention is $\mathbb{F}^n$.

(3) $L(\mathcal{A})$ is ranked, with rank function $r(X) = n - \dim X$. To see this, observe that each covering relation $X \lessdot Y$ comes from intersecting an affine linear subspace X with a hyperplane H that neither contains nor is disjoint from X, so that $\dim(X \cap H) = \dim X - 1$.

(4) $L(\mathcal{A})$ has a $\hat{\mathbf{1}}$ element if and only if $\bigcap_{H \in \mathcal{A}} H \neq \emptyset$. Such an arrangement is called **central**. In this case $L(\mathcal{A})$ is a lattice (and is referred to as the **intersection lattice** of $\mathcal{A}$). Since translation does not affect whether an arrangement is central (or indeed any of its combinatorial structure), we will typically assume that $\mathrm{Cent}(\mathcal{A}) := \bigcap_{H \in \mathcal{A}} H$ contains the zero vector, which is to say that every hyperplane in $\mathcal{A}$ is a *linear* hyperplane in $\mathbb{F}^n$. (So an arrangement is central if and only if it is a translation of an arrangement of linear hyperplanes.)

(5) When $\mathcal{A}$ is central, the lattice $L(\mathcal{A})$ is in fact geometric. It is atomic by definition, and it is submodular because it is a sublattice of the chain-finite modular lattice $L(\mathbb{F}^n)^*$ (the lattice of all subspaces of $\mathbb{F}^n$ ordered by reverse inclusion). The associated matroid $M(\mathcal{A}) = M(L(\mathcal{A}))$ is represented over $\mathbb{F}$ by the normal vectors $\mathbf{v}_H$ to the hyperplanes $H \in \mathcal{A}$. (Note that any normals will do, since the matroid is unchanged by scaling the $\mathbf{v}_H$ independently.)

Therefore, all of the tools we have developed for looking at posets, lattices and matroids can be applied to study hyperplane arrangements.

The dimension of an arrangement cannot be extracted from its intersection poset. If Br_4 were a "genuine" 4-dimensional arrangement then we would not be able to represent it in $\mathbb{R}^3$. The reason we could do so is that the hyperplanes of Br_4 intersected in a positive-dimensional space that we could effectively "get rid of". This observation motivates the following definition.

Definition 6.5. Let $\mathcal{A} \subset \mathbb{F}^n$ be an arrangement and let

$$N(\mathcal{A}) = \mathbb{F}\langle \mathbf{v}_H \mid H \in \mathcal{A}\rangle.$$

The **essentialization** of $\mathcal{A}$ is the arrangement

$$\operatorname{ess}(\mathcal{A}) = \{H \cap N(\mathcal{A}) \mid H \in \mathcal{A}\} \;\subset\; N(\mathcal{A}).$$

If $N(\mathcal{A}) = \mathbb{F}^n$ then we say that $\mathcal{A}$ is **essential**. So $\operatorname{ess}(\mathcal{A})$ is always essential, and $L(\operatorname{ess}(\mathcal{A})) \cong L(\mathcal{A})$ as posets. The **rank** of $\mathcal{A}$ is the dimension of its essentialization.

For example, $\operatorname{rank}\mathrm{Br}_n = \dim\operatorname{ess}(\mathrm{Br}_n) = n-1$, and $\operatorname{rank}\mathcal{A}_G = r(G) = |V(G)| - c$, where c is the number of connected components of G.

If $\mathcal{A}$ is linear, then we could define the essentialization by setting $V = N(\mathcal{A})^\perp = \operatorname{Cent}(\mathcal{A})$ and define $\operatorname{ess}(\mathcal{A}) = \{H/V \mid H \in \mathcal{A}\} \;\subset\; \mathbb{F}^n/V$. Thus $\mathcal{A}$ is essential if and only if $\operatorname{Cent}(\mathcal{A}) = 0$. Moreover, if $\mathcal{A}$ is linear then $\operatorname{rank}(\mathcal{A})$ is the rank of its intersection lattice — so rank is a combinatorial invariant, unlike dimension.

Definition 6.6. Let $\mathcal{A}$ be an arrangement. Its **characteristic polynomial** is defined as

$$\chi_{\mathcal{A}}(k) = \sum_{x \in L(\mathcal{A})} \mu(\hat{\mathbf{0}}, x) k^{\dim x} = k^{\dim\mathcal{A} - \operatorname{rank}\mathcal{A}} \chi_{L(\mathcal{A})}(k). \tag{6.1}$$

The power of k in front is a correction factor that allows us to keep track of dimension as well as rank. Of course, the correction factor $k^{\dim\mathcal{A} - \operatorname{rank}\mathcal{A}}$ vanishes when $\mathcal{A}$ is essential.

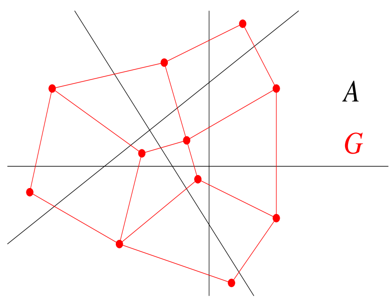
6.2. **Counting Regions of Hyperplane Arrangements: Examples.** Let $\mathcal{A} \subset \mathbb{R}^n$ be a real hyperplane arrangement. The **regions** of $\mathcal{A}$ are the connected components of $\mathbb{R}^n \setminus \mathcal{A}$. Each component is the interior of a (bounded or unbounded) polyhedron; in particular, it is homeomorphic to $\mathbb{R}^n$. We call a region **relatively bounded** if the corresponding region in $\operatorname{ess}(\mathcal{A})$ is bounded. (Note that if $\mathcal{A}$ is not essential then every region is unbounded.) Let

$$\begin{aligned} r(\mathcal{A}) &= \text{ number of regions of } \mathcal{A}, \\ b(\mathcal{A}) &= \text{ number of relatively bounded regions of } \mathcal{A}. \end{aligned}$$

Example 6.7. 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 6.8. The Boolean arrangement Bool_n consists of the n coordinate hyperplanes in $\mathbb{R}^n$. It is a central, essential arrangement whose intersection lattice is the Boolean lattice of rank n; accordingly, $\chi_{\mathrm{Bool}_n}(k) = (k-1)^n$. The complement $\mathbb{R}^n \setminus \mathrm{Bool}_n$ 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, $r(\mathrm{Bool}_n) = 2^n$ and $b(\mathrm{Bool}_n) = 0$.

Example 6.9. 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, 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

$$v = r, \tag{6.2a}$$

$$f = 1 + \binom{m}{2} = \frac{m^2 - m + 2}{2} \tag{6.2b}$$

$$4(f-1) = 2e - (r - b). \tag{6.2c}$$

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 (6.2c) and (6.2b) we obtain

$$e = m + 2(f-1) = m^2. \tag{6.2d}$$

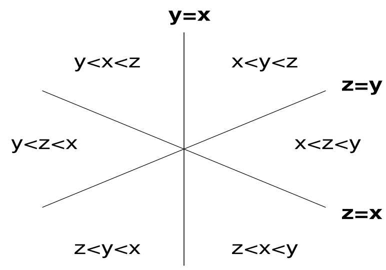
Euler's formula for planar graphs says that $v - e + f = 2$. Substituting in (6.2a), (6.2b) and (6.2d) and solving for r gives

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

and therefore

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

Example 6.10. The braid arrangement Br_n consists of the $\binom{n}{2}$ hyperplanes $H_{ij} = \{\mathbf{x} \mid x_i = x_j\}$ in $\mathbb{R}^n$. The complement $\mathbb{R}^n \setminus \mathrm{Br}_n$ 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, $r(\mathrm{Br}_n) = n!$. (Stanley: "Rarely is it so easy to compute the number of regions!") Furthermore,

$$\chi_{\mathrm{Br}_n}(k) = k(k-1)(k-2)\cdots(k-n+1).$$

Note that the braid arrangement is central but not essential; its center is the line $x_1 = x_2 = \cdots = x_n$, so its rank is $n-1$.

Example 6.11. Let $G = (V, E)$ be a simple graph with $V = [n]$. The **graphic arrangement** $\mathcal{A}_G \subset \mathbb{R}^n$ is the subarrangement of Br_n containing the hyperplanes H_{ij}. Thus $L(\mathcal{A}_G)$ is a geometric lattice, and the corresponding matroid is just the graphic matroid of G, since that matroid is represented by the vectors $\{\mathbf{e}_i - \mathbf{e}_j \mid ij \in E\}$, which are normal to the hyperplanes H_{ij}. In particular, the characteristic polynomial of $L(\mathcal{A}_G)$ is precisely the chromatic polynomial of G (see Section 4.3). We will see another explanation for this fact later; see Example 6.22.

The regions of $\mathbb{R}^n \setminus \mathcal{A}_G$ are the open polyhedra whose defining inequalities include either $x_i < x_j$ or $x_i > x_j$ for each edge $ij \in E$. Those inequalities give rise to an orientation of G, and it is not hard to check that this correspondence is a bijection between regions and acyclic orientations. Hence

$$r(\mathcal{A}_G) \;=\; \text{number of acyclic orientations of } G \;=\; |\chi_{L(\mathcal{A}_G)}(-1)|.$$

6.3. **Zaslavsky's Theorem.** This last example motivates the main result of this section, which was historically the first major theorem about hyperplane arrangements.

Theorem 6.12 (Zaslavsky 1975)**.** *Let $\mathcal{A}$ be a real hyperplane arrangement, and let $\chi_{\mathcal{A}}$ be the characteristic polynomial of its intersection poset. Then*

$$r(\mathcal{A}) = (-1)^{\dim \mathcal{A}} \chi_{\mathcal{A}}(-1) \;\; \textit{and} \tag{6.3}$$

$$b(\mathcal{A}) = (-1)^{\operatorname{rank} \mathcal{A}} \chi_{\mathcal{A}}(1). \tag{6.4}$$

The proof will take some time, and will combine both geometric and combinatorial techniques.

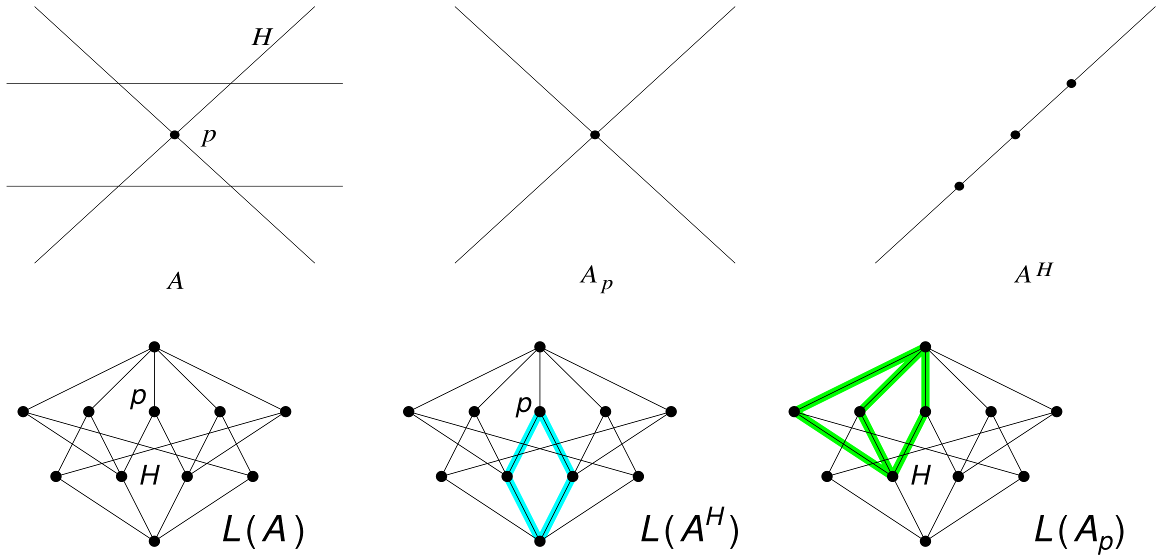
Let $x \in L(\mathcal{A})$; recall that this means that x is an affine space formed by some intersection of the hyperplanes in $\mathcal{A}$. Define

$$\begin{aligned} \mathcal{A}_x &= \{H \in \mathcal{A} \mid H \supseteq x\}, \\ \mathcal{A}^x &= \{W \mid W = H \cap x,\ H \in \mathcal{A} \setminus \mathcal{A}_x\}. \end{aligned} \tag{6.5a}$$

Thus $\mathcal{A}_x$ is an arrangement with the same ambient space as $\mathcal{A}$, and $\mathcal{A}^x$ is an arrangement with ambient space x. The reason for this notation is that the posets $L(\mathcal{A}_x)$ and $L(\mathcal{A}^x)$ are isomorphic respectively to the principal order ideal and principal order filter generated by x in $L(\mathcal{A})$. I.e.,

$$L(\mathcal{A}_x) = \{y \in L(\mathcal{A}) \mid y \leq x\}, \qquad L(\mathcal{A}^x) = \{y \in L(\mathcal{A}) \mid y \geq x\}.$$

Example 6.13. 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.



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.

In what follows, we abbreviate $\mathcal{A} \setminus \{H\}$ by $\mathcal{A} \setminus H$.

Proposition 6.14. *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*

$$r(\mathcal{A}) \;=\; r(\mathcal{A}') + r(\mathcal{A}'') \tag{6.6}$$

and

$$b(\mathcal{A}) \;=\; \begin{cases} b(\mathcal{A}') + b(\mathcal{A}'') & \text{if } \operatorname{rank}\mathcal{A} = \operatorname{rank}\mathcal{A}', \\ 0 & \text{if } \operatorname{rank}\mathcal{A} = \operatorname{rank}\mathcal{A}' + 1. \end{cases} \tag{6.7}$$

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 numbers 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 $\mathcal{A}''$ (see, e.g., Example 6.13). So

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

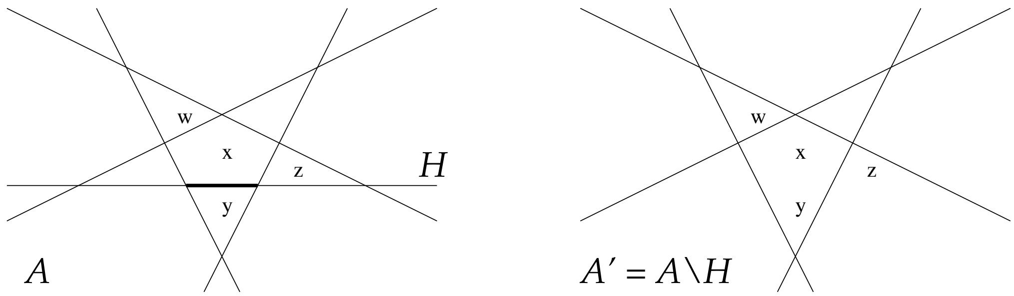
proving (6.6). By the way, if (and only if) H is a coloop then it borders *every* region of $\mathcal{A}$, so $r(\mathcal{A}) = 2r(\mathcal{A}')$ in this case.

Now we count bounded regions. If $\operatorname{rank}\mathcal{A} = \operatorname{rank}\mathcal{A}' + 1$, then in particular $N(\mathcal{A}') \neq \mathbb{R}^n$, and in this case every region of $\mathcal{A}'$ must contain a line orthogonal to $N(\mathcal{A}')$. Therefore every region of $\mathcal{A}$ contains a ray, and $b(\mathcal{A}) = 0$. Otherwise, the bounded regions of $\mathcal{A}$ come in a few different flavors.

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

Second, for each bounded 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). (E.g., for region x below, $\bar{R} \cap H$ is the line segment separating regions x and y, indicated in boldface.) 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 x, y, which together contribute two to each side of (6.7); no in the case of z, which contributes one to each side of (6.7).)



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 $b(\mathcal{A}) = b(\mathcal{A}') + b(\mathcal{A}'')$ as desired. $\square$

Proposition 6.14 looks a lot like a Tutte polynomial deletion/contraction recurrence. This suggests that we should be able to interpret $r(\mathcal{A})$ and $b(\mathcal{A})$ in terms of the characteristic polynomial $\chi_{\mathcal{A}}$. The first step is to find a more convenient form for the characteristic polynomial, using the crosscut theorem.

Proposition 6.15 (Whitney's formula)**.** *For any real hyperplane arrangement $\mathcal{A}$, we have*

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

Proof. Consider the interval $[\hat{\mathbf{0}}, x]$. The atoms in this interval are the hyperplanes of $\mathcal{A}$ containing x, and they form a lower crosscut of $[\hat{\mathbf{0}}, x]$. Therefore, Rota's crosscut theorem (Thm. 5.23) says that

$$\mu(\hat{\mathbf{0}}, x) \;=\; \sum_{\mathcal{B}\subseteq\mathcal{A}:\ \bigcap\mathcal{B}=x} (-1)^{|\mathcal{B}|}. \tag{6.8}$$

Plugging (6.8) into the definition of the characteristic polynomial, we get

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

as desired. Note that $\mathcal{B} = \emptyset$ should be considered a central subarrangement for the purpose of this formula, corresponding to the summand $x = \hat{\mathbf{0}}$ and giving rise to the leading term $k^{\dim \mathcal{A}}$ of $\chi_{\mathcal{A}}(k)$. $\square$

Proposition 6.16 (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*

$$\chi_{\mathcal{A}}(k) \;=\; \chi_{\mathcal{A}'}(k) - \chi_{\mathcal{A}''}(k). \tag{6.9}$$

Sketch of proof. Splitting up Whitney's formula gives

$$\chi_{\mathcal{A}}(k) \;=\; \underbrace{\sum_{\text{central } \mathcal{B}\subseteq\mathcal{A}:\ H\notin\mathcal{B}} (-1)^{|\mathcal{B}|} k^{n-\operatorname{rank}\mathcal{B}}}_{\Sigma_1} \;+\; \underbrace{\sum_{\text{central } \mathcal{B}\subseteq\mathcal{A}:\ H\in\mathcal{B}} (-1)^{|\mathcal{B}|} k^{n-\operatorname{rank}\mathcal{B}}}_{\Sigma_2}.$$

We have $\Sigma_1 = \chi_{\mathcal{A}'}(k)$ by definition, so we need only show that $\Sigma_2 = -\chi_{\mathcal{A}''}(k)$. This is a little trickier. Every summand $\mathcal{B}$ of Σ_2 gives rise to a central subarrangement $\mathcal{B}''$ of $\mathcal{A}''$, defined by

$$\mathcal{B}'' = \{H' \cap H \mid H' \in \mathcal{B}\},$$

but since different hyperplanes in $\mathcal{A}$ can have the same intersections with H, different subarrangements of $\mathcal{A}$ can induce the same subarrangement of $\mathcal{A}''$. So we need to be careful about the bookkeeping. Let $K_1, \dots, K_s$ be the distinct spaces of the form $H' \cap H$ for $H' \in \mathcal{A} \setminus H$, and let

$$\mathcal{A}_i = \{H' \in \mathcal{A} \setminus H \mid H' \cap H = K_i\},$$

so that $\{\mathcal{A}_1, \dots, \mathcal{A}_s\}$ is a set partition of $\mathcal{A} \setminus H$ into nonempty blocks.

For each summand $\mathcal{B}$ of Σ_2, define

$$\operatorname{Ind}(\mathcal{B}'') = \operatorname{Ind}(\mathcal{B}) = \{j \in [s] \mid \mathcal{A}_j \cap \mathcal{B} \neq \emptyset\},$$

which we call the *index set* of $\mathcal{B}$. Then two subsets of $\mathcal{A} \setminus H$ give rise to the same central subarrangement of $\mathcal{A}''$ if and only if they have the same index set; moreover, $|\operatorname{Ind}(\mathcal{B})| = |\mathcal{B}''|$. Note that $\operatorname{Cent}(\mathcal{B}'') = \operatorname{Cent}(\mathcal{B})$ and so

$$n - \operatorname{rank}\mathcal{B} \;=\; \dim(\cap\mathcal{B}) \;=\; \dim \bigcap_{j \in \operatorname{Ind}(\mathcal{B})} K_j \;=\; \dim \operatorname{Cent}(\mathcal{B}') \;=\; \dim H - \operatorname{rank}\mathcal{B}''. \tag{6.10}$$

We now calculate Σ_2 by breaking up its summands by their index sets. Note that the arrangements $\mathcal{B}$ with index set J are precisely of the form

$$\mathcal{B} = \{H\} \cup \bigcup_{j \in H} \mathcal{C}_j$$

where each $\mathcal{C}_j$ is a nonempty subset of $\mathcal{A}_j$. Therefore,

$$\begin{aligned}
\Sigma_2 &= \sum_{J\colon\, \emptyset \neq J \subseteq [s]} \; \sum_{\mathcal{B}\colon\, \operatorname{Ind}(\mathcal{B}) = J} (-1)^{|\mathcal{B}|} k^{n - \operatorname{rank}\mathcal{B}} \\
&= \sum_J k^{\dim H - \operatorname{rank}\mathcal{A}''(J)} \sum_{\mathcal{B}\colon\, \operatorname{Ind}(\mathcal{B}) = J} (-1)^{|\mathcal{B}|} && \text{(by (6.10))} \\
&= \sum_{\text{central } \mathcal{B}'' \subseteq \mathcal{A}'} k^{\dim H - \operatorname{rank}\mathcal{B}''} \sum_{\substack{(\mathcal{C}_j)_{j \in \operatorname{Ind}(\mathcal{B}'')}: \\ \emptyset \subsetneq \mathcal{C}_j \subseteq \mathcal{A}_j}} (-1)^{1 + \sum |\mathcal{C}_j|} \\
&= -\sum_{\mathcal{B}''} k^{\dim H - \operatorname{rank}\mathcal{B}''} \prod_{j \in \operatorname{Ind}(\mathcal{B}'')} \sum_{\mathcal{C}_j} (-1)^{|\mathcal{C}_j|} && \text{(``unexpanding'' the sum)} \\
&= -\sum_{\mathcal{B}''} (-1)^{|\mathcal{B}''|} k^{\dim H - \operatorname{rank}\mathcal{B}''}
\end{aligned}$$

(because the jth sum is just the binomial expansion of $(1-1)^{|\mathcal{C}_j|} = 0$, with one $+1$ term removed)

$$\begin{aligned}
&= \sum_{\mathcal{B}''} (-1)^{|\mathcal{B}''|} k^{\dim H - \operatorname{rank}\mathcal{B}''} \\
&= -\chi_{\mathcal{A}''}(k). && \text{(by Whitney's formula).} \qquad \square
\end{aligned}$$

Remark 6.17. This recurrence is strongly reminiscent of the chromatic recurrence (4.7). Indeed, if $\mathcal{A} = \mathcal{A}_G$ is a graphic arrangement in $\mathbb{R}^n$, e is an edge of G, and H_e is the corresponding hyperplane in $\mathcal{A}_G$, then it is clear that $\mathcal{A}_{G \setminus e} = \mathcal{A}_G \setminus \{H_e\}$. In addition, two hyperplanes $H_f, H_{f'}$ will have the same intersection with H_e if and only if f, f' become parallel upon contracting e, so $\mathcal{A}_{G/e}$ can be identified with $(\mathcal{A}_G)^{H_e}$ (where the coordinates on $H_e \cong \mathbb{R}^{n-1}$ are given by equating the coordinates for the two endpoints of e).

We now restate and prove Zaslavsky's theorem.

Theorem 6.12. Let $\mathcal{A} \subset \mathbb{R}^n$ be a real hyperplane arrangement, and let $\chi_{\mathcal{A}}$ be the characteristic polynomial of its intersection poset. Then $r(\mathcal{A}) = \tilde{r}(\mathcal{A})$ and $b(\mathcal{A}) = \tilde{b}(\mathcal{A})$, where

$$\tilde{r}(\mathcal{A}) := (-1)^{\dim \mathcal{A}} \chi_{\mathcal{A}}(-1), \qquad \tilde{b}(\mathcal{A}) = (-1)^{\operatorname{rank} \mathcal{A}} \chi_{\mathcal{A}}(1).$$

Proof. We have already done the hard work; this is just putting the pieces together.

If $|\mathcal{A}| = 1$, then $L(\mathcal{A})$ is the lattice with two elements, namely $\mathbb{R}^n$ and a single hyperplane H, and its characteristic polynomial is $k^n - k^{n-1}$. Thus $\tilde{r}(\mathcal{A}) = (-1)^n((-1)^n - (-1)^{n-1}) = 2$ and $\tilde{b}(\mathcal{A}) = -(1-1) = 0$, which match $r(\mathcal{A})$ and $b(\mathcal{A})$.

For the general case, we just need to show that $\tilde{r}$ and $\tilde{b}$ satisfy the same recurrences as r and b (see Prop. 6.14). First,

$$\begin{aligned}
\tilde{r}(\mathcal{A}) &= (-1)^{\dim \mathcal{A}} \chi_{\mathcal{A}}(-1) \\
&= (-1)^{\dim \mathcal{A}} (\chi_{\mathcal{A}'}(-1) - \chi_{\mathcal{A}''}(-1)) \\
&= (-1)^{\dim \mathcal{A}'} \chi_{\mathcal{A}'}(-1) + (-1)^{\dim \mathcal{A}''} \chi_{\mathcal{A}''}(-1) \\
&= \tilde{r}(\mathcal{A}') + \tilde{r}(\mathcal{A}'').
\end{aligned}$$

As for $\tilde{b}$, if $\operatorname{rank} \mathcal{A} = \operatorname{rank} \mathcal{A}' + 1$, then in fact $\mathcal{A}'$ and $\mathcal{A}''$ have the same essentialization, hence the same rank, and their characteristic polynomials only differ by a factor of k. The deletion/restriction recurrence (Prop. 6.16) therefore implies $\tilde{b}(\mathcal{A}) = 0$.

On the other hand, if $\operatorname{rank} \mathcal{A} = \operatorname{rank} \mathcal{A}'$, then $\operatorname{rank} \mathcal{A}'' = \operatorname{rank} \mathcal{A} - 1$ and a calculation similar to that for $\tilde{r}$ (replacing dimension with rank) shows that $\tilde{b}(\mathcal{A}) = \tilde{b}(\mathcal{A}') + \tilde{b}(\mathcal{A}'')$. $\square$

Corollary 6.18. *Let $\mathcal{A} \subset \mathbb{R}^n$ be a central hyperplane arrangement with at least one hyperplane, so that $L(\mathcal{A})$ is a nontrivial 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 6.19. Let $m \geq n$, and let $\mathcal{A}$ be an arrangement of m linear hyperplanes in general position in $\mathbb{R}^n$. "General position" means precisely that the corresponding matroid M is $U_n(m)$, whose rank function

$r : 2^{[m]} \to \mathbb{N}$ is given by $r(A) = \min(n, |A|)$. 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

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

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 $\mathbb{R}^2$. The calculation of $r(\mathcal{A})$ and $b(\mathcal{A})$ for the arrangement of m affine hyperplanes in general position in $\mathbb{R}^n$ is left as an exercise.

6.4. **The finite field method.** Let $\mathbb{F}_q$ be the finite field of order q, and let $\mathcal{A} \subset \mathbb{F}_q^n$ be a hyperplane arrangement. The "regions" of $\mathbb{F}_q^n \setminus \mathcal{A}$ are just its points (assuming, if you wish, that we endow $\mathbb{F}_q^n$ 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 6.20. $|\mathbb{F}_q^n \setminus \mathcal{A}| = \chi_{\mathcal{A}}(q)$.

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. 6.15). $\square$

This fact has a much more general application, which was systematically mined by Athanasiadis, e.g., [Ath96]. Let $\mathcal{A} \subset \mathbb{R}^n$ be an arrangement of hyperplanes whose defining equations have integer coefficients. For a prime p, let $\mathcal{A}_p = \mathcal{A} \otimes \mathbb{F}_p$ be the arrangement in $\mathbb{F}_p^n$ defined by regarding the equations in $\mathcal{A}$ as lying over $\mathbb{F}_p$. If p is sufficiently large, then in fact $L(\mathcal{A}_p) \cong L(\mathcal{A})$ (in this case we say that $\mathcal{A}$ **reduces correctly modulo** p). Specifically, if M is the matrix of normal vectors, then it suffices to choose p larger than any minor of M, so that a set of columns of M is linearly independent over $\mathbb{F}_p$ iff it is independent over $\mathbb{Q}$. Since there are infinitely many such primes, the following highly useful result follows:

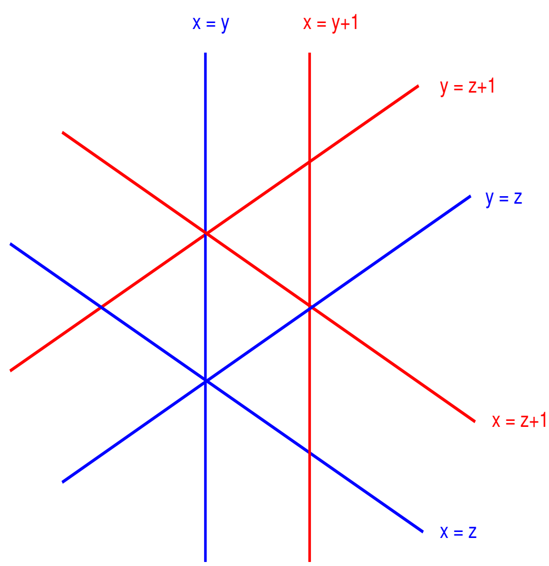
Corollary 6.21. *Let $\mathcal{A} \subset \mathbb{R}^n$ be a hyperplane arrangement defined over the integers. Then $\chi_{\mathcal{A}}(p)$ is the polynomial that counts points in the complement of $\mathcal{A}_p$, for large enough primes p.*

Example 6.22. Let $G = ([n], E)$ be a simple graph and let $\mathcal{A}_G$ be the corresponding graphic arrangement in $\mathbb{R}^n$. Note that $\mathcal{A}_G$ reduces correctly over every finite field $\mathbb{F}_q$ (because graphic matroids are regular). A point $(x_1, \ldots, x_n) \in \mathbb{F}_q^n$ can be regarded as the q-coloring of G that assigns color x_i to vertex i. The proper q-colorings are precisely the points of $\mathbb{F}_q^n \setminus \mathcal{A}_G$. The number of such colorings is $p_G(k)$ (the chromatic polynomial of G evaluated at q). On the other hand, by Proposition 6.20, it is also the characteristic polynomial $\chi_{\mathcal{A}_G}(q)$. Since $p_G(k) = \chi_{\mathcal{A}_G}(q)$ for infinitely many q (namely, all integer prime powers), the polynomials must be equal.

Example 6.23. The *Shi arrangement* is the arrangement of $n(n-1)$ hyperplanes in $\mathbb{R}^n$ defined by

$$\mathrm{Shi}_n = \{x_i = x_j,\ x_i = x_j + 1 \mid 1 \leq i < j \leq n\}.$$

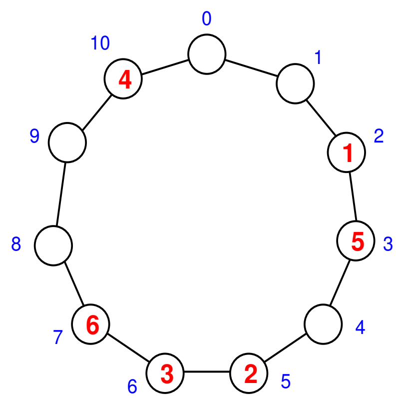
In other words, take the braid arrangement, clone it, and nudge each of the cloned hyperplanes a little bit in the direction of the bigger coordinate. The Shi arrangement has rank $n-1$ (every hyperplane in it contains a line parallel to the all-ones vector), so we may project along that line to obtain the essentialization in $\mathbb{R}^{n-1}$. Thus $\mathrm{ess}(\mathrm{Shi}_2)$ consists of two points on a line, while $\mathrm{ess}(\mathrm{Shi}_3)$ is shown below.



Proposition 6.24. *The characteristic polynomial of the Shi arrangement is $\chi_{\mathrm{Shi}_n}(q) = q(q-n)^{n-1}$. In particular, the numbers of regions and bounded regions are respectively $r(\mathrm{Shi}_n) = |\chi(-1)| = (n+1)^{n-1}$ and $b(\mathrm{Shi}_n) = |\chi(1)| = (n-1)^{n-1}$.*

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

Proof. It suffices to count the points in $\mathbb{F}_q^n \setminus \mathrm{Shi}_n$ for a large enough prime q. Let $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{F}_q^n \setminus \mathrm{Shi}_n$. 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 $\boldsymbol{i}$ 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 clockwise, starting at **1** and putting in a divider sign | for each bead without a red number, we get

$$15 \mid 236 \mid \mid 4 \mid$$

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 are n choices for the bead containing the red **1** goes, we obtain $\left|\mathbb{F}_q^n \setminus \mathrm{Shi}_n\right| = q(q-n)^{n-1}$ as desired. □

6.5. **Supersolvable Lattices and Arrangements.** We have seen that for a simple graph $G = ([n], E)$, the chromatic polynomial $p_G(k)$ is precisely the characteristic polynomial of the graphic arrangement $\mathcal{A}_G$. (This is another reason for the correction factor $k^{\dim \mathcal{A} - \operatorname{rank} \mathcal{A}}$ in (6.1).) For some graphs, the chromatic polynomial factors into linear terms. If $G = K_n$, then $p_G(k) = k(k-1)(k-2)\cdots(k-n+1)$, and if G is a forest with n vertices and c components, then $p_G(k) = k^c(k-1)^{n-c}$. This property does not hold for all graphs.

Example 6.25. 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:

$$\chi_{C_4}(k) \ = \ k(k-1)(k^2 - 3k + k).$$

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.*

Let L be a lattice. Recall from (2.3) 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 2.28).

Definition 6.26. 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 $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ 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 $y \wedge x = \hat{\mathbf{0}}$ 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 $r(x \vee y) = r(\hat{\mathbf{1}}) = 3$ and $r(x \wedge y) = r(\hat{\mathbf{0}}) = 0$. So x is not a modular element.

Proposition 6.27. *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

$$\pi \wedge \sigma = \Big\{C \cap B \mid C \in X\Big\} \cup \Big\{\{i\} \mid i \notin B\Big\}, \qquad \pi \vee \sigma = \left\{\bigcup_{C \in X} C\right\} \cup Y$$

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, suppose B, C are nonsingleton blocks of π, with $i, j \in B$ and $k, \ell \in C$. Let σ be the partition with exactly two nonsingleton blocks $\{i,k\}, \{j,\ell\}$. Then $r(\sigma) = 2$ and $r(\pi \wedge \sigma) = r(\hat{\mathbf{0}}) = 0$, but

$$r(\pi \vee \sigma) \;=\; r(\pi) + 1 \;<\; r(\pi) + r(\sigma) - r(\pi \wedge \sigma)$$

so π is not a modular element. ∎

The usefulness of a modular element is that if one exists, we can factor the characteristic polynomial of L.

Theorem 6.28. *Let L be a geometric lattice of rank n, and let $z \in L$ be a modular element. Then*

$$\chi_L(k) \;=\; \chi_{[\hat{\mathbf{0}},z]}(k) \sum_{y:\; y \wedge z = \hat{\mathbf{0}}} \mu_L(\hat{\mathbf{0}}, y) k^{n - r(z) - r(y)}. \tag{6.11}$$

Here is a sketch of the proof; for the full details, see [Sta07, pp. 440–441]. We work in the dual Möbius algebra $A^*(L) = A(L^*)$; that is, the vector space of $\mathbb{C}$-linear combinations of elements of L, with multiplication given by join (rather than meet as in §5.5). Thus the "algebraic" basis of $A^*(L)$ is

$$\{\sigma_y \overset{\text{def}}{\equiv} \sum_{x:\; x \geq y} \mu(y,x) x \mid y \in L\}.$$

First, one shows by direct calculation that

$$\sigma_{\hat{\mathbf{0}}} \;=\; \sum_{x \in L} \mu(x) x \;=\; \left(\sum_{v:\; v \leq z} \mu(v) v\right) \left(\sum_{y:\; y \wedge z = \hat{\mathbf{0}}} \mu(y) y\right) \tag{6.12}$$

for any $z \in L$. Second, for $z, y, v \in L$ such that z is modular, $v \leq z$, and $y \wedge z = 0$, one shows first that $z \wedge (v \vee y) = v$ (by rank considerations) and then that $\operatorname{rank}(v \vee y) = \operatorname{rank}(v) + \operatorname{rank}(y)$. Third, make the substitutions $v \mapsto k^{\operatorname{rank} z - \operatorname{rank} v}$ and $y \mapsto k^{n - \operatorname{rank} y - \operatorname{rank} z}$ in the two sums on the RHS of (6.12). Since

$vy = v \vee y$, the last observation implies that substituting $x \mapsto k^{n-\operatorname{rank} x}$ on the LHS preserves the product, and the equation becomes (6.11).

In particular, every atom a is modular, so

$$\chi_L(k) \;=\; (k-1) \sum_{x:\; x \not\geq a} \mu_L(\hat{\mathbf{0}}, x) k^{r(L)-1-r(x)}.$$

This does not really tell us anything new, because we already knew that $k-1$ had to be a factor of $\chi_L(k)$, because $\chi_L(1) = \sum_{x \in L} \mu_L(\hat{\mathbf{0}}, x) = 0$. Also, the sum in the expression is not the characteristic polynomial of a lattice. On the other hand, a modular *coatom* lets us peel off a linear factor:

Corollary 6.29. *Let L be a geometric lattice, and let $z \in L$ be a coatom that is a modular element. Then*

$$\chi_L(k) \;=\; (k-e)\chi_{[\hat{\mathbf{0}},z]}(k),$$

where e is the number of atoms $a \in L$ such that $a \not\leq z$.

If we are extremely lucky, then L will have a saturated chain of modular elements

$$\hat{\mathbf{0}} = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_{n-1} \lessdot x_n = \hat{\mathbf{1}}.$$

In this case, we can apply Corollary 6.29 successively with $z = x_{n-1}$, $z = x_{n-2}$, $\ldots$, $z = x_1$ to split the characteristic polynomial completely into linear factors:

$$\begin{aligned}
\chi_L(k) \;&=\; (k-e_{n-1})\chi_{[\hat{\mathbf{0}},x_{n-1}]}(k)\\
&=\; (k-e_{n-1})(k-e_{n-2})\chi_{[\hat{\mathbf{0}},x_{n-2}]}(k)\\
&=\; \ldots\\
&=\; (k-e_{n-1})(k-e_{n-2})\cdots(k-e_0),
\end{aligned}$$

where

$$\begin{aligned}
e_i \;&=\; \#\{\text{atoms } a \text{ of } [\hat{\mathbf{0}}, x_{i+1}] \mid a \not\leq x_i\}\\
&=\; \#\{a \in A \mid a \leq x_{i+1},\; a \not\leq x_i\}.
\end{aligned}$$

Definition 6.30. A geometric lattice L is **supersolvable** if it has a *modular saturated chain*, that is, a saturated chain $\hat{\mathbf{0}} = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_n = \hat{\mathbf{1}}$ such that every x_i is a modular element. A central hyperplane arrangement $\mathcal{A}$ is called supersolvable if $L(\mathcal{A})$ is supersolvable.

Example 6.31. Every modular lattice is supersolvable, because every saturated chain is modular.

Example 6.32. The partition lattice Π_n (and therefore the associated hyperplane arrangement Br_n) is supersolvable by induction. Let z be the coatom with blocks $[n-1]$ and $\{n\}$, which is a modular element by Proposition 6.27. 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

$$\chi_{\Pi_n}(k) \;=\; (k-1)(k-2)\cdots(k-n+1).$$

Example 6.33. 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 $r(x \wedge y) = 0$. (We have already seen that the characteristic polynomial of L does not split.)

Theorem 6.34. *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 $v_1, \dots, v_n$ 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.

Such an ordering is called a **perfect elimination ordering**. The proof of Theorem 6.34 is left as an exercise (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. This equivalence is sometimes known as *Dirac's theorem*. It is fairly easy to prove that supersolvable graphs are chordal, but the converse is somewhat harder; see, e.g., [Wes96, pp. 224–226]. There are other graph-theoretic formulations of this property; see, e.g., [Dir61].

If G satisfies the condition of Theorem 6.34, then we can see directly why its 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

$$\chi(G; k) \ = \ \prod_{i=1}^{n} (k - |C_i|).$$

6.6. **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 a connected topological space, but not simply connected. Thus instead of counting regions, we should count holes, as expressed by the homology groups. Brieskorn [Bri73] solved this problem completely:

Theorem 6.35 (Brieskorn [Bri73])**.** *The homology groups $H_i(X, \mathbb{Z})$ are free abelian, and the Poincáre polynomial of X is the characteristic polynomial backwards:*

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

In a very famous paper, Orlik and Solomon [OS80] 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 [Ryb11] constructed arrangements with isomorphic lattices of flats but different fundamental groups. There is much more to say on this topic!

6.7. **Exercises.**

Exercise 6.1. Let $m > n$, and let $\mathcal{A}$ be the arrangement of m affine hyperplanes in general position in $\mathbb{R}^n$. Calculate $\chi_{\mathcal{A}}(k)$, $r(\mathcal{A})$, and $b(\mathcal{A})$.

Exercise 6.2. (Stanley, HA, 2.5) Let G be a graph on n vertices, let $\mathcal{A}_G$ be its graphic arrangement in $\mathbb{R}^n$, and let $\mathcal{B}_G = \operatorname{Bool}_n \cup \mathcal{A}_G$. (That is, $\mathcal{B}$ consists of the coordinate hyperplanes $x_i = 0$ in $\mathbb{R}^n$ together with the hyperplanes $x_i = x_j$ for all edges ij of G.) Calculate $\chi_{\mathcal{B}_G}(q)$ in terms of $\chi_{\mathcal{A}_G}(q)$.

Exercise 6.3. (Stanley, EC2, 3.115(d)) The *type B braid arrangement* $\mathcal{O}_n \subseteq \mathbb{R}^n$, is the arrangement with two hyperplanes $x_i = x_j$, $x_i = -x_j$ for every pair i, j with $1 \leq i < j \leq n$ (so $n(n-1)$ hyperplanes in all). Calculate the characteristic polynomial and the number of regions of $\mathcal{O}_n$.

Exercise 6.4. Recall that each permutation $w = (w_1, \ldots, w_n) \in \mathfrak{S}_n$ corresponds to a region of the braid arrangement Br_n, namely the open cone $C_w = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_{w_1} < x_{w_2} < \cdots < x_{w_n}\}$. Denote its closure by $\overline{C_w}$. For any set $W \subseteq \mathfrak{S}_n$, consider the closed fan

$$F(W) = \bigcup_{w \in W} \overline{C_w} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_{w_1} \leq \cdots \leq x_{w_n} \ \text{ for some } w \in W\}.$$

Prove that $F(W)$ is a convex set if and only if W is the set of linear extensions of some poset P on $[n]$. (A *linear extension* of P is a total ordering $\prec$ consistent with the ordering of P, i.e., if $x <_P y$ then $x \prec y$.)

Exercise 6.5. The runners in a race are seeded $1, \ldots, n$ (stronger runners are assigned higher numbers). To even the playing field, the rules specify that you earn one point for each higher-ranked opponent you beat, and one point for each lower-ranked opponent you beat *by at least one second*. (If a higher-ranked runner beats a lower-ranked runner by less than 1 second, no one gets a the point for that matchup.) Let s_i be the number of points scored by the i^{th} player and let $s = (s_1, \ldots, s_n)$ be the *score vector*.

(a) Show that the possible score vectors are in bijection with the regions of the Shi arrangement.
(b) Work out all possible score vectors in the cases of 2 and 3 players. Conjecture a necessary and sufficient condition for $(s_1, \ldots, s_n)$ to be a possible score vector for n players. Prove it if you can.

Exercise 6.6. Prove Theorem 6.34.

7. More Topics

7.1. Ehrhart Theory (contributed by Margaret Bayer).

How many integer or rational points are in a convex polytope?

Definition 7.1. A polytope $P \subseteq \mathbb{R}^N$ is **integral** (resp. **rational**) if and only if all vertices of P have integer (resp. rational) coordinates.

For a set $P \subseteq \mathbb{R}^N$ and a positive integer n, let $nP = \{nx : x \in P\}$. (nP is called a **dilation** of P.)

The **(relative) boundary** of P, written ∂P, is the union of proper faces of P, that is, the set of points $x \in P$ such that for every $\varepsilon > 0$, the ball of radius ε (its intersection with $\mathrm{aff}(P)$) contains both points of P and points not in P. The **(relative) interior** of P, $\operatorname{int} P$, is $P \setminus \partial P$.

For a polytope $P \subseteq \mathbb{R}^N$ define sequences

$$\begin{aligned} i(P,n) &= |nP \cap \mathbb{Z}^N| \\ i^*(P,n) &= |n(\operatorname{int} P) \cap \mathbb{Z}^N| \end{aligned}$$

$i(P,n)$ is the number of integer points in nP or, equivalently, the number of rational points in P of the form $\left(\frac{a_0}{n}, \frac{a_1}{n}, \ldots, \frac{a_N}{n}\right)$. Our goal is to understand the functions $i(P,n)$ and $i^*(P,n)$.

We start with P a simplex, and with an easy example. Let

$$P = \operatorname{conv}\{(0,0,0), (1,1,0), (1,0,1), (0,1,1)\} \in \mathbb{R}^3.$$

Then

$$nP = \operatorname{conv}\{(0,0,0), (n,n,0), (n,0,n), (0,n,n)\}.$$

Each point in nP can be written as $\beta_1(n,n,0) + \beta_2(n,0,n) + \beta_3(0,n,n) + \beta_4(0,0,0)$, with $0 \leq \beta_i \leq 1$ and $\sum \beta_i = 1$; or, alternatively, as $\alpha_1(1,1,0) + \alpha_2(1,0,1) + \alpha_3(0,1,1)$, with $0 \leq \alpha_i \leq n$ and $\sum \alpha_i \leq n$.

Case 1. If the α_i are all integers, the resulting points are integer points and the sum of the coordinates is even. How many such points are there? It equals the number of monomials in four variables of degree n, that is, $\binom{n+3}{3}$. However there are other integer points in nP.

Case 2. We can allow the fractional part of α_i to be 1/2. If any one of the α_i has fractional part 1/2, the others must also. Writing $\gamma_i = \alpha_i - 1/2$, we get points of the form

$$(\gamma_1 + 1/2)(1,1,0) + (\gamma_2 + 1/2)(1,0,1) + (\gamma_3 + 1/2)(0,1,1) = \\ \gamma_1(1,1,0) + \gamma_2(1,0,1) + \gamma_3(0,1,1) + (1,1,1).$$

Note here that $\sum \gamma_i = (\sum \alpha_i) - 3/2 \leq n - 3/2$. Since the γ_i are integers, $\sum \gamma_i \leq n-2$. So the number of these points equals the number of monomials in four variables of degree $n-2$, that is, $\binom{n+1}{3}$.

Adding these we get

$$i(P,n) = \binom{n+3}{3} + \binom{n+1}{3} = \frac{1}{3}n^3 + n^2 + \frac{5}{3}n + 1.$$

Note, in particular, that this is a polynomial in n.

And what about the number of integer points in the interior of P?

Note that all the points in Case 2 are interior points because each $\alpha_i = \gamma_i + 1/2 > 0$ and their sum is at most $n - 2 + 3/2$ (less than n). A point in Case 1 is an interior point if and only if all the $\alpha_i > 0$ and $\sum \alpha_i < n$. The four-tuples $(\alpha_1 - 1, \alpha_2 - 1, \alpha_3 - 1, n - 1 - \sum \alpha_i)$ correspond to monomials in four variables of degre $n-4$; there are $\binom{n-1}{3}$ of them. Thus we get

$$i^*(P,n) = \binom{n+1}{3} + \binom{n-1}{3} = \frac{1}{3}n^3 - n^2 + \frac{5}{3}n - 1,$$

another polynomial. (Anything else you notice? Is it a coincidence?)

It is convenient to visualize the dilations nP of P in a cone. For $P \subseteq \mathbb{R}^N$ an integral N-simplex, let $\tilde{P} = \{(x,1) \in \mathbb{R}^{N+1} : x \in P\}$, and let C be the simplicial cone generated by $\tilde{P}$:

$$C = C(\tilde{P}) = \{ry : y \in \tilde{P},\, r \in \mathbb{R},\, r \geq 0\}.$$

The boundary and interior of C are $\partial C = \{ry : y \in \partial\tilde{P}\}$ and $\operatorname{int} C = C \setminus \partial C$. Then the polytope nP can be identified with a cross-section of C:

$$C \cap \{(z,n) \in \mathbb{R}^{N+1} : z \in \mathbb{R}^N\} = \{(z,n) \in \mathbb{R}^{N+1} : z \in nP\}.$$

The integer point functions are then

$$\begin{aligned} i(P,n) &= |C \cap \{(z,n) \in \mathbb{R}^{N+1} : z \in \mathbb{Z}^N\}| \\ i^*(P,n) &= |\operatorname{int} C \cap \{(z,n) \in \mathbb{R}^{N+1} : z \in \mathbb{Z}^N\}|. \end{aligned}$$

We can represent all points in the cone in terms of the vertices of P.

Proposition 7.2. *Let P be a rational N-simplex in $\mathbb{R}^N$, with vertices $v_0, v_1, \ldots, v_N$, and let $C = C(\tilde{P})$. A point $z \in \mathbb{R}^{N+1}$ is a rational point in C if and only if $z = \sum_{i=0}^N c_i(v_i,1)$ for some nonnegative rational numbers c_i. Furthermore, this representation of z is unique.*

A slightly different representation is more useful. Let

$$Q = \{\sum_{i=0}^N r_i(v_i,1) : \text{for each } i,\, 0 \leq r_i < 1\}.$$

Q is a half-open parallelipiped containing 0 and $\tilde{P}$.

Proposition 7.3. *Let P be an integral N-simplex in $\mathbb{R}^N$, with vertices $v_0, v_1, \ldots, v_N$, and let $C = C(\tilde{P})$. A point $z \in \mathbb{Z}^{N+1}$ is an integer point in C if any only if $z = y + \sum_{i=0}^N r_i(v_1,1)$ for some $y \in Q \cap \mathbb{Z}^{N+1}$ and some nonnegative integers r_i. Futhermore, this representation of z is unique.*

So to count integer points in C (and hence to determine $i(P,n)$), we only need to know how many integer points are in Q with each fixed (integer) last coordinate. We call the last coordinate of $z \in Q$ the degree of z. Note that for $z \in Q$, $\deg z = \sum_{i=0}^N r_i$ for some r_i, $0 \le r_i < 1$, so if $\deg z$ is an integer, $0 \le \deg z \le N$.

Theorem 7.4. *Let P be an integral N-simplex in $\mathbb{R}^N$, with vertices $v_0, v_1, \ldots, v_N$, let $C = C(\tilde{P})$, and let $Q = \{\sum_{i=0}^N r_i(v_i, 1) : \text{for each } i,\ 0 \le r_i < 1\}$. Let δ_j be the number of points of degree j in $Q \cap \mathbb{Z}^{N+1}$. Then*

$$\sum_{n=0}^{\infty} i(P,n)\lambda^n = \frac{\delta_0 + \delta_1\lambda + \cdots + \delta_N\lambda^N}{(1-\lambda)^{N+1}}.$$

Corollary 7.5. *For P an integral N-simplex, $i(P,n)$ is a polynomial in n.*

Proof.

$$\begin{aligned}\sum_{n=0}^{\infty} i(P,n)\lambda^n &= (\delta_0 + \delta_1\lambda + \cdots + \delta_N\lambda^N)(1+\lambda+\lambda^2+\cdots)^{N+1} \\ &= (\delta_0 + \delta_1\lambda + \cdots + \delta_N\lambda^N)\left(\sum_{k-0}^{\infty}\binom{k+N}{N}\lambda^k\right).\end{aligned}$$

The coefficient of λ^n on the right hand side is $\sum_{j=0}^N \delta_j \binom{n-j+N}{N}$. □

For the interior of P (and of C) we use an analogous construction, but with the opposite half-open parallelipiped. Let

$$Q^* = \left\{\sum_{i=0}^N r_i(v_i,1) : \text{for each } i,\ 0 < r_i \le 1\right\}.$$

Proposition 7.6. *Let P be an integral N-simplex in $\mathbb{R}^N$, with vertices $v_0, v_1, \ldots, v_N$, and let $C = C(\tilde{P})$. A point $z \in \mathbb{Z}^{N+1}$ is an integer point in $\operatorname{int} C$ if and only if $z = y + \sum_{i=0}^N c_i(v_1, 1)$ for some $y \in Q^* \cap \mathbb{Z}^{N+1}$ and some nonnegative integers c_i. Furthermore, this representation of z is unique.*

So to count integer points in $\operatorname{int} C$ (and hence to determine $i^*(P,n)$), we only need to know how many integer points are in Q^* with each fixed (integer) last coordinate. Note that for $z \in Q^*$, $0 < \deg z \le N+1$.

Theorem 7.7. *Let P be an integral N-simplex in $\mathbb{R}^N$, with vertices $v_0, v_1, \ldots, v_N$, let $C = C(\tilde{P})$, and let $Q^* = \{\sum_{i=0}^N r_i(v_i, 1) : \text{for each } i,\ 0 < r_i \le 1\}$. Let δ_j^* be the number of points of degree j in $Q^* \cap \mathbb{Z}^{N+1}$. Then*

$$\sum_{n=0}^{\infty} i^*(P,n)\lambda^n = \frac{\delta_1^*\lambda + \delta_2^*\lambda^2 + \cdots + \delta_{N+1}^*\lambda^{N+1}}{(1-\lambda)^{N+1}}.$$

Corollary 7.8. *For P an integral N-simplex, $i^*(P,n)$ is a polynomial in n.*

Now the punchline is that there is an easy relationship between the δ_i and the δ_i^*. Note that

$$\begin{aligned}Q^* &= \left\{\sum_{i=0}^N r_i(v_i,1) : \text{for each } i,\ 0 < r_i \le 1\right\} \\ &= \left\{\sum_{i=0}^N (1-t_i)(v_i,1) : \text{for each } i,\ 0 \le t_i < 1\right\} \\ &= \left\{\sum_{i=0}^N (v_i,1) - \sum_{i=0}^N t_i(v_i,1) : \text{for each } i,\ 0 \le t_i < 1\right\} \\ &= \sum_{i=0}^N (v_i,1) - Q = \left(\sum_{i=0}^N v_i, N+1\right) - Q\end{aligned}$$

An element of $Q^* \cap \mathbb{Z}^{N+1}$ of degree k corresponds to an element of $Q \cap \mathbb{Z}^{N+1}$ of degree $N+1-k$. Thus $\delta_k^* = \delta_{N+1-k}$.

Theorem 7.9. *If P is an integral N-simplex in $\mathbb{R}^N$, then*

$$F(P,\lambda) := \sum_{n=0}^{\infty} i(P,n)\lambda^n = \frac{\delta_0 + \delta_1\lambda + \cdots + \delta_N\lambda^N}{(1-\lambda)^{N+1}}$$

$$F^*(P,\lambda) := \sum_{n=0}^{\infty} i^*(P,n)\lambda^n = \frac{\delta_N\lambda + \delta_{N-1}\lambda^2 + \cdots + \delta_0\lambda^{N+1}}{(1-\lambda)^{N+1}}.$$

Thus

$$F^*(P,\lambda) = (-1)^{N+1}F(P,1/\lambda).$$

This relationship is known as **Ehrhart reciprocity**.

So far I have considered only integral simplices. To extend the result to integral polytopes requires triangulation of the polytope, that is, subdivision of the polytope into simplices. The extension is nontrivial. We cannot just add up the functions i and i^* for the simplices in the triangulation, since interior points of the polytope can be contained in the boundary of a simplex of the triangulation, and in fact in the boundary of more than one simplex of the triangulation. But it works in the end.

Theorem 7.10. *Let $P \subset \mathbb{R}^N$ be an integral polytope of dimension N. Then*

$$(1-\lambda)^{N+1}\sum_{i=0}^{\infty} i(P,n)\lambda^n$$

is a polynomial in λ of degree at most N.

As before, write this polynomial as $\sum_{j=0}^{N} \delta_j \lambda^j$. What can we say about the coefficients δ_j?

$\delta_0 = i(P,0) = 1$, since this is the number of integer points in the polytope $0P = \{0\}$.

$\delta_1 + (N+1)\delta_0 = i(P,1)$, so $\delta_1 = |P \cap \mathbb{Z}^N| - (N+1)$.

Also, recall that $i(P,n) = \sum_{j=0}^{N} \delta_j \binom{n-j+N}{N}$. Let C be the leading coefficient of $i(P,n)$ as a polynomial in n, i.e.,

$$C = \frac{1}{N!}\sum_{j=0}^{N} \delta_j = \lim_{n\to\infty} \frac{i(P,n)}{n^N}.$$

I claim C is the volume of P. How can you see this? Note that $\mathrm{vol}(nP) = n^N \mathrm{vol}(P)$ (if P is of full dimension N). Now the volume of nP can be estimated by the number of lattice points in nP, that is, by $i(P,n)$. In fact,

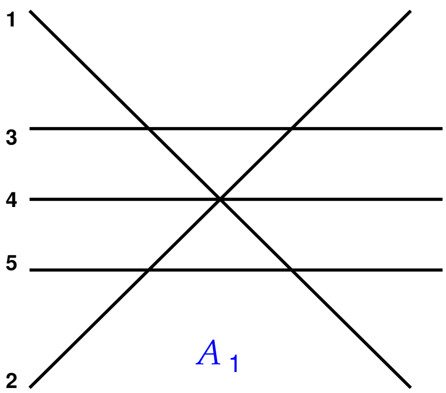
$$0 = \lim_{n\to\infty} \frac{i(P,n) - \mathrm{vol}(nP)}{n^N} = \lim_{n\to\infty} \frac{i(P,n)}{n^N} - \mathrm{vol}(P).$$

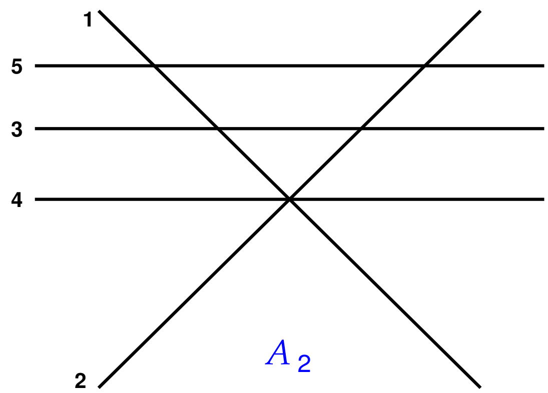
So $C = \lim_{n\to\infty} \dfrac{i(P,n)}{n^N} = \mathrm{vol}(P)$.

One last comment. The Ehrhart theory can be generalized to rational polytopes. In the more general case, the functions $i(P,n)$ and $i^*(P,n)$ need not be polynomials, but are **quasipolynomials**—restricted to a congruence class in some modulus (depending on the denominators occurring in the coordinates of the vertices) they are polynomials. An equivalent description is that the function $i(P,n)$ is a polynomial in n and expressions of the form $\gcd(n,k)$, e.g.,

$$i(P,n) \;=\; \begin{cases} (n+1)^2 & n \text{ even} \\ n^2 & n \text{ odd} \end{cases} \;=\; (n+\gcd(n,2)-1)^2.$$

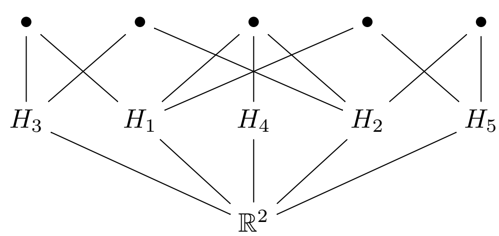
7.2. **Oriented Matroids.** Consider the following two hyperplane arrangements $\mathcal{A}_1, \mathcal{A}_2$ in $\mathbb{R}^2$, with lines $H_1, \dots, H_5$ (abbreviated $1, \dots, 5$ in the diagrams):





It is easy to check that their intersection posets are isomorphic:

$L(\mathcal{A}_1) \;\cong\; L(\mathcal{A}_2) \;=$



However, there is good reason not to consider $\mathcal{A}_1, \mathcal{A}_2$ isomorphic as arrangements. For example, both two bounded regions in $\mathcal{A}_1$ are triangles, while in $\mathcal{A}_2$ there is a triangle and a trapezoid. Also, the point $H_1 \cap H_2 \cap H_4$ lies between H_3 and H_5 in $\mathcal{A}_1$, while it lies below both of them in $\mathcal{A}_2$. However, the intersection poset lacks the power to model geometric data like "between," "below," and "trapezoid," so we will define a stronger combinatorial invariant called the **big face lattice**[8] of $\mathcal{A}$.

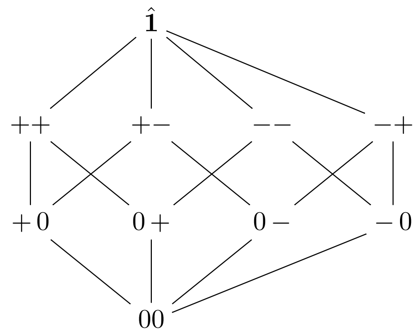
In general, let $\mathcal{A} = \{H_1, \dots, H_n\}$ be a hyperplane arrangement in $\mathbb{R}^d$, with normal vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. For each i, let ℓ_i be an affine linear functional on $\mathbb{R}^n$ such that $H_i = \{\mathbf{x} \in \mathbb{R}^d \mid \ell_i(\mathbf{x}) = 0\}$. That is, $\ell_i : \mathbb{R}^2 \to \mathbb{R}$ is defined by $\ell_i(\mathbf{x}) = \|\operatorname{proj}_{\mathbf{v}_i}(\mathbf{x})\|$.

The intersections of hyperplanes in $\mathcal{A}$, together with its regions, decompose $\mathbb{R}^d$ as a *polyhedral cell complex* — a disjoint union of polyhedra, each homeomorphic to $\mathbb{R}^e$ for some $e \leq d$. We can encode each cell by recording whether the linear functionals $\ell_1, \dots, \ell_n$ are positive, negative or zero on it. Specifically, for $c = (c_1, \dots, c_n) \in \{+, -, 0\}^n$, let

$$F = F(c) = \left\{ \mathbf{x} \in \mathbb{R}^d \;\middle|\; \begin{array}{ll} \ell_i(\mathbf{x}) > 0 & \text{if } c_i = + \\ \ell_i(\mathbf{x}) < 0 & \text{if } c_i = - \\ \ell_i(\mathbf{x}) = 0 & \text{if } c_i = 0 \end{array} \right\}$$

[8] I.e., the big lattice of faces, not the lattice of big faces.

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 $\hat{\mathscr{F}}(\mathcal{A}) = \mathscr{F}(\mathcal{A}) \cup \{\hat{\mathbf{0}}, \hat{\mathbf{1}}\}$ (ordered by $F \leq F'$ if $\bar{F} \subseteq \bar{F}'$) is a lattice, called the **big face lattice of** $\mathcal{A}$. (If $\mathcal{A}$ is central, then $\mathscr{F}(\mathcal{A})$ already has a unique minimal element and we don't bother adding an additional $\hat{0}$.) For example, if $\mathcal{A} = \mathrm{Bool}_2 = \{y = 0, x = 0\} \subset \mathbb{F}^2$, then the big face lattice is:



For example,

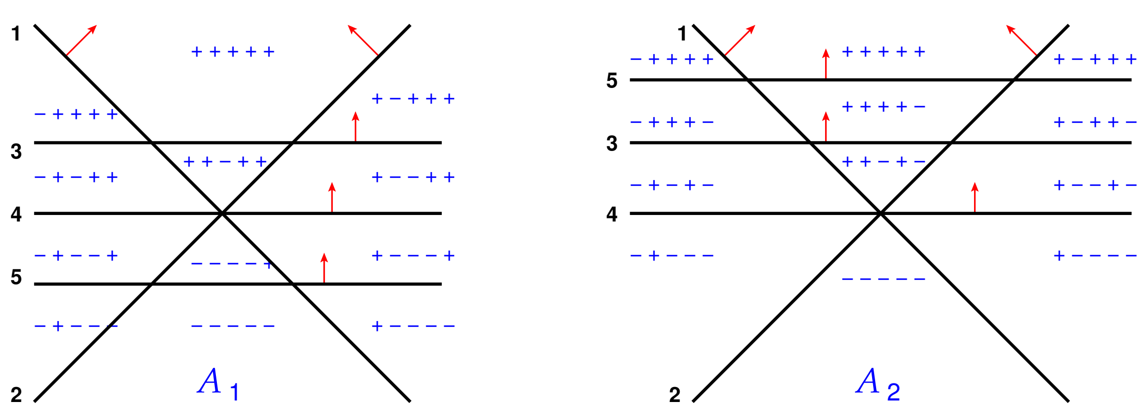
$$
\begin{aligned}
+- &= \{(x,y) \in \mathbb{R}^2 : \ y > 0, \ x < 0\} \cong \mathbb{R}^2 \quad \text{and} \\
+0 &= \{(x,y) \in \mathbb{R}^2 : \ y > 0, \ x = 0\} \cong \mathbb{R}
\end{aligned}
$$

are respectively the (open) second quadrant and the positive half of the y-axis.

The big face lattice does capture the geometry of $\mathcal{A}$. For instance, the two arrangements $\mathcal{A}_1, \mathcal{A}_2$ shown above have isomorphic intersection posets but non-isomorphic face lattices. (This may be clear to you now; there are lots of possible explanations and we'll see one soon.)

Properties of the big face lattice. Consider the linear forms ℓ_i that were used in representing each face by a covector. Recall that specifying ℓ_i is equivalent to specifying a normal vector $\mathbf{v}_i$ to the hyperplane H_i (with $\ell_i(\mathbf{x}) = \mathbf{v}_i \cdot \mathbf{x}$). As we know, the vectors $\mathbf{v}_i$ represent a matroid whose lattice of flats is precisely $L(\mathcal{A})$. Scaling $\mathbf{v}_i$ (equivalently, ℓ_i) by a nonzero constant $\lambda \in \mathbb{R}$ has no effect on the matroid represented by the $\mathbf{v}_i$'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 $\mathbf{v}_i$, but an **orientation** for each one — hence the term "oriented matroid". Equivalently, for each hyperplane H_i, we are designating one of the two corresponding halfspaces (i.e., connected components of $\mathbb{R}^d \setminus H_i$) as positive and the other as negative.

Example: Consider our running example of $\mathcal{A}_1$ and $\mathcal{A}_2$. In the following figure, the red arrows point from each hyperplane to its positive halfspace, and the blue strings are the covectors of maximal faces.



Proposition 7.11. *The covectors whose negatives are also covectors are precisely those that correspond to unbounded faces. In particular, $\mathcal{A}$ is central if and only if every negative of a covector is a covector.*

In the arrangement $\mathcal{A}_1$, consider the point $p = \ell_1 \cap \ell_2 \cap \ell_4$. That three lines intersect at p means that there is a linear dependence among the corresponding normal vectors. Suitably scaled, this will look like

$$\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_4 = 0,$$

or on the level of linear forms,

$$\ell_1 + \ell_2 - \ell_4 = 0. \tag{7.1}$$

Of course, knowing which subsets of V are linearly dependent is equivalent to knowing the matroid M represented by V. Indeed, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is a circuit of M.

However, (7.1) tells us more than that: there exists no $x \in \mathbb{R}^3$ such that

$$\ell_1(x) > 0, \qquad \ell_2(x) > 0, \quad \text{and} \quad \ell_4(x) - 0.$$

That is, $\mathcal{A}$ has no covector of the form $+ + * - *$ (for any $* \in \{+,-,0\}$). We say that $+ + 0 - 0$ is the corresponding **oriented circuit**.

For $c \in \{+,-,0\}^n$, write

$$c_+ = \{i \mid c_i = +\}, \qquad c_- = \{i \mid c_i = -\}.$$

Definition 7.12. 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) $00\cdots0 \notin \mathscr{C}$.
(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 $\{\ell_i, \dots, \ell_n\}$ 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

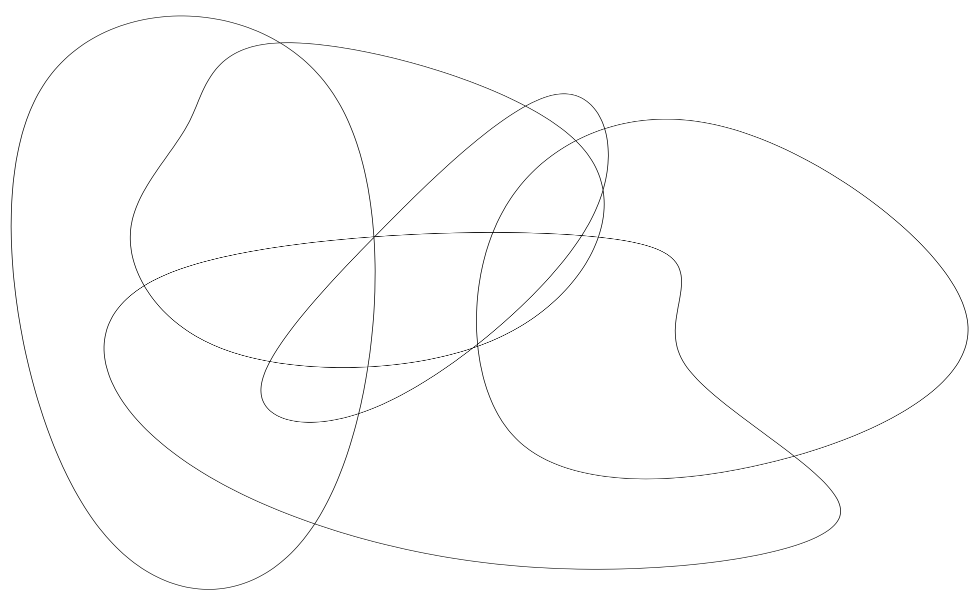
$$\{c_+ \cup c_- \mid c \in \mathscr{C}\}$$

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!)

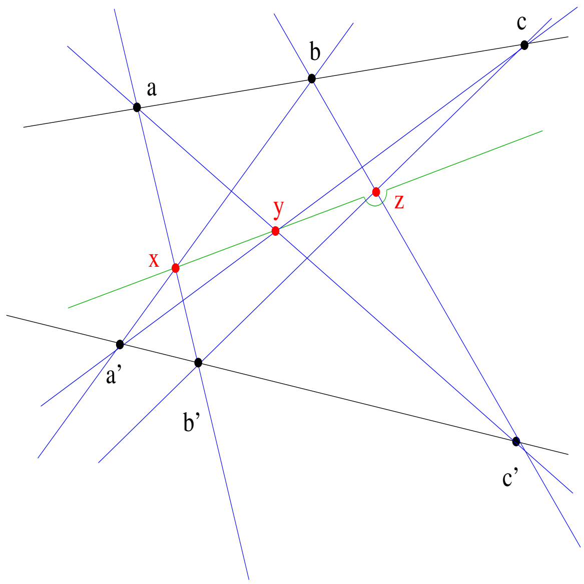
As in the unoriented setting, the circuits of an oriented matroid represent minimal obstructions to being a covector. That is, for every real hyperplane arrangement $\mathcal{A}$, we can construct a circuit system $\mathscr{C}$ for an oriented matroid such that if k is a covector of $\mathcal{A}$ and 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 7.13. Recall the construction of the non-Pappus matroid (Example 3.28). 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.

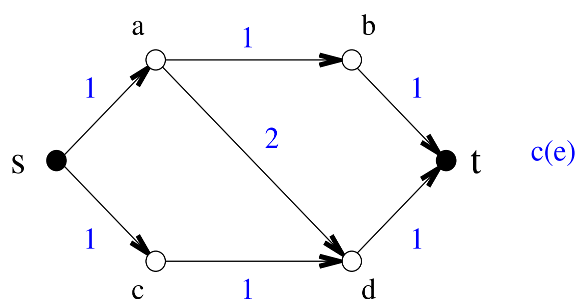


7.3. **The Max-Flow/Min-Cut Theorem.** The main theorem of this section is the Max-Flow/Min-Cut Theorem of Ford and Fulkerson. Strictly speaking, it probably belongs to graph theory or combinatorial optimization rather than algebraic combinatorics, but it is a wonderful theorem and has applications to posets and algebraic graph theory, so I can't resist including it.

Definition: A **network** is a directed graph $N = (V, E)$ with the following additional data:

- A distinguished **source** $s \in V$ and **sink** $t \in V$.
- A **capacity** function $c : E \to \mathbb{R}_{\geq 0}$.

If $c(e) \in \mathbb{N}$ for all $e \in E$, we say the network is **integral**. In what follows, we will only consider integral networks.



We want to think of an s, t-network as modeling a situation where stuff—data, traffic, liquid, electrical current, etc.—is flowing from s to t. The capacity of an edge is the amount of stuff that can flow through it (or perhaps the amount of stuff per unit time). This is a very general model that can be specialized to describe cuts, connectivity, matchings and other things in directed and undirected graphs.

A **flow** on N is a function $f : E \to \mathbb{N}$ that satisfies the **capacity constraints**

$$0 \leq f(e) \leq c(e) \qquad \forall e \in E \tag{7.2}$$

and the **conservation constraints**

$$f^-(v) = f^+(v) \qquad \forall v \in V \setminus \{s, t\} \tag{7.3}$$

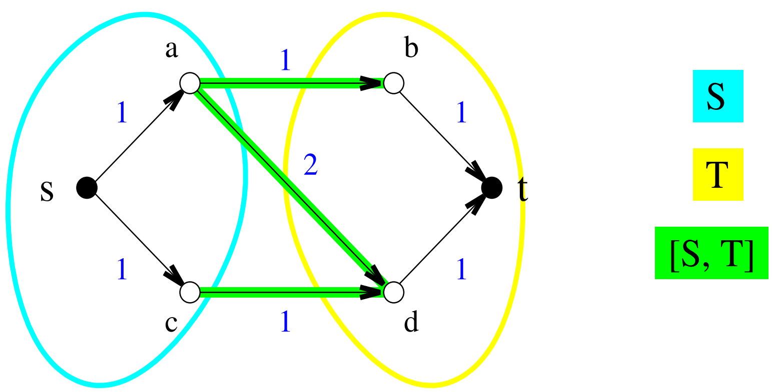
where
$$f^-(v) = \sum_{e=\overrightarrow{uv}} f(e), \qquad f^+(v) = \sum_{e=\overrightarrow{vw}} f(e).$$
That is, stuff cannot accumulate at any internal vertex of the network, nor can it appear out of nowhere.

The **value** $|f|$ of a flow f is the net flow into the sink:
$$|f| = f^-(t) - f^+(t) = f^+(s) - f^-(s).$$
The second equality follows from the conservation constraints. Typically, we'll assume that there are no edges into the source or out of the sink, so that
$$|f| = f^-(t) = f^+(s).$$

The **max-flow problem** is to find a flow of maximum value. The dual problem is the **min-cut problem**, which we now describe.

Definition: Let $N = (V, E)$ be an s,t-network. Let $S, T \subset V$ with $S \cup T = V$, $S \cap T = \emptyset$, $s \in S$, and $t \in T$. The corresponding **cut** is
$$[S, T] \;=\; \{\overrightarrow{st} \in E \mid s \in S, t \in \bar{S}\}$$
and the **capacity** of that cut is
$$c(S, T) = \sum_{e \in E} c(e).$$

A cut can be thought of as a bottleneck through which all flow must pass. For example, in the network at which we have been looking, we could take $S = \{s, a, c\}$, $T = \{b, d, t\}$, as follows. Then $[S, T] = \{\overrightarrow{ab}, \overrightarrow{ad}, \overrightarrow{cd}\}$, and $c(S, T) = 1 + 2 + 1 = 4$.



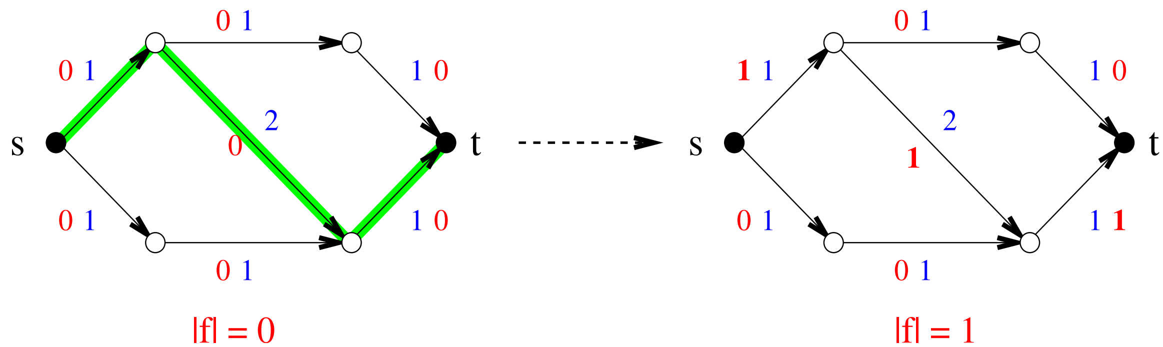
For $A \subset V$, define $f^-(A) = \sum_{e \in [\bar{A}, A]} f(e)$, $f^+(A) = \sum_{e \in [A, \bar{A}]} f(e)$.

Proposition 7.14. *Let f be a flow, and let $A \subseteq V$. Then:*
$$f^+(A) - f^-(A) \;=\; \sum_{v \in A} (f^+(v) - f^-(v)). \tag{7.4a}$$
In particular, if $[S, T]$ is a cut, then
$$f^+(S) - f^-(S) = f^-(T) - f^+(T) \;=\; |f|, \tag{7.4b}$$
$$|f| \leq \; c(S, T). \tag{7.4c}$$

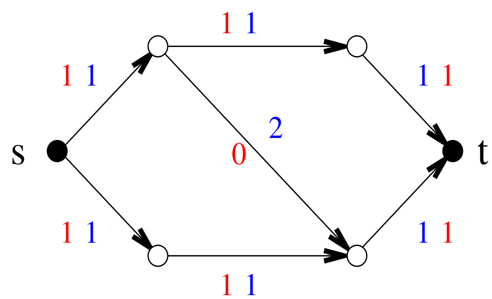
The proof (which requires little more than careful bookkeeping) is left as an exercise.

The inequality (7.4c) is known as **weak duality**; it says that the maximum value of a flow is less than or equal to the minimum capacity of a cut, a principle known as **weak duality**.

Observation: If we can find a path from s to t in which no edge is being used to its full capacity, then we can increase the flow along every edge on that path, and thereby increase the value of the flow by the same amount.

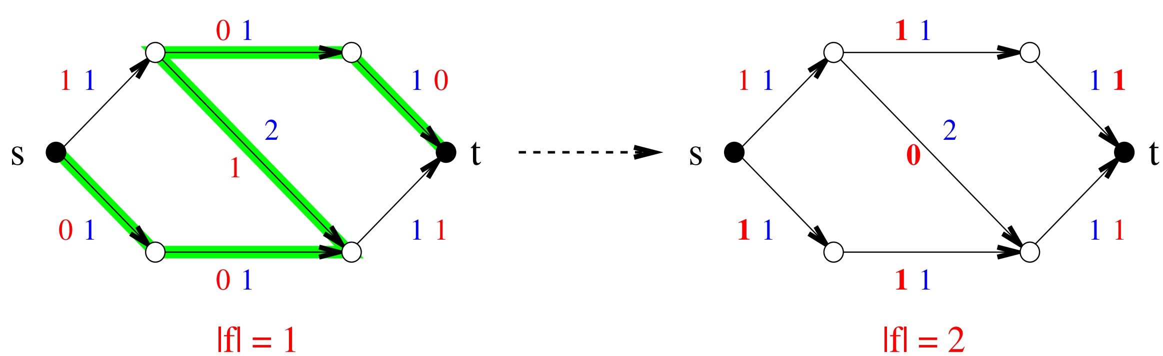


The problem is that some flows cannot be increased in this way, but are nevertheless not maximal. The flow on the right above is an example. In every path from s to t, there is some edge e with $f(e) = c(e)$. However, the flow shown below evidently has a greater value, namely 2:



Observation: Flow along an edge $\overrightarrow{xy}$ can be regarded as *negative* flow from y to x.

Accordingly, there is another way to increase flow. Look for a path from s to t in which each edge e is either pointed forward and has $f(e) < c(e)$, or is pointed backward and has $f(e) > 0$. Then, increasing flow on the forward edges and decreasing flow on the backward edges will increase the value of the flow. This is called an **augmenting path** for f.



The **Ford-Fulkerson Algorithm** is a systematic way to construct a maximum flow by looking for augmenting paths. The wonderful feature of the algorithm is that if a flow f has no augmenting path, the algorithm will automatically find a source/sink cut of capacity equal to $|f|$ — thus certifying immediately that the flow is maximum and that the cut is minimum.

Input: A network $N = (V, E)$ with source s, sink t, and capacity function $c : E \to \mathbb{N}$.

(1) Let f be the zero flow: $f(e) = 0$ for all edges e.
(2) Find an *augmenting path*, i.e., a sequence of vertices
$$P: \quad x_0 = s,\ x_1,\ x_2,\ \ldots,\ x_{n-1},\ x_t$$
such that for every i, $i = 0, \ldots, n-1$, we have either
  - $e_i = \overrightarrow{x_i x_{i+1}} \in E$, and $f(e_i) < c(e_i)$ ("e_i is a forward edge"); or
  - $e_i = \overrightarrow{x_{i+1} x_i} \in E$, and $f(e_i) > 0$ ("e_i is a backward edge").

(3) Define $\tilde{f} : E \to \mathbb{N}$ by $\tilde{f}(e) = f(e)+1$ if e appears forward in P; $\tilde{f}(e) = f(e)-1$ if e appears backward in P; and $\tilde{f}(e) = f(e)$ if $e \notin P$. Then it is easy to verify $\tilde{f}$ satisfies the capacity and conservation constraints, and that $|\tilde{f}| = |f| + 1$.
(4) Repeat steps 2–3 until no augmenting path can be found.

The next step is to prove that this algorithm actually works. That is, when it terminates, it will have computed a flow of maximum possible value.

Proposition 7.15. *Suppose that f is a flow that has no augmenting path. Let*
$$S = \{v \in V \mid \text{there is an augmenting path from } s \text{ to } v\}, \qquad T = V \setminus S.$$
Then $s \in S$, $t \in T$, and $c(S,T) = |f|$. In particular, f is a maximum flow and $[S,T]$ is a minimum cut.

Proof. Note that $t \notin S$ precisely because f has no augmenting path. By (7.4b), we have
$$|f| \;=\; f^+(S) - f^-(S) \;=\; \sum_{e \in [S,\bar{S}]} f(e) - \sum_{e \in [\bar{S},S]} f(e) \;=\; \sum_{e \in [S,\bar{S}]} f(e).$$

But $f(e) = c(e)$ for every $e \in [S,T]$ (otherwise S would be bigger than what it actually is), so this last quantity is just $c(S,T)$. The final assertion follows by weak duality. ☐

We have proven:

Theorem 7.16 (Max-Flow/Min-Cut Theorem for Integral Networks)**.** *For any integral network N, the maximum value of a flow equals the minimum value of a cut.*

In fact the theorem is true for non-integral networks as well, although the Ford-Fulkerson algorithm may not work.

Definition 7.17. Let N be a source/sink network. A flow f in N is *acyclic* if, for every directed cycle $C \subseteq D$ (i.e., every set of edges $x_1 \to x_2 \to \cdots \to x_n \to x_1$), there is some $e \in C$ for which $f(e) = 0$. The flow f is *partitionable* if there is a collection of s,t-paths $P_1, \ldots, P_{|f|}$ such that for every $e \in E$,
$$f(e) = \#\{i \mid e \in P_i\}.$$
(Here "s,t-path" means "path from s to t".) In this sense f can be regarded as the "sum" of the paths P_i, each one contributing a unit of flow.

Proposition 7.18. *Let N be a source/sink network.*

(1) For every flow in N, there exists an acyclic flow with the same value.
(2) Every acyclic integral flow is partitionable.

Proof. Suppose that some directed cycle C has positive flow on every edge. Let $k = \min\{f(e) \mid e \in C\}$. Define $\tilde{f} : E \to \mathbb{N}$ by

$$\tilde{f}(e) = \begin{cases} f(e) - k \text{ if } e \in C, \\ f(e) \text{ if } e \notin C. \end{cases}$$

Then it is easy to check that $\tilde{f}$ is a flow, and that $|\tilde{f}| = |f|$. If we repeat this process, it must eventually stop (because the positive quantity $\sum_{e \in E} f(e)$ decreases with each iteration), which means that the resulting flow is acyclic. This proves (1).

Given an acyclic flow f, find an s,t-path P_1 along which all flow is positive. Decrement the flow on each edge of P_1; doing this will also decrement $|f|$. Now repeat this for an s,t-path P_2, etc. Eventually, we partition f into a collection of s,t-paths of cardinality $|f|$. □

This result has many applications in graph theory: Menger's theorems, the König-Egerváry theorem, etc.

7.4. **Min-Max Theorems on Posets.** The basic result in this area is Dilworth's Theorem, which resembles the Max-Flow/Min-Cut Theorem (and can indeed be derived from it; see the exercises).

Definition 7.19. 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 7.20 (Dilworth's Theorem). *Let P be a finite poset. Then*

$$\text{width}(P) \;=\; \max\big\{s \mid P \textit{ has an antichain of size } s\big\}.$$

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 7.21 (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 whose cardinality equals the length of the longest chain in P. $\square$

There is a marvelous generalization of Dilworth's Theorem (and its trivial cousin) due to Curtis Greene and Daniel Kleitman [GK76, Gre76].

Theorem 7.22 (Greene-Kleitman)**.** *Let P be a finite poset. Define two sequences of positive integers*

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell), \qquad \mu = (\mu_1, \mu_2, \ldots, \mu_m)$$

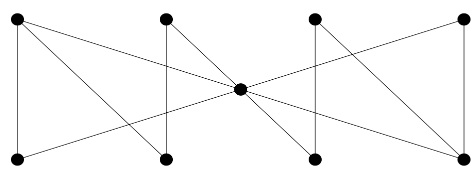
by

$$\lambda_1 + \cdots + \lambda_k = \max\left\{|C_1 \cup \cdots \cup C_k| : \ C_i \subseteq P \ \textit{chains} \right\},$$
$$\mu_1 + \cdots + \mu_k = \max\left\{|A_1 \cup \cdots \cup A_k| : \ A_i \subseteq P \ \textit{disjoint antichains} \right\}.$$

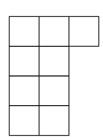
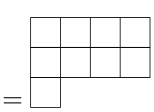
Then:

(1) λ and μ are both partitions *of $|P|$, i.e., weakly decreasing sequences whose sum is $|P|$.*
(2) λ and μ are conjugates *(written $\mu = \tilde{\lambda}$): the row lengths of λ are the column lengths in μ, and vice versa.*

For example, consider the following poset:



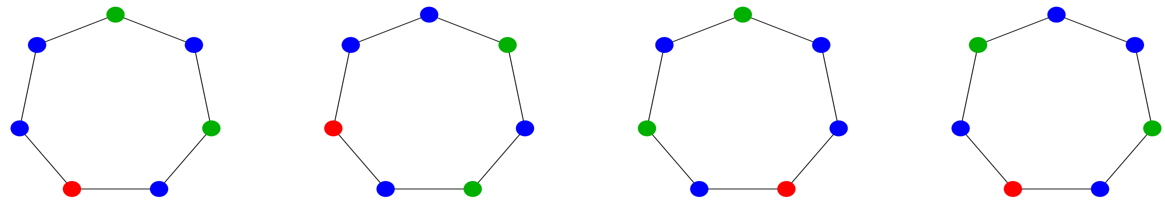
Then you can check that

$\lambda = (3, 2, 2, 2) =$  and $\mu = (4, 4, 1) =$  $= \tilde{\lambda}$.

Note that Dilworth's Theorem is just the special case $\mu_1 = \ell$.

7.5. **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 7.23. *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 7.24 (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 7.25. 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 7.26. 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

$$\frac{k^7 + 7k^4 + 6k}{14}. \tag{7.5}$$

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 7.27. *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,

$$\#\textit{equivalence classes of colorings} = \frac{1}{|G|} \sum_{g \in G} k^{\ell(g)}. \tag{7.6}$$

Let's rephrase Example 7.26 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 (7.5) is an example of the general formula (7.6).

Example 7.28. 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

$$\frac{k^4 + 11k^2}{12}.$$

7.6. **Grassmannians.** A standard reference for everything in this and the following section is Fulton [Ful97].

Part of the motivations for the combinatorics of partitions and tableaux comes from classical enumerative geometric questions like this:

Problem 7.29. *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?

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 $\mathrm{Gr}(k, V)$ (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 Ω_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 Ω_λ is given by the formula
$$\overline{\Omega_\lambda} = \bigcup_{\mu \leq \lambda} \Omega_\mu$$
where $\leq$ is the usual partial order on Young's lattice (i.e., containment of Ferrers diagrams).
- Consequently, the Poincaré polynomial of $\mathrm{Gr}(k, \mathbb{C}^n)$ (i.e., the Hilbert series of its cohomology ring)[9] 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.

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:
$$M = \begin{bmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nk} \end{bmatrix}.$$

[9] 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.

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

$$\mathrm{Gr}(k,\mathbb{F}^n) = \mathbb{F}^{n\times k}/GL_k(\mathbb{F}). \tag{7.7}$$

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 (7.7),

$$\dim \mathrm{Gr}(k,\mathbb{F}^n) = \dim \mathbb{F}^{n\times k} - \dim GL_k(\mathbb{F}) = nk - k^2 = k(n-k)$$

where "dim" means dimension as a manifold over $\mathbb{F}$. Technically, this dimension calculation does not follow from (7.7) 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: the 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, ending 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:

$$\begin{bmatrix}1&0\\0&1\\0&0\\0&0\end{bmatrix}\qquad \begin{bmatrix}1&0\\0&*\\0&1\\0&0\end{bmatrix}\qquad \begin{bmatrix}1&0\\0&*\\0&*\\0&1\end{bmatrix}\qquad \begin{bmatrix}*&*\\1&0\\0&1\\0&0\end{bmatrix}\qquad \begin{bmatrix}*&*\\1&0\\0&*\\0&1\end{bmatrix}\qquad \begin{bmatrix}*&*\\*&*\\1&0\\0&1\end{bmatrix} \tag{7.8}$$

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 7.30. 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* $\Omega_{\mathbf{p}}$ 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 7.31. *(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 $\Omega_{\mathbf{p}}$.*

(2) For every $\mathbf{p} \in \binom{[n]}{k}$, there is a diffeomorphism

$$\Omega_{\mathbf{p}} \xrightarrow{\sim} \mathbb{F}^{|\mathbf{p}|}$$

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 $\mathbf{p} \ge \mathbf{q}$ if $p_i \ge q_i$ for every i. Then

$$\mathbf{p} \ge \mathbf{q} \implies \overline{\Omega_{\mathbf{p}}} \supseteq \Omega_{\mathbf{q}}. \tag{7.9}$$

(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* $\Omega_{(n-k+1,n-k+2,\dots,n)}$*, 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 $\overline{\Omega_{\mathbf{p}}}$ are called **Schubert varieties**.

Proof. (1) is immediate from the definition.

For (2), the map $\Omega_{\mathbf{p}} \to \mathbb{F}^{|\mathbf{p}|}$ 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 (7.8) 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 (7.8):

$$M \;=\; \begin{bmatrix} 1 & 0 \\ 0 & z \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

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 (7.8). Therefore, the Schubert cell $\Omega_{1,2}$ is in the closure of the Schubert cell $\Omega_{1,3}$. 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 (7.9).

For (4), the elements of $Y_{k,n}$ are partitions $\lambda = (\lambda_1, \dots, \lambda_k)$ such that $n - k \geq \lambda_1 > \dots > \lambda_k \geq 0$. 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 $|\mathbf{p}| = k(n-k)$. $\square$

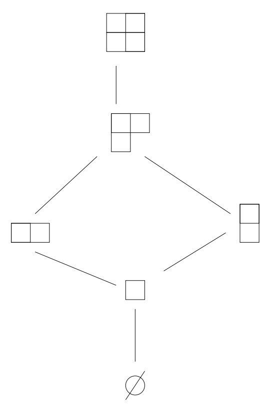
This theorem amounts to a description of $\mathrm{Gr}(k,\mathbb{F}^n)$ as a cell complex. (If you have not heard the term "cell complex" before, now you know what it means: a topological space that is the disjoint union of cells — that is, of copies of vector spaces — such that the closure of every cell 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 $\mathrm{Gr}(1, \mathbb{F}^n)$ 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 (7.8). 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 7.32. What does all this have to do with enumerative geometry questions such as Problem 7.29? The answer (modulo technical details) is that the cohomology ring $H^*(X)$ encodes intersections of subvarieties[10] 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 7.29 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 7.29 is (drum roll, please) **two**, which is hardly obvious! To carry out this calculation, one needs 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*.

[10]If you are more comfortable with differential geometry than algebraic geometry, feel free to think "submanifold" instead of "subvariety".

7.7. **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 $F_k = \mathbb{F}\langle F_1, \dots, F_k\rangle$ 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

$$F\ell(n) = \coprod_{w \in \mathfrak{S}_n} X_w.$$

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 (7.9)). 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

$$\overline{X_w} = \bigcup_{v \in \mathfrak{S}_n:\ v \leq w} X_v$$

where $\leq$ means (strong) Bruhat order. 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*. t 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.

7.8. **Exercises.**

Exercise 7.1. Prove Proposition 7.14.

Exercise 7.2. Let $G(V, E)$ be a graph. A **matching** on G is a collection of edges no two of which share an endpoint. A **vertex cover** is a set of vertices that include at least one endpoint of each edge of G. Let $\mu(G)$ denote the size of a maximum matching, and let $\beta(G)$ denote the size of a minimum vertex cover.

(a) (Warmup) Show that $\mu(G) \leq \beta(G)$ for every graph G. Exhibit a graph for which the inequality is strict.
(b) The *König-Egerváry Theorem* asserts that $\mu(G) = \beta(G)$ whenever G is bipartite, i.e., the vertices of G can be partitioned as $X \cup Y$ so that every edge has one endpoint in each of X, Y. Derive the König-Egerváry Theorem as a consequence of the Max-Flow/Min-Cut Theorem.
(c) Prove that the König-Egerváry Theorem and Dilworth's Theorem imply each other.

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 $\rho : G \to GL(V)$ (or the pair (ρ, 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 **faithful** if it is injective as a group homomorphism.

Example 8.2. Let G be any group. The **trivial representation** is the map $\rho_{\text{triv}} : G \to GL(\mathbb{F}^1) \cong \mathbb{F}^\times$ sending $g \mapsto 1$ for all $g \in G$.

Example 8.3. 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,

$$\mathbb{F}G = \left\{ \sum_{h \in G} a_h h \mid a_h \in \mathbb{F} \right\}.$$

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.

The vector space $\mathbb{F}G$ is a ring, with multiplication given by multiplication in G and extended $\mathbb{F}$-linearly. In this context it is called the **group algebra** of G over $\mathbb{F}$. A representation of G is equivalent to a left module over the group algebra: the map ρ makes V into a $\mathbb{F}G$-module, and every module arises in such a way.

Example 8.4. Let $G = \mathfrak{S}_n$, the symmetric group on n elements. There is a representation $\rho_{\text{def}} : G \to \mathbb{F}^n$ that maps each permutation $\sigma \in G$ to the permutation matrix with 1's in the positions $(i, \sigma(i))$ for every $i \in [n]$, and 0's elsewhere. This is called the **defining representation** of $\mathfrak{S}_n$.

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 natural representations of G.

(1) We can regard s as a reflection and r as a rotation in $\mathbb{R}^2$, to obtain a faithful 2-dimensional representation defined by
$$\rho(s) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \rho(r) = \begin{bmatrix} \cos(2\pi/n) & \sin(2\pi/n) \\ -\sin(2\pi/n) & \cos(2\pi/n) \end{bmatrix}.$$
(2) The permutation representations of G on vertices or on edges are faithful n-dimensional representations.
(3) The n-gon has n diameters (lines of reflection symmetry)., the dihedral group acts on diameters and thus gives rise to another n-dimensional permutation representation. This representation is faithful if and only if n is odd. If n is even, then $r^{n/2}$ acts by rotation by $180°$ and fixes all diameters.

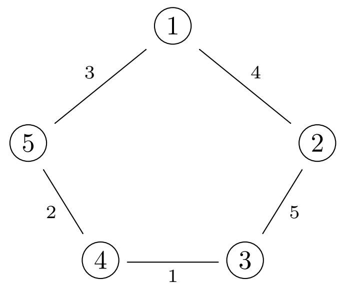
Example 8.8. The symmetric group $\mathfrak{S}_n$ has a nontrivial 1-dimensional representation, the **sign representation**, given by
$$\rho_{\text{sign}}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{if } \sigma \text{ is odd.} \end{cases}$$
Note that $\rho_{\text{sign}}(g) = \det \rho_{\text{def}}(g)$, 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 are two representations 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**, or a **homomorphism**, if $\rho'(g) \cdot \phi(v) = \phi(\rho(g) \cdot v)$ for all $g \in G$ and $v \in V$. More concisely, $g\pi = \phi g$ for all $g \in G$, i.e., the following diagram commutes
$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "\rho(g)"'] & V' \arrow[d, "\rho'(g)"] \\ V \arrow[r, "\phi"] & V' \end{tikzcd} \tag{8.1}$$
We sometimes use the notation $\phi : \rho \to \rho'$. An **isomorphism** of representations is a G-equivariant map that is a vector space isomorphism. Note that 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, as in the figure on the left. Then the permutation action on vertices is identical to that on edges. In other words, the diagram on the right commutes for all $g \in D_n$.



$$\begin{tikzcd} \text{vertices} \arrow[r, "\text{opp}"] \arrow[d, "g"] & \text{edges} \arrow[d, "g"] \\ \text{vertices} \arrow[r, "\text{opp}"] & \text{edges} \end{tikzcd}$$

Example 8.12. One way in which G-equivariant transformations occur is when a group action "naturally" induces another action. For instance, consider the permutation action of $\mathfrak{S}_4$ on the vertices of K_4, which induces a representation ρ_V on the space $V = \mathbb{F}\langle \mathbf{v}_1, \dots, \mathbf{v}_4 \rangle \cong \mathbb{F}^4$. This action naturally determines an action on the six edges of K_4, which in turn induces a permutation representation ρ_E on $E = \mathbb{F}\langle \mathbf{e}_{12}, \dots, \mathbf{v}_{34} \rangle \cong \mathbb{F}^6$. The "inducement" can be described by a G-equivariant transformation — but be careful: it is not a map $V \to E$ but a map $E \to V$, namely

$$\phi(\mathbf{e}_{ij}) = \mathbf{v}_i + \mathbf{v}_j$$

which is easily seen to be a G-equivariant transformation $\rho_E \to \rho_V$. In the same vein, let $v_1, \dots, v_n$ are the points of a regular n-gon in $\mathbb{R}^2$ centered at the origin, e.g., $\mathbf{v}_j = (\cos 2\pi j/n, \sin 2\pi j/n)$. Then the map $\mathbb{R}^n \to \mathbb{R}^2$ sending the j^{th} standard basis vector to $\mathbf{v}_j$ is D_n-equivariant.

Example 8.13. Consider the representations ρ_{triv} and ρ_{sign} of $\mathfrak{S}_2$. Recall that these are the representations on $V = \mathbb{F}$ given by

$$\rho_{\text{triv}}(\text{id}) = \rho_{\text{triv}}((1\ 2)) = 1, \qquad \rho_{\text{sign}}(\text{id}) = 1, \quad \rho_{\text{sign}}((1\ 2)) = -1.$$

If $\rho = \rho' \in \{\rho_{\text{triv}}, \rho_{\text{sign}}\}$, then any linear transformation $\phi : \mathbb{F} \to \mathbb{F}$ (i.e., any map $\phi(x) = cx$ for some $c \in \mathbb{F}$) satisfies the commutative diagram (8.1), and if $c = 0$ then ϕ is an isomorphism. So in these cases the set of G-equivariant homomorphisms is actually isomorphic to $\mathbb{F}$. On the other hand, if $\phi : \rho_{\text{triv}} \to \rho_{\text{sign}}$ is G-equivariant, then for any $x \in \mathbb{F}$ we have

$$\rho_{\text{sign}}(\phi(v)) = -\phi(v) = -cv, \qquad \phi(\rho_{\text{triv}}(v)) = \phi(v) = cv,$$

and if $\operatorname{char} \mathbb{F} \neq 2$ then these things are only equal if $c = 0$, so there is no nontrivial G-homomorphism $\rho_{\text{triv}} \to \rho_{\text{sign}}$. This example is important because we are going to use the set (in fact, G-representation!) of G-homomorphisms $\phi : \rho \to \rho'$ to measure how similar ρ and ρ' are.

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 $\mathbb{F}\langle e_1 + e_2 \rangle$ and acts by -1 on $\mathbb{F}\langle e_1 - e_2 \rangle$. These two vectors form a basis of V, and one can check that the change-of-basis map ϕ is an isomorphism $\rho \to \sigma$., i.e.,

$$\phi = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

8.2. **Irreducibility, Indecomposability and Maschke's Theorem.**

Definition 8.15. Let (ρ, V) be a representation of G. A vector subspace W is G**-invariant** if $\rho(g)W \subseteq W$ for every $g \in G$. In this case each map $\rho(g)$ restricts to a linear automorphism $\rho|_W(g)$ of W, so in fact $\rho(g)W = W$, and thus $(\rho|_W, W)$ is a representation of G. Equivalently, W is a $\mathbb{F}G$-submodule of V.

A representation (ρ, V) of G is **decomposable** if there are G-invariant subspaces $W, W^\perp$ with $W \cap W^\perp = 0$ and $W + W^\perp = V$. (Here $W^\perp$ is the **complement** of G; the notation does not presuppose the existence of a scalar product.) Otherwise, V is **indecomposable**.

The representation 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**. A semisimple representation is determined up to isomorphism by the multiplicity with which each isomorphism type of irrep appears.

Clearly, every representation can be written as the direct sum of indecomposable representations, and irreducible representations are indecomposable. On the other hand, there exist indecomposable representations that are not irreducible.

Example 8.16. As in Example 8.14, let $V = \{e_1, e_2\}$ be the standard basis for $\mathbb{F}^2$, where $\operatorname{char}\mathbb{F} \neq 2$. 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 the change-of-basis map

$$\phi = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}^{-1}$$

is a G-equivariant isomorphism $\rho_{\text{def}} \to \rho_{\text{triv}} \oplus \rho_{\text{sign}}$. On the other hand, if $\operatorname{char}\mathbb{F} = 2$, then the matrix ϕ is not invertible and this argument breaks down. 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.

Fortunately, we can rule out this kind of pathology most of the time!

Theorem 8.17 (Maschke's Theorem)**.** *Let G be a finite group, let $\mathbb{F}$ be a field whose characteristic does not divide $|G|$, and let (ρ, V) be a representation of G over $\mathbb{F}$. Then every G-invariant subspace has a G-invariant complement. In particular, (ρ, V) is semisimple.*

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 a *projection*, i.e., a linear surjection 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 turn π into a G-equivariant projection by "averaging over G". Specifically, define a map $\tilde{\pi} : V \to V$ by

$$\tilde{\pi}(v) \stackrel{\text{def}}{\equiv} \frac{1}{|G|} \sum_{g \in G} g\pi(g^{-1}v). \tag{8.2}$$

Note first that $\tilde{\pi}(v) \in W$ 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 $\pi(g^{-1}w) = g^{-1}w$ and

$$\tilde{\pi}(w) = \frac{1}{|G|} \sum_{g \in G} gg^{-1}w = w. \tag{8.3}$$

Therefore $\tilde{\pi}$ is also a projection $V \to W$. Third, we check that $\tilde{\pi}$ is G-equivariant. Let $h \in G$; then

$$\begin{aligned}\tilde{\pi}(hv) &= \frac{1}{|G|}\sum_{g\in G} g\pi(g^{-1}hv)\\ &= \frac{1}{|G|}\sum_{k\in G:\ hk=g}(hk)\pi((hk)^{-1}hv)\\ &= \frac{1}{|G|}h\sum_{k\in G}k\pi(k^{-1}v) \quad = \quad h\tilde{\pi}(v)\end{aligned}$$

as desired.

Define $W^\perp = \ker\tilde{\pi}$. Certainly $V \cong W \oplus W^\perp$ as vector spaces. By G-equivariance, if $v \in W^\perp$ and $g \in G$, then $\tilde{\pi}(gv) = g\tilde{\pi}(v) = 0$, i.e., $gv \in W^\perp$. That is, $W^\perp$ is G-invariant. □

Note that if $\operatorname{char}\mathbb{F}$ does divide $|G|$, then the proof breaks down because the map $v \mapsto \sum_{g\in G} g\pi(g^{-1}v)$ will kill everything in W instead of preserving it (consider the effect of multiplying both sides of (8.3) by $|G|$).

Maschke's Theorem implies that, when the conditions hold, 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 (which does not require any assumption on $\operatorname{char}\mathbb{F}$):

Proposition 8.18. *Any G-equivariant map has a G-equivariant kernel and G-equivariant image.*

This is again a familiar fact from the module-theoretic standpoint — every kernel or image of a module homomorphism is also a module.

8.3. **Characters.** The first miracle of representation theory is that we can detect the isomorphism type of a representation ρ without knowing every coordinate of every matrix $\rho(g)$: it turns out that all we need to know is the *traces* of the $\rho(g)$.

Definition 8.19. 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 8.20. Some simple facts and some characters we've seen before:

- A one-dimensional representation is its own character.
- For any representation ρ, we have $\chi_\rho(1) = \dim\rho$, because $\rho(1)$ is the $n \times n$ identity matrix.
- The defining representation ρ_{def} of $\mathfrak{S}_n$ has character

  $$\chi_{\text{def}}(\sigma) = \text{number of fixed points of } \sigma.$$

  Indeed, this is true for any permutation representation of any group.
- 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 8.21. 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:

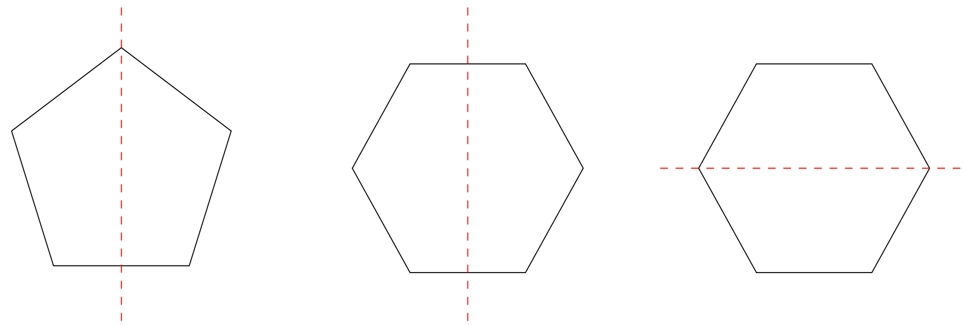
$$\rho(s) = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \qquad \rho(r) = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}$$

Its character is

$$\chi_\rho(r^j) = 2\cos\theta \quad (0 \le j < n), \qquad \chi_\rho(sr^j) = 0 \quad (0 \le j < n).$$

On the other hand, if ρ' is the n-dimensional permutation representation on the vertices, then

$$\chi_{\rho'}(g) = \begin{cases} n & \text{if } g = 1, \\ 0 & \text{if } g \text{ is a nontrivial rotation,} \\ 1 & \text{if } n \text{ is odd and } g \text{ is a reflection,} \\ 0 & \text{if } n \text{ is even and } g \text{ is a reflection through two edges,} \\ 2 & \text{if } n \text{ is even and } g \text{ is a reflection through two vertices.} \end{cases}$$



One fixed point　　No fixed points　　Two fixed points

Proposition 8.22. *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. $\square$

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 $\mathbb{C}$.

8.4. **New Characters from Old.** The basic vector space functors of direct sum, duality, tensor product and Hom carry over naturally to representations, and behave well on their characters. Throughout this section, let (ρ, V), (ρ', V') be representations of G, with $V \cap V' = 0$.

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, we can define the direct sum $(\rho \oplus \rho', V \oplus V')$ 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 $\chi_{\rho\oplus\rho'}(h) = \chi_\rho(h) + \chi_{\rho'}(h)$.

2. **Duality**. Recall that the *dual space* V^* of V consists of all $\mathbb{F}$-linear transformations $\phi : V \to \mathbb{F}$. A G-representation (ρ, V) gives rise to an 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 8.23. *For every $h \in G$,*

$$\chi_{\rho^*}(h) = \overline{\chi_\rho(h)}. \tag{8.4}$$

where the bar denotes complex conjugate.

Proof. Choose a basis $\{v_1, \dots, v_n\}$ of V consisting of eigenvectors of h (we can do this 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.** Let $V = \mathbb{F}\langle \mathbf{v}_1, \dots, \mathbf{v}_n\rangle$ and $V' = \mathbb{F}\langle \mathbf{v}'_1, \dots, \mathbf{v}'_m\rangle$. As a vector space, we define

$$V \otimes V' = \mathbb{F}\left\langle v_i \otimes v'_j \mid 1 \le i \le n,\ 1 \le j \le m\right\rangle,$$

equipped with a multilinear action of $\mathbb{F}$ (that is, $c(x \otimes y) = cx \otimes y = x \otimes cy$ for $c \in \mathbb{F}$). In particular, $\dim(V \otimes V') = (\dim V)(\dim V')$. We can 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 $nm \times nm$ matrix in block form

$$\begin{bmatrix} a_{11}\rho'(h) & \cdots & a_{1n}\rho'(h) \\ \vdots & & \vdots \\ a_{n1}\rho'(h) & \cdots & a_{nn}\rho'(h) \end{bmatrix}$$

where $\rho(h) = [a_{ij}]_{i,j=1\dots n}$. Taking the trace, we see that

$$\chi_{\rho\otimes\rho'}(h) = \chi_\rho(h)\chi_{\rho'}(h). \tag{8.5}$$

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

$$(h \cdot \phi)(v) = h(\phi(h^{-1}v)) = \rho'(h)\Big(\phi(\rho(h^{-1})(v))\Big). \tag{8.6}$$

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

$$\chi_{\operatorname{Hom}(\rho,\rho')}(h) = \overline{\chi_\rho(h)}\ \chi_{\rho'}(h). \tag{8.7}$$

The other kind of Hom is

$$\operatorname{Hom}_G(V, W) = \{G\text{-equivariant transformations } \phi : V \to W\}.$$

Evidently $\operatorname{Hom}_G(V, W) \subseteq \operatorname{Hom}_{\mathbb{C}}(V, W)$, but equality need not hold. For example, if V and W are the trivial and sign representations of $\mathfrak{S}_n$ $(n \ge 2)$, then $\operatorname{Hom}_{\mathbb{C}}(V, W) \cong \mathbb{C}$ but $\operatorname{Hom}_G(V, W) = 0$. See Example 8.13.

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,

$$\mathrm{Hom}_G(V, W) = \mathrm{Hom}_{\mathbb{C}}(V, W)^G. \tag{8.8}$$

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 8.24. *Let (V, ρ) be a representation of G. Then*

$$\dim_{\mathbb{C}} V^G = \frac{1}{|G|} \sum_{h \in G} \chi_\rho(h) = \langle \chi_{\mathrm{triv}},\, \chi_\rho \rangle_G \,.$$

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. □

Example 8.25. Suppose that ρ is a permutation representation. Then V^G is the space of functions that are constant on the orbits, and its dimension is the number of orbits. Therefore, the formula of Proposition 8.24 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 8.26. *For any two representations ρ, ρ' of G, we have $\langle \chi_\rho,\, \chi_{\rho'} \rangle_G = \dim_{\mathbb{C}} \mathrm{Hom}_G(\rho, \rho')$.*

Proof.

$$\begin{aligned}
\langle \chi_\rho,\, \chi_{\rho'} \rangle_G &= \frac{1}{|G|} \sum_{h \in G} \overline{\chi_\rho(h)} \chi_{\rho'}(h) \\
&= \frac{1}{|G|} \sum_{h \in G} \chi_{\mathrm{Hom}(\rho,\rho')}(h) && \text{(by (8.7))} \\
&= \dim_{\mathbb{C}} \mathrm{Hom}(\rho, \rho')^G && \text{(by Proposition 8.24)} \\
&= \dim_{\mathbb{C}} \mathrm{Hom}_G(\rho, \rho') && \text{(by (8.8))}.
\end{aligned}$$

□

The inner product is the tool that lets us prove the following useful fact, known as *Schur's lemma.*

Proposition 8.27 (Schur's Lemma)**.** *Let G be a group, and let (ρ, V) and (ρ', V') be finite-dimensional irreps of G over a field $\mathbb{F}$.*

(1) Every G-equivariant map $\phi : V \to V'$ is either zero or an isomorphism.
(2) If in addition $\mathbb{F}$ is algebraically closed, then

$$\operatorname{Hom}_G(V, V') \cong \begin{cases} \mathbb{F} & \text{if } \rho \cong \rho' \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the only G-equivariant maps from an G-irrep to itself are multiplication by a scalar.

Proof. For (1), recall from Proposition 8.18 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 $\phi \in \operatorname{Hom}_G(V, V')$. If $\rho \ncong \rho'$ 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. ∎

8.6. **The Fundamental Theorem of Representation Theory for Finite Groups.** We are ready to prove the following big theorem, which says that all information about a representation

Theorem 8.28. *Let (ρ, V) and (ρ', V') be finite-dimensional representations of G over $\mathbb{C}$.*

(1) If ρ and ρ' are irreducible, then

$$\langle \chi_\rho,\, \chi_{\rho'} \rangle_G \;=\; \begin{cases} 1 & \text{if } \rho \cong \rho' \\ 0 & \text{otherwise.} \end{cases}$$

(2) If $\rho_1, \dots, \rho_n$ are distinct irreducible representations and

$$\rho \;=\; \bigoplus_{i=1}^{n} (\underbrace{\rho_i \oplus \cdots \oplus \rho_i}_{m_i}) \;=\; \bigoplus_{i=1}^{n} \rho_i^{\oplus m_i}$$

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.
(3) If $\chi_\rho = \chi_{\rho'}$ then $\rho \cong \rho'$.
(4) 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

$$\sum_{i=1}^{n} (\dim \rho_i)^2 = |G|. \tag{8.9}$$

(5) 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. Assertion (1) follows from part (2) of Schur's Lemma together with Proposition 8.26, and (2) follows because the inner product is bilinear on direct sums. For (3), 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 (2).

For (4), 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.

For (5), 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

$$Z = \left\{ \phi \in C\ell(G) \mid \langle \phi,\, \chi_\rho \rangle_G = 0 \text{ for every irreducible character } \rho \right\}.$$

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 $h \in G$,

$$\begin{aligned}
T_\rho(hv) &= \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} (gh)(v) \\
&= \frac{h}{|G|} \sum_{g \in G} \overline{\phi(g)} (h^{-1} g h v) \\
&= \frac{h}{|G|} \sum_{k \in G} \overline{\phi(hkh^{-1})} (kv) && \text{(setting } k = h^{-1} g h,\ hkh^{-1} = g\text{)} \\
&= \frac{h}{|G|} \sum_{k \in G} \overline{\phi(k)} (kv) && \text{(because } \phi \in C\ell(G)\text{)} \\
&= h T_\rho(v).
\end{aligned}$$

Suppose for the moment that ρ is irreducible. By Schur's Lemma, T_ρ is multiplication by a scalar (possibly zero). On the other hand

$$\begin{aligned}
\operatorname{tr}(T_\rho) &= \operatorname{tr}\left(\frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} \rho(g) \right) \\
&= \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} \chi_\rho(g) && \text{(note that each } \phi(g) \text{ is a scalar)} \\
&= \langle \phi,\, \chi_\rho \rangle_G \\
&= 0
\end{aligned}$$

because of the assumption $\phi \in Z$. But if T_ρ has trace zero and is multiplication by a scalar, then that scalar must be zero. Therefore, $T_\rho = 0$ for every irreducible ρ.

It is clear from the definition that 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.

Now, 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. We have now shown that the linear span of the irreducible characters has trivial orthogonal complement as a subspace of $C\ell(G)$, so it is all of $C\ell(G)$, completing the proof. $\square$

Example 8.29. 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 will notate a character χ by the bracketed triple $[\chi(C_1), \chi(C_2), \chi(C_3)]$.

We already know two irreducible 1-dimensional characters of $\mathfrak{S}_3$, namely the trivial character $\chi_{\text{triv}} = [1, 1, 1]$ and the sign character $\chi_{\text{sign}} = [1, -1, 1]$. Equation (8.9) says that $\mathfrak{S}_3$ has three irreps, the squares of whose dimensions add up to $|\mathfrak{S}_3| = 6$ so we are looking for one more irreducible character χ_{other} of dimension 2. By (e) of Theorem 8.28, we have

$$\begin{aligned}
\chi_{\text{reg}} &= \chi_{\text{triv}} + \chi_{\text{sign}} + 2\chi_{\text{other}}\\
\chi_{\text{other}} &= \frac{1}{2}\left(\chi_{\text{reg}} - \chi_{\text{triv}} - \chi_{\text{sign}}\right)\\
&= \frac{1}{2}\left([6, 0, 0] - [1, 1, 1] - [1, -1, 1]\right)\\
&= \frac{1}{2}[4, 0, -2] \;=\; [2, 0, -1].
\end{aligned}$$

Example 8.30. Let's kick it up a notch and calculate all the irreducible characters of $\mathfrak{S}_4$. There are five conjugacy classes, 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 8.28.) 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 8.31. 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 8.28 (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

$$G \text{ abelian} \quad \Longrightarrow \quad \mathrm{Hom}(G, \mathbb{C}^\times) \cong G,$$

an isomorphism known as *Pontrjagin duality*. More generally, for any group G (not necessarily abelian) we have

$$\operatorname{Hom}(G, \mathbb{C}^\times) \cong G^{ab}. \tag{8.10}$$

This is quite useful when computing irreducible characters, because it tells us right away about the one-dimensional characters of an arbitrary group.

Example 8.32. 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 (8.10) gives

$$\operatorname{Hom}(\mathfrak{S}_n, \mathbb{C}^\times) \;\cong\; \mathfrak{S}_n/\mathfrak{A}_n \;\cong\; \mathbb{Z}/2\mathbb{Z}$$

and therefore χ_{triv} and χ_{sign} are the only one-dimensional characters of $\mathfrak{S}_n$. A more elementary way of seeing this is that a one-dimensional character must map the conjugacy class of 2-cycles to either 1 or -1, and the 2-cycles generate all of $\mathfrak{S}_n$, hence determine the character completely.

If you want to compute the characters of $\mathfrak{S}_5$ with your bare hands (exercise!), note that there are 21 multisets of positive integers whose squares add up to $|\mathfrak{S}_5| = 5! = 120$, but only four of them that contain exactly two 1's:

$$1,1,2,2,2,5,9, \quad 1,1,2,2,5,6,7, \quad 1,1,2,3,4,5,8, \quad 1,1,4,4,5,5,6.$$

By examining the defining representation and using the tensor product, you should be able to figure out which one of these is the actual list of dimensions of irreps.

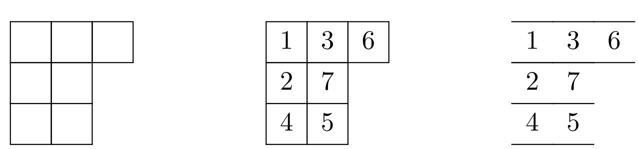
8.8. Characters of the Symmetric Group.

We worked out the irreducible characters of $\mathfrak{S}_4$ and $\mathfrak{S}_5$ *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 set of 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 8.33. 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. We write $\operatorname{sh}(T) = \mu$ to indicate that a tableau or tabloid T is of shape μ.



Ferrers diagram Young tableau Young tabloid

[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.

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. This is the **μ-tabloid representation** of $\mathfrak{S}_n$. Its character is denoted χ_μ.

Example 8.34. For $n = 3$, the characters of the tabloid representations ρ_μ are as follows.

		cycle shape λ		
		111	21	3
Tabloid shape μ	3	1	1	1
	21	3	1	0
	111	6	0	0
	$\lvert C_\lambda \rvert$	1	3	2

(8.11)

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 $\mu = (1, 1, \dots, 1)$ are just the permutations of $[n]$. Therefore
$$\rho_{(1,1,\dots,1)} \cong \rho_{\text{reg}}.$$
- 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 (8.11) is a triangular matrix; in particular it is invertible, so the characters ρ_μ are linearly independent — hence a basis — in the space of class functions. In fact, this is the case for all n. To prove this, we first need to define two orders on the set $\mathrm{Par}(n)$.

Definition 8.35. 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 $k > 0$,

$$\lambda_1 = \mu_1, \quad \lambda_2 = \mu_2, \quad \dots, \quad \lambda_{k-1} = \mu_{k-1}, \quad \lambda_k < \mu_k.$$

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 8.36. The *dominance order* on $\mathrm{Par}(n)$ is defined by

$$\lambda \trianglelefteq \mu \qquad \iff \qquad \sum_{i=1}^{k} \lambda_i \;\le\; \sum_{i=1}^{k} \mu_i \quad \forall k.$$

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 $n \le 5$.) Lexicographic order is a linear extension of dominance order: that is,

$$\lambda \lhd \mu \qquad \Longrightarrow \qquad \lambda < \mu. \tag{8.12}$$

Since the tabloid representations ρ_μ are permutation representations, we can calculate χ_μ by counting fixed points. That is,

$$\chi_\mu(C_\lambda) = \#\{\text{tabloids } T \mid \text{sh}(T) = \mu,\ w(T) = \lambda\} \tag{8.13}$$

for any $w \in C_\lambda$.

Proposition 8.37. *Let $\lambda, \mu \vdash n$. Then:*

(1) $\chi_\lambda(C_\lambda) \neq 0$.
(2) $\chi_\mu(C_\lambda) \neq 0$ *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:

$$\begin{array}{|ccc|}\hline 1 & 3 & 6 \\ \hline 2 & 7 & \\ \hline 4 & 5 & \\ \hline \end{array} \qquad\qquad \begin{array}{|ccc|}\hline 1 & 3 & 6 \\ \hline 4 & 5 & \\ \hline 2 & 7 & \\ \hline \end{array}$$

It follows from (8.13) that $\chi_\lambda(C_\lambda) \neq 0$. (In fact $\chi_\lambda(C_\lambda) = \prod_j r_j!$, where r_j is the number of occurrences of j in λ.)

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 8.38. *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 $\text{Par}(n)$. By Proposition 8.37, 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

$$\begin{aligned} \langle \sigma_\nu,\ \chi_\nu \rangle_G &\neq 0, \\ \langle \sigma_\nu,\ \chi_\mu \rangle_G &= 0 \qquad \text{if } \nu < \mu. \end{aligned} \tag{8.14}$$

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**):

$$V^\mu = \bigoplus_\lambda (\mathsf{Sp}_\lambda)^{\oplus K_{\lambda,\mu}}$$

for some nonnegative integers $K_{\lambda,\mu} = \langle \sigma_\lambda,\ \chi_\mu \rangle_G$.

Example 8.39. Recall the table of characters (8.11) of the tabloid representations for $n = 3$. We will use this to (re-)produce the table of irreducible characters from Example 8.29. Abbreviate $\chi_{(1,1,1)}$ by χ_{111}, etc.

First, $\chi_3 = [1, 1, 1] = \chi_{\text{triv}}$ 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 produces a character orthonormal to σ_3, namely

$$\sigma_{21} \;=\; \chi_{21} - \sigma_3 \;=\; [2, 0, -1].$$

Notice that σ_{21} is irreducible. Finally, for the character χ_{111}, we have

$$\langle \chi_{111},\ \sigma_3 \rangle_G = 1, \qquad\qquad \langle \chi_{111},\ \sigma_{21} \rangle_G = 2.$$

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 (8.14), 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

$$\Big[\chi_\mu\Big]_{\mu\vdash 3} \;=\; \begin{bmatrix}1&1&1\\3&1&0\\6&0&0\end{bmatrix} \;=\; \underbrace{\begin{bmatrix}1&0&0\\1&1&0\\1&2&1\end{bmatrix}}_{K}\begin{bmatrix}1&1&1\\2&0&-1\\1&-1&1\end{bmatrix} \;=\; \Big[K_{\lambda,\mu}\Big]_{\lambda,\mu\vdash 3}\;\;\Big[\sigma_\lambda\Big]_{\lambda\vdash 3}$$

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. Restriction, Induction, and Frobenius Reciprocity.

Definition 8.40. 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 8.41. 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 (the Specht module $\mathsf{Sp}_{(2,1)}$) whose character ψ 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. Indeed

$$\langle\psi,\,\psi\rangle_G = \frac{1}{6}(2^2 + 3\cdot 0^2 + 2\cdot(-1)^2) = 1, \qquad \langle\psi,\,\psi\rangle_H = \frac{1}{3}(2^2 + 2\cdot(-1)^2) = 2.$$

Let $\omega = e^{2\pi i/3}$. Then 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 8.42. Let $H \subset G$ be finite groups, and let $\rho : H \to GL(W)$ be a representation of H. Define the **induced representation** $\mathrm{Ind}_H^G(\rho)$ (alternate notation: $\rho\uparrow_H^G$) as follows. First, choose a set of left coset representatives for H in G, that is, a set $B = \{b_1, \ldots, b_r\}$ such that $G = \bigsqcup_{j=1}^r b_j H$ (here the square cup means disjoint union). Let $\mathbb{C}B$ be the $\mathbb{C}$-vector space with basis B, and let $V = \mathbb{C}B \otimes W = \bigoplus_{i=1}^r b_i \otimes W$.

Now let $g \in G$ act on $b_i \otimes w \in V$ as follows. Find the unique $b_j \in B$ 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 to a representation of G on V by linearity.

Proposition 8.43. $\mathrm{Ind}_H^G(\rho)$ *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

$$gb_i = b_j h \tag{8.15}$$

and in turn there is a unique $b_\ell \in B$ and $h' \in H$ such that

$$g' b_j = b_\ell h'. \tag{8.16}$$

We need to verify that $g' \cdot (g \cdot (b_i \otimes w)) \;=\; (g'g) \cdot (b_i \otimes w)$. 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 (8.15) and (8.16), $gb_i = b_j h b_i^{-1}$ 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 $g \in G$. Using the decomposition $V = \bigoplus_i b_i \otimes W$, we can regard $\mathrm{Ind}_H^G(\rho)(g)$ as a $r \times r$ block matrix with blocks of size $\dim W$, where 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,

$$\begin{aligned}
\chi_{\mathrm{Ind}_H^G(\rho)}(g) &= \sum_{i \in [r]:\ gb_i = b_i h \text{ for some } h \in H} \chi_\rho(h) \\
&= \sum_{i \in [r]:\ b_i^{-1} g b_i \in H} \chi_\rho(b_i^{-1} g b_i) \\
&= \frac{1}{|H|} \sum_{i \in [r]:\ b_i^{-1} g b_i \in H} \ \sum_{h \in H} \chi_\rho(h^{-1} b_i^{-1} g b_i h)
\end{aligned}$$

(because χ_ρ is constant on conjugacy classes of H)

$$= \frac{1}{|H|} \sum_{k \in G:\ k^{-1}gk \in H} \chi_\rho(k^{-1}gk).$$

Here we have put $k = b_i h$, which runs over all elements of G. Note that $k^{-1}gk = h^{-1}b_i^{-1}gb_ih \in H$ if and only if $b_i^{-1}gb_i \in H$, simply because H is a group. The character of $\operatorname{Ind}_H^G(\rho)$ is independent of the choice of B; therefore, so is the representation itself. □

Corollary 8.44. *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 8.45. *Let $H \subset G$ and let ρ be the trivial representation of H. Then*

$$\operatorname{Ind}_H^G \chi_{\text{triv}}(g) = \frac{\#\{k \in G : \ k^{-1}gk \in H\}}{|H|}. \tag{8.17}$$

Example 8.46. 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 8.47. 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 (8.17); 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 8.30, the decomposition into irreducible characters is $\chi_1 + \chi_2 + 2\chi_5$.

Restriction and inducing are related by the following useful formula.

Theorem 8.48 (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.

$$\begin{aligned} \left\langle \operatorname{Ind}_H^G \chi, \ \psi \right\rangle_G &= \frac{1}{|G|} \sum_{g \in G} \overline{\operatorname{Ind}_H^G \chi(g)} \cdot \psi(g) \\ &= \frac{1}{|G|} \sum_{g \in G} \frac{1}{|H|} \sum_{k \in G: \ k^{-1}gk \in H} \overline{\chi(k^{-1}gk)} \cdot \psi(g) && \text{(by Prop. 8.43)} \\ &= \frac{1}{|G||H|} \sum_{h \in H} \sum_{k \in G} \sum_{g \in G: \ k^{-1}gk = h} \overline{\chi(h)} \cdot \psi(k^{-1}gk) \\ &= \frac{1}{|G||H|} \sum_{h \in H} \sum_{k \in G} \overline{\chi(h)} \cdot \psi(h) && \text{(i.e., } g = khk^{-1}) \\ &= \frac{1}{|H|} \sum_{h \in H} \overline{\chi(h)} \cdot \psi(h) \ = \ \left\langle \chi, \ \operatorname{Res}_H^G \psi \right\rangle_H. \end{aligned}$$

□

Example 8.49. Sometimes, Frobenius reciprocity suffices to calculate the isomorphism type of an induced representation. Let ψ, χ_1 and χ_2 be as in Example 8.41. 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.

8.10. **Exercises.** All representations should be assumed to be over $\mathbb{C}$.

Exercise 8.1. Let χ be an irreducible character of G and let ψ be a one-dimensional character. Prove that $\omega := \chi \otimes \psi$ is an irreducible character.

Exercise 8.2. Consider the permutation action of the symmetric group $G = \mathfrak{S}_4$ on the vertices of the complete graph K_4, whose corresponding representation is the defining representation ρ_{def}. Let σ be the 3-dimensional representation corresponding to the action of $\mathfrak{S}_4$ on pairs of opposite edges of K_4.

(a) Compute the character of σ.
(b) Explicitly describe all G-equivariant linear transformations $\phi : \sigma \to \rho_{\text{def}}$. (Hint: Schur's lemma should be useful.)

Exercise 8.3. Work out the character table of $\mathfrak{S}_5$ without using any of the material in Section 8.8. (Hint: To construct another irreducible character, start by considering the action of $\mathfrak{S}_5$ on the edges of the complete graph K_5 induced by the usual permutation action on the vertices.)

Exercise 8.4. Work out the character table of the quaternion group Q. (Recall that Q is the group of order 8 whose elements are $\{\pm 1, \pm i, \pm j, \pm k\}$ with relations $i^2 = j^2 = k^2 = -1$, $ij = k$, $jk = i$, $ki = j$.)

Exercise 8.5. Work out the characters of the Specht modules Sp_λ for all $\lambda \vdash 4$. (This should not be hard — start with the characters of the tabloid representations, then do linear algebra; feel free to use a computer algebra system if you want.) Compare your result to the character table of $\mathfrak{S}_4$ calculated ad hoc in Example 8.30. Can you make any observations or conjectures about the character of Sp_λ in terms of λ?

Exercise 8.6. Recall that the *alternating group* $\mathfrak{A}_n$ consists of the $n!/2$ even permutations in $\mathfrak{S}_n$, that is, those with an even number of even-length cycles.

(a) Show that the conjugacy classes in $\mathfrak{A}_4$ are not simply the conjugacy classes in $\mathfrak{S}_4$. (Hint: Consider the possibilities for the dimensions of the irreducible characters of $\mathfrak{A}_4$.)
(b) Determine the conjugacy classes in $\mathfrak{A}_4$, and the complete list of irreducible characters.
(c) Use this information to determine $[\mathfrak{A}_4, \mathfrak{A}_4]$ without actually computing any commutators.

9. Symmetric Functions

9.1. Prelude: Symmetric polynomials.

Definition 9.1. Let R be a commutative ring, typically $\mathbb{Q}$ or $\mathbb{Z}$. A **symmetric polynomial** is a polynomial in $R[x_1, \dots, x_n]$ that is invariant under permuting the variables.

Note that the symmetric polynomials that are homogeneous of degree d form an R-module.

For example, if $n = 3$, then up to scalar multiplication, the only symmetric polynomial 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 polynomial 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 polynomial forms a vector space.

Here is a basis for the space of degree 3 symmetric polynomials:
$$\begin{aligned}
&x_1^3 + x_2^3 + x_3^3,\\
&x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2,\\
&x_1x_2x_3.
\end{aligned}$$

Each member of this basis is a sum of the monomials in a single orbit under the action of $\mathfrak{S}_3$. Accordingly, we can index them 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 *countably infinite* set of variables $\{x_1,x_2,\dots\}$, which means that we need to work not with polynomials, but with. . .

9.2. **Formal power series.** Let R be a ring, and let $\mathbf{x}=\{x_i:\ i\in I\}$ be a set of commuting indeterminates indexed by I. (Typically $I=\mathbb{P}$, but we may as well make the definition general.) A **monomial** is an object $\mathbf{x}^\alpha=\prod_{i\in I}x_i^{\alpha_i}$, where $\alpha=(\alpha_i)_{i\in I}\in\mathbb{N}^I$ and $\sum_{i\in I}\alpha_i$ is finite (equivalently, all but finitely many of the α_i are zero). A **formal power series** is an object of the form

$$\sum_{\alpha\in\mathbb{N}^I} c_\alpha\mathbf{x}^\alpha$$

with $c_\alpha\in R$ for all α. The set $R[[\mathbf{x}]]$ of all formal power series is easily seen to be an abelian group under addition. In fact, it is a ring as well, with multiplication given by

$$\left(\sum_{\alpha\in\mathbb{N}^I}c_\alpha\mathbf{x}^\alpha\right)\left(\sum_{\beta\in\mathbb{N}^I}d_\beta\mathbf{x}^\beta\right)=\sum_{\gamma\in\mathbb{N}^I}\left(\sum_{(\alpha,\beta):\ \alpha+\beta=\gamma}c_\alpha d_\beta\right)\mathbf{x}^\gamma,$$

because the inner sum on the right-hand side has only finitely many terms for each γ, and is thus a well-defined element of R.

In the ring of formal power series, a function can be identified with its Taylor expansion. For example, it makes sense to say

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

because the identity $(1-x)(1+x+x^2+\cdots)=1$ holds in $\mathbb{Z}[[x]]$. Note that the interval of convergence does not matter, since x is merely an indeterminate, not a variable that takes values in (say) $\mathbb{R}$.

Familiar functions from analysis (like exp and ln) can be equated with their formal power series, which frequently carry combinatorial meaning. In addition, identities from analysis hold in $R[[x]]$, often for combinatorial reasons (see Exercise 9.2 for an example).

9.3. **Symmetric functions.** We can now define symmetric functions properly, as elements of the ring of formal power series.

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. Equivalently,

$$m_\lambda=\sum_\alpha\mathbf{x}^\alpha$$

where the sum ranges over all rearrangements α of λ (regarding λ as a countably infinite sequence in which all but finitely many terms are 0).

We then define

$$\Lambda_d = \Lambda_{R,d}(\mathbf{x}) = \{\text{degree-}d \text{ symmetric functions in indeterminates } \mathbf{x} \text{ with coefficients in } R\},$$
$$\Lambda = \Lambda_R(\mathbf{x}) = \bigoplus_{d\geq 0} \Lambda_d.$$

Each Λ_d is a finitely generated free R-module, with basis $\{m_\lambda \mid \lambda \vdash d\}$. Moreover, Λ is a graded R-algebra. 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 a graded 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.

9.4. **Elementary symmetric functions.** For $k \in \mathbb{N}$ we define

$$e_k \quad = \quad \sum_{\substack{S\subset\mathbb{P} \\ |S|=k}} \prod_{s\in S} x_s \quad = \quad \sum_{0<i_1<i_2<\cdots<i_k} x_{i_1} x_{i_2} \cdots x_{i_k} \quad = \quad m_{11\cdots 1}$$

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

$$\begin{array}{rcl@{\quad}rcrcr}
e_3 & = & \sum_{i<j<k} x_i x_j x_k & = & & & & & m_{111}, \\
e_{21} & = & (x_1+x_2+x_3+\cdots)(x_1x_2+x_1x_3+x_2x_3+x_1x_4+\cdots) & = & & & m_{21} & + & 3m_{111}, \\
e_{111} & = & (x_1+x_2+x_3+\cdots)^3 & = & m_3 & + & 3m_{21} & + & 6m_{111}.
\end{array}$$

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 $R[[t, x_1, x_2, \dots]]$.

Recall the dominance partial order $\trianglerighteq$ on partitions from Definition 8.36.

Theorem 9.3. *Let $b_{\lambda,\mu}$ be the coefficient of e_λ when expanded in the monomial basis, that is,*

$$e_\lambda = \sum_\mu b_{\lambda,\mu} m_\mu.$$

Then $b_{\lambda,\tilde{\lambda}} = 1$, and $b_{\lambda,\mu} = 0$ unless $\tilde{\lambda} \trianglerighteq \mu$. In particular $\{e_\lambda \mid \lambda \vdash n\}$ is a vector space basis for Λ_n.

Proof. It suffices to figure out the coefficient of the monomial $\mathbf{x}^\mu = \prod_i x_i^{\mu_i}$ in the expansion of e_λ. That coefficient will be the number of ways to factorize $\mathbf{x}^\mu$ into squarefree pieces whose degrees are the parts of λ.

I claim that $b_{\lambda\mu}$ is the number of ways to put balls in boxes as follows. Suppose that we have λ_1 balls labeled 1, λ_2 balls labeled 2, etc. We also have countably infinitely many boxes. Place the balls in the boxes so that box i receives μ_i balls, all with different labels.

For example, let $\mu = (3,2,2)$ and $\lambda = (3,2,1,1)$, so that we are trying to count the ways of factoring $\mathbf{x}^\mu = x_1^2 x_2^2 x_3^2$ into four squarefree monomials of degrees 3, 2, 1, 1. One such factorization is

$$x_1^3 x_2^2 x_3^2 \;=\; x_1x_2x_3 \cdot x_1x_3 \cdot x_2 \cdot x_1$$

which corresponds to the balls-and-boxes picture

Box 1	Box 2	Box 3
1 2 4	1 3	1 2

in which a ball labeled j in box i indicates that the j^{th} factor included x_i.

Now observe that the first k boxes collectively can contain no more than k balls with each label. That is,

$$\mu_1 + \cdots + \mu_k \leq \tilde{\lambda}_1 + \cdots + \tilde{\lambda}_k \qquad (\forall k)$$

which is precisely the statement that $b_{\lambda,\mu} = 0$ unless $\tilde{\lambda} \trianglerighteq \mu$. Moreover, if $\tilde{\lambda} = \mu$ then there is exactly one way to do this: place balls labeled $1, \dots, \tilde{\lambda}_i$ in the i^{th} box. Hence $b_{\lambda,\tilde{\lambda}} = 1$.

Therefore, if we order partitions of n by any linear extension of dominance (such as the lexicographic order), then the matrix $[b_{\lambda,\mu}]$ will be seen to be upper unitriangular. In particular it is invertible over $\mathbb{Z}$, and it follows that the $\mathbb{Z}$-module spanned by the e_λ's is the same as that spanned by the m_μ's, namely $\Lambda_{\mathbb{Z},d}$. □

Corollary 9.4 ("Fundamental Theorem of Symmetric Functions")**.** *The elementary symmetric functions $e_1, e_2, \dots$ are algebraically independent. Therefore, $\Lambda = \mathbb{C}[e_1, e_2, \dots]$ as rings.*

Proof. Given any nontrivial polynomial relation among the e_i's, extracting the homogeneous pieces would give a nontrivial linear relation among the e_λ's, which does not exist. □

9.5. **Complete homogeneous symmetric functions.** For $k \in \mathbb{N}$, we define h_k to be the sum of *all* monomials of degree k:

$$h_k = \sum_{0<i_1\leq i_2 \leq \cdots \leq i_k} x_{i_1} x_{i_2} \cdots x_{i_k} = \sum_{\lambda \vdash k} m_\lambda.$$

(So $h_0 = 1$.) 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

$$\begin{aligned}
h_{111} &= m_1^3 \;=\; (x_1 + x_2 + x_3 + \cdots)^3 \;=\; 6m_{111} + 3m_{21} + m_3,\\
h_{21} &= h_1 h_2 \;=\; e_1(m_{11} + m_2) \;=\; e_1(e_{11} - e_2) \;=\; e_{111} - e_{21} = m_3 + 2m_{21} + 3m_{111},\\
h_3 &= m_{111} + m_{21} + m_3.
\end{aligned}$$

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,

$$h_1 - e_1 = 0, \quad h_2 - e_1h_1 + e_2 = 0, \quad h_3 - e_1h_2 + e_2h_1 - e_3 = 0, \quad \dots$$

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. So again $\{h_\lambda \mid \lambda \vdash n\}$ is a vector space basis for Λ_n, and $h_1, h_2, \dots$ generate Λ as a $\mathbb{C}$-algebra.

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 9.5. *$\omega(\omega(f)) = f$ 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 fir every $n \geq 1$

$$\begin{aligned} 0 &= \sum_{k=0}^{n} (-1)^{n-k} \omega(e_k)\omega(h_{n-k}) \\ &= \sum_{k=0}^{n} (-1)^{n-k} h_k \omega(h_{n-k}) \\ &= \sum_{k=0}^{n} (-1)^{k} h_{n-k} \omega(h_k) \\ &= (-1)^n \sum_{k=0}^{n} (-1)^{n-k} h_{n-k} \omega(h_k) \end{aligned}$$

and comparing this last expression with the original Jacobi-Trudi relations gives $\omega(h_k) = e_k$. $\square$

Corollary 9.6. *$\{h_\lambda\}$ is a graded $\mathbb{Z}$-basis for Λ. Moreover, $\Lambda_R \cong R[h_1, h_2, \dots]$ as rings.*

9.6. **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,

$$p_2 = m_2,$$
$$p_{11} = (x_1 + x_2 + \cdots)^2 = m_2 + 2m_{11}.$$

While the p_λ are a vector space basis for $\Lambda_{\mathbb{Q}}$ (the proof is left as an exercise), they are not a $\mathbb{Z}$-module basis for $\Lambda_{\mathbb{Z}}$. In other words, not every symmetric function with integer coefficients can be expressed as an integer combination of the power-sums; for example, $m_{11} = (p_{11} - p_2)/2$.

Expanding the power-sum symmetric functions in the monomial basis will provide the first example of the deep connection between representation theory and symmetric functions.

Proposition 9.7. *For $\lambda \vdash n$, we have*

$$p_\lambda = \sum_{\mu \vdash n} \chi_\mu(C_\lambda) m_\mu$$

where χ_μ is the character of the tabloid representation of shape μ (see Section 8.8) and C_λ is the conjugacy class of cycle-shape λ.

The proof is left as an exercise.

9.7. **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 9.8. 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 $\mathsf{CST}(\lambda)$. 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 $\mathbf{x}^T = \mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots$. For example:

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 3 \\ \hline 2 & 3 & \\ \cline{1-2}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 4 & 8 & \\ \cline{1-2}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 1 & 4 & \\ \cline{1-2}\end{array}$$

$$x_1^2 x_2 x_3^2 \qquad\qquad x_1^3 x_4 x_8 \qquad\qquad \text{Not a CST}$$

Definition 9.9. The **Schur function** corresponding to a partition λ is

$$s_\lambda \;=\; \sum_{T \in \mathsf{CST}(\lambda)} \mathbf{x}^T.$$

It is far from obvious that s_λ is symmetric, but in fact it is! We will prove this shortly.

Example 9.10. 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 $\mathsf{CST}(\lambda)$ 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.

$$\begin{array}{|c|c|}\hline a & b \\ \hline c & \\ \cline{1-1}\end{array} \qquad\qquad \begin{array}{|c|c|}\hline a & c \\ \hline b & \\ \cline{1-1}\end{array}$$

Therefore, the coefficient of m_{111} is 1.

Finally, for every $a \neq b \in \mathbb{P}$, 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 $a > b$.

$$\begin{array}{|c|c|}\hline a & b \\ \hline b & \\ \cline{1-1}\end{array} \qquad\qquad \begin{array}{|c|c|}\hline b & b \\ \hline a & \\ \cline{1-1}\end{array}$$

Therefore, $s_{(2,1)} \;=\; 2m_{111} + m_{21}$.

Proposition 9.11. *s_λ is a symmetric function for all λ.*

Proof. First, observe that the number

$$c(\lambda,\alpha) = |\{T \in \mathsf{CST}(\lambda) \mid \mathbf{x}^T = \mathbf{x}^\alpha\}| \tag{9.7}$$

depends only on the ordered sequence of nonzero exponents[12] 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

$$x_1^3\,x_2^2\,x_3^3\,x_4^3\,\underline{\boldsymbol{x_5^4}\,\boldsymbol{x_6^7}}\,x_7^3 \quad \text{and} \quad x_1^3\,x_2^2\,x_3^3\,x_4^3\,\underline{\boldsymbol{x_5^7}\,\boldsymbol{x_6^4}}\,x_7^3.$$

Let T be the following λ-CST:

[12] This is precisely the statement that s_λ is a **quasisymmetric** function.

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							

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 9.12. *The Schur functions $\{s_\lambda \mid \lambda \vdash n\}$ are a $\mathbb{Z}$-basis for $\Lambda_\mathbb{Z}$.*

Proof. 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 if T is a CST of shape λ and content μ (so in particular $K_{\lambda\mu} > 0$), then

- 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 $\sum_{i=1}^k \mu_i \leq \sum_{i=1}^k \lambda_i$, i.e., which is just the statement that $\lambda \trianglerighteq \mu$. But that means that the matrix $[K_{\lambda\mu}]_{\lambda,\mu \vdash n}$ is unitriangular, and the Schur functions are a vector space basis for $\Lambda_\mathbb{Q}$ and a free module basis for $\Lambda_\mathbb{Z}$, just as in the proof of Theorem 9.3. ∎

9.8. **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

$$\Omega = \prod_{i,j \geq 1} (1 - x_i y_j)^{-1}, \qquad \Omega^* = \prod_{i,j \geq 1} (1 + x_i y_j).$$

These series are symmetric with respect to each of the variable sets $\mathbf{x} = \{x_1, x_2, \dots\}$ and $\mathbf{y} = \{y_1, y_2, \dots\}$, so they can be regarded as elements of $\Lambda(\mathbf{x}) \otimes \Lambda(\mathbf{y})$. 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[13]

$$z_\lambda = 1^{r_1} r_1! \, 2^{r_2} r_2! \, \cdots, \qquad \varepsilon_\lambda = (-1)^{r_2 + r_4 + \cdots}. \tag{9.9}$$

For example, if $\lambda = (3,3,2,1,1,1)$ then $z_\lambda = 1^3 3! \, 2^1 1! \, 3^2 2! = 216$. Note that ε_λ is just the sign of the conjugacy class C_λ, and

$$|C_\lambda| = n!/z_\lambda. \tag{9.10}$$

To see this, take any permutation w written in one-line notation and interpret it as a permutation $\sigma \in C_\lambda$ by breaking it into cycles: e.g., if $\lambda = (3,3,2,1,1,1)$ as above, then the one-line permutation $w = (8,1,6,4,9,2,11,5,10,7,3)$ becomes the cycle-notation permutation $(8\ 1\ 6)(4\ 9\ 2)(11\ 5)(10)(7)(3)$. Then z_λ is the number of w's that give rise to a given σ, since permuting cycles of the same length, or cyclically permuting the numbers within a single cycle, does not change σ. Thus z_λ itself 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) — consider the transitive action of $\mathfrak{S}_n$ on C_λ by conjugation.

[13]Stanley uses m_i where I am using r_i. I want to avoid conflict with the notation for monomial symmetric functions.

Proposition 9.13. *We have*

$$\prod_{i,j\geq 1}(1-x_iy_j)^{-1} = \sum_\lambda h_\lambda(\mathbf{x})m_\lambda(\mathbf{y}) = \sum_\lambda \frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}, \tag{9.11}$$

$$\prod_{i,j\geq 1}(1+x_iy_j) = \sum_\lambda e_\lambda(\mathbf{x})m_\lambda(\mathbf{y}) = \sum_\lambda \varepsilon_\lambda \frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}, \tag{9.12}$$

where the sums run over all partitions λ.

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

$$\log(1+q) = \sum_{n\geq 1}(-1)^{n+1}\frac{q^n}{n}, \qquad \exp(q) = \sum_{n\geq 0}\frac{q^n}{n!}. \tag{9.14}$$

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,

$$\begin{aligned}\log\Omega = \log\prod_{i,j\geq 1}(1-x_iy_j)^{-1} &= -\log\prod_{i,j\geq 1}(1-x_iy_j) = -\sum_{i,j\geq 1}\log(1+(-x_iy_j))\\ &= \sum_{i,j\geq 1}\sum_{n\geq 1}\frac{x_i^ny_j^n}{n} = \sum_{n\geq 1}\frac{1}{n}\left(\sum_{i\geq 1}x_i^n\right)\left(\sum_{j\geq 1}y_j^n\right)\\ &= \sum_{n\geq 1}\frac{p_n(\mathbf{x})p_n(\mathbf{y})}{n}\end{aligned}$$

and now exponentiating both sides and applying the power series expansion for exp, we get

$$\begin{aligned}\Omega &= \exp\left(\sum_{n\geq 1}\frac{p_n(\mathbf{x})p_n(\mathbf{y})}{n}\right) = \sum_{k\geq 0}\frac{1}{k!}\left(\sum_{n\geq 1}\frac{p_n(\mathbf{x})p_n(\mathbf{y})}{n}\right)^k\\ &= \sum_{k\geq 0}\frac{1}{k!}\left[\sum_{\lambda:\,\ell(\lambda)=k}\binom{k}{r_1!\,r_2!\,\dots}\left(\frac{p_1(\mathbf{x})p_1(\mathbf{y})}{1}\right)^{r_1(\lambda)}\left(\frac{p_2(\mathbf{x})p_2(\mathbf{y})}{2}\right)^{r_2(\lambda)}\cdots\right]\\ &= \sum_\lambda \frac{\prod_{i=1}^\infty (p_i(\mathbf{x})p_i(\mathbf{y}))^{r_i(\lambda)}}{\prod_{i=1}^\infty i^{r_i}r_i!} = \sum_\lambda \frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}.\end{aligned}$$

□

The proofs of the identities for the dual Cauchy kernel are analogous, and are left as an exercise.

Corollary 9.14. *For all n, we have* *(i)* $h_n \;=\; \sum_{\lambda\vdash n} p_\lambda/z_\lambda$;

(ii) $e_n \;=\; \sum_{\lambda\vdash n} \varepsilon_\lambda p_\lambda/z_\lambda$;

(iii) $\omega(p_\lambda) \;=\; \varepsilon_\lambda p_\lambda$.

Proof. (i) 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 the desired expression for h_n.

(ii) Start with (9.12) and do the same thing.

(iii) Let ω act on symmetric functions in $\mathbf{x}$ while fixing those in $\mathbf{y}$. (Technically we should write $\omega\otimes\mathrm{id}$, since we are mapping $\Lambda\otimes\Lambda$ to itself.) 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. ☐

Definition 9.15. The **Hall inner product** on symmetric functions is defined by

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

and extending by linearity to all of Λ.

By linear algebra, if $\{u_\lambda\}$ and $\{v_\mu\}$ are graded bases for Λ indexed by partitions such that $\Omega = \sum_\lambda u_\lambda(\mathbf{x})v_\lambda(\mathbf{y})$, then they are orthogonal to each other with respect to the Hall inner product; i.e., $\langle u_\lambda, v_\mu\rangle = \delta_{\lambda\mu}$. Thus $\{p_\lambda\}$ and $\{p_\lambda/z_\lambda\}$ are orthogonal, and $\{p_\lambda/\sqrt{z_\lambda}\}$ 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.)

The involution ω is an isometry with respect to the Hall inner product, 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 (iii) of Corollary 9.14, we have

$$\langle\omega p_\lambda, \omega p_\mu\rangle = \langle\varepsilon_\lambda p_\lambda, \varepsilon_\lambda p_\mu\rangle = \varepsilon_\lambda^2\langle p_\lambda, p_\mu\rangle = \langle p_\lambda, p_\mu\rangle$$

because $\varepsilon_\lambda \in \{1, -1\}$ for all λ.

The basis $\{p_\lambda/\sqrt{z_\lambda}\}$ is orthonormal, but it is not at all nice from a combinatorial point of view, because it involves irrational coefficients. There is a much better orthonormal basis, namely the Schur functions. The next goal will be to prove that

$$\Omega \;=\; \prod_{i,j\geq 1}\frac{1}{1-x_iy_j} \;=\; \sum_\lambda s_\lambda(\mathbf{x})s_\lambda(\mathbf{y}). \tag{9.15}$$

The proof requires a marvelous bijection called the **RSK correspondence** (for Robinson, Schensted and Knuth).

9.9. **The RSK Correspondence.**

Definition 9.16. 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 $\mathsf{SYT}(\lambda)$ for the set of all standard tableaux of shape λ, and set $f^\lambda = |\,\mathsf{SYT}(\lambda)|$.

For example, if $\lambda = (3,3)$, then $f^\lambda = 5$; the members of $\mathsf{SYT}(\lambda)$ 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 $\mathsf{RSK}(w) = (P, Q)$ of standard tableaux of the same shape $\lambda \vdash n$, using the following *row-insertion* operation defined as follows.

Definition 9.17. Let T be a column-strict tableau and let $x \in \mathbb{P}$. 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 leftmost 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 9.18. Let $w = 57214836 \in \mathfrak{S}_8$. We start with a pair (P, Q) of empty tableaux.

<u>Step 1: Row-insert $w_1 = 5$ into P.</u> We do this in the obvious way. Since it's the first cell added, we add a cell containing 1 to Q.

$$P = \begin{array}{|c|}\hline 5\\\hline\end{array} \qquad Q = \begin{array}{|c|}\hline 1\\\hline\end{array} \tag{9.16a}$$

<u>Step 2: Row-insert $w_2 = 7$ into P.</u> 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.

$$P = \begin{array}{|c|c|}\hline 5&7\\\hline\end{array} \qquad Q = \begin{array}{|c|c|}\hline 1&2\\\hline\end{array} \tag{9.16b}$$

<u>Step 3: Row-insert $w_3 = 2$ into P.</u> 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.

$$P = \begin{array}{|c|c|}\hline 2 & 7 \\ \hline 5 \\ \cline{1-1}\end{array} \qquad Q = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \tag{9.16c}$$

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.

$$P = \begin{array}{|c|c|}\hline 1 & 7 \\ \hline 2 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array} \qquad Q = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \tag{9.16d}$$

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.

$$P = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 7 \\ \hline 5 \\ \cline{1-1}\end{array} \qquad Q = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 \\ \cline{1-1}\end{array} \tag{9.16e}$$

Step 6: Row-insert $w_6 = 8$ into P. The 8 just goes at the end of the first row.

$$P = \begin{array}{|c|c|c|}\hline 1 & 4 & 8 \\ \hline 2 & 7 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array} \qquad Q = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \tag{9.16f}$$

Step 7: Row-insert $w_7 = 3$ into P. 3 bumps 4, and then 4 bumps 7.

$$P = \begin{array}{|c|c|c|}\hline 1 & 3 & 8 \\ \hline 2 & 4 \\ \cline{1-2} 5 & 7 \\ \cline{1-2}\end{array} \qquad Q = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 \\ \cline{1-2} 4 & 7 \\ \cline{1-2}\end{array} \tag{9.16g}$$

Step 8: Row-insert $w_8 = 6$ into P. 6 bumps 8 into the second row.

$$P = \begin{array}{|c|c|c|}\hline 1 & 3 & 6 \\ \hline 2 & 4 & 8 \\ \hline 5 & 7 \\ \cline{1-2}\end{array} \qquad Q = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 & 8 \\ \hline 4 & 7 \\ \cline{1-2}\end{array} \tag{9.16h}$$

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 9.19. Suppose that we were given the pair of tableaux in (9.16h). 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 (9.16g).

Example 9.20. 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 indicated in boldface 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	**10**	18
3	4	11	**12**		
6	7	**13**			
9	**15**	17			
14	**16**				

$\longleftarrow$

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 on the right.

Iterating this "de-insertion" allows us to recover w from the pair (P, Q). We have proved the following fact:

Theorem 9.21. *The RSK correspondence is a bijection*

$$\mathfrak{S}_n \xrightarrow{\mathsf{RSK}} \bigcup_{\lambda \vdash n} \mathsf{SYT}(\lambda) \times \mathsf{SYT}(\lambda).$$

Corollary 9.22. $\sum_{\lambda \vdash n} (f^\lambda)^2 = n!$.

To confirm this for $n = 3$, here are all the SYT's with 3 boxes:

$$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array} \qquad \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array} \qquad \begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}\,.$$

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 9.23. *Let $w \in \mathfrak{S}_n$. If $\mathsf{RSK}(w) = (P, Q)$, then $\mathsf{RSK}(w^{-1}) = (Q, P)$. In particular, the number of involutions in $\mathfrak{S}_n$ is* $\displaystyle\sum_{\lambda \vdash n} f^\lambda$.

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 [Sta99, §7.13] from which the symmetry is more apparent.

9.10. **Knuth equivalence and jeu de taquin.**

Definition 9.24. Let $\mathbf{b}, \mathbf{b}'$ be finite ordered lists of positive integers (or "words in the alphabet $\mathbb{P}$"). We say that $\mathbf{b}, \mathbf{b}'$ are *Knuth equivalent*, written $\mathbf{b} \sim \mathbf{b}'$, if one can be obtained from the other by a sequence of transpositions as follows:

1. If $x \leq y < z$, then $\cdots xzy \cdots \ \sim \ \cdots zxy \cdots$.
2. If $x < y \leq z$, then $\cdots yxz \cdots \ \sim \ \cdots yzx \cdots$.

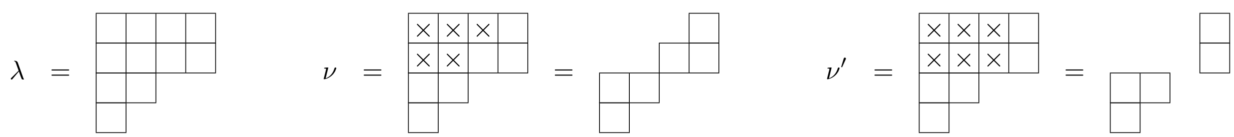
(Here the notation $\cdots xzy \cdots$ means a word that contains the letters x, z, y consecutively.)

For example, $2122\underline{131}2 \sim 2122\underline{311}2$ by Rule 1, and $212\underline{231}12 \sim 212\underline{213}12$ by Rule 2 (applied in reverse).

This definition looks completely unmotivated at first, but hold that thought!

Definition 9.25. 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 9.26. 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 9.8.) For example, here are two column-strict skew tableaux:

Again, there is no requirement that a skew tableau be connected.

We now define an equivalence relation on column-strict skew tableaux, called *jeu de taquin*[14]. The rule is as follows:

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.

[14] French for "sliding game", roughly; it refers to the 15-square puzzle with sliding tiles that used to come standard on every Macintosh in about 1985.

For example, the following is a sequence of jeu de taquin moves. The bullets $\bullet$ denote the inner corner that is being slid into.

$$\begin{array}{ccc} & \bullet & 1 & 4 \\ & 1 & 2 & \\ 2 & 3 & 4 & \end{array} \to \begin{array}{cccc} & 1 & 1 & 4 \\ & \bullet & 2 & \\ 2 & 3 & 4 & \end{array} \to \begin{array}{cccc} & 1 & 1 & 4 \\ & 2 & \bullet & \\ 2 & 3 & 4 & \end{array} \to \begin{array}{cccc} & 1 & 1 & 4 \\ \bullet & 2 & 4 & \\ 2 & 3 & & \end{array} \to \begin{array}{cccc} & 1 & 1 & 4 \\ 2 & 2 & 4 & \\ \bullet & 3 & & \end{array}$$

$$\to \begin{array}{cccc} \bullet & 1 & 1 & 4 \\ 2 & 2 & 4 & \\ 3 & & & \end{array} \to \begin{array}{cccc} 1 & \bullet & 1 & 4 \\ 2 & 2 & 4 & \\ 3 & & & \end{array} \to \begin{array}{cccc} 1 & 1 & \bullet & 4 \\ 2 & 2 & 4 & \\ 3 & & & \end{array} \to \begin{array}{cccc} 1 & 1 & 4 & 4 \\ 2 & 2 & & \\ 3 & & & \end{array} \tag{9.17}$$

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 $T \approx T'$. 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 (9.17) above. In fact, the rectification is unique; the order in which we choose inner corners does not matter.

Definition 9.27. Let T be a column-strict skew tableau. The *row-reading word* of T, denoted $\mathrm{row}(T)$, is obtained by reading the rows left to right, bottom to top.

For example, the reading words of the skew tableaux in (9.17) 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 $\mathrm{row}(T)$. For skew tableaux, this is not the case. Note that some of the slides in (9.17) do not change the row reading word; as a simpler example, the following skew tableaux both have reading word 122:

$$\begin{array}{ccc} 1 & 2 & 2 \end{array} \qquad \begin{array}{ccc} & & 2 \\ 1 & 2 & \end{array} \qquad \begin{array}{ccc} & 2 & 2 \\ 1 & & \end{array}$$

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 $\mathbb{P}$, 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.)

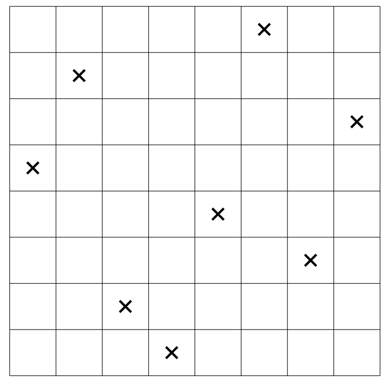
Theorem 9.28. *(Knuth–Schützenberger) For two words $\mathbf{b}, \mathbf{b}'$, the following are equivalent:*

(1) $P(\mathbf{b}) = P(\mathbf{b}')$.
(2) $\mathbf{b} \sim \mathbf{b}'$.
(3) $T \approx T'$, 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 [Ful97]) as the equivalence of "bumping" (the RSK algorithm as presented in Section 9.9) and "sliding" (jeu de taquin).

9.11. **Yet another 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 $\underline{\text{it must be the case that } \mu_i = \lambda_i + 1 \text{ for some } i}$. Obtain ρ from μ by incrementing μ_{i+1}.

Rule 3: If $\mu \neq \nu$, then set $\rho = \mu \vee \nu$ (where $\vee$ means the join in Young's lattice: i.e., take the componentwise maximum of the elements of μ and ν).

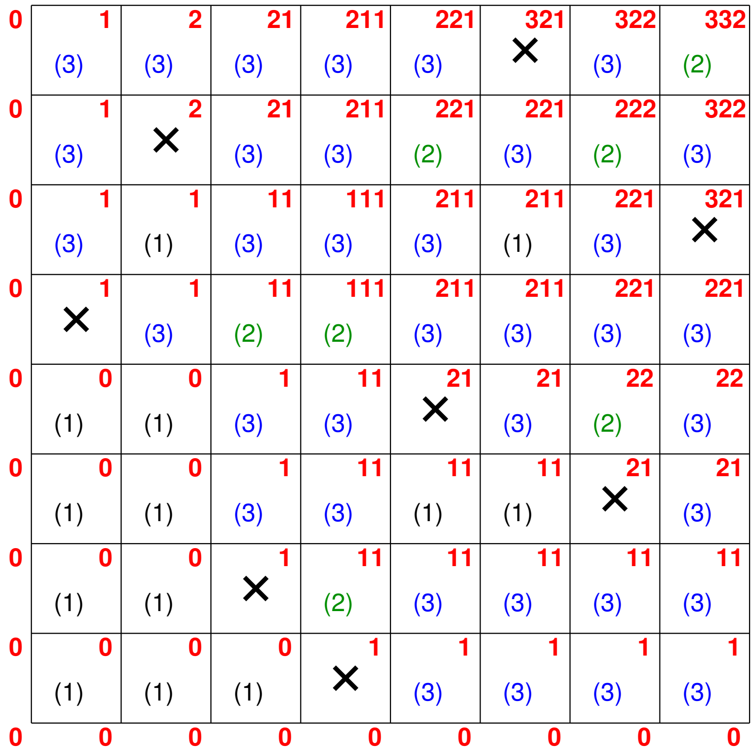
Rule X: If there is an X in the box, then $\underline{\text{it must be the case that } \lambda = \mu = \nu}$. Obtain ρ from λ by incrementing λ_1.

Note that the underlined assertions need to be proved; this can be done by induction.

Example 9.29. Let $n = 8$ and $w = 57214836$. In Example 9.18, we found that $\mathsf{RSK}(w) = (P, Q)$, where

$$P = \begin{array}{|c|c|c|}\hline 1 & 3 & 6 \\ \hline 2 & 4 & 8 \\ \hline 5 & 7 \\ \cline{1-2} \end{array} \qquad \text{and} \qquad Q = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 & 8 \\ \hline 4 & 7 \\ \cline{1-2} \end{array}\,.$$

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.



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.
- It follows that 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 SYT's are the P and Q of the RSK correspondence!

9.12. **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

$$w = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \cdots & q_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \tag{9.18}$$

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 9.30. 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:

Step #	P	Q
1	$\begin{ytableau} 2 \end{ytableau}$	$\begin{ytableau} 1 \end{ytableau}$
2	$\begin{ytableau} 2 & 4 \end{ytableau}$	$\begin{ytableau} 1 & 1 \end{ytableau}$
3	$\begin{ytableau} 1 & 4 \\ 2 \end{ytableau}$	$\begin{ytableau} 1 & 1 \\ 2 \end{ytableau}$
4	$\begin{ytableau} 1 & 1 \\ 2 & 4 \end{ytableau}$	$\begin{ytableau} 1 & 1 \\ 2 & 4 \end{ytableau}$
	$\vdots$	$\vdots$
9	$\begin{ytableau} 1 & 1 & 2 & 2 & 4 \\ 2 & 3 & 3 \\ 4 \end{ytableau}$	$\begin{ytableau} 1 & 1 & 4 & 4 & 5 \\ 2 & 4 & 5 \\ 5 \end{ytableau}$

The tableaux P, Q we get in this way will always have the same shape as each other, and will 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

$$\left\{\text{generalized permutations } \binom{\mathbf{q}}{\mathbf{p}} \text{ of length } n\right\} \xrightarrow{\mathsf{RSK}} \bigcup_{\lambda \vdash n} \{(P, Q) \mid P, Q \in \mathsf{CST}(\lambda)\}\,. \tag{9.19}$$

in which the weights of the tableaux P, Q are $\mathbf{x}^P = x_{p_1} \cdots x_{p_n}$ and $\mathbf{x}^Q = x_{q_1} \cdots x_{q_n}$.

On the other hand, a generalized permutation $w = \binom{\mathbf{q}}{\mathbf{p}}$ as in (9.18) can be specified by an infinite matrix $M = [m_{ij}]_{i,j \in \mathbb{P}}$ with finitely many nonzero entries, in which m_{ij} is the number of occurrences of (q_i, p_i) in w. For example, the generalized permutation w of Example 9.30 corresponds to the integer matrix

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 2 & 0 & 0 & \cdots \\ 0 & 2 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Under this bijection, the monomial $\mathbf{x}^P\mathbf{y}^Q$ equals

$$\prod_{i,j}(x_iy_j)^{m_{ij}} = \prod_{i=1}^{\infty} x_i^{\text{sum of } i^{th} \text{ row}} \prod_{j=-1}^{\infty} y_j^{\text{sum of } j^{th} \text{ column}}.$$

The generating function for matrices by these weights is...

$$\sum_{M=[m_{ij}]\in\mathbb{N}^{n\times n}} \prod_{i,j}(x_iy_j)^{m_{ij}} \;=\; \prod_{i,j\geq 1}\sum_{m_{ij}=0}^{\infty}(x_iy_j)^{m_{ij}} \;=\; \prod_{i,j\geq 1}\frac{1}{1-x_iy_j} \;=\; \Omega.$$

...the Cauchy kernel! On the other hand,

$$\begin{aligned}
\Omega = \prod_{i,j\geq 1}\frac{1}{1-x_iy_j} &= \sum_{M=[m_{ij}]\in\mathbb{N}^{n\times n}} \prod_{i,j}(x_iy_j)^{m_{ij}} = \sum_{n\in\mathbb{N}}\sum_{w=\binom{\mathbf{q}}{\mathbf{p}}} x_{a_1}\cdots x_{a_n}\, y_{b_1}\cdots y_{b_n} \\
&= \sum_{\lambda}\sum_{P,Q\,\in\,\mathsf{CST}(\lambda)} x^P y^Q && \text{(by RSK)} \\
&= \sum_{\lambda}\left(\sum_{P\in\mathsf{CST}(\lambda)} x^P\right)\left(\sum_{Q\in\mathsf{CST}(\lambda)} y^Q\right) \quad = \quad \sum_{\lambda} s_\lambda(x)s_\lambda(y).
\end{aligned}$$

We have proven:

Theorem 9.31. $\Omega = \sum_\lambda s_\lambda(x)s_\lambda(y)$. *Consequently, the Schur functions form an orthonormal $\mathbb{Z}$-basis for Λ under the Hall inner product.*

9.13. **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 Proposition 8.43:

$$f_1\cdot f_2 \;=\; \operatorname{Ind}_{\mathfrak{S}_{n_1,n_2}}^{\mathfrak{S}_n}(f_1\otimes f_2) \;=\; \frac{1}{n_1!\,n_2!}\sum_{\substack{g\in\mathfrak{S}_n:\\ g^{-1}wg\in\mathfrak{S}_{n_1,n_2}}}(f_1\otimes f_2)(g^{-1}wg).$$

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

$$\mathfrak{S}_\lambda \;=\; \mathfrak{S}_{\{1,\ldots,\lambda_1\}} \;\times\; \mathfrak{S}_{\{\lambda_1+1,\ldots,\lambda_1+\lambda_2\}} \;\times\;\cdots\;\times\; \mathfrak{S}_{\{n-\lambda_\ell+1,\ldots,n\}} \;\subset\; \mathfrak{S}_n.$$

For $w\in\mathfrak{S}_n$, denote by $\lambda(w)$ the cycle-shape of w, expressed as a partition.

Definition 9.32. 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
$$\psi(w) = p_{\lambda(w)}. \tag{9.20}$$
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 9.33. *The Frobenius characteristic* $\mathbf{ch} : C\ell(\mathfrak{S}) \to \Lambda$ *has the following properties:*

(1) If $\lambda \vdash n$, then $\mathbf{ch}(1_\lambda) = p_\lambda/z_\lambda$.
(2) $\mathbf{ch}$ is an isometry, i.e., it preserves inner products:
$$\langle f, g\rangle_{\mathfrak{S}_n} \;=\; \langle \mathbf{ch}(f), \mathbf{ch}(g)\rangle_\Lambda.$$
(3) $\mathbf{ch}$ is a ring isomorphism.
(4) $\mathbf{ch}(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}\chi_{\mathrm{triv}}) = h_\lambda$.
(5) $\mathbf{ch}(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}\chi_{\mathrm{sign}}) = e_\lambda$.
(6) The irreducible characters of $\mathfrak{S}_n$ are $\{\mathbf{ch}^{-1}(s_\lambda) \mid \lambda \vdash n\}$.
(7) For all characters χ, we have $\mathbf{ch}(\chi\otimes\chi_{\mathrm{sign}}) = \omega(\mathbf{ch}(\chi))$.

Proof. (1): Recall from (9.9) that $|C_\lambda| = n!/z_\lambda$, where $z_\lambda = 1^{r_1}r_1!\,2^{r_2}r_2!\ldots$, 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.

(2): Let $\lambda, \mu \vdash n$; then
$$\langle 1_\lambda,\, 1_\mu\rangle_{\mathfrak{S}_n} \;=\; \frac{1}{n!}\sum_{w\in\mathfrak{S}_n}\overline{1_\lambda(w)}1_\mu(w) \;=\; \frac{1}{n!}|C_\lambda|\delta_{\lambda\mu} \;=\; \delta_{\lambda\mu}/z_\lambda,$$
$$\left\langle \frac{p_\lambda}{z_\lambda},\, \frac{p_\mu}{z_\mu}\right\rangle_\Lambda \;=\; \frac{1}{\sqrt{z_\lambda z_\mu}}\left\langle \frac{p_\lambda}{\sqrt{z_\lambda}},\, \frac{p_\mu}{\sqrt{z_\mu}}\right\rangle_\Lambda \;=\; \frac{1}{\sqrt{z_\lambda z_\mu}}\delta_{\lambda\mu} \;=\; \delta_{\lambda\mu}/z_\lambda.$$

(3) Let $n = j+k$ and let $f \in C\ell(\mathfrak{S}_{[j]})$ and $g \in C\ell(\mathfrak{S}_{[j+1,n]})$ (so that elements of these two groups commute, and the cycle-type of a product is just the multiset union of the cycle-types.

$$\begin{aligned}
\mathbf{ch}(fg) &= \left\langle \mathrm{Ind}_{\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_n}(f\otimes g),\ \psi \right\rangle_{\mathfrak{S}_n} && \text{(where } \psi(w) = p_{\lambda(w)} \text{ for } w \in \mathfrak{S}_n \text{ as in (8.43))}\\
&= \left\langle f\otimes g,\ \mathrm{Res}_{\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_n}\psi \right\rangle_{\mathfrak{S}_j\times\mathfrak{S}_k} && \text{(by Frobenius reciprocity)}\\
&= \frac{1}{j!\,k!}\sum_{(w,x)\in\mathfrak{S}_j\times\mathfrak{S}_k} \overline{f\otimes g(w,x)}\cdot p_{\lambda(wx)} && \text{(inner product in } \mathfrak{S}_j\times\mathfrak{S}_k\text{)}\\
&= \left(\frac{1}{j!}\sum_{w\in\mathfrak{S}_j}\overline{f(w)}\ p_{\lambda(w)}\right)\left(\frac{1}{k!}\sum_{x\in\mathfrak{S}_k}\overline{g(x)}\ p_{\lambda(x)}\right)\\
&= \mathbf{ch}(f)\ \mathbf{ch}(g).
\end{aligned}$$

(4,5): 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 **ch** is a ring homomorphism, we obtain

$$h_\lambda = \prod_{i=1}^{\ell} h_{\lambda_i} = \prod_{i=1}^{\ell}\mathbf{ch}(\chi_{\mathrm{triv}}(\mathfrak{S}_n)) = \mathbf{ch}\left(\prod_{i=1}^{\ell}\chi_{\mathrm{triv}}(\mathfrak{S}_n)\right) = \mathbf{ch}(\mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}\chi_{\mathrm{triv}})$$

and likewise $e_\lambda = \mathbf{ch}(\mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}\chi_{\mathrm{sign}})$.

(6): To prove this we need to do some more work.

(7): Left as an exercise.

□

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:

$$s_\mu s_\nu = \sum_\lambda c_{\mu,\nu}^\lambda s_\lambda \tag{9.21}$$

with $c_{\mu,\nu}^\lambda \in \mathbb{Z}$. These numbers $c_{\mu,\nu}^\lambda \in \mathbb{Z}$ (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 Schur functions are a $\mathbb{Z}$-basis.) Equivalently, we can define the $c_{\mu,\nu}^\lambda$ in terms of the Hall inner product:

$$c_{\mu,\nu}^\lambda = \langle s_\mu s_\nu,\ s_\lambda\rangle_\Lambda.$$

Via the Frobenius characteristic, we can interpret the $c_{\mu,\nu}^\lambda$ in terms of representations of the symmetric group:

$$c_{\mu,\nu}^\lambda = \left\langle \mathrm{Ind}_{\mathfrak{S}_q\times\mathfrak{S}_r}^{\mathfrak{S}_n}(\chi^\mu\otimes\chi^\nu),\ \chi^\lambda\right\rangle_{\mathfrak{S}_n} = \left\langle \chi^\mu\otimes\chi^\nu,\ \mathrm{Res}_{\mathfrak{S}_q\times\mathfrak{S}_r}^{\mathfrak{S}_n}(\chi^\lambda)\right\rangle_{\mathfrak{S}_q\times\mathfrak{S}_r}$$

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.

(1) Find a combinatorial interpretation for the coefficients $c^\lambda_{\mu,\nu}$. (Answer: The Littlewood-Rchardson Rule, in all its various guises.)
(2) Find a combinatorial interpretation for the value of the irreducible characters χ_λ on the conjugacy class C_μ. (Answer: The Murnaghan-Nakayama Rule.)
(3) Determine the dimension f^λ of the irreducible character χ_λ. (Answer: The Hook-Length Formula.)

9.14. **Skew Tableaux and the Littlewood-Richardson Rule.** Let $\nu = \lambda/\mu$ be a skew shape, and let $\mathsf{CST}(\lambda/\mu)$ denote the set of all column-strict skew tableaux of shape λ/μ. It is natural to define the *skew Schur function*

$$s_{\lambda/\mu}(x_1, x_2, \dots) = \sum_{T \in \mathsf{CST}(\lambda/\mu)} x_T.$$

For example, suppose that $\lambda = (2,2)$ and $\mu = (1)$, so that

$$\nu = \begin{array}{|c|c|}\cline{2-2} \multicolumn{1}{c|}{} & \\ \hline & \\ \hline\end{array}.$$

What are the possibilities for $T \in \mathsf{CST}(\nu)$? 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).

$$\begin{array}{|c|c|}\cline{2-2} \multicolumn{1}{c|}{} & a \\ \hline b & c \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{2-2} \multicolumn{1}{c|}{} & b \\ \hline a & c \\ \hline\end{array} \qquad\qquad \begin{array}{|c|c|}\cline{2-2} \multicolumn{1}{c|}{} & a \\ \hline a & b \\ \hline\end{array} \qquad\qquad \begin{array}{|c|c|}\cline{2-2} \multicolumn{1}{c|}{} & a \\ \hline b & b \\ \hline\end{array}$$

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 (9.21). Better yet, they are symmetric in μ and ν.

Proposition 9.34. *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 $T \in \mathsf{CST}(\lambda)$ 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 9.35. *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

$$\begin{aligned}\Omega(\mathbf{x},\mathbf{z})\Omega(\mathbf{y},\mathbf{z}) &= \left(\sum_\mu s_\mu(\mathbf{x})s_\mu(\mathbf{z})\right)\left(\sum_\nu s_\nu(\mathbf{y})s_\nu(\mathbf{z})\right) = \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y})s_\mu(\mathbf{z})s_\nu(\mathbf{z}) \\ &= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y})\sum_\lambda c^\lambda_{\mu,\nu}s_\lambda(\mathbf{z}). \end{aligned} \tag{9.22}$$

On the other hand, we also have (formally setting $s_{\lambda/\mu}=0$ if $\mu\not\subseteq\lambda$)

$$\begin{aligned}\Omega(\mathbf{x},\mathbf{z})\Omega(\mathbf{y},\mathbf{z}) &= \sum_\lambda s_\lambda(\mathbf{x}\,\mathbf{y})s_\lambda(\mathbf{z}) = \sum_\lambda\sum_{\mu\subseteq\lambda} s_\mu(\mathbf{x})s_{\lambda/\mu}(\mathbf{y})s_\lambda(\mathbf{z}) \\ &= \sum_\lambda\sum_\mu s_\mu(\mathbf{x})s_\lambda(\mathbf{z})\sum_\nu \tilde{c}_{\lambda/\mu,\nu}s_\nu(\mathbf{y}) \\ &= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y})\sum_\lambda s_\lambda(\mathbf{z})\tilde{c}_{\lambda/\mu,\nu}. \end{aligned} \tag{9.23}$$

(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 9.34.) Now the theorem follows from the equality of (9.22) and (9.23). $\square$

There are a lot of combinatorial interpretations of the Littlewood-Richardson numbers. Here is one:

Theorem 9.36 (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 9.37 (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 μ 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 μ 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.

9.15. **The Murnaghan-Nakayama Rule.** We know from Theorem 9.33 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 9.38 (Murnaghan-Nakayama Rule (1937))**.**

$$\chi^\lambda(C_\mu) = \sum_{\substack{\text{rim-hook tableaux } T\\ \text{of shape } \lambda \text{ and content } \mu}} \prod_{i=1}^{n} (-1)^{1+\mathrm{ht(H_i(T))}}.$$

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 9.39. $\dim S^\lambda = f^\lambda$.

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 elegant piece of symmetric-function combinatorics.

9.16. **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 [Sag01, §4.5]. Define $h_0 = e_0 = 1$ and $h_k = e_k = 0$ for $k < 0$.

Theorem 9.40. *For any $\lambda = (\lambda_1 \dots, \lambda_\ell)$ we have*

$$s_\lambda = \left| h_{\lambda_i - i + j} \right|_{i,j=1\dots,\ell} \tag{9.24}$$

and

$$s_{\tilde\lambda} = \left| e_{\lambda_i - i + j} \right|_{i,j=1\dots,\ell}. \tag{9.25}$$

For example,

$$s_{311} = \begin{vmatrix} h_3 & h_4 & h_5 \\ h_0 & h_1 & h_2 \\ h_{-1} & h_0 & h_1 \end{vmatrix} = \begin{vmatrix} h_3 & h_4 & h_5 \\ 1 & h_1 & h_2 \\ 0 & 1 & h_1 \end{vmatrix} = h_{311} + h_5 - h_{41} - h_{32}.$$

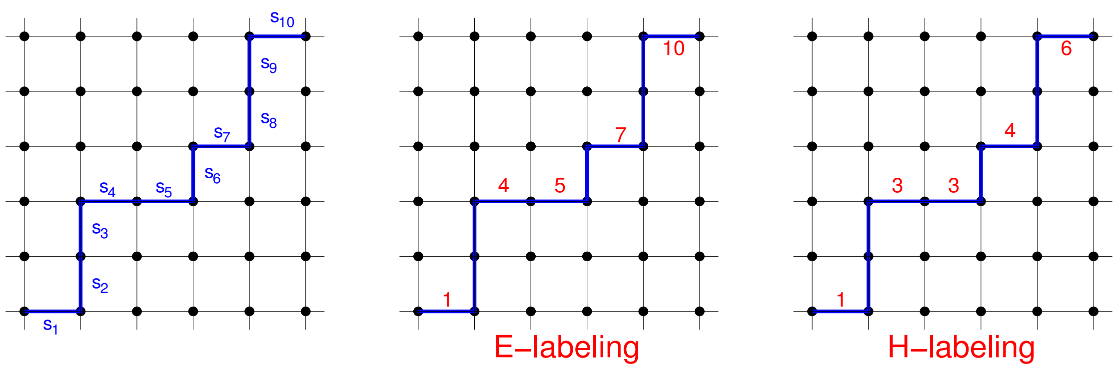
We will just prove the h case

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.

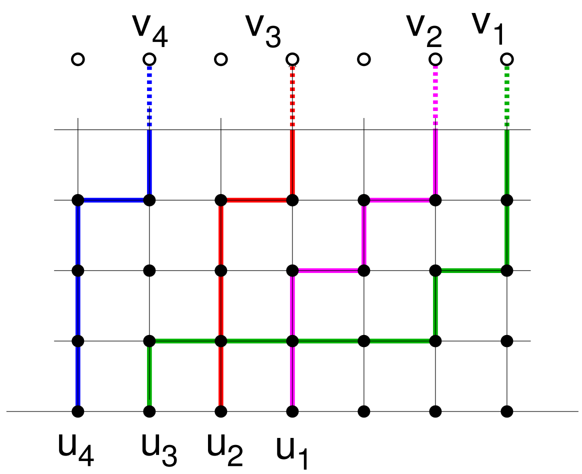
For a partition λ of length ℓ, a **λ-path family** is a (U, V)-path family, where U, V are defined by

$$u_i = (1 - i, 0), \qquad v_i = (\lambda_i - i + 1, \infty).$$

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 $(-1)^{\mathbf{P}} = -1$.



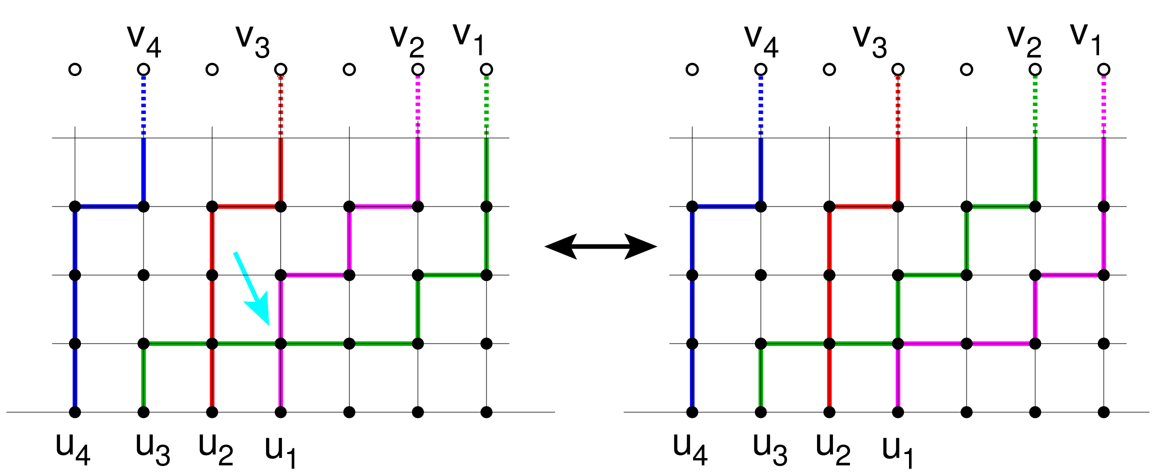
Expanding the determinant on the right-hand side of (9.24), we see that

$$\begin{aligned} |h_{\lambda_i - i + j}|_{i,j=1\ldots,\ell} &= \sum_{\pi \in \mathfrak{S}_\ell} \varepsilon(\pi) \prod_{i=1}^{\ell} h_{\lambda_i - i + \pi(i)} \\ &= \sum_{\pi \in \mathfrak{S}_\ell} \varepsilon(\pi) \sum_{\mathbf{P}=(\pi, P_1, \ldots, P_\ell)} \hat{x}^{P_1} \cdots \hat{x}^{P_\ell} \\ &= \sum_{\mathbf{P}} (-1)^{\mathbf{P}} \hat{x}^{\mathbf{P}}. \end{aligned} \tag{9.26}$$

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

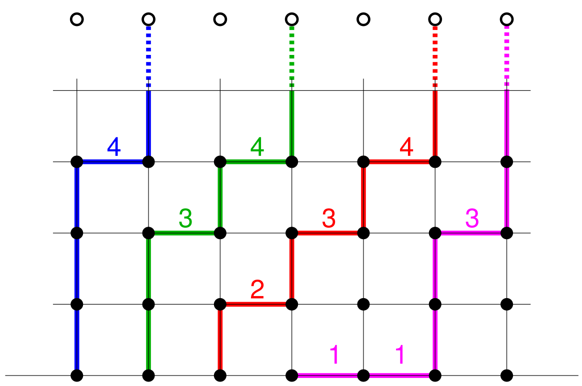
- this operation is an involution on non-good path families;
- $x^{\mathbf{P}} = x^{\mathbf{P}^\sharp}$; and
- $(-1)^{\mathbf{P}} = -(-1)^{\mathbf{P}^\sharp}$.

Therefore by (9.26) we have

$$|h_{\lambda_i - i + j}|_{i,j=1\dots,\ell} = \sum_{\mathbf{P}} (-1)^{\mathbf{P}} \hat{x}^{\mathbf{P}} = \sum_{\mathbf{P}\text{ good}} \hat{x}^{\mathbf{P}}. \tag{9.27}$$

Step 4: Enumerate weights of good path families.

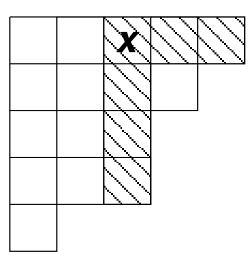
Now, for each good path family, label each path using the H-labeling. Clearly the labels weakly increase as we move north along each path. Moreover, for each j and $i < i'$, the j^{th} step of the path P_i is strictly southeast of the j^{th} step of $P_{i'}$. What this means is that we can construct a column-strict tableau of shape λ by reading off the labels of each path. For example:



1	1	3
2	3	4
3	4	
4		

Therefore, (9.27) implies that $|h_{\lambda_i - i + j}|_{i,j=1\dots,\ell} = s_\lambda$ as desired. $\square$

9.17. **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

$$h(i,j) = \lambda_i - (i-1) + \tilde{\lambda}_j - (j-1) - 1 = \lambda_i + \tilde{\lambda}_j - i - j + 1. \tag{9.28}$$

In particular,

$$h(i,1) = \lambda_i + \tilde{\lambda}_1 - i = \lambda_i + \ell - i. \tag{9.29}$$

Theorem 9.41 (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 9.42. 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 [Aig07]. We need a couple of lemmas about hook numbers.

Lemma 9.43. *Let $\lambda \vdash n$, $\ell = \ell(\lambda)$, and $1 \leq i \leq \ell$. Then the sequence*

$$\underbrace{h(i,1), \quad \ldots, \quad h(i,\lambda_i)}_{A}, \quad \underbrace{h(i,1) - h(i+1,1), \quad \ldots, \quad h(i,1) - h(\ell,1)}_{B}. \tag{9.30}$$

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:

$$\begin{array}{llll} i=1: & 9,\ 7,\ 6,\ 3,\ 1, & 9-7,\ 9-5,\ 9-4,\ 9-1 & = \ 9,\ 7,\ 6,\ 3,\ 1,\ 2,\ 4,\ 5,\ 8 \\ i=2: & 7,\ 5,\ 4,\ 1, & 7-5,\ 7-4,\ 7-1 & = \ 7,\ 5,\ 4,\ 1,\ 2,\ 3,\ 6 \\ i=3: & 5,\ 3,\ 2, & 5-4,\ 5-1 & = \ 5,\ 3,\ 2,\ 1,\ 4 \\ i=4: & 4,\ 2,\ 1, & 4-1 & = \ 4,\ 2,\ 1,\ 3 \\ i=5: & 1 & & = \ 1. \end{array}$$

Proof of Lemma 9.43. From the definition, it is immediate that A and B are respectively strictly decreasing and strictly increasing sequences of positive integers $\leq h(i,1)$. 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 (9.28) in the first line.) Indeed, either λ has a box in the position (j,k) or it doesn't. If it does, then $\lambda_k \geq j$ and $\tilde{\lambda}_j \geq k$, so $z > 0$. If it doesn't, then $\lambda_k \leq j-1$ and $\tilde{\lambda}_j \leq k-1$, so $z < 0$. $\square$

Remark 9.44. Here is a pictorial way to see Lemma 9.43. 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 9.43 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 9.43 All this can probably be proved by induction on the number of boxes.

Corollary 9.45. *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

$$\prod_{j=1}^{\lambda_i} h(i,j) \;=\; \frac{h(i,1)!}{\displaystyle\prod_{k=i+1}^{r} h(i,1) - h(k,1)}$$

and therefore

$$\prod_{x\in\lambda} h(x) \;=\; \prod_{i=1}^{r}\prod_{j=1}^{\lambda_i} h(i,j) \;=\; \frac{\displaystyle\prod_{i=1}^{r} h(i,1)!}{\displaystyle\prod_{i=1}^{r}\prod_{k=i+1}^{r} (h(i,1) - h(k,1))} \;=\; \frac{\displaystyle\prod_{i=1}^{r} h(i,1)!}{\displaystyle\prod_{1\leq i<j\leq r} (h(i,1) - h(j,1))}. \tag{9.31}$$

Proof of the Hook-Length Formula. A standard tableau is just a column-strict tableau with entries $1, \dots, n$, each occurring once. Therefore, if we expand $s_\lambda = \sum_{T\in\mathsf{CST}(\lambda)} x^T$ 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 $\Lambda_n \to \mathbb{C}$.

$$\Phi(h_{(\nu_1,\dots,\nu_r)}) = \frac{n!}{\nu_1!\cdots\nu_r!} \tag{9.32}$$

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 $\nu_1, \dots \nu_r$.

For $1 \le i \le \ell$, let
$$\mu_i = h(i,1) = \lambda_i + \ell - i.$$
Now, the Jacobi-Trudi identity (9.24) gives

$$\begin{aligned}
f^\lambda &= \Phi(s_\lambda) = \Phi\Big(\det[h_{\lambda_i - i + j}]\Big) \\
&= \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) \Phi(h_{\lambda_1+\pi(1)-1,\ \dots,\ \lambda_\ell+\pi(\ell)-\ell}) \\
&= \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) \frac{n!}{(\lambda_1+\pi(1)-1)! \cdots (\lambda_\ell+\pi(\ell)-\ell)!} && \text{(by (9.32))} \\
&= n!\ \det\left[\frac{1}{(\lambda_i+j-i)!}\right] && \text{(this is the tricky part)} \\
&= \frac{n!}{(\lambda_1+\ell-1)! \cdots (\lambda_\ell)}\ \det\left[\frac{(\lambda_i+\ell-i)!}{(\lambda_i+j-i)!}\right] && \text{(this is just multiplying by 1)} \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!}\ \det\left[\frac{\mu_i!}{(\mu_i-\ell+j)!}\right] \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!} \begin{vmatrix}
\mu_1(\mu_1-1)\cdots(\mu_1-\ell+1) & \mu_1(\mu_1-1)\cdots(\mu_1-\ell+2) & \cdots & \mu_1 & 1 \\
\mu_2(\mu_2-1)\cdots(\mu_2-\ell+1) & \mu_2(\mu_2-1)\cdots(\mu_2-\ell+2) & \cdots & \mu_2 & 1 \\
\vdots & \vdots & & \vdots & \vdots \\
\mu_\ell(\mu_\ell-1)\cdots(\mu_\ell-\ell+1) & \mu_\ell(\mu_\ell-1)\cdots(\mu_\ell-\ell+2) & \cdots & \mu_\ell & 1
\end{vmatrix}
\end{aligned}$$

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 $i \ne j$ (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,

$$\begin{aligned}
f^\lambda &= \frac{n!}{\mu_1! \cdots \mu_\ell!} \prod_{1 \le i < j \le \ell} (\mu_i - \mu_j) \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!} \prod_{1 \le i < j \le \ell} (h(i,1) - h(j,1)) \\
&= \frac{n!}{\mu_1! \cdots \mu_\ell!} \ \frac{\prod_{i=1}^{\ell} h(i,1)!}{\prod_{x \in \lambda} h(x)} && \text{(by (9.31))} \\
&= \frac{n!}{\prod_{x \in \lambda} h(x)}.
\end{aligned}$$

□

9.18. **Quasisymmetric functions and Hopf algebras.**

Definition 9.46. 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 $QSym \subset \mathbb{C}[[\mathbf{x}]]$. 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 $\operatorname{Comp}(n)$. 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, $|\operatorname{Comp}(n)| = 2^{n-1}$. We can define a partial order on $\operatorname{Comp}(n)$ 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 9.47. Let $\mathcal{M}$ be a matroid on ground set E of size n. Consider weight functions $f : E \to \mathbb{P}$; 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?

9.19. **Exercises.**

Exercise 9.1. Prove Proposition 9.7.

Exercise 9.2. Give a purely combinatorial proof that $\exp\ln(1+x) = 1+x$. In other words, expand the composition $\exp\ln x$ as a formal power series, using the definitions of exp and ln in (9.14), and compute the coefficient of x^k for each k. Hint: Interpret the coefficients as counting permutations.

Exercise 9.3. Supply the proofs for the identities (9.12).

Exercise 9.4. Prove part (7) of Theorem 9.33.

Exercise 9.5. Fill in the proofs of the underlined assertions in Rule 2 and Rule X for the alternate RSK algorithm in Section 9.11.

Exercise 9.6. For this problem, you will probably want to use one of the alternate RSK algorithms from Sections 9.10 and 9.11.

(a) For $w \in \mathfrak{S}_n$, let $(P(w), Q(w))$ be the pair of tableaux produced by the RSK algorithm from w. Denote by w^* the reversal of w in one-line notation (for instance, if $w = 57214836$ then $w^* = 63841275$). Prove that $P(w^*) = P(w)^T$ (where ^{T} means transpose).

(b) *(Open problem)* For which permutations does $Q(w^*) = Q(w)$? Computation indicates that the number of such permutations is

$$\begin{cases} \dfrac{2^{(n-1)/2}(n-1)!}{((n-1)/2)!^2} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even,} \end{cases}$$

but I don't know a combinatorial (or even an algebraic) reason.

(c) *(Open problem)* For which permutations does $Q(w^*) = Q(w)^T$? I have no idea what the answer is. The sequence $(q_1, q_2, \dots) = (1, 2, 2, 12, 24, 136, 344, 2872, 7108, \dots)$, where $q_n = \#\{w \in \mathfrak{S}_n \mid Q(w^*) = Q(w)^T\}$, does not seem to appear in the Online Encyclopedia of Integer Sequences.

Exercise 9.7. (*Open problem*) Let $n \geq 2$ and for $\sigma \in \mathfrak{S}_n$, let $f(\sigma)$ denote the number of fixed points. As a warmup, prove that $\sum_{\sigma\in\mathfrak{S}_n} f(\sigma)^2 = 2\cdot n!$. Open problem (to the best of my knowledge): Prove that for any n, k, the number $\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} f(\sigma)^k$ is an integer. It appears to be A203647 in OEIS. Find a formula and/or a representation-theoretic interpretation.

10. 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

$$\begin{tikzcd} A\otimes A\otimes A \arrow[r, "\mu\otimes I"] \arrow[d, "I\otimes\mu"'] & A\otimes A \arrow[d, "\mu"] \\ A\otimes A \arrow[r, "\mu"'] & A \end{tikzcd} \qquad \begin{tikzcd} a\otimes b\otimes c \arrow[r, mapsto] \arrow[d, mapsto] & ab\otimes c \arrow[d, mapsto] \\ a\otimes bc \arrow[r, mapsto] & abc \end{tikzcd}$$

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:

$$\begin{tikzcd}
Z & & Z \otimes Z \arrow[ll, "\Delta"'] \\
Z \otimes Z \arrow[u, "\Delta"] & & Z \otimes Z \otimes Z \arrow[u, "I \otimes \Delta"'] \arrow[ll, "\Delta \otimes I"]
\end{tikzcd}$$

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 10.1 (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*[15]. 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 10.2 (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 the chains of lengths 0 and 1), 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

$$\Delta[P] = \sum_{x \in P} [\hat{\mathbf{0}}, x] \otimes [x, \hat{\mathbf{1}}].$$

Coassociativity is checked by the following calculation:

$$\begin{aligned}
\Delta \otimes I(\Delta(P)) &= \Delta \otimes I \left(\sum_{x \in P} [\hat{\mathbf{0}}, x] \otimes [x, \hat{\mathbf{1}}] \right) \\
&= \sum_{x \in P} \Delta([\hat{\mathbf{0}}, x]) \otimes [x, \hat{\mathbf{1}}] \\
&= \sum_{x \in P} \left(\sum_{y \in [\hat{\mathbf{0}}, x]} [\hat{\mathbf{0}}, y] \otimes [y, x] \right) \otimes [x, \hat{\mathbf{1}}] \\
&= \sum_{x \leq y \in P} [\hat{\mathbf{0}}, y] \otimes [y, x] \otimes [x, \hat{\mathbf{1}}] \\
&= \sum_{y \in P} [\hat{\mathbf{0}}, y] \otimes \left(\sum_{x \in [y, \hat{\mathbf{1}}]} [y, x] \otimes [x, \hat{\mathbf{1}}] \right) \\
&= \sum_{y \in P} [\hat{\mathbf{0}}, y] \otimes \Delta([y, \hat{\mathbf{1}}]) \;=\; I \otimes \Delta(\Delta(P)).
\end{aligned}$$

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 $\Delta(F) = \sum F_1 \otimes F_2$.

[15]There are those who call this "mmutative".

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 10.3. A Hopf algebra is a bialgebra $\mathcal{H}$ with a **antipode** $S : \mathcal{H} \to \mathcal{H}$, which satisfies the commutative diagram

$$\begin{tikzcd}
\mathcal{H} \arrow[r, "\Delta"] \arrow[drr, "\varepsilon"'] & \mathcal{H} \otimes \mathcal{H} \arrow[rr, "S \otimes I"] & & \mathcal{H} \otimes \mathcal{H} \arrow[r, "\mu"] & \mathcal{H} \\
& & \mathbb{C} \arrow[urr, "u"'] & &
\end{tikzcd} \tag{10.1}$$

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 10.4. *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* (10.1) *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., $S = (-1)^k \omega$.

10.1. **Characters of Hopf algebras.** Let $\mathcal{H}$ be a Hopf algebra. A **character** on $\mathcal{H}$ is a $\mathbb{C}$-linear map $\zeta : \mathcal{H} \to \mathbb{C}$ that is multiplicative, i.e., $\zeta(1_{\mathcal{H}}) = 1_{\mathbb{C}}$ and $\zeta(h \cdot h') = \zeta(h)\zeta(h')$.

For example, if $\mathcal{H}$ is the graph Hopf algebra, then we can define a character by

$$\zeta(G) = \begin{cases} 1 & \text{if } G \text{ has no edges,} \\ 0 & \text{if } G \text{ has one or more edges,} \end{cases}$$

for a graph G, and then extending by linearity to all of $\mathcal{G}$. This map is multiplicative (because $G \cdot H$ has an edge iff either G or H does); it also looks completely silly to define such a thing.

The reason this is interesting is that characters can be multiplied by the **convolution product** defined as follows: if $h \in \mathcal{H}$ and $\Delta(h) = \sum h_1 \otimes h_2$ in Sweedler notation, then

$$(\zeta * \eta)(h) = \sum \zeta(h_1)\eta(h_2).$$

One can check that convolution is associative (the calculation resembles checking that the incidence algebra of a poset is an algebra). The counit ε is a two-sided identity for convolution, i.e., $\zeta * \varepsilon = \varepsilon * \zeta = \zeta$ for all characters ζ. Moreover, the definition (10.1) of the antipode implies that

$$\zeta * (\zeta \circ S) = \varepsilon$$

(check this too). Therefore, the set of all characters forms a group.

Why would you want to convolve characters? Consider the graph Hopf algebra with the character ζ, and let $k \in \mathbb{N}$. The k^{th} convolution power of ζ is given by

$$\begin{aligned} \underbrace{\zeta * \cdots * \zeta}_{k \text{ times}}(G) &= \sum_{V(G) = V_1 \cup\!\!\!\cdot \cdots \cup\!\!\!\cdot V_k} \zeta(G|_{V_1}) \cdots \zeta(G|_{V_k}) \\ &= \sum_{V(G) = V_1 \cup\!\!\!\cdot \cdots \cup\!\!\!\cdot V_k} \begin{cases} 1 & \text{if } V_1, \dots, V_k \text{ are all cocliques,} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

(here $\cup\!\!\!\cdot$ denotes disjoint union; recall that a *coclique* is a set of vertices of which no two are adjacent). In other words, $\zeta^n(G)$ counts the number of functions $f : V \to [k]$ so that $f(x) \neq f(y)$ whenever x, y are adjacent. But such a thing is precisely a proper k-coloring! I.e.,

$$\zeta^n(G) = p(G;\, k)$$

where p is the chromatic polynomial (see Section 4.3). This turns out to be true as a polynomial identity in k — for instance, $\zeta^{-1}(G)$ is the number of acyclic orientations. One can even view the Tutte polynomial $T(G; x, y)$ as a character $\tau_{x,y}(G)$ with parameters x, y; it turns out that $\tau^k_{x,y}(G)$ is itself a Tutte polynomial evaluation — see Brandon Humpert's Ph.D. thesis [Hum11].

10.2. **Exercises.**

Exercise 10.1. 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 10.4.

Exercise 10.2. 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....)

11. Appendix: Notational Index

Basics

$\mathbb{N}$	nonnegative integers $0, 1, 2, \dots$
$\mathbb{P}$	positive integers $1, 2, \dots$
$[n]$	$\{1, \dots, n\}$
2^S	power set of a set S (or the associated poset)
$\mathfrak{S}_n$	symmetric group on n letters
$\binom{S}{k}$	set of k-element subsets of a set S
$\mathbb{F}\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$	$\mathbb{F}$-vector space with basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$

Posets

$\lessdot$	"is covered by"
2^S	power set of a set S (or the associated poset)
$\hat{\mathbf{0}}, \hat{\mathbf{1}}$	unique minimum and maximum elements of a poset
$[x, y]$	interval in a poset
P^*	dual poset to P
Π_n	partition lattice (the lattice of all set partitions of $[n]$)
$S(n, k)$	Stirling number of the second kind
Y	Young's lattice of integer partitions
$\tilde{\lambda}$	conjugate of a partition λ
$K(G)$	clique poset of a graph G

Simplicial Complexes and Polytopes

$\langle \cdots \rangle$	simplicial complex generated by a set of faces
$\lvert\Delta\rvert$	(standard) geometric realization of Δ
$\Delta(P)$	order complex of a poset P
$\Bbbk[\Delta]$	Stanley-Reisner (face) ring of Δ over a field $\Bbbk$
$C_k(\Delta, R)$	simplicial chain groups over a ring R
$\tilde{H}_k(\Delta, R)$	reduced simplicial homology groups
$\mathbb{S}^d$	d-dimensional sphere
P^*	dual of a polytope P

Lattices

$\wedge, \vee$	meet, join
$\mathbb{F}_q$	finite field of order q
$L_n(q)$	lattice of vector subspaces of $\mathbb{F}_q^n$
$J(P)$	lattice of order ideals of a poset P
$\mathrm{Irr}(L)$	poset of join-irreducible elements in a lattice L
N_5	nonmodular, nonranked 5-element lattice
M_5	modular, ranked, nondistributive 5-element lattice
$L(E)$	geometric lattice represented by a set of vectors E
$L^{\mathrm{aff}}(E)$	geometric lattice represented by a set of affine points E

Matroids and the Tutte Polynomial

$\bar{A}$	closure operator applied to A
$M(G)$	graphic matroid of a graph G
$\mathscr{I}$	matroid independence system
$\mathscr{B}$	matroid basis system
$\mathscr{C}$	matroid circuit system
M^*	dual of a matroid M
$M \oplus M'$	direct sum of matroids
$M - e$	matroid deletion
M/e	matroid contraction
$U_k(n)$	uniform matroid of rank k on ground set of size n
T_M, $T_M(x, y)$	Tutte polynomial of M
$p_G(k)$	chromatic polynomial of a graph G
$C(e, B)$	fundamental circuit of e with respect to basis B
$C^*(e, B)$	fundamental cocircuit of e with respect to basis B

Poset Algebra

$\mathrm{Int}(P)$	set of intervals of poset P
$I(P)$	incidence algebra of P
$f * g$	convolution product in $I(P)$
δ	Kronecker delta function (as an element of $I(P)$)
ζ	zeta function in $I(P)$
μ	Möbius function in $I(P)$
$\chi_P(x)$	characteristic polynomial of poset P
$A(L)$	Möbius algebra of a lattice L

Hyperplane Arrangements

Bool_n	Boolean arrangement
Br_n	braid arrangement
$L(\mathcal{A})$	intersection poset of arrangement $\mathcal{A}$
$\mathrm{ess}(\mathcal{A})$	essentialization of $\mathcal{A}$
$r(\mathcal{A})$	number of regions of a real arrangement $\mathcal{A}$
$b(\mathcal{A})$	number of bounded regions of a real arrangement $\mathcal{A}$
$\mathcal{A}_x$, $\mathcal{A}^x$	See Eqn. 6.5a
Shi_n	Shi arrangement
$\mathbb{P}^d\mathbb{F}$	d-dimensional projective space over field $\mathbb{F}$
$\mathrm{proj}(\mathcal{A})$	projectivization of a central arrangement $\mathcal{A}$
$c\mathcal{A}$	cone over $\mathcal{A}$
$\mathcal{A}_G$	arrangement associated with a graph G

Combinatorial Algebraic Varieties

$\mathrm{Gr}(k, V)$	Grassmannian of k-dimensional vector subspaces of V
Ω_λ	Schubert cell in a Grassmannian
$F\ell(n)$	(complete) flag variety in dimension n
X_w	Schubert cell in a flag variety

Representation Theory

D_n	dihedral group of order $2n$ (symmetries of a regular n-gon)
ρ_{triv}, ρ_{reg}	trivial and regular representations of a group
ρ_{sign}, ρ_{def}	sign and defining representations of $\mathfrak{S}_n$
χ_ρ or χ	character of a representation ρ
$\langle \rho, \rho' \rangle_G$	inner product on class functions of G (see Thm. 8.28)
$\text{Hom}_{\mathbb{C}}(V, W)$	$\mathbb{C}$-linear maps $V \to W$
$\text{Hom}_G(V, W)$	G-equivariant $\mathbb{C}$-linear maps $V \to W$
V^G	space of invariants of a G-action on V
$[a, b]$	commutator: $aba^{-1}b^{-}1$
$[G, G]$	commutator subgroup of G
$\mathfrak{A}_n$	alternating group on n letters
Par_n	partitions of n
C_λ	conjugacy class of $\mathfrak{S}_n$ consisting of permutations with cycle shape λ
$\lambda < \mu$	lexicographic (total) order on partitions
$\lambda \lhd \mu$	dominance (partial) order on partitions
$\text{sh}(T)$	shape of a tabloid T
(ρ_μ, V^μ)	tabloid representation of shape μ
χ_μ	character of tabloid representation
Sp_λ	Specht module
$K_{\lambda,\mu}$	Kostka numbers
$\text{Res}^G_H(\rho)$, $\text{Res}^G_H(\chi)$	restricted representation/character
$\text{Ind}^G_H(\rho)$, $\text{Ind}^G_H(\chi)$	induced representation/character

Symmetric Functions

$\mathbf{x}^\alpha$	monomial in variables $\mathbf{x}$ with exponent vector α
$R[[\mathbf{x}]$	ring of formal power series in variables $\mathbf{x}$ with coefficients in R
m_λ	monomial symmetric function
e_λ	elementary symmetric function
h_λ	(complete) homogeneous symmetric function
p_λ	power-sum symmetric function
Λ_d, $\Lambda_{R,d}(\mathbf{x})$	R-module of degree-d symmetric functions in $\mathbf{x}$
Λ, $\Lambda_R(\mathbf{x})$	R-algebra of symmetric functions in $\mathbf{x}$
ω	involutory automorphism $\Lambda \to \Lambda$ swapping e's and h's
$\mathsf{CST}(\lambda)$	set of column-strict tableaux of shape λ
s_λ	Schur function
Ω, Ω^*	Cauchy kernel and dual Cauchy kernel
z_λ	size of centralizer of a partition of shape λ (see (9.9))
ε_λ	sign of a partition of shape λ (see (9.9))
$\mathsf{SYT}(\lambda)$	set of standard tableaux of shape λ
f^λ	number of standard tableaux of shape λ
$T \leftarrow x$	row-insertion (§9.9)
$\mathbf{ch}$	Frobenius characteristic (§9.13)
$c^\lambda_{\mu,\nu}$	Littlewood-Richardson coefficients
$h(x)$	Length of the hook with corner x

Hopf Algebras

To be added.

References

[Aig07] Martin Aigner, *A course in enumeration*, Graduate Texts in Mathematics, vol. 238, Springer, Berlin, 2007.

[Ath96] Christos A. Athanasiadis, *Characteristic polynomials of subspace arrangements and finite fields*, Adv. Math. **122** (1996), no. 2, 193–233. MR 1409420 (97k:52012)

[BH93] Winfried Bruns and Jürgen Herzog, *Cohen-Macaulay rings*, Cambridge Studies in Advanced Mathematics, vol. 39, Cambridge University Press, Cambridge, 1993. MR 1251956 (95h:13020)

[BM71] H. Bruggesser and P. Mani, *Shellable decompositions of cells and spheres*, Math. Scand. **29** (1971), 197–205 (1972). MR 0328944 (48 #7286)

[BO92] Thomas Brylawski and James Oxley, *The Tutte polynomial and its applications*, Matroid applications, Encyclopedia Math. Appl., vol. 40, Cambridge Univ. Press, Cambridge, 1992, pp. 123–225. MR 1165543 (93k:05060)

[Bol98] Béla Bollobás, *Modern graph theory*, Graduate Texts in Mathematics, vol. 184, Springer-Verlag, New York, 1998. MR 1633290 (99h:05001)

[BR07] Matthias Beck and Sinai Robins, *Computing the continuous discretely*, Undergraduate Texts in Mathematics, Springer, New York, 2007, Integer-point enumeration in polyhedra. MR 2271992 (2007h:11119)

[Bri73] 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.

[Dir61] G.A. Dirac, *On rigid circuit graphs*, Abh. Math. Sem. Univ. Hamburg **25** (1961), 71–76. MR 0130190 (24 #A57)

[Ful97] William Fulton, *Young tableaux*, London Mathematical Society Student Texts, vol. 35, Cambridge University Press, Cambridge, 1997.

[GK76] Curtis Greene and Daniel J. Kleitman, *The structure of Sperner k-families*, J. Combin. Theory Ser. A **20** (1976), no. 1, 41–68.

[Gre76] Curtis Greene, *Some partitions associated with a partially ordered set*, J. Combin. Theory Ser. A **20** (1976), no. 1, 69–79.

[Grü03] 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.

[Hat02] Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR 1867354 (2002k:55001)

[Hum11] Brandon Humpert, *Polynomials associated with graph coloring and orientations*, Ph.D. thesis, University of Kansas, 2011.

[MS05] Ezra Miller and Bernd Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics, vol. 227, Springer-Verlag, New York, 2005. MR 2110098 (2006d:13001)

[OS80] Peter Orlik and Louis Solomon, *Combinatorics and topology of complements of hyperplanes*, Invent. Math. **56** (1980), no. 2, 167–189.

[Oxl92] James G. Oxley, *Matroid theory*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1992.

[Ryb11] 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.

[S$^+$14] W. A. Stein et al., *Sage Mathematics Software (Version 6.4.1)*, The Sage Development Team, 2014, `http://www.sagemath.org`.

[Sag01] Bruce E. Sagan, *The symmetric group*, second ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001.

[Sch13] Alexander Schrijver, *A course in combinatorial optimization*, Available online at `http://homepages.cwi.nl/~lex/files/dict.pdf` (retrieved 1/21/15), 2013.

[Sta96] Richard P. Stanley, *Combinatorics and commutative algebra*, second ed., Progress in Mathematics, vol. 41, Birkhäuser Boston Inc., Boston, MA, 1996. MR 1453579 (98h:05001)

[Sta99] ———, *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)

[Sta07] ———, *An introduction to hyperplane arrangements*, Geometric combinatorics, IAS/Park City Math. Ser., vol. 13, Amer. Math. Soc., Providence, RI, 2007, Also available online: `http://www-math.mit.edu/~rstan/arrangements/arr.html`, pp. 389–496.

[Sta12] ———, *Enumerative combinatorics. Volume 1*, second ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012. MR 2868112

[Wes96] Douglas B. West, *Introduction to graph theory*, Prentice Hall, Inc., Upper Saddle River, NJ, 1996. MR 1367739 (96i:05001)