NOTES FOR MATH 264C (SPRING 2021)

STEVEN V SAM

Contents

Date: May 8, 2021.

1. Coxeter groups

1.1. **Definition and first examples.** Let S be a finite set (it is possible to allow S to be infinite, but we assume this for simplicity). Let $m\colon S \times S \to \mathbf{Z}_{>0} \cup \{\infty\}$ be a function such that $m(s,s) = 1$ and $m(s,s') = m(s',s) \ge 2$ whenever $s \ne s'$. The data of S and m is a **Coxeter system**, and the **Coxeter group** W associated to this data is the free group generated by $s \in S$ subject to the relations

$$(ss')^{m(s,s')} = 1 \text{ if } m(s,s') < \infty.$$

Most importantly, this implies that $s^2 = 1$ for all $s \in S$, so that W is generated by involutions. We denote the pair of Coxeter group together with its generators by (W,S) and call $|S|$ the **rank**.

We have a sign homomorphism $\mathrm{sgn}\colon W \to \{1,-1\}$ given by $\mathrm{sgn}(s) = -1$ for all $s \in S$.

Example 1.1. If $S = \{s,t\}$ and $m(s,t) < \infty$, then W has the presentation

$$W = \langle s,t \mid s^2 = t^2 = (st)^m = 1\rangle$$

and is isomorphic to a dihedral group of order $2m$, i.e., the symmetries of a regular n-gon. If $m = \infty$, then we can think of this as an infinite dihedral group. □

Example 1.2. Let $S = \{s_1,\dots,s_{n-1}\}$ and $m(s_i,s_j) = 3$ if $|i-j| = 1$ and 2 otherwise (and $i \ne j$). Then we get a surjective homomorphism from W to the symmetric group $\mathfrak{S}_n$ by sending s_i to the transposition $(i,i+1)$. To see this is well-defined, it is enough to know that $(i,i+1)$ and $(j,j+1)$ commute if $|i-j| \ne 1$ and that $(i,i+1)(i+1,i+2) = (i,i+1,i+2)$ is a 3-cycle (so has order 3). In fact, they are isomorphic, but this is easier to prove after we develop a few general results. We'll use this case as a running example and call this Coxeter system A_{n-1}. □

Example 1.3. We will see later that any finite subgroup of $\mathbf{GL}_n(\mathbf{R})$ generated by reflections has the structure of a Coxeter group. This includes the previous two cases. □

At the moment, it is unclear that this data is unique, i.e., it could be that different S,m give isomorphic groups. Our first goal is to construct a faithful matrix representation of W.

1.2. **The geometric representation.** Let V be the real vector space with basis $\{\alpha_s \mid s \in S\}$. We define a symmetric bilinear form B on V via

$$B(\alpha_s,\alpha_{s'}) = \begin{cases} -\cos\frac{\pi}{m(s,s')} & \text{if } m(s,s') < \infty \\ -1 & \text{if } m(s,s') = \infty \end{cases}.$$

In particular, the α_s are unit vectors with respect to B. For $s \in S$, define $\sigma_s\colon V \to V$ by

$$\sigma_s(v) = v - 2B(\alpha_s,v)\alpha_s.$$

Note that $\sigma_s(\alpha_s) = -\alpha_s$.

Lemma 1.4. *For $s \in S$ and $v, w \in V$, we have $B(v,w) = B(\sigma_s v, \sigma_s w)$.*

Proof. We have

$$\begin{aligned} B(\sigma_s v, \sigma_s w) &= B(v - 2B(\alpha_s, v)\alpha_s, w - 2B(\alpha_s, w)\alpha_s) \\ &= B(v,w) - 2B(\alpha_s, w)B(v, \alpha_s) - 2B(\alpha_s, v)B(\alpha_s, w) + 4B(\alpha_s, v)B(\alpha_s, w)B(\alpha_s, \alpha_s) \\ &= B(v,w) \end{aligned}$$

since B is symmetric and $B(\alpha_s, \alpha_s) = 1$. □

Lemma 1.5. *For $s, t \in S$, $\sigma_s \sigma_t$ has order $m(s,t)$.*

Proof. If $s = t$, then

$$\begin{aligned} \sigma_s^2(v) &= \sigma_s(v - 2B(\alpha_s, v)\alpha_s) \\ &= v - 2B(\alpha_s, v)\alpha_s + 2B(\alpha_s, v)\alpha_s = v. \end{aligned}$$

so $\sigma_s^2 = 1$ (and clearly $\sigma_s \neq 1$).

Now assume $s \neq t$ and set $m = m(s,t)$. Let U be the span of $\{\alpha_s, \alpha_t\}$.

First consider the case $m < \infty$. For a general element $v = c_s\alpha_s + c_t\alpha_t \in U$, we have

$$\begin{aligned} B(v,v) &= c_s^2 - 2c_s c_t \cos(\pi/m) + c_t^2 \\ &= c_s^2 - 2c_s c_t \cos(\pi/m) + c_t^2(\cos^2(\pi/m) + \sin^2(\pi/m)) \\ &= (c_s - c_t\cos(\pi/m))^2 + (c_t \sin(\pi/m))^2. \end{aligned}$$

Since $m \geq 2$, we have $\sin(\pi/m) > 0$. Hence, the last quantity is positive if $c_t \neq 0$. If $c_t = 0$, then $B(v,v) = c_s^2$, so we see that $B|_U$ is positive definite, and in particular nondegenerate. Let $U^\perp = \{v \in V \mid B(\alpha_s, v) = B(\alpha_t, v) = 0\}$. Then $U \cap U^\perp = 0$ since $B|_U$ is nondegenerate, and $U + U^\perp = V$ by generalities on bilinear forms. In particular, the matrix of $\sigma_s\sigma_t$ with respect to this block decomposition has an identity component on $U^\perp$, and so it suffices to show that its order on U is m.

Since

$$B(\alpha_s, \alpha_t) = -\cos(\pi/m) = \cos(\pi - \pi/m),$$

the angle between α_s and α_t is $\theta = \pi - \pi/m$, and hence we have an isometry between U and $\mathbf{R}^2$ (with the standard inner product) such that $\alpha_s \mapsto (1,0)$ and $\alpha_t \mapsto (\cos\theta, \sin\theta)$. Under this isomorphism, σ_s becomes $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ and σ_t becomes (using the double-angle formulas)

$$\begin{bmatrix} 1 - 2\cos^2\theta & -2\cos\theta\sin\theta \\ -2\cos\theta\sin\theta & 1 - 2\sin^2\theta \end{bmatrix} = \begin{bmatrix} -\cos(2\theta) & -\sin(2\theta) \\ -\sin(2\theta) & \cos(2\theta) \end{bmatrix}.$$

In particular, $\sigma_s\sigma_t$ is rotation by -2θ, i.e., rotation by $2\pi/m$, so has order m.

Now consider the case $m = \infty$ so that $B(\alpha_s, \alpha_t) = -1$. Set $u = \alpha_s + \alpha_t$. Then $B(u, \alpha_s) = B(u, \alpha_t) = 0$, so that $\sigma_s(u) = \sigma_t(u) = u$. Thus,

$$\sigma_s\sigma_t\alpha_s = \sigma_s(\alpha_s + 2\alpha_t) = -\alpha_s + 2\alpha_t + 4\alpha_s = 2u + \alpha_s$$

which means $(\sigma_s\sigma_t)^k(\alpha_s) = 2ku + \alpha_s$, so $\sigma_s\sigma_t$ has infinite order. □

Corollary 1.6. *We have a unique homomorphism $\sigma\colon W \to \mathbf{GL}(V)$ given by $\sigma(s) = \sigma_s$ for all $s \in S$.*

Example 1.7. Consider the Coxeter system A_{n-1}. We saw before that W surjects onto the symmetric group. Consider the geometric realization; let $\alpha_i = \alpha_{s_i}$. Then

$$B(\alpha_i, \alpha_j) = \begin{cases} -1/2 & \text{if } |i-j| = 1 \\ 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}.$$

Consider the hyperplane in $\mathbf{R}^n$ (with standard basis $e_1, \dots, e_n$) of vectors whose coordinate sum is 0. We have an isometry with V and this hyperplane which sends α_i to $(e_i - e_{i+1})/\sqrt{2}$. It is convenient to multiply the inner product by 2, so that we can simply take $e_i - e_{i+1}$. If we extend the action of W to all of $\mathbf{R}^n$ by acting trivially on $e_1 + \cdots + e_n$, then σ_i is simply the matrix which swaps e_i and e_{i+1}, so we see immediately that the image is the symmetric group $\mathfrak{S}_n$. □

1.3. **The geometric representation is faithful.** Our next goal is to show that σ is injective. We need a few ingredients for that. First, we introduce the **length function** $\ell\colon W \to \mathbf{Z}_{\geq 0}$. For $w \in W$, we let $\ell(w)$ be the minimum number n such that there exists an expression $w = s_1 s_2 \cdots s_n$ for $s_i \in S$. By convention, $\ell(1) = 0$. If $\ell(w) = n$, then any n-tuple $(s_1, \dots, s_n) \in S^n$ such that $w = s_1 \cdots s_n$ is a **reduced expression** for w.

The following properties follow from the definition:

Lemma 1.8. *For all $w, w' \in W$, we have*

(1) $\ell(w) = \ell(w^{-1})$,
(2) $\ell(w) = 1$ if and only if $w \in S$,
(3) $\ell(ww') \leq \ell(w) + \ell(w')$,
(4) $\ell(ww') \geq \ell(w) - \ell(w')$,
(5) For all $s \in S$, we have $\ell(ws), \ell(sw) \in \{\ell(w) - 1, \ell(w) + 1\}$.

Proof.
(1) $s_1 \cdots s_r$ is a reduced expression for w if and only if $s_r \cdots s_1$ is a reduced expression for w^{-1}.
(2) Clear
(3) Multiply reduced expressions for w and w' to get a (not necessarily reduced) expression for ww'.
(4) From (3), $\ell(ww') + \ell(w'^{-1}) \geq \ell(w)$ and from (1), $\ell(w'^{-1}) = \ell(w')$.
(5) The inequality $\ell(w) - 1 \leq \ell(ws) \leq \ell(w) + 1$ follows from the above points. The sign homomorphism tells us that $\ell(ws) \neq \ell(w)$. The proof for $\ell(sw)$ is the same. □

Our next task is to understand $\ell(ws)$ compared to $\ell(w)$. For that, we need the notion of roots. Define the **root system** of W to be

$$\Phi = \{w(\alpha_s) \mid w \in W,\ s \in S\}.$$

The elements are called **roots**. A root α is **positive** (notation: $\alpha > 0$) if it is a non-negative linear combination of the α_s, and it is **negative** (notation: $\alpha < 0$) if it is a non-positive linear combination of the α_s. The set of positive roots is Φ^+ and the set of negative roots is Φ^-. At the moment, there could be roots which are neither positive nor negative, but we will prove that such roots don't exist. In other words, we will show that every root is either positive or negative.

Example 1.9. We consider the case of a finite dihedral group W, i.e., $S = \{s, t\}$ and $m = m(s,t) < \infty$. From the previous proof, the geometric representation of W is isometric to $\mathbf{R}^2$ with the standard inner product in such a way that $\alpha_s = (1,0)$ and $\alpha_t = (\cos\theta, \sin\theta)$ with $\theta = \pi - \pi/m$. Then the roots are precisely the points $(\cos(j\pi/m), \sin(j\pi/m))$, which we can think of as the vertices of a regular $2m$-gon. We will use the following observation later: if $0 \le j < m$, then $(\cos(j\pi/m), \sin(j\pi/m))$ is a non-negative linear combination of α_s and α_t (by convexity), and otherwise it is a non-positive linear combination of them.

Furthermore, there are $2m$ elements of W, and each one is an alternating product of s and t of length $\le m$. All such expressions are different except the two length m expressions are the same since $(st)^m = 1$. Since we'll need it later, we claim that if w has no reduced expression ending in s, then $w(\alpha_s) > 0$. Such an element is either of the form $(st)^k$ or $t(st)^k$ for $0 \le k < m/2$. From the proof of Lemma 1.5, st is rotation by $2\pi/m$, so rotating $(1,0)$ less than $m/2$ times gives a positive root. Reflecting any of these points across the line perpendicular to α_t can again be seen to give a positive root. □

Finally, given a subset $I \subset S$, we let $W_I \subset W$ be the subgroup generated by the elements of I and we let ℓ_I be the length function on W_I with respect to the generators I, i.e., for $w \in W_I$, $\ell_I(w)$ is the minimal number of elements of I needed to generate w. (We will see later that this is nothing more than the restriction of ℓ to W_I, but at the moment this is not clear, so we need to distinguish them.)

Theorem 1.10. *Pick $w \in W$ and $s \in S$. If $\ell(ws) > \ell(w)$, then $w(\alpha_s)$ is a positive root. If $\ell(ws) < \ell(w)$, then $w(\alpha_s)$ is a negative root.*

Proof. The second statement follows from the first statement by using ws in place of w.

We prove the first statement by induction on $\ell(w)$. If $\ell(w) = 0$, there is nothing to prove, so assume $\ell(w) > 0$ and pick $t \in S$ such that $\ell(wt) < \ell(w)$. Then $t \ne s$, and we let $I = \{s, t\}$. Define

$$A = \{(x, x_I) \in W \times W_I \mid w = xx_I,\ \ell(w) = \ell(x) + \ell_I(x_I)\}.$$

Note that $(wt, t) \in A$, so $A \ne \varnothing$. Pick $(v, v_I) \in A$ so that $\ell(v)$ is minimized. Then $\ell(v) \le \ell(wt) = \ell(w) - 1$. We claim that $\ell(vs) > \ell(v)$ (and then by the same argument, we will also have $\ell(vt) > \ell(v)$). If not, then

$$\begin{aligned}\ell(w) &\le \ell(vs) + \ell(sv_I) = (\ell(v) - 1) + \ell(sv_I)\\ &\le (\ell(v) - 1) + \ell_I(sv_I)\\ &\le (\ell(v) - 1) + (\ell_I(v_I) + 1) = \ell(v) + \ell_I(v_I) = \ell(w).\end{aligned}$$

So then all of the inequalities must be equalities. In particular, $(vs, sv_I) \in A$, which contradicts the choice of (v, v_I), and hence the claim is proven.

Hence by induction, $v(\alpha_s)$ and $v(\alpha_t)$ are positive roots. We claim that $v_I(\alpha_s)$ is a non-negative linear combination of α_s and α_t. This will finish the proof.

First, a reduced expression for v_I is an alternating product of s and t. It must end in t (if it ends in s, then since $\ell(w) = \ell(v) + \ell(v_I)$ and $w = vv_I$, we would have $\ell(ws) < \ell(w)$), and so $\ell_I(v_I s) > \ell_I(v_I)$.

Let $m = m(s,t)$. To prove the claim, we consider whether m is finite or not. If m is finite, the claim follows from Example 1.9. So assume $m = \infty$. We have $v_I = (st)^k$ or $t(st)^k$ for some k. As shown in the previous proof, $(\sigma_s\sigma_t)^k(\alpha_s) = (2k+1)\alpha_s + 2k\alpha_t$, which has positive coefficients. Next,

$$\sigma_t(\sigma_s\sigma_t)^k(\alpha_s) = \sigma_t((2k+1)\alpha_s + 2k\alpha_t) = (2k+1)\alpha_s + (2k+2)\alpha_t.$$ □

Corollary 1.11. *$\Phi = \Phi^+ \amalg \Phi^-$ and $\Phi^- = -\Phi^+$.*

Proof. The first claim follows from the fact that $\ell(ws) \neq \ell(w)$. For the second, if $\alpha = w(\alpha_s)$ is a root, then $-\alpha = ws(\alpha_s)$ is also a root. □

Corollary 1.12. *$\sigma\colon W \to \mathbf{GL}(V)$ is faithful, i.e., $\ker\sigma$ is trivial.*

Proof. Suppose not and pick $w \in \ker\sigma$ with $\ell(w) > 1$. We can find $s \in S$ such that $\ell(ws) < \ell(w)$. By the theorem, $w(\alpha_s) = \alpha_s$ is a negative root, which is a contradiction. □

Example 1.13. Together with our previous examples, this gives a proof that the Coxeter group of the A_{n-1} Coxeter system is isomorphic to $\mathfrak{S}_n$. In particular, $\mathfrak{S}_n$ has the following presentation:

$$\mathfrak{S}_n \cong \langle s_1, \dots, s_{n-1} \mid s_i^2 = 1,\ (s_is_{i+1})^3 = 1,\ (s_is_j)^2 = 1 \text{ if } |i-j| > 1\rangle. \qquad \square$$

1.4. **Parabolic subgroups.** Previously, for $I \subseteq S$ we defined $W_I \subseteq W$ to be the subgroup generated by I. We could also consider the Coxeter group $\overline{W}_I$ generated by I with the function m restricted to $I \times I$. It is clear that we have a surjective map $\overline{W}_I \to W_I$, and a priori this may have a kernel (in principle, allowing extra generators $S \setminus I$ and relations *could* cause certain elements to be the same). Fortunately, this does not happen:

Lemma 1.14. *The map $\overline{W}_I \to W_I$ is an isomorphism.*

Proof. The subspace V_I of V generated by $\{\alpha_s \mid s \in I\}$ is canonically identified with the geometric representation of $\overline{W}_I$. So the map $\sigma_I\colon \overline{W}_I \to \mathbf{GL}(V_I)$ factors through W_I. However we know that σ_I is injective, so $\overline{W}_I \to W_I$ must also be injective. □

We call W_I a **parabolic subgroup**. From the above, it is also a Coxeter group. Here are some other important properties:

Theorem 1.15. *(1) Pick $I \subseteq S$. If $w = s_1 \cdots s_r$ is a reduced expression (with $s_i \in S$) for $w \in W_I$, then $s_i \in I$ for all i. In particular, $\ell_I = \ell|_{W_I}$.*
(2) For subsets $I, J \subseteq S$, we have $W_I \cap W_J = W_{I\cap J}$ and $W_{I\cup J}$ is the subgroup generated by W_I and W_J.
(3) S is a minimal generating set for W, i.e., no proper subset of S generates W.

Proof. (1) We prove this by induction on $\ell(w)$. The base case is clear. Otherwise, assume $r > 0$ and let $s = s_r$. By Theorem 1.10, $w(\alpha_s)$ is a negative root. Since $w \in W_I$, we have an expression $w = t_1 \cdots t_p$ with $t_j \in I$. Hence

$$w(\alpha_s) = t_1 \cdots t_p(\alpha_s) = \alpha_s + \sum_{j=1}^{p} c_j \alpha_{t_j}$$

just using what the formula for generators in the geometric representation looks like. If $s \notin \{t_1, \dots, t_p\}$, then this contradicts that all coefficients of $w(\alpha_s)$ are negative, so we conclude that $s \in I$ and hence $ws \in W_I$ as well. By induction, we conclude that $s_1, \dots, s_{r-1} \in I$.

(2) These immediately follow from (1).

(3) Suppose a proper subset I generates W, i.e., $W = W_I$ and pick $s \in S \setminus I$. Then since $s \in W_I$, from (1) we must have that $s \in I$, which is a contradiction. □

1.5. **More on roots.**

Lemma 1.16. *For every $s \in S$, σ_s preserves the set $\Phi^+ \setminus \{\alpha_s\}$. Similarly, σ_s preserves the set $\Phi^- \setminus \{-\alpha_s\}$.*

Proof. Pick $\alpha \in \Phi^+$ with $\alpha \neq \alpha_s$. By Lemma 1.4, $B(\alpha, \alpha) = 1$, and so α is not a scalar multiple of α_s. Hence we have

$$\alpha = \sum_{i \in S} c_i \alpha_i$$

where $c_i \geq 0$ for all i and $c_t > 0$ for some $t \neq s$. So the coefficient of α_t in $s(\alpha)$ is still c_t, and so $s(\alpha)$ is not a negative root. By Corollary 1.11, $s(\alpha) \in \Phi^+$. But also $s(\alpha) \neq \alpha_s$ since it has a positive coefficient for α_t.

The second statement follows from $\Phi^- = -\Phi^+$. □

Theorem 1.17. *For every $w \in W$, $\ell(w) = |\{\alpha \in \Phi^+ \mid w(\alpha) < 0\}|$.*

Proof. We prove the statement by induction on $\ell(w)$. This is clear if $\ell(w) = 0$, so assume $\ell(w) > 0$ and pick $s \in S$ such that $\ell(ws) = \ell(w) - 1$. By Theorem 1.10, we have $w(\alpha_s) < 0$. In particular, $ws(\alpha_s) > 0$. If $\alpha \in \Phi^+ \setminus \{\alpha_s\}$, then $s(\alpha) \in \Phi^+ \setminus \{\alpha_s\}$ by the previous lemma. This shows that

$$\{\alpha \in \Phi^+ \mid w(\alpha) < 0\} = \{\alpha_s\} \amalg \{\alpha \in \Phi^+ \mid ws(\alpha) < 0\}.$$

By induction, the size of the last set is $\ell(ws) = \ell(w) - 1$, so the statement of the theorem holds for w. □

Example 1.18. For the Coxeter system A_{n-1}, we have identified its geometric realization with the action of $\mathfrak{S}_n$ on $\mathbf{R}^n$ (more precisely, the zero sum hyperplane) by permutations. The positive roots are $e_i - e_j$ for $i < j$ while the negative roots are $e_i - e_j$ for $i > j$. Hence for $w \in \mathfrak{S}_n$, a positive root $e_i - e_j$ becomes a negative root if and only if $w(i) > w(j)$. A pair (i, j) such that $i < j$ and $w(i) > w(j)$ is an **inversion** of w, and so $\ell(w)$ is the number of inversions: $\ell(w) = |\{i < j \mid w(i) > w(j)\}|$. □

Given a root $\alpha = w(\alpha_s)$, we define $s_\alpha = wsw^{-1}$.

Lemma 1.19. *$\sigma(s_\alpha)(v) = v - 2B(v, \alpha)\alpha$ and hence the definition of s_α only depends on α and not the choice of s, w. Furthermore, $s_\alpha = s_{-\alpha}$ and for $\alpha, \beta \in \Phi^+$, we have $s_\alpha = s_\beta$ if and only if $\alpha = \beta$.*

Proof. For the first statement, write $\alpha = w(\alpha_s)$. Then

$$wsw^{-1}(v) = w(w^{-1}v - 2B(w^{-1}v, \alpha_s)\alpha_s) = v - 2B(v, \alpha)\alpha,$$

where in the second equality we used Lemma 1.4. Since σ is injective (Corollary 1.12), we see that a different choice of w, s would lead to the same element s_α.

The equality $s_\alpha = s_{-\alpha}$ is obvious from the formula for $\sigma(s_\alpha)$. Now suppose that $s_\alpha = s_\beta$. Since $B(\beta, \beta) = 1$, we have $-\beta = s_\beta(\beta) = s_\alpha(\beta) = \beta - 2B(\alpha, \beta)\alpha$, and so $\beta = B(\alpha, \beta)\alpha$. However, β and α are both unit vectors which means $\beta = \pm\alpha$. Since they both assumed to be positive roots, they must be equal. □

We call the elements s_α **reflections**, and we let T denote the set of all reflections. From above, T is the union of the conjugacy classes containing the elements of S, and T is in natural bijection with the set of positive roots Φ^+.

Note that $\mathrm{sgn}(t) = -1$ for any $t \in T$, so $\ell(wt) \neq \ell(w)$ for all $w \in W$.

Example 1.20. For $\mathfrak{S}_n$, the reflections are the transpositions (i, j). □

We now generalize Theorem 1.10.

Theorem 1.21. *Pick $w \in W$ and $t \in T$. Then $\ell(wt) > \ell(w)$ if and only if $w(\alpha_t)$ is a positive root.*

Proof. We first prove that $\ell(wt) > \ell(w)$ implies that $w(\alpha_t) > 0$ by induction on $\ell(w)$. If $\ell(w) = 0$ there is nothing to show, so assume $\ell(w) > 0$ and pick $s \in S$ such that $\ell(sw) = \ell(w) - 1$. Then we have

$$\ell(swt) \geq \ell(wt) - 1 > \ell(w) - 1 = \ell(sw).$$

So by induction, $sw(\alpha_t) > 0$. If $w(\alpha_t) < 0$, then by Lemma 1.16, we must have $w(\alpha_t) = -\alpha_s$. But then $\alpha_t = w^{-1}s(\alpha_s)$, which means $t = (w^{-1}s)s(sw) = w^{-1}sw$ and hence $wt = sw$. In particular, $\ell(wt) > \ell(w) > \ell(sw)$ gives a contradiction. So we conclude that $w(\alpha_t) > 0$.

To finish, we need to prove that $\ell(wt) < \ell(w)$ implies that $w(\alpha_t) < 0$. To do that, we can use what we have just shown using wt in place of w. □

1.6. **Strong exchange condition.**

Theorem 1.22. *Let $w = s_1 \cdots s_r$ with $s_i \in S$ (not necessarily a reduced expression). Pick $t \in T$ such that $\ell(wt) < \ell(w)$. Then there exists $1 \leq j \leq r$ such that $wt = s_1 \cdots s_{j-1}s_{j+1} \cdots s_r$. If the expression is a reduced expression, then j is unique.*

Proof. Write $t = s_\alpha$ where α is a positive root. By Theorem 1.21, we have $w(\alpha) < 0$. In particular, there exists j such that $s_js_{j+1} \cdots s_r(\alpha) < 0$ and $s_{j+1} \cdots s_r(\alpha) > 0$ (if $j = r$ we interpret $s_{j+1} \cdots s_r$ to be the identity). By Lemma 1.16, $s_{j+1} \cdots s_r(\alpha) = \alpha_{s_j}$. Then by definition, $t = (s_r \cdots s_{j+1})s_j(s_{j+1} \cdots s_r)$ and so

$$wt = s_1 \cdots s_r(s_r \cdots s_{j+1})s_j(s_{j+1} \cdots s_r) = s_1 \cdots s_{j-1}s_{j+1} \cdots s_r.$$

Now assume there is $i < j$ satisfying the theorem, so that

$$wt = s_1 \cdots s_{i-1}s_{i+1} \cdots s_r = s_1 \cdots s_{j-1}s_{j+1} \cdots s_r.$$

Cancelling off $s_1 \cdots s_{i-1}$ from the beginning and $s_{j+1} \cdots s_r$ from the end gives $s_{i+1} \cdots s_j = s_i \cdots s_{j-1}$, which implies $s_i \cdots s_j = s_{i+1} \cdots s_{j-1}$. In particular, we can reduce the length of the original expression for w by 2, which means it was not reduced. □

Corollary 1.23. *If $w = s_1 \cdots s_r$ is not a reduced expression, then we can find $1 \leq i < j \leq r$ such that $w = s_1 \cdots s_{i-1}s_{i+1} \cdots s_{j-1}s_{j+1} \cdots s_r$. In particular, given any expression for w as a product of elements of S, it is always possible to remove an even number of the s to get a reduced expression.*

Let $I \subseteq S$ and define $W^I = \{w \in W \mid \ell(ws) > \ell(w) \text{ for all } s \in I\}$.

Proposition 1.24. *Given $w \in W$, there is a unique $u \in W^I$ and $v \in W_I$ such that $w = uv$. Furthermore, $\ell(w) = \ell(u) + \ell(v)$. Moreover, u is the unique element of minimal possible length in the coset wW_I.*

Proof. Pick $u \in wW_I$ of minimal possible length and let $v = u^{-1}w \in W_I$. Then $u \in W^I$: if $s \in I$, then $us \in wW_I$, and hence by our choice of u, we have $\ell(u) < \ell(us)$. Now pick reduced expressions $u = s_1 \cdots s_p$ and $v = s'_1 \cdots s'_q$. By Theorem 1.15(1), we have $s'_i \in I$ for all i. Suppose that $s_1 \cdots s_ps'_1 \cdots s'_q$ is not a reduced expression. Then by Corollary 1.23, it is possible to delete two of the generators without affecting the product. If one of them is an s_i, then we have found a coset representative with smaller length than u, which is a problem. So then both must be of the form s'_i, which contradicts that we picked a reduced expression for v. Hence $\ell(w) = \ell(u) + \ell(v)$.

Suppose that $u' \in wW_I$ satisfies $\ell(u') = \ell(u)$. Then $u' = uv'$ with $v' \in W_I$ and the same argument as above shows that $\ell(u') = \ell(u) + \ell(v')$, so in particular, $v' = 1$ and $u' = u$. Hence u is the unique element of minimal length in wW_I.

Finally suppose there is another $u' \in W^I$ such that $u' \in wW_I$ with $u' \neq u$. By the previous paragraph, $\ell(u') - \ell(u) = r > 0$, and we can write $u' = us_1 \cdots s_r$ with $s_i \in I$. But then $\ell(u's_r) < \ell(u')$ which contradicts that $u' \in W^I$. Hence u is unique element of $W^I \cap wW_I$. ☐

Hence, the elements of W^I are called **minimal length coset representatives**.

Example 1.25. Consider the symmetric group case $W = \mathfrak{S}_n$. Fix $i < n$ and let $I = S \setminus \{s_i\}$. Then $W_I \cong \mathfrak{S}_i \times \mathfrak{S}_{n-i}$ is the subgroup that preserves the sets $\{1, \dots, i\}$ and $\{i+1, \dots, n\}$. Hence, the minimal length coset representatives are those permutations w that satisfy $w(1) < w(2) < \cdots < w(i)$ and $w(i+1) < \cdots < w(n)$. This is equivalent to a choice of the i-element subset $\{w(1), \dots, w(i)\}$, which is consistent with the fact that there are $\frac{n!}{i!(n-i)!}$ cosets. Every parabolic subgroup of $\mathfrak{S}_n$ is of the form $\mathfrak{S}_{n_1} \times \cdots \times \mathfrak{S}_{n_d}$ where $n_1 + \cdots + n_d = n$, and the description of minimal length coset representatives is similar. ☐

1.7. **Bruhat order.** We now introduce a partial ordering on W. Given elements $v, w \in W$, write $v \to w$ to mean that $\ell(w) > \ell(v)$ and $v^{-1}w \in T$. Clearly, $\to$ is an antisymmetric relation; we let $\leq$ be the partial ordering generated by $\to$, i.e., $w' \leq w$ if there exists $w_0, \dots, w_n$ such that $w' = w_0 \to w_1 \to \cdots \to w_n = w$ (we allow $n = 0$ which just says that $w \leq w$). This is the **Bruhat order** on W. It turns out to have a lot of important uses, though many are beyond the scope of this course.

Proposition 1.26. *Let $w' \leq w$ and $s \in S$. Then at least one of the following hold:*

(1) $w's \leq w$,
(2) $w's \leq ws$.

Proof. We first handle the case $w' \to w$. So $w = w't$ where $t \in T$ and $\ell(w) > \ell(w')$. If $s = t$, we're done, so assume that $s \neq t$. We consider how $\ell(w's)$ compares to $\ell(w')$.

If $\ell(w's) = \ell(w') - 1$, then we have $w's \to w' \to w$ and hence $w's \leq w$.

Otherwise, we have $\ell(w's) = \ell(w') + 1$. If we pick a reduced expression $w' = s_1 \cdots s_r$, then $s_1 \cdots s_r s_{r+1}$ $(s_{r+1} = s)$ is a reduced expression for $w's$. Let $t' = sts$ so that $ws = (w's)t'$. We claim that $\ell(w's) < \ell(ws)$. If not, then the strong exchange condition (Theorem 1.22) implies that there is a unique $1 \leq j \leq r+1$ such that $s_1 \cdots s_{j-1} s_{j+1} \cdots s_{r+1}$ is a reduced expression for ws. Since $s \neq t$, we must have $j < r+1$, so $w = s_1 \cdots s_{j-1} s_{j+1} \cdots s_r$. But then $\ell(w) = r - 1 < \ell(w')$, which contradicts our original assumption, and our claim is proven. Thus $w's \to ws$ since $(ws)^{-1}(w's) = sts = t'$.

For the general case, there are $w_0, \dots, w_n$ such that $w' = w_0 \to w_1 \to \cdots \to w_n = w$. We handle it by induction on n, where the base case $n = 0$ is obvious. From what we've shown, we have $w's \leq w_1$ or $w's \leq w_1 s$. In the first case, we use that $w_1 \leq w$ to conclude that $w's \leq w$. In the second case, we use induction to conclude that $w_1 s \leq w$ (and hence $w's \leq w$) or $w_1 s \leq ws$ (and hence $w's \leq ws$). ☐

Given a product $s_1 \cdots s_r$, a **subword** is $s_{i_1} \cdots s_{i_p}$ where $1 \leq i_1 < \cdots < i_p \leq r$.

Theorem 1.27. *Let $w \in W$ and pick a reduced expression $w = s_1 \cdots s_r$. Then $v \leq w$ if and only if v is a subword of $s_1 \cdots s_r$. In particular, the number of v below w is finite.*

Proof. First suppose that $v \leq w$ so we have $v = w_0 \to w_1 \to \cdots \to w_n = w$. We show that v is a subword of $s_1 \cdots s_r$ by induction on n. The base case $n = 0$ is vacuous, so assume $n \geq 1$.

We have $w_{n-1}^{-1}w \in T$ and $\ell(w_{n-1}) < \ell(w)$. The strong exchange condition (Theorem 1.22) and its corollary then implies that there is a subword which is a reduced expression for w_{n-1}. Hence, by induction on n, some subword of this reduced expression multiplies to v; this in particular is a subword of $s_1 \cdots s_r$.

Conversely, suppose we have $v = s_{i_1} \cdots s_{i_p}$ where $1 \le i_1 < \cdots < i_p \le r$. We need to show that $v \le w$, and we proceed by induction on $r = \ell(w)$. Again, $\ell(w) = 0$ is vacuous, so we assume $\ell(w) > 0$. If $i_p < r$, then v is a subword of $s_1 \cdots s_{r-1} = ws_r$, and so by induction, $v \le ws_r$. But also $ws_r \to w$, so $v \le w$. Otherwise, we have $i_p = r$. Then $s_{i_1} \cdots s_{i_{p-1}}$ is a subword of $s_1 \cdots s_{r-1}$, so again by induction, $vs_r \le ws_r$. By Proposition 1.26, we have either $v \le w$ or $v \le ws_r$. Since $ws_r \to w$, we conclude either way that $v \le w$. □

As a partially ordered set (poset), W satisfies a lot of strong properties. For example, it is *graded*: between any two elements $v \le w$, the length of any maximal chain $v < x_1 < \cdots < x_{n-1} < w$ has the same size $n = \ell(w) - \ell(v)$ and *shellable* (definition omitted). We won't explore them in this course, but see [BB, §2.7] for a starting point.

Example 1.28. Consider the dihedral group of order $2m$ with Coxeter generators s, t. For $i = 1, \dots, m-1$, there are exactly 2 elements of length i, which are alternating products of s and t. The two alternating products of length m are equal: if $m = 2u$ is even, then multiplying the relation $(st)^m = 1$ on the right by $(ts)^u$ gives $(st)^u = (ts)^u$, and if $m = 2u+1$ is odd, then multiplying the relation $(st)^m = 1$ on the right by $(ts)^u t$ gives $(st)^u s = (ts)^u t$. We see from the subword description of the Bruhat order that $v \le w$ if and only if $\ell(v) < \ell(w)$. This also works if $m = \infty$. □

Recall that for $I \subseteq S$, W_I is a Coxeter group. Hence it has a Bruhat order, in addition to the restriction of the Bruhat order of W to it. Fortunately these agree:

Corollary 1.29. *For $I \subseteq S$, the restriction of the Bruhat order from W to W_I agrees with the Bruhat order on W_I.*

Proof. Pick $v, w \in W_I$. By Theorem 1.15, every reduced expression for w uses only generators from I. Fix one. Then $v \le w$ if and only if v is a subword of this expression (Theorem 1.27), and the truth of this latter condition does not depend on whether we consider v, w as elements of W_I or W. □

Remark 1.30. When W is the Weyl group of a semisimple complex Lie algebra (or more generally a Kac–Moody algebra), the Bruhat order plays a fundamental role in the theory of Schubert varieties and controls when one contains another. □

1.8. **Fundamental domain.** We continue to let V denote the geometric representation of (W, S). The goal now is to describe a topological space that W acts on together with a fundamental domain for that action. This will be used to show that $W \subset \mathbf{GL}(V)$ is a discrete subspace (where $\mathbf{GL}(V)$ is given the induced topology from $\mathbf{R}^{n^2}$), a fact that we will use later.

Let V^* denote the dual space of V; this has an action of W via $(wf)(v) = f(w^{-1}v)$ for $w \in W$, $f \in V^*$ and $v \in V$.

For each $s \in S$, define

$$\begin{aligned} Z_s &= \{f \in V^* \mid f(\alpha_s) = 0\}, \\ A_s &= \{f \in V^* \mid f(\alpha_s) > 0\}, \\ A'_s &= \{f \in V^* \mid f(\alpha_s) < 0\}. \end{aligned}$$

The closure $\overline{A}_s$ of A_s is $A_s \cup Z_s$, we let $C = \bigcap_{s \in S} A_s$ and $D = \overline{C} = \bigcap_{s\in S} \overline{A}_s$. In other words,

$$D = \{f \mid f(\alpha_s) \geq 0 \text{ for all } s \in S\}$$

is the nonnegative orthant.

For each subset $I \subseteq S$, we define

$$C_I = \bigcap_{s \in I} Z_s \cap \bigcap_{s \notin I} A_s = \{f \in D \mid f(\alpha_s) = 0 \text{ if and only if } s \in S\}.$$

In particular, the C_I are disjoint and partition D.

For $w \in W$, define $w(D) = \{w(x) \mid x \in D\}$. The union $\mathcal{C} = \bigcup_{w \in W} w(D)$ is the **Tits cone** of W.

Theorem 1.31. *We have the following properties.*

(a) *Pick $w \in W$ and $I, J \subseteq S$. If $w(C_I) \cap C_J \neq \varnothing$, then $I = J$ and $w \in W_I$.*
(b) *The stabilizer of any point of C_I is W_I.*
(c) *Every W-orbit in $\mathcal{C}$ intersects D in exactly one point.*

Proof. (a) We prove the statement by induction on $\ell(w)$. Since the case $\ell(w) = 0$ is obvious, we assume $\ell(w) > 0$ and pick $s \in S$ such that $\ell(sw) < \ell(w)$. Then $\ell(w^{-1}s) < \ell(w^{-1})$, so by Theorem 1.10, we see that $w^{-1}(\alpha_s)$ is a negative root. In particular, if $f \in D$, then $(wf)(\alpha_s) = f(w^{-1}(\alpha_s)) \leq 0$, so that $wf \in \overline{A}'_s$, i.e., $w(D) \subset \overline{A}'_s$. Hence

$$w(C_I) \cap C_J \subset w(D) \cap D \subset \overline{A}'_s \cap \overline{A}_s = Z_s.$$

Pick $f \in w(C_I) \cap C_J$ (which is nonempty by assumption). Then $f \in Z_s$ so that $f(\alpha_s) = 0$ which means that $s \in J$ by definition of C_J. For any $g \in Z_s$, we have

$$(sg)(v) = g(sv) = g(v - 2B(\alpha_s, v)\alpha_s) = g(v)$$

so $sg = g$. This implies that $s(C_J) = C_J$ and $f \in sw(C_I) \cap C_J$. By induction, we must have $I = J$ and $sw \in W_I$, and so $w \in W_I$.

(b) Pick $x \in C_I$. If $wx = x$, then $w(C_I) \cap C_I \neq \varnothing$ and so by (a), we have $w \in W_I$. Conversely, from above, each $s \in I$ fixes C_I, and so the stabilizer of any point of C_I is W_I.

(c) By definition, every W-orbit in $\mathcal{C}$ intersects D in at least one point. Suppose that $f, g \in D$ are in the same W-orbit, i.e., $f = w(g)$. Then $f \in C_J$ and $g \in C_I$ for some subsets $I, J \subset S$ and so $w(C_I) \cap C_J \neq \varnothing$. By (a), $I = J$ and $w \in W_I$. By (b), $w(g) = g$, which means $f = g$. □

If we pick a basis for V, then we can identify $\mathbf{GL}(V)$ with the set of non-invertible $n \times n$ matrices. The set of all $n \times n$ matrices is a Euclidean space of dimension n^2, so has a natural topology. Using this, $\mathbf{GL}(V)$ inherits the subspace topology (the topology we get does not depend on the choice of basis because a different identification amounts to conjugation by a change of basis matrix which is a homeomorphism on $\mathbf{R}^{n^2}$). We note that $\mathbf{GL}(V) \times V \to V$ given by $(g, v) \mapsto g(v)$ is given by linear functions and hence is continuous. Similarly, $\mathbf{GL}(V) \times V^* \to V^*$ is continuous.

We say that a subset A of $\mathbf{GL}(V)$ is discrete if the subspace topology on it is discrete, i.e., for every $x \in A$, there is an open set $U \subset \mathbf{GL}(V)$ such that $U \cap A = \{x\}$.

Corollary 1.32. *$W \subset \mathbf{GL}(V)$ is a discrete subset.*

Proof. Pick $f \in C$ and $w \in W$. The function $\mathbf{GL}(V) \to V^*$ given by $g \mapsto gw^{-1}f$ is continuous. Since $C \subset V^*$ is open, its preimage C' in $\mathbf{GL}(V)$ is also open. If $v \in C'$, then $vw^{-1}f \in C$. By Theorem 1.31(c), we must have $vw^{-1}f = f$. Since $C = C_\varnothing$, by Theorem 1.31(b), we have $v = w$. In particular, for each $w \in W$ there is an open set of $\mathbf{GL}(V)$ which only contains w, and hence the induced topology on W is discrete. □

1.9. **Poincaré series.** As before (W, S) is a Coxeter group. For any subset $U \subseteq W$, we define a formal power series

$$U(t) = \sum_{w \in U} t^{\ell(w)} = \sum_{n \geq 0} |\{w \in U \mid \ell(w) = n\}| t^n.$$

Note that since S is a finite set, the coefficient of t^n is always finite. We will primarily be interested in the case $U = W$, and for $I \subseteq S$, the cases $U = W_I$ and $U = W^I$, where the latter is the set of minimal length coset representatives of W_I. By Proposition 1.24, we have

$$W(t) = W_I(t) W^I(t).$$

Lemma 1.33. *If W is finite, then there is a unique element w_0 of maximal possible length, and w_0 sends every positive root to a negative root.*

Proof. First, the number of roots is finite (each one gives a reflection in W and they coincide if and only if the roots differ by a sign). Let w_0 be an element of maximal possible length. By Theorem 1.10, we must have $w_0(\alpha_s) < 0$ for all $s \in S$, and hence w_0 sends every positive root to a negative root. By Theorem 1.17, $\ell(w_0)$ is the number of positive roots. This implies that w_0^2 sends every positive root to a positive root, so $w_0^2 = 1$. Similarly, if w' is any element with $\ell(w_0) = \ell(w')$, then w' must also send all positive roots to negative roots. This implies that $w'w_0$ sends all positive roots to positive roots, and so $\ell(w'w_0) = 0$, which implies that $w' = w_0^{-1} = w_0$. □

Given $w \in W$, we define its **(right) descent set** to be $D_R(w) = \{s \in S \mid \ell(ws) < \ell(w)\}$. Then $W^I = \{w \in W \mid D_R(w) \subseteq S \setminus I\}$.

Proposition 1.34. *We have*

$$\sum_{I \subseteq S} (-1)^{|I|} W^I(t) = \begin{cases} t^{\ell(w_0)} & \textit{if } W \textit{ is finite} \\ 0 & \textit{otherwise} \end{cases}.$$

Proof. In the sum above, the contribution of $w \in W$ is

$$t^{\ell(w)} \sum_{I \subseteq S \setminus D_R(w)} (-1)^{|I|} = \begin{cases} t^{\ell(w)} & \text{if } D_R(w) = S \\ 0 & \text{otherwise} \end{cases}.$$

By Theorem 1.10, $D_R(w) = S$ if and only if $w(\alpha_s) < 0$ for all $s \in S$, which is equivalent to $w(\alpha) < 0$ for every positive root α.

If W is infinite, no such w exists, as $\ell(w) < \infty$ is the number of positive roots that become negative under w (Theorem 1.17) and ℓ is unbounded (since the number of elements of a given length is finite). Hence $\sum_{I \subseteq S} (-1)^{|I|} W^I(t) = 0$ in this case.

If W is finite, then $D_R(w) = S$ implies that $\ell(w)$ is the number of positive roots, which forces $w = w_0$ by Lemma 1.33. □

Note that $W^S(t) = 1$. Subtract this term from both sides and divide both sides by $W(t)$ (using that $W(t) = W_I(t)W^I(t)$) to get

$$\sum_{I \subsetneq S} \frac{(-1)^{|I|}}{W_I(t)} = \frac{f(t)}{W(t)} \tag{1.35}$$

where $f(t) = \begin{cases} t^{\ell(w_0)} - (-1)^{|S|} & \text{if } W \text{ is finite} \\ -(-1)^{|S|} & \text{otherwise} \end{cases}$.

This allows us to calculate $W(t)$ by induction on $|S|$ as soon as we can determine $\ell(w_0)$ for W finite. We will do this later.

Example 1.36. If $|S| = 0$, then $W(t) = 1$.

If $|S| = 1$, there is also nothing to do since $W \cong \mathbf{Z}/2$, so $W(t) = 1 + t$. Note that this is consistent with the recursion: $1 = \frac{t+1}{W(t)}$. ☐

In principle, this almost gives us an algorithm to compute $W(t)$ in general. However, we need to know when W is finite, and when that happens, how to compute $\ell(w_0)$. From what we've discussed, this is the number of positive roots. We'll address these questions in the next section.

Example 1.37. We can compute $W(t)$ for $W = \mathfrak{S}_n$ by induction on n. For this, we write permutations in 1-line notation, i.e., $w(1)w(2)\cdots w(n)$. Given such an element, we can insert $n+1$ in one of $n+1$ places. If we insert it right after $w(i)$, each of $(j, n+1)$ for $j > i$ is an inversion. If we insert at the beginning, then $(j, n+1)$ is an inversion for all j. In particular, we get

$$\mathfrak{S}_{n+1}(t) = \mathfrak{S}_n(t) \cdot (1 + t + \cdots + t^n) = \mathfrak{S}_n(t)\frac{1 - t^{n+1}}{1-t} = \prod_{i=1}^{n} \frac{1 - t^{i+1}}{1-t}. \qquad \square$$

2. Finite Coxeter groups

2.1. **Group representations.** This section is mostly a survey of introductory representation theory that we will use throughout the course, so it makes more sense to just refer to it as needed.

Let G be a finite group. The identity element of G will be called 1_G or just 1. A **(linear) representation** of G over $\mathbf{C}$ is a homomorphism

$$\rho_V \colon G \to \mathbf{GL}(V)$$

for some $\mathbf{C}$-vector space V, where $\mathbf{GL}(V)$ is the group of invertible linear operators on V. Equivalently, giving a representation is the same as giving a linear action of G on V, i.e., a function $G \times V \to V$ which we think of as a multiplication $g \cdot v$ for $g \in G$ and $v \in V$ such that:

- $g \cdot (v + v') = g \cdot v + g \cdot v'$,
- $(gg') \cdot v = g \cdot (g' \cdot v)$,
- $1_G \cdot v = v$, and
- $g \cdot (\lambda v) = \lambda(g \cdot v)$ for any $\lambda \in \mathbf{C}$.

The multiplication is obtained by setting $g \cdot v = \rho_V(g)(v)$. We will usually assume that V is finite-dimensional, or built out of finite-dimensional pieces in a controlled way.

We will generally take the perspective that V "is" the representation, and the information ρ_V is implicit but not always mentioned. So properties of a representation such as dimension, being nonzero, etc. come from the vector space V.

Let V and V' be two representations of G. A linear map $f\colon V \to V'$ is **G-equivariant** if for all $g \in G$, we have

$$f \circ \rho_V(g) = \rho_{V'}(g) \circ f,$$

or more compactly: $f(g \cdot v) = g \cdot f(v)$ for all $v \in V$. An **isomorphism** is a G-equivariant map which is invertible; if an isomorphism exists we write $V \cong V'$.

A subspace $V' \subseteq V$ is a **subrepresentation** if $g \cdot v \in V'$ for all $g \in G$ and $v \in V'$. A representation V is **irreducible** if the only subrepresentations are V and 0.

Example 2.1. (1) For any vector space V we can define $\rho_V(g)$ to be the identity on V. This is clearly a representation. When $\dim V = 1$, this is called the **trivial representation**.

(2) Let X be a finite set with a G-action. Recall this means that we have a function $G \times X \to X$ denoted $(g, x) \mapsto g \cdot x$ such that $1_G \cdot x = x$ for all $x \in X$, and $g \cdot (g' \cdot x) = (gg') \cdot x$ for all $g, g' \in G$ and $x \in X$. Let $V = \mathbf{C}[X]$ be the $\mathbf{C}$-vector space with basis $\{e_x \mid x \in X\}$ and define ρ_V by $\rho_V(g)e_x = e_{g\cdot x}$. This is the **permutation representation** of X.

A special case of a group action is when $X = G$ and $g \cdot x = gx$ is given by the group operation. In that case, $\mathbf{C}[G]$ is called the **regular representation**. □

$\mathbf{C}[G]$ has a natural multiplication: on basis vectors it is $e_g e_{g'} = e_{gg'}$, and then extend it linearly using the distributive law. A G-representation V is equivalent to a left $\mathbf{C}[G]$-module: we define $e_g v = \rho_V(g)v$ (details omitted). This is the **group algebra**.

Lemma 2.2. *All of the eigenvalues of $\rho(g)$ are roots of unity, and $\rho(g)$ is diagonalizable for all $g \in G$.*

Proof. Let λ be an eigenvalue of $\rho(g)$ with eigenvector v. Then $\rho(g)^{|G|} = 1$ but also $\rho(g)^{|G|}v = \lambda^{|G|}v$, so $\lambda^{|G|} = 1$.

Consider the Jordan normal form of $\rho(g)$, which recall is an upper-triangular matrix whose diagonal entries are the eigenvalues of $\rho(g)$ and whose superdiagonal (the entries in positions $(i, i+1)$) are either 0 or 1. Then $\rho(g)$ is diagonalizable if and only if the superdiagonal is 0. Furthermore, if any of those entries are 1, then no positive power of $\rho(g)$ is equal to the identity, but we know that $\rho(g)^{|G|} = 1$. □

We define

$$\operatorname{Tr}(g \mid V) = \operatorname{Tr} \rho_V(g).$$

This gives a function $\chi_V\colon G \to \mathbf{C}$ defined by $\chi_V(g) = \operatorname{Tr}(g \mid V)$. We will also denote it by $\operatorname{char}(V)$. Define a **class function** to be a function $G \to \mathbf{C}$ which is invariant under conjugation. They form a vector space over $\mathbf{C}$. An example of a class function is χ_V.

We define an inner product on the space of class functions on G via

$$\langle \varphi, \psi \rangle_G = \frac{1}{|G|} \sum_{g \in G} \overline{\varphi(g)} \psi(g).$$

where the bar is complex conjugation.

Let V and W be representations of G. There are a few basic operations we will make use of:

- (Direct sum) The direct sum $V \oplus W$ is a representation with multiplication given by $g \cdot (v, w) = (g \cdot v, g \cdot w)$. Then $\chi_{V \oplus W} = \chi_V + \chi_W$.
- (Dual) Recall that the dual space V^* is the vector space of linear functionals $V \to \mathbf{C}$. It is a representation with multiplication given as follows: if $f \in V^*$, then $g \cdot f$ is the functional defined by $(g \cdot f)(v) = f(g^{-1} \cdot v)$. Then $\chi_{V^*} = \overline{\chi}_V$.
- (Tensor product) Recall that the tensor product $V \otimes W$ is a vector space which is spanned by symbols of the form $v \otimes w$ with $v \in V$ and $w \in W$ subject to the relations
  - $(v + v') \otimes w = v \otimes w + v' \otimes w$,
  - $v \otimes (w + w') = v \otimes w + v \otimes w'$,
  - $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$ for any $\lambda \in \mathbf{C}$.

  Then $V \otimes W$ is a representation of G via $g \cdot \sum_i (v_i \otimes w_i) = \sum_i (g \cdot v_i) \otimes (g \cdot w_i)$. Then $\chi_{V \otimes W} = \chi_V \chi_W$.
- (Symmetric powers) The symmetric group $\mathfrak{S}_n$ acts on $V^{\otimes n} = V \otimes \cdots \otimes V$ by permuting tensor factors. The quotient by this action, i.e., identifying $\sigma(x) = x$ for all $x \in V^{\otimes n}$, is the symmetric power $\mathrm{Sym}^n V$, which inherits an action of G. We write the coset of $x_1 \otimes \cdots \otimes x_n$ by $x_1 \cdots x_n$. If $v_1, \dots, v_r$ is a basis for V, then $\{v_{i_1} \cdots v_{i_n} \mid 1 \le i_1 \le \cdots \le i_n \le r\}$ is a basis for $\mathrm{Sym}^n V$.
- (Exterior powers) Modify the previous construction by identifying $\sigma(x) = \mathrm{sgn}(\sigma) x$ for all $x \in V^{\otimes n}$ to get the exterior power $\bigwedge^n V$, which inherits an action of G. We write the coset of $x_1 \otimes \cdots \otimes x_n$ by $x_1 \wedge \cdots \wedge x_n$. If $v_1, \dots, v_r$ is a basis for V, then $\{v_{i_1} \wedge \cdots \wedge v_{i_n} \mid 1 \le i_1 < \cdots < i_n \le r\}$ is a basis for $\bigwedge^n V$.
- (Invariants) $V^G = \{v \in V \mid g \cdot v = v \text{ for all } g \in G\}$ is the space of G-invariants and is clearly a subrepresentation of V.

Proposition 2.3. *The following properties hold:*

1. *If $\chi_V = \chi_{V'}$ then $V \cong V'$.*
2. *If $\chi_V(g) = 0$ for all $g \neq 1$, then V is isomorphic to a direct sum of copies of the regular representation, i.e., $V \cong \mathbf{C}[G]^{\oplus N}$ where $N = (\dim V)/|G|$.*
3. *χ_V is real-valued if and only if $V \cong V^*$.*
4. *$\langle \chi_V, \chi_{V'} \rangle_G = \dim(V^* \otimes V')^G$.*

Given a subgroup $H \subseteq G$, any representation ρ of G becomes a representation of H by restricting the map. This is called the **restriction** of ρ, and is denoted $\mathrm{Res}^G_H \rho$. In fact, restriction makes sense for any class function.

On the other hand, given a representation $\rho\colon H \to \mathbf{GL}(V)$, one can define the **induced representation** $\mathrm{Ind}^G_H V$ which is a representation of G. This is conceptually clearest to define using tensor products. First, as before, V is a left $\mathbf{C}[H]$-module. Second, $\mathbf{C}[G]$ can be made into a *right* $\mathbf{C}[H]$-module via $g \cdot h = gh$ for $g \in G$ and $h \in H$. We can then define the tensor product over $\mathbf{C}[H]$: $\mathbf{C}[G] \otimes_{\mathbf{C}[H]} V$. More generally, if R is a (not necessarily commutative) ring and M is a right R-module and N is a left R-module, then $M \otimes_R N$ is the abelian group spanned by symbols $m \otimes n$ with $m \in M$ and $n \in N$ subject to the relations:

- $(m + m') \otimes n = m \otimes n + m' \otimes n$,
- $m \otimes (n + n') = m \otimes n + m \otimes n'$,
- $mr \otimes n = m \otimes rn$ for any $r \in R$.

In general, there is no further structure on $M \otimes_R N$. In our case, we can make $\mathbf{C}[G] \otimes_{\mathbf{C}[H]} V$ into a left $\mathbf{C}[G]$-module by $g \cdot (\sum_i e_{g_i} \otimes v_i) = \sum_i e_{gg_i} \otimes v_i$. Note that the tensor product will be a $\mathbf{C}$-vector space. If $v_1, \dots, v_n$ is a basis for V and $g_1, \dots, g_r$ are representatives for the left cosets G/H, then a basis for $\mathrm{Ind}_H^G V$ is $\{e_{g_i} \otimes v_j\}$, so in particular, $\dim(\mathrm{Ind}_H^G V) = |G/H| \dim V$.

Example 2.4. Let X be a set with a transitive G-action, i.e., for all $x, y \in X$, there exists $g \in G$ so that $gx = y$. Pick any point $x \in X$ and let H be its stabilizer. Then the left action of G on G/H is the same as the action of G on X under the identification $gH \mapsto gx$. Hence $\mathbf{C}[G/H] \cong \mathbf{C}[X]$. Furthermore, we can identify $\mathbf{C}[G/H]$ with $\mathrm{Ind}_H^G \mathbf{C}$ where $\mathbf{C}$ is the trivial representation of H. We denote its character by $1 \uparrow_H^G$. ☐

Theorem 2.5 (Frobenius reciprocity)**.** *Given groups $H \subseteq G$ and representations U, V of H and G respectively, we have*

$$\langle \mathrm{Ind}_H^G U, V \rangle_G = \langle U, \mathrm{Res}_H^G V \rangle_H.$$

This is a corollary of the usual adjunction between hom and tensor, but we won't go into the details.

In particular, we have $\langle 1 \uparrow_H^G, V \rangle_G = \dim V^H$, which is how it will be used later.

2.2. **Finite groups.** It will be convenient to encode a Coxeter group by a finite graph Γ. The vertex set is S. For $s \neq s'$, we connect s and s' with $m(s, s') - 2$ many edges. The standard convention to denote multiple edges is to draw a single edge with the number $m(s, s')$ above it. (Hence if $m(s, s') = 3$ we do not decorate the edge.) We will call Γ a Coxeter graph.

We call (W, S) **irreducible** if Γ is connected. Taking the disjoint union of two graphs amounts to taking the direct product of Coxeter groups (generators without edges commute with each other) and also taking the direct sum of geometric realizations. So if we want to classify the finite Coxeter groups, we can focus our attention on irreducible Coxeter systems.

Recall that $G \subset \mathbf{GL}(V)$ acts **irreducibly**, or that V is an **irreducible representation** of G, if the only G-invariant subspaces of V are 0 and V. An element $g \in \mathbf{GL}(V)$ is a **reflection** if $\ker(1 - g)$ is a hyperplane and g has finite order (this agrees with our previous definition when g belongs to a Coxeter group in its geometric representation).

Lemma 2.6. *Let $G \subset \mathbf{GL}_n(\mathbf{R})$ be a finite group acting irreducibly on $\mathbf{R}^n$ and that contains at least one reflection. Then there is a unique bilinear form (up to scalar multiple) which is preserved by G. Furthermore, this bilinear form is, up to a sign, symmetric positive definite.*

Proof. If B is a nonzero bilinear form preserved by G, then it gives a G-equivariant linear map $\mathbf{R}^n \to (\mathbf{R}^n)^*$ via $v \mapsto B(v, -)$. The kernel is a G-invariant subspace, so since G acts irreducibly, this map is an isomorphism. Let B' be another nonzero bilinear form preserved by G. Then for every $v \in \mathbf{R}^n$, $B(v, -)$ must equal $B'(v', -)$ for some unique vector v', call it $\varphi(v)$. Then $\varphi \colon \mathbf{R}^n \to \mathbf{R}^n$ is a linear map. By G-equivariance of B and B', we have $B(gv, -) = B'(gv', -)$, so φ is G-equivariant.

Let $s \in G$ be a reflection. Then image$(1 - s)$ is 1-dimensional, let $(1 - s)v$ be a nonzero vector. Since φ is G-equivariant, we have $\varphi(1 - s)v = (1 - s)\varphi(v)$, which is a multiple of $(1 - s)v$, so $(1 - s)v$ is an eigenvector of φ; call its eigenvalue α. Then $\varphi - \alpha I$ is also G-equivariant, so $\ker(\varphi - \alpha I)$ is a nonzero G-invariant subspace. By irreducibility, it must be the whole space, so $\varphi = \alpha I$ and hence $B' = \alpha B$.

For the last part, pick any symmetric positive definite bilinear form B' and define $B(v, w) = \sum_{g \in G} B'(gv, gw)$. Then B is preserved by G and clearly symmetric. Suppose that $v \neq 0$. Then for any $g \in G$, we have $gv \neq 0$ and hence $B'(gv, gv) > 0$. So $B(v, v) > 0$ and B is positive definite. By uniqueness, every G-invariant form is a scalar multiple of B. ☐

Remark 2.7. The second part of the proof is essentially Schur's lemma, except we cannot apply it over the real numbers since a G-invariant operator may not have an eigenvalue. Here we are using the existence of a reflection to force the existence of an eigenvalue. □

Theorem 2.8 (Maschke)**.** *Let U be a vector space over a field of characteristic 0, and let G be a finite subgroup of $\mathbf{GL}(U)$. If $U' \subseteq U$ is a G-invariant subspace, then there is a complementary G-invariant subspace U'' (i.e., $U' \cap U'' = 0$ and $U' + U'' = U$).*

Proof. Pick any complementary subspace X of U' and let $p\colon U \to U'$ be the projection operator whose kernel is X. Then $\frac{1}{|G|}\sum_{g\in G} gp$ is also a projection operator and its kernel U'' is complementary to U' and preserved by G. □

Lemma 2.9. *Let (W, S) be a Coxeter group. If the Coxeter graph of W is connected and W is finite, then the geometric representation V is an irreducible representation of W.*

Proof. Let $V' \subseteq V$ be a subrepresentation and set $S' = \{s \in S \mid \alpha_s \in V'\}$. If $v \in V'$ and $t \notin S'$, then $\sigma_t(v) - v = -2B_W(v, \alpha_t)\alpha_t$ is an element of V' which means $B_W(v, \alpha_t) = 0$. In particular, the induced subgraph on S' is a connected component, so either $S' = S$ (in which case $V' = V$) or $S' = \varnothing$ (in which case $V' \subseteq \ker B_W$). In other words, any proper subrepresentation must be a subspace of $\ker B_W$. To finish, we show that $\ker B_W = 0$.

Since $\ker B_W$ is a W-invariant subspace, it has a W-invariant complement V'' by Maschke's theorem (Theorem 2.8). If $V'' = V$, then $\ker B_W = 0$, and we are done. Otherwise, V'' is a proper subrepresentation, so by the above, $V'' \subset \ker B_W$. But $V'' \cap \ker B_W = 0$ which means that $V'' = 0$ and $\ker B_W = V$. This means that $B_W = 0$ which is false by definition. □

Remark 2.10. This fails when W is infinite. For example, when W is the infinite dihedral group with $S = \{s, t\}$, we've seen that $\alpha_s + \alpha_t$ is fixed by every element of W. □

Proposition 2.11. *W is finite if and only if the bilinear form B_W is positive definite.*

Proof. Suppose that W is finite. Consider the geometric representation $W \to \mathbf{GL}(V)$. It suffices to consider the case that the Coxeter graph is connected. In that case, W acts irreducibly on V by Lemma 2.9, and hence by Lemma 2.6, either B_W or $-B_W$ is positive definite. But $B_W(\alpha_s, \alpha_s) = 1$ for any $s \in S$, so in fact B_W is positive definite.

Conversely, suppose that B_W is positive definite. Then the orthogonal group preserving it is compact (with respect to an orthonormal basis, the condition for A to be in the orthogonal group is $AA^T = I$, which means it is closed; each column of an orthogonal matrix lies on the unit sphere, and hence the orthogonal group is bounded). By Corollary 1.32, W is a discrete subgroup. A discrete subset of a compact space must be finite. □

In the next sections, we will actually classify all possible cases when B_W is positive definite and try to describe the groups as explicitly as possible.

The following isn't logically needed for what follows in the notes, but is a nice fact to know.

Theorem 2.12. *Any finite subgroup $G \subset \mathbf{GL}_n(\mathbf{R})$ generated by reflections is isomorphic to a Coxeter group.*

Proof. First, $\mathbf{R}^n$ has a positive definite G-invariant symmetric form: $(v, w) = \frac{1}{|G|}\sum_{g\in G} gv\cdot gw$ where $\cdot$ is the usual dot product. Without loss of generality, we may assume the standard basis is orthonormal with respect to this form. Given a hyperplane $H \subset \mathbf{R}^n$, let s_H be the reflection that fixes H and negates a normal vector to H. Define $\mathcal{H}_G = \{H \mid s_H \in G\}$.

$\mathbf{R}^n \setminus \bigcup_{H \in \mathcal{H}_G} H$ has finitely many connected components which we call chambers. Let C be one of them and let $\overline{C}$ be its closure in $\mathbf{R}^n$. Now let S be the set of s_H where H are the hyperplanes bounding $\overline{C}$, and let G' be the subgroup generated by S.

We claim that the G'-orbit of every vector in $\mathbf{R}^n$ has nonempty intersection with $\overline{C}$. To see this, fix $a \in C$, and $v \in \mathbf{R}^n$. Since G' is finite, the orbit of v has an element v' of minimal distance to v. If $v' \notin C$, there is a hyperplane $s_H \in S$ separating a and v'. Then $s_H(v')$ is an element in the G'-orbit that is closer to a, which is a contradiction.

If we apply an element of G' to a chamber, we get another chamber since G fixes $\mathcal{H}_G$ and acts continuously. Hence, from the claim, for any reflection $s_H \in G$, there exists $g \in G'$ such that gH bounds $\overline{C}$, so that $gs_Hg^{-1} = s_{gH} \in G'$. In particular, G' contains all the reflections of G, and hence $G' = G$.

For each $s \in S$, let α_s be a unit vector normal to its fixed hyperplane that points towards C (i.e., $(\alpha_s, a) > 0$ for any $a \in C$). We claim that the α_s are linearly independent. First, we have $(\alpha_s, \alpha_t) \le 0$ for each $s \ne t$ (the angle between α_s and α_t is at most $\pi/2$). Suppose that there is a dependency $\sum_{s \in S} c_s \alpha_s = 0$. Let v be the sum of the terms with $c_s > 0$. If $v = 0$, then there are no c_s that are positive: otherwise, pairing with some $a \in C$, we have $(a, v) > 0$. That implies there are no c_s that are negative either, and so the dependency is trivial. Finally, if $v \ne 0$, then

$$0 < (v, v) = \sum_{s \in S,\ c_s > 0} \sum_{t \in S,\ c_t < 0} -c_s c_t (\alpha_s, \alpha_t)$$

and all of the terms are ≤ 0, which is a contradiction, and the claim is proven.

For $s, t \in S$, let $m(s,t)$ be the order of st. By restricting to the span of $\{\alpha_s, \alpha_t\}$, and using the calculation in the proof of Lemma 1.5, we see that $(\alpha_s, \alpha_t) = -\cos(\pi/m(s,t))$. We have a corresponding Coxeter group W and a surjective homomorphism $W \to G$. Finally, $\mathbf{R}^n$ can be identified with the geometric representation of W, and so $W \to G$ is injective by Corollary 1.12. ☐

2.3. **Classification.** As we saw in Proposition 2.11, if W is finite, then B_W is positive definite, so our first goal is to classify the connected graphs whose associated bilinear form is positive definite, i.e., $B_W(v, v) > 0$ whenever $v \ne 0$. It will also be convenient to classify the positive semidefinite bilinear forms (i.e., $B_W(v, v) \ge 0$ for all v). By choosing an ordering on S, we can encode B_W as a symmetric matrix A_W whose (s,t) entry is $B_W(\alpha_s, \alpha_t) = -\cos(\frac{\pi}{m(s,t)})$. Given a subset $S' \subseteq S$, the determinant of a submatrix of A_W whose rows and columns are indexed by S' is called a **principal minor**.

Recall that a bilinear form B is positive definite (respectively, positive semidefinite) if and only if all of the principal minors of the corresponding symmetric matrix A are positive (respectively, nonnegative).

Call a symmetric $n \times n$ matrix A decomposable if there is a nonempty proper subset $S \subsetneq [n]$ such that $A_{ij} = 0$ whenever $i \in S$ and $j \notin S$. A matrix is **indecomposable** if it is not decomposable. The matrix of B_W is indecomposable if and only if Γ is connected.

Theorem 2.13 (Perron–Frobenius). *Let A be a real $n \times n$ symmetric matrix which is positive semidefinite and indecomposable, and such that $A_{ij} \le 0$ for $i \ne j$. Then:*

(a) *Let $N = \{x \in \mathbf{R}^n \mid x^T A x = 0\}$. Then $N = \ker A$ and has dimension at most 1.*
(b) *The smallest eigenvalue (all eigenvalues are real since A is symmetric) has multiplicity 1, and has an eigenvector whose coordinates are all strictly positive.*

Proof. Since A is symmetric, we can orthogonally diagonalize it: $A = P^T DP$ where D is diagonal with real entries $d_1 \geq \cdots \geq d_n$. Since A is positive semidefinite, all of the eigenvalues are nonnegative: $(P^{-1}e_n)^T A(P^{-1}e_n) = e_n^T De_n = d_n \geq 0$.

(a) If $x \in N$, then for $y = Px$, we have $\sum_{i=1}^n d_i y_i^2 = 0$. Since $d_i \geq 0$ for all i, this implies that for each i, either $d_i = 0$ or $y_i = 0$, or simply $d_i y_i = 0$. This means that $DPx = 0$, which implies $x \in \ker A$. Clearly $\ker A \subseteq N$, so we get equality.

Suppose $N \neq 0$, and pick nonzero $x \in N$. Let $z = \sum_{i=1}^n |x_i|e_i$. Since $A_{i,j} \leq 0$ for $i \neq j$, we have

$$0 \leq z^T Az = \sum_{i=1}^n A_{i,i}x_i^2 + \sum_{i\neq j} A_{i,j}|x_i x_j| \leq x^T Ax = 0$$

which forces the first $\leq$ to be an equality, i.e., $z \in N$. Let $J = \{j \mid z_j \neq 0\}$. We claim that $J = [n]$. If not, pick $i \notin J$. Since $N = \ker A$, we have $Az = 0$, which means that $\sum_{j\in J} A_{ij}z_j = 0$. Since $A_{ij} \leq 0$ and $z_j > 0$, we must have $A_{ij} = 0$ for all $j \in J$. In particular, A is decomposable, which contradicts our assumption, so $J = [n]$. We conclude that if $x \in N$ is nonzero, then every coordinate of x is nonzero. Now suppose we have two nonzero vectors $x, y \in N$. Then $y_1 \neq 0$ and $x - \frac{x_1}{y_1} y_1$ has a coordinate equal to 0, and hence must be 0. We conclude that $\dim N \leq 1$.

(b) We have just shown this to be true if $d_n = 0$. In general, we can consider the matrix $A - d_n I$ which is again positive semidefinite and indecomposable and has nonnegative off diagonal entries. Then its kernel is the eigenspace for the smallest eigenvalue of A. □

Since every Coxeter graph is in bijection with some Coxeter group (W, S), we can unambiguously define it to be positive (semi)definite if its associated bilinear form is. For the remainder of the section, given a graph Γ, a subgraph is the result of possibly deleting vertices and edges. A proper subgraph is a subgraph different from Γ.

Corollary 2.14. *If Γ is a positive semidefinite connected graph, then all of its proper subgraphs are positive definite.*

Proof. Let A be the symmetric matrix of Γ. Let Γ' be a subgraph with symmetric matrix A'. Order the vertices $1, \dots, n$ so that Γ' uses the vertices $1, \dots, k$. Assume that A' is not positive definite. Then we can find nonzero $x \in \mathbf{R}^k$ such that $x^T A'x \leq 0$. Let $y = (|x_1|, \dots, |x_k|, 0, \dots, 0)^T \in \mathbf{R}^n$. Since $-\cos$ is increasing on the interval $[0, \pi/2]$, we have $A'_{i,j} \geq A_{i,j}$ for all $i, j \leq k$. Hence

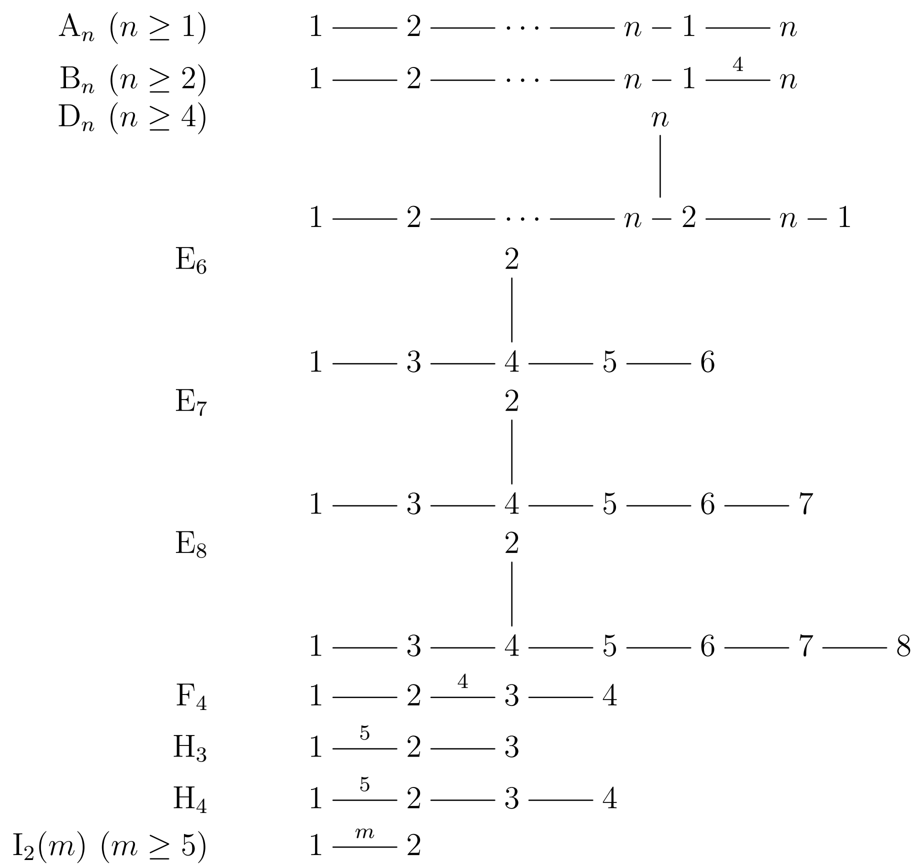
$$0 \leq y^T Ay = \sum_{i,j} A_{i,j}|x_i x_j| \leq \sum_{i,j} A'_{i,j}|x_i x_j| \leq \sum_{i,j} A'_{i,j}x_i x_j = x^T A'x \leq 0.$$

So the first inequality is an equality, so all of the coordinates of y are positive by Theorem 2.13, which means that $k = n$. The second inequality is also an equality, which means that $A_{i,j} = A'_{i,j}$ for all i, j, i.e., $\Gamma' = \Gamma$. □

The next result gives a list of positive definite connected Coxeter graphs (which we will prove is the complete list). It will be helpful to know the following values of cosine:

$$\cos\left(\frac{\pi}{3}\right) = \frac{1}{2}, \qquad \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}, \qquad \cos\left(\frac{\pi}{5}\right) = \frac{1+\sqrt{5}}{4}, \qquad \cos\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{2}.$$

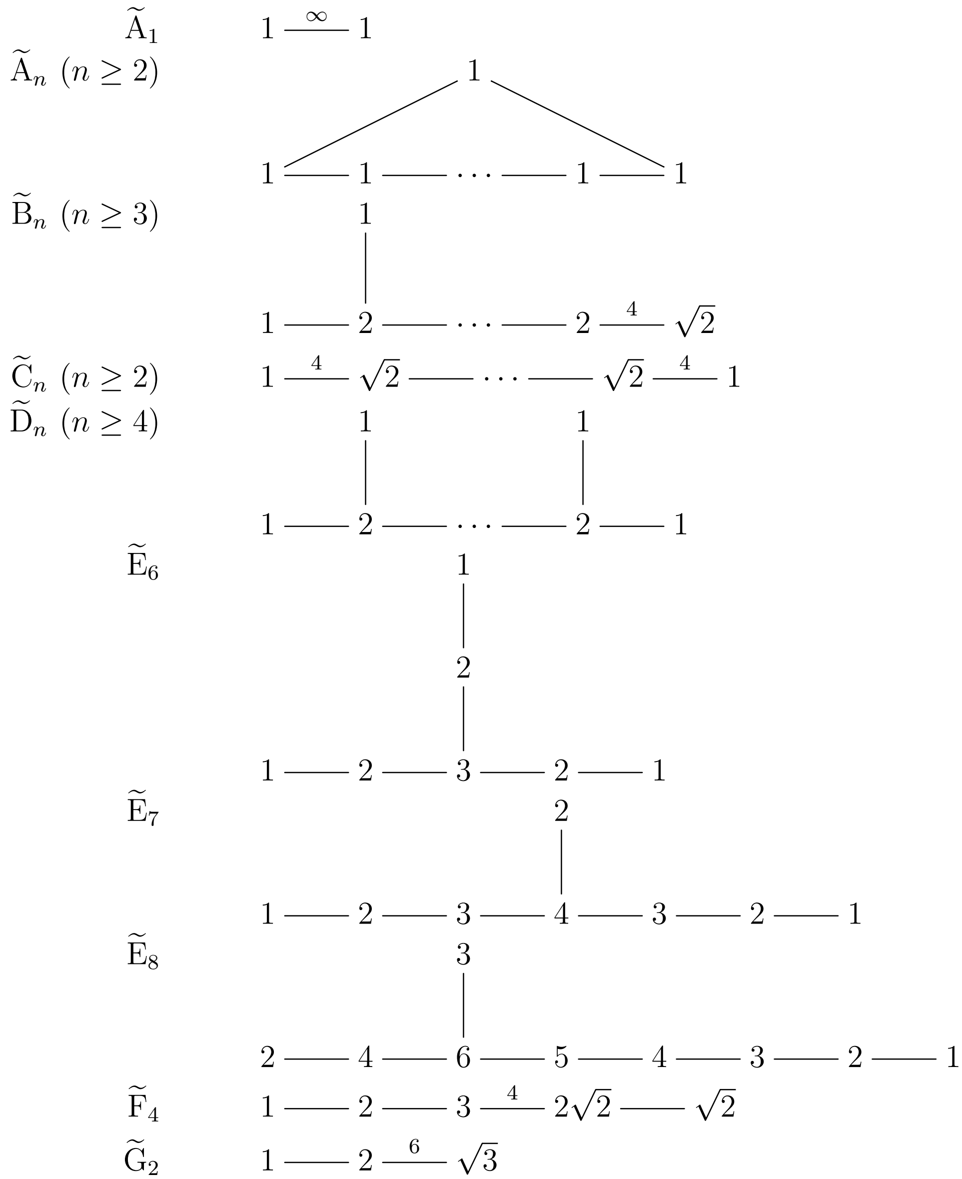
Proposition 2.15. *The following Coxeter graphs are all positive definite (the vertex numbering will be used later when we construct the root systems):*



Proof. Note that a principal minor of A_W is the determinant of the matrix corresponding to a graph Γ' obtained from Γ by deleting some nodes. By inspection, we see that deleting nodes from any graph in this list gives a disjoint union of graphs that are still in the list. In particular, we just need to check that $\det A_W > 0$ for each graph; we omit this check. ☐

Remark 2.16. The restrictions on m, n are imposed so that there are no coincidences. For example, $\mathrm{A}_2 = \mathrm{I}_2(3)$ and $\mathrm{B}_2 = \mathrm{I}_2(4)$. For reasons coming from Lie theory, the case $\mathrm{I}_2(6)$ is also known as G_2. ☐

Proposition 2.17. *The following Coxeter graphs are positive semidefinite, but not positive definite. In each case, the subscript on the name is the rank of the matrix, which is one less* *than the number of nodes. The labels are the coordinates of a kernel vector.*



Proof. In each example above, every graph obtained by deleting nodes appears on the previous list, and hence is positive definite. The coordinates of a kernel element are listed on the nodes, so we see that each graph is positive semidefinite, but not positive definite. □

Theorem 2.18. *Every positive semidefinite connected graph is either listed in Proposition 2.15 or Proposition 2.17.*

See [H1, §2.7].

2.4. **Construction of finite root systems.** An additive subgroup L of $\mathbf{R}^n$ is called a **lattice** if $L \cong \mathbf{Z}^n$ and $(v, w) \in \mathbf{Z}$ for all $v, w \in L$. When W is finite, we want to know when W preserves a lattice in the geometric representation V. If this is the case, we say that W is crystallographic, or that W is a **Weyl group**.

Proposition 2.19. *If W preserves a lattice, then $m(s,t) \in \{2, 3, 4, 6\}$ for all $s \neq t$.*

Proof. Suppose that W preserves a lattice L. Then we can pick a basis for V consisting of elements of L. In this basis, the matrix of every element $w \in W$ has integer entries, and in particular, we see that the trace (which is basis independent) of every $w \in W$ is an integer. If $|S| \le 1$, the result is clear so suppose otherwise. Now pick $s, t \in S$ with $s \neq t$. We have seen in the proof of Lemma 1.5 that, in some basis, st is the direct sum of an identity matrix of size $\dim V - 2$ and the 2×2 rotation matrix by the angle $2\pi/m(s,t)$, so the trace is $\dim V - 2 + 2\cos(2\pi/m(s,t))$. Hence $\cos(2\pi/m(s,t)) \in \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$. Since $m(s,t) < \infty$, this forces $m(s,t) \in \{2, 3, 4, 6\}$. □

In terms of our classification, this only allows the possibilities A_n, B_n, D_n, E_6, E_7, E_8, F_4, and $\mathrm{I}_2(6) = \mathrm{G}_2$. For each case, we give the geometric representation explicitly. In each case, the $\mathbf{Z}$-span of the α_i is a lattice preserved by W.

In each case, we should normalize the α_i to have unit length if we want the inner products to match the values of B_W, but we present it as follows to match other conventions in Lie theory. In all cases, $e_1, \dots, e_n$ are the standard basis vectors for $\mathbf{R}^n$. Furthermore, we can partially order roots by $\alpha \ge \beta$ if $\alpha - \beta$ is a positive root. There is always a unique maximal element, which we denote by $\widetilde{\alpha}$.

- For the A_n root system, take $V = \{x \in \mathbf{R}^{n+1} \mid x_1 + \cdots + x_{n+1} = 0\}$ and take the set of vectors $e_i - e_j$ for $i \neq j$. For simple roots, take $\alpha_i = e_i - e_{i+1}$ for $i = 1, \dots, n$. The positive roots are $e_i - e_j$ for $i < j$. Then $W \cong \mathfrak{S}_{n+1}$ is the symmetric group and $|W| = (n+1)!$. We have $\widetilde{\alpha} = e_1 - e_{n+1}$.
- For the B_n root system, take $V = \mathbf{R}^n$ and take the set of vectors $\pm e_i \pm e_j$ for $i \neq j$ and $\pm e_i$ for $i = 1, \dots, n$. For simple roots, take $\alpha_i = e_i - e_{i+1}$ for $i = 1, \dots, n-1$ and $\alpha_n = e_n$. The positive roots are $e_i \pm e_j$ for $i < j$ and e_i for $i = 1, \dots, n$. Then $W \cong \mathfrak{S}_n \ltimes (\mathbf{Z}/2)^n$ is the hyperoctahedral group (signed permutation matrices) and $|W| = 2^n n!$. We have $\widetilde{\alpha} = e_1 + e_2$.
- There is a "dual" root system C_n. We take $\mathbf{R}^n$ and the set of vectors are $\pm e_i \pm e_j$ for $i \neq j$ and $\pm 2e_i$. For simple roots, take $\alpha_i = e_i - e_{i+1}$ for $i = 1, \dots, n-1$ and $\alpha_n = 2e_n$. The positive roots are $e_i \pm e_j$ for $i < j$ and $2e_i$ for $i = 1, \dots, n$. W is the same as in type B_n, and $\widetilde{\alpha} = 2e_1$.
- For the D_n root system, take $V = \mathbf{R}^n$ and take the set of vectors $\pm e_i \pm e_j$ for $i \neq j$. For simple roots, take $\alpha_i = e_i - e_{i+1}$ for $i = 1, \dots, n-1$ and $\alpha_n = e_{n-1} + e_n$. The positive roots are $e_i \pm e_j$ for $i < j$. Let $(\mathbf{Z}/2)^{n-1}$ denote the subgroup of $(\mathbf{Z}/2)^n$ of vectors whose coordinate sum is 0. Then $W \cong \mathfrak{S}_n \ltimes (\mathbf{Z}/2)^{n-1}$ is the demihyperoctahedral group (signed permutation matrices with an even number of negative signs) and $|W| = 2^{n-1} n!$. We have $\widetilde{\alpha} = e_1 + e_2$.
- We now give a description of the E_8 root system. Take $V = \mathbf{R}^8$ and take the set of vectors x such that
  - either all coordinates of x are integers, or all coordinates of x are half-integers, i.e., all coordinates of $2x$ are odd integers,
  - $x_1 + \cdots + x_8$ is an even integer, and
  - $x_1^2 + \cdots + x_8^2 = 2$.

There are 240 such vectors, and we set

$$\alpha_1 = \frac{1}{2}(e_1 - e_2 - e_3 - e_4 - e_5 - e_6 - e_7 + e_8), \qquad \alpha_2 = e_1 + e_2$$
$$\alpha_3 = e_2 - e_1, \qquad \alpha_4 = e_3 - e_2, \qquad \alpha_5 = e_4 - e_3,$$
$$\alpha_6 = e_5 - e_4, \qquad \alpha_7 = e_6 - e_5, \qquad \alpha_8 = e_7 - e_6.$$

We have $|W| = 2^{14}3^5 5^2 7$ and $\widetilde{\alpha} = e_7 + e_8$.

- To construct E_n for $n = 6, 7$, we replace V by the span of $\alpha_1, \dots, \alpha_n$ and only take those roots lying in that span. We have $|W(\mathrm{E}_6)| = 51840$ and $|W(\mathrm{E}_7)| = 2903040$. The number of roots are 72 and 126, respectively, and the highest roots are $\frac{1}{2}(e_1 + e_2 + e_3 + e_4 + e_5 - e_6 - e_7 + e_8)$ and $e_8 - e_7$, respectively.

Remark 2.20. The Weyl groups of types E_n for $n = 6, 7, 8$ have very close connections to matrix groups over finite fields. For example, one has an isomorphism $W(\mathrm{E}_7) \cong \mathbf{Z}/2 \times \mathbf{Sp}_6(\mathbf{F}_2)$ where $\mathbf{Sp}_6(\mathbf{F}_2)$ is the symplectic group that preserves a non-degenerate symplectic form on the finite vector space $\mathbf{F}_2^6$. See [H1, §2.12] for some other descriptions. □

- For the F_4 root system, take $V = \mathbf{R}^4$ and take the set of vectors x such that
  - either all coordinates of x are integers, or all coordinates of x are half-integers, i.e., all coordinates of $2x$ are odd integers, and
  - $x_1^2 + x_2^2 + x_3^2 + x_4^2$ is 1 or 2.

  There are 48 such vectors and we set

$$\alpha_1 = e_2 - e_3, \qquad \alpha_2 = e_3 - e_4, \qquad \alpha_3 = e_4, \qquad \alpha_4 = \frac{1}{2}(e_1 - e_2 - e_3 - e_4).$$

  Then $|W| = 1152$ and $\widetilde{\alpha} = e_1 + e_2$.
- For the G_2 root system, take $V = \{x \in \mathbf{R}^3 \mid x_1 + x_2 + x_3 = 0\}$ and take the set of vectors $x \in \mathbf{Z}^3 \cap V$ such that $x_1^2 + x_2^2 + x_3^2$ is 2 or 6. There are 12 such vectors and we set

$$\alpha_1 = e_1 - e_2, \qquad \alpha_2 = -2e_1 + e_2 + e_3.$$

  W is the dihedral group of order 12 and $\widetilde{\alpha} = -e_1 - e_2 + 2e_3$.

As for the rest of the finite irreducible Coxeter groups:

- $W(\mathrm{I}_2(m))$ is the dihedral group of order $2m$, i.e., the symmetries of a regular m-gon.
- $W(\mathrm{H}_3)$ is the symmetry group of the regular icosahedron: a 3-dimensional polytope built out of 20 triangles which meet 5 at a time at each of the 12 vertices. Abstractly, the group has size 120 and is isomorphic to $\mathbf{Z}/2 \times \mathfrak{A}_5$, where $\mathfrak{A}_5$ is the alternating group.
- $W(\mathrm{H}_4)$ is the symmetry group of the 600-cell, which is a 4-dimensional polytope that is analogous to the icosahedron: its boundary consists of 600 tetrahedra which meet 5 at a time in a common edge. The group has size 14400.

In [H1, §2.13], $W(\mathrm{H}_4)$ is realized as a quaternionic reflection group. In the next section, we'll consider finite complex reflection groups, which are generally not Coxeter groups, but share many algebraic properties with real reflection groups.

2.5. **Hyperbolic Coxeter groups.** Our next goals are to discuss the Coxeter groups whose graph is positive semidefinite. Before that, we describe the class of hyperbolic Coxeter groups which are close to positive semidefinite and can be classified by considering subgraphs.

We now assume that the bilinear form B_W is nondegenerate and set $n = |S|$. In that case, we define the dual basis ω_s to α_s via the conditions $B_W(\omega_s, \alpha_t) = \delta_{s,t}$. We define

$$C = \{v \in V \mid B_W(v, \alpha_s) > 0 \text{ for all } s \in S\}$$
$$= \{\sum_s c_s\omega_s \mid c_s > 0\}.$$

Recall that every real bilinear form can be diagonalized into the form $x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$ for a unique choice of (p, q), which is called its **signature**.

Then (W, S) is defined to be **hyperbolic** if B_W has signature $(n-1, 1)$ and $B_W(v, v) < 0$ for all $v \in C$.

Lemma 2.21. *Let E be an n-dimensional real vector space and B a symmetric bilinear form of signature $(n-1, 1)$. For nonzero $v \in E$, define $v^\perp = \{w \in E \mid B(v, w) = 0\}$. Then B restricted to $v^\perp$ is positive semidefinite if and only if $B(v, v) \le 0$.*

Proof. First suppose that $B(v, v) \neq 0$. In particular, $v \notin v^\perp$, so we have an orthogonal direct sum $E = \langle v\rangle \oplus v^\perp$. Since B is the direct sum of its restrictions to each of $\langle v\rangle$ and $v^\perp$, its restriction to $v^\perp$ has signature $(n-2, 1)$ if $B(v, v) > 0$ (and hence is not positive semidefinite) and has signature $(n-1, 0)$ if $B(v, v) < 0$ (and hence is positive definite).

Finally suppose that $B(v, v) = 0$. Since B has signature $(n-1, 1)$, there is a hyperplane H so that the restriction of B to H is positive definite. Then $H \neq v^\perp$ since $v \in v^\perp$, so B is positive definite on $H \cap v^\perp$, which means its signature on $v^\perp$ is $(n-2, 0)$, and hence B is positive semidefinite on $v^\perp$. ☐

Proposition 2.22. *Let Γ be the Coxeter graph for (W, S). Then (W, S) is hyperbolic if and only if*

(a) *B_W is nondegenerate but not positive definite, and*
(b) *For each $s \in S$, the subgraph $\Gamma \setminus s$ is positive semidefinite.*

For the proof below, for $s \in S$, we define L_s to be the span of $\{\alpha_t \mid t \neq s\}$.

Proof. First suppose that (W, S) is hyperbolic. Then (a) holds by definition. Now pick $s \in S$. Since $x \mapsto B_W(x, x)$ is continuous and ω_s is in the closure of C, we have $B_W(\omega_s, \omega_s) \le 0$. By definition, $\omega_s^\perp = L_s$. By Lemma 2.21, the restriction of B_W to L_s is positive semidefinite. By construction of B_W, this restriction is just the bilinear form of $\Gamma \setminus s$, and hence (b) holds.

Now suppose that (a) and (b) hold. Define $N = \{v \in V \mid B_W(v, v) < 0\}$. By (a), $N \neq \varnothing$, and by (b), $N \cap L_s = \varnothing$ for all $s \in S$. In particular, $N \subset V \setminus \bigcup_{s\in S} L_s$. The latter is disconnected and has one connected component U_T for each subset $T \subset S$, namely U_T is the set of vectors $\sum_s c_s\alpha_s$ where $c_s > 0$ if $s \in T$ and $c_s < 0$ if $s \notin T$. Suppose that B_W has signature (p, q) with $q \ge 2$. Then there is a 2-dimensional subspace Z on which B_W is negative definite, which means that $Z \setminus 0 \subset N$. However, $Z \setminus 0 \cong \mathbf{R}^2 \setminus 0$ is connected, which means it is a subset of some U_T. But $Z \setminus 0$ is closed under negation while U_T is not, which is a contradiction. Hence the signature of B_W must be $(n-1, 1)$.

Hence in some choice of basis, $B_W(x, x) = x_1^2 + \cdots + x_{n-1}^2 - x_n^2$, so $N = \{x \mid x_1^2 + \cdots + x_{n-1}^2 < x_n^2\}$. In particular, N has two connected components N_+ and N_- corresponding to the sign of x_n, and by the triangle equality, each component is closed under addition, i.e., if $x, y \in N_\pm$, then $x + y \in N_\pm$. Finally, we have to show that $C \subset N$. By (b), for each s, B_W is positive semidefinite on $\omega_s^\perp = L_s$, so that $B_W(\omega_s, \omega_s) \le 0$ by Lemma 2.21, and hence each ω_s belongs to the closure of N. We claim that they all belong to the closure of a single connected component.

Suppose that $B_W(v, v) < 0$ and write $v = \sum_s c_s \alpha_s$. Let v_+ be the sum of the terms for which $c_s > 0$ and let $v_- = v - v_+$. We have

$$B_W(v, v) = \sum_s c_s^2 + \sum_{s \neq t} c_s c_t B_W(\alpha_s, \alpha_t),$$

and hence $B_W(v_+, v_+) + B_W(v_-, v_-) \le B_W(v, v)$ since the difference comes from terms in the second sum (which are non-negative since c_s and c_t have different signs and $B_W(\alpha_s, \alpha_t) \le 0$), and hence either $B_W(v_+, v_+) < 0$ or $B_W(v_-, v_-) < 0$.

If $v \in N$, then either $v_+ \in N$ or $v_- \in N$, and in either case, its negative is also in N. We conclude that the two connected components of N correspond to when all the coefficients (in the α basis) are all positive or all negative.

So if we write $\omega_s = \sum_t w_{s,t} \alpha_t$, either all coefficients are non-negative or non-positive. We claim it is the latter. The matrix $(w_{s,t})$ is the inverse (and transpose, depending on indexing conventions) of the matrix $\beta = (B_W(\alpha_s, \alpha_t))$ and $w_{s,s}$ is the determinant of the matrix for $\Gamma \setminus s$ divided by $\det \beta < 0$. By (b), the first quantity is non-negative. If it is positive, we are done. Otherwise, if it is 0, then we have

$$1 = B_W(\omega_s, \alpha_s) = \sum_{t \neq s} w_{s,t} B_W(\alpha_s, \alpha_t).$$

Since $B_W(\alpha_s, \alpha_t) \le 0$, we conclude that $w_{s,t} \le 0$.

Hence the closure $\overline{C}$ belongs to $\overline{N}$. By considering interiors, we have $C \subset N$. □

I don't have much to say about the connections with hyperbolic geometry, but see the references in [H1, §6.8] for further information.

A list of all of the hyperbolic Coxeter groups of ranks ≥ 4 are given in [H1, §6.9]. In rank 2, every graph is positive semidefinite, so there are no hyperbolic examples.

Corollary 2.23. *Every connected Coxeter graph Γ on 3 vertices is either positive semidefinite or hyperbolic.*

Proof. By the above comments, Γ automatically satisfies condition (b) above. So we need to see the possibilities for the signature of B_W. There are two possibilities. First, if Γ is a path, in which case the matrix for B_W is of the form

$$\begin{pmatrix} 1 & -a & 0 \\ -a & 1 & -b \\ 0 & -b & 1 \end{pmatrix}$$

where $a, b \ge 0$. Its eigenvalues are $1, 1 \pm \sqrt{a^2 + b^2}$ and in particular is either positive semidefinite or has signature $(2, 1)$. Otherwise, Γ is a cycle, and the matrix for B_W has the form

$$\begin{pmatrix} 1 & -a & -c \\ -a & 1 & -b \\ -c & -b & 1 \end{pmatrix}$$

where $a, b, c \geq 1/2$. The determinant is $1 - (a^2 + b^2 + c^2 + 2abc) \leq 0$, and equality is achieved when $a = b = c = 1/2$, in which case this is $\widetilde{\mathrm{A}}_2$ and positive semidefinite. Otherwise, the determinant is negative (and hence B_w is nondegenerate) and the trace is positive, which means the only possibility for the signature is $(2, 1)$. $\square$

3. Polynomial invariants

3.1. **Some commutative algebra.** We will need some basic results in commutative algebra. Let $\mathbf{k}$ be a field. First, given a vector space V of dimension n, pick a basis $x_1, \dots, x_n$. The ring of polynomials $A = \mathbf{k}[x_1, \dots, x_n]$ identifies with the symmetric algebra $\mathrm{Sym}(V) = \bigoplus_{d \geq 0} \mathrm{Sym}^d V$. In particular, if a group G acts on V, then it acts on $\mathbf{k}[x_1, \dots, x_n]$. Since $(V^*)^* = V$, we can interpret A as the ring of polynomial *functions* on V^*.

We recall that a nonempty subset $I \subset A$ is an ideal if it is an A-submodule of A, i.e., $a + b \in I$ if $a, b \in I$ and $ab \in I$ if $a \in A$ and $b \in I$. A set S generates I if every element of I is an A-linear combination of elements of S.

A finitely generated $\mathbf{k}$-algebra is a quotient ring of A by an ideal. We have the following basic result (which states that R is noetherian):

Theorem 3.1 (Hilbert basis theorem)**.** *Let R be a finitely generated algebra over a field. Every ideal of R has a finite generating set. More generally, every submodule of a finitely generated R-module is also finitely generated.*

We will primarily interested in graded situations. In particular, let $d_1, \dots, d_n$ be positive integers and make the convention that $\deg(x_i) = d_i$. In many cases, $d_i = 1$ for all i, but we will use the more general case as well. This introduces a grading $A = \bigoplus_{d \geq 0} A_d$ where A_d is the linear span of all monomials $x_1^{p_1} \cdots x_n^{p_n}$ such that $\sum_{i=1}^n p_i d_i = d$. (Note $A_0 = \mathbf{k}$ is the span of the constant polynomial since all $d_i > 0$.)

An A-module M is graded if we have a decomposition $M = \bigoplus_{d \geq 0} M_d$ such that if $f \in A_d$ and $m \in M_e$, then $fm \in M_{d+e}$. An element in M_d is called homogeneous. In particular, a graded ideal I is called **homogeneous** and means that $I = \bigoplus_{d \geq 0}(I \cap A_d)$. For graded modules, we can always find a generating set that consists of homogeneous elements.

For a graded module M, its **Hilbert series** is the formal power series

$$\mathrm{H}_M(t) = \sum_{d \geq 0} (\dim M_d) t^d.$$

The graded tensor product (over $\mathbf{k}$) is given by $(M \otimes N)_d = \bigoplus_{e=0}^d M_e \otimes N_{d-e}$. The Hilbert series is multiplicative in the sense that

$$\mathrm{H}_{M \otimes N}(t) = \mathrm{H}_M(t)\mathrm{H}_N(t).$$

An important case is $M = \mathbf{k}[x_1, \dots, x_n] = \mathbf{k}[x_1] \otimes \cdots \otimes \mathbf{k}[x_n]$, in which case we get

$$\mathrm{H}_{\mathbf{k}[x_1, \dots, x_n]}(t) = \prod_{i=1}^n \frac{1}{1 - t^{\deg(x_i)}}.$$

Finally, we recall that polynomials $f_1, \dots, f_k$ are **algebraically independent** if for any nonzero polynomial $h(y_1, \dots, y_k)$ (in new variables), we have $h(f_1, \dots, f_k) \neq 0$. An algebraically independent set can be extended to a transcendence basis of $\mathbf{k}(x_1, \dots, x_n)$ over $\mathbf{k}$, so in particular, any algebraically independent set has size at most n.

3.2. **Molien's formula.** From now on, we assume $\mathbf{k} = \mathbf{C}$ is the field of complex numbers.

Lemma 3.2. *Let V be a finite-dimensional vector space. Define $\varphi\colon V \to V$ by*

$$\varphi(v) = \frac{1}{|G|}\sum_{g\in G} g\cdot v.$$

φ is a projection and its image is V^G. In particular,

$$\dim V^G = \mathrm{Tr}(\varphi) = \frac{1}{|G|}\sum_{g\in G}\mathrm{Tr}(g|V).$$

Proof. First we prove the image is contained in V^G: for any $h \in G$, we have

$$h\cdot\varphi(v) = \frac{1}{|G|}\sum_{g\in G} hg\cdot v = \frac{1}{|G|}\sum_{g\in G} g\cdot v = \varphi(v).$$

where in the second equality, we reindexed the sum since $\{hg \mid g \in G\} = \{g \mid g \in G\}$.

On the other hand, given $w \in V^G$, we have $\varphi(w) = w$, and so the image is all of V^G.

These two facts imply that φ is a projection: $\varphi^2(v) = \varphi(\varphi(v))$ and $\varphi(v) \in V^G$ which implies that $\varphi^2(v) = \varphi(v)$.

For the last statement, note that the eigenvalues of a projection are either 0 or 1 (since it's a root of the polynomial $t^2 - t$) and the multiplicity of 1 is its rank. □

We now apply this formula to the action of G on $A = \mathbf{C}[x_1,\dots,x_n] = \mathrm{Sym}(V)$. While A is infinite-dimensional, it is a direct sum of finite-dimensional G-representations $\mathrm{Sym}^d(V)$, so we can apply the formula to each piece individually.

Theorem 3.3 (Molien's formula)**.** *For $g \in V$, let $\rho_V(g)$ be the linear operator on V corresponding to multiplication by g. We have*

$$\sum_{d\geq 0}\dim(\mathrm{Sym}^d V)^G t^d = \frac{1}{|G|}\sum_{g\in G}\frac{1}{\det(1-\rho_V(g)t)}.$$

If we assign all elements of V to have degree 1, this quantity is $\mathrm{H}_{A^G}(t)$.

Proof. Pick $g \in G$. Let $z_1,\dots,z_n$ be the eigenvalues (with multiplicity) of $\rho_V(g)$. If we pick an eigenbasis $v_1,\dots,v_n$ for $\rho_V(g)$, then $v_{i_1}\cdots v_{i_d}$ with $1 \le i_1 \le \cdots \le i_d \le n$ is an eigenbasis for g acting on $\mathrm{Sym}^d(V)$ with eigenvalues $z_{i_1}\cdots z_{i_d}$. We see that

$$\sum_{d\geq 0}\mathrm{Tr}(g|\,\mathrm{Sym}^d V)t^d = \prod_{i=1}^n \frac{1}{1-z_i t} = \frac{1}{\det(1-\rho_V(g)t)}.$$

The claimed formula now follows from Lemma 3.2. □

3.3. **Ring of invariants.** Recall the projection operator given by the formula $\varphi = \frac{1}{|G|}\sum_{g\in G}\rho(g)$. We are now interested in the case when the representation is the symmetric algebra $A = \mathrm{Sym}(V) \cong \mathbf{C}[x_1,\dots,x_n]$. In that case, we write $f^\#$ instead of $\varphi(f)$. Note that if $f_1 \in A^G$ and $f_2 \in A$, then $(f_1f_2)^\# = f_1 f_2^\#$. In particular, $\#\colon A \to A^G$ is a surjective A^G-module homomorphism which preserves degrees.

Proposition 3.4. *Let I be the ideal of A generated by A^G. Suppose that $f_1, \dots, f_k \in A^G$ are positive degree homogeneous elements that generate I. Then $f_1, \dots, f_k$ generate A^G as a* $\mathbf{C}$*-algebra.*

In particular, A^G is a finitely generated $\mathbf{C}$*-algebra.*

Proof. We prove that $f \in A^G$ is a polynomial in the $f_1, \dots, f_k$ by induction on $\deg f$. We may assume that f is homogeneous. Then $f \in I$ and hence there exist homogeneous $h_1, \dots, h_k \in A$ such that $f = h_1 f_1 + \cdots + h_k f_k$. Now apply #: $f = h_1^\# f_1 + \cdots + h_k^\# f_k$. In particular

$$\deg h_i^\# = \deg h_i = \deg f - \deg f_i < \deg f.$$

By induction on degree, $h_i^\#$ is a polynomial in the $f_1, \dots, f_k$. Substituting these expressions gives the desired result.

The last statement follows from Hilbert's basis theorem since I is finitely generated. □

Remark 3.5. The above proof required the use of #, which relies on being able to divide by $|G|$, and hence does not extend when $\mathbf{k}$ is a field of positive characteristic p where p divides $|G|$. However, the conclusion that A^G is finitely generated, due to Noether, still holds as we now show. So now let $\mathbf{k}$ be any field.

Let t be a new indeterminate. For each $i = 1, \dots, r$, consider the polynomial $p_i(t) \in A[t]$ defined by $p_i(t) = \prod_{g \in G}(t - gx_i)$. In fact, the coefficients are symmetric in the gx_i and hence are elements of A^G. Let B be the $\mathbf{k}$-subalgebra of A^G generated by the coefficients of the $p_i(t)$ for $i = 1, \dots, n$. By definition, B is finitely generated. Furthermore, A is a generated as a B-module by $\{x_i^j \mid 1 \le i \le n,\ 0 \le j < |G|\}$: $p_i(t)$ is a monic polynomial of degree $|G|$ and has x_i as a root, so $x_i^{|G|}$ can be rewritten as a linear combination of lower powers of x_i with coefficients in B. Next, A^G is a B-submodule of A, and hence by noetherianity, it is also a finitely generated B-module. The set of these B-module generators together with the generators for B as an algebra then give a set of algebra generators for A^G. □

For any integral domain R, we let $\mathrm{Frac}(R)$ denote its field of fractions.

Proposition 3.6. *We have $\mathrm{Frac}(A^G) = \mathrm{Frac}(A)^G$. In particular, if G acts faithfully on V, then $\mathrm{Frac}(A)$ is a degree $|G|$ extension over $\mathrm{Frac}(A^G)$. The transcendence degree of $\mathrm{Frac}(A^G)$ is n.*

Proof. It is clear that $\mathrm{Frac}(A^G) \subseteq \mathrm{Frac}(A)^G$. For the other inclusion, suppose that $p/q \in \mathrm{Frac}(A)^G$ with $p, q \in A$. Let $p' = \prod_{g \in G,\ g \ne 1} gp$. Then $p/q = \frac{pp'}{qp'}$ and pp' is G-invariant. Hence qp' is also G-invariant, so $p/q \in \mathrm{Frac}(A^G)$.

In particular, $\mathrm{Frac}(A)$ is a degree $|G|$ extension over $\mathrm{Frac}(A^G)$. The last statement follows from the fact that $\mathrm{Frac}(A)$ has transcendence degree n, and that it is constant for finite extensions. □

3.4. Chevalley's theorem.

We say that $g \in \mathbf{GL}_n(\mathbf{C})$ is a **(complex) reflection** if $\mathrm{rank}(g - I) = 1$ and g has finite order. Throughout, $W \subset \mathbf{GL}_n(\mathbf{C})$ is a finite group generated by reflections and $A = \mathbf{C}[x_1, \dots, x_n]$.

Lemma 3.7. *Let $f_1, \dots, f_k \in A^W$ be homogeneous polynomials with f_1 not in the ideal of A^W generated by $f_2, \dots, f_k$. Suppose that we have homogeneous polynomials $h_1, \dots, h_k \in A$ and a linear relation*

$$f_1 h_1 + \cdots + f_k h_k = 0. \tag{3.7a}$$

Then $h_1 \in I$.

Proof. We prove this by induction on $\deg h_1$. If $\deg h_1 = 0$, then h_1 is a constant. Apply # to get

$$h_1 f_1 = -(f_2 h_2^{\#} + \cdots + f_k h_k^{\#}),$$

If $h_1 \neq 0$, this contradicts our assumption that f_1 is not in the ideal of A^W generated by $f_2, \ldots, f_k$. Hence we conclude that $h_1 = 0 \in I$.

Now assume that $\deg h_1 > 0$. Let $s \in W$ be a reflection and let ℓ_s be a nonzero linear equation defining its fixed hyperplane. For each i, $sh_i - h_i$ is identically 0 on this hyperplane, and hence $sh_i - h_i = h_i' \ell$ for some homogeneous polynomial h_i' (extend $\{\ell\}$ to a basis for the linear polynomials and write $sh_i - h_i$ as a polynomial in these basis elements; the condition is that this polynomial is identically 0 when we substitute $\ell = 0$). If we apply s to (3.7a), we get $f_1(sh_1) + \cdots + f_k(sh_k) = 0$, and their difference gives

$$\ell(f_1 h_1' + \cdots + f_k h_k') = 0.$$

Since A is an integral domain, we conclude that $f_1 h_1' + \cdots + f_k h_k' = 0$. Next, $\deg h_1' = \deg h_1 - 1$, so by induction, we conclude that $h_1' \in I$ and hence $sh_1 - h_1 \in I$ or that $sh_1 \equiv h_1 \pmod I$. This last statement holds for each reflection, which means that $gh_1 \equiv h_1 \pmod I$ for all $g \in W$. Unpacking the definition of #, we see that $h_1^{\#} \equiv h_1 \pmod I$. Finally, $h_1^{\#} \in A^W \subset I$, which means that $h_1 \in I$. □

Lemma 3.8. *If f is a homogeneous polynomial in $x_1, \ldots, x_n$, then*

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = (\deg f) f.$$

Proof. If $f = x_1^{p_1} \cdots x_n^{p_n}$, then the left side is $\sum_{i=1}^n p_i f$, which is $(\deg f)f$. By linearity, the equation is true for any sum of monomials of the same total degree. □

Theorem 3.9 (Chevalley)**.** *A^W has a generating set consisting of n algebraically independent homogeneous elements.*

Proof. As in Proposition 3.4, let I be the ideal of A generated by A^W. Let $f_1, \ldots, f_k \in A^W$ be a minimal set of positive degree homogeneous elements that generate I. Then $f_1, \ldots, f_k$ generate A^W as a $\mathbf{C}$-algebra, and our goal is to show that they are algebraically independent. Let $y_1, \ldots, y_k$ be new indeterminates with $\deg(y_i) = \deg(f_i)$, and define a ring homomorphism

$$\psi \colon \mathbf{C}[y_1, \ldots, y_k] \to A^W$$
$$\psi(h(y_1, \ldots, y_k)) = h(f_1, \ldots, f_k).$$

Then $f_1, \ldots, f_k$ are algebraically independent if and only if $\ker \psi = 0$. Suppose otherwise. Since ψ preserves degrees, $\ker \psi$ is a homogeneous ideal; pick homogeneous nonzero $h \in \ker \psi$ of minimal degree (it must have positive degree). Apply $\frac{\partial}{\partial x_i}$ to $h(y_1, \ldots, y_k) = 0$ and the chain rule to get ($h_j = \frac{\partial h}{\partial y_j}(f_1, \ldots, f_k)$)

$$\sum_{j=1}^k h_j \frac{\partial f_j}{\partial x_i} = 0.$$

Renumber the $h_1, \dots, h_k$ if necessary so that $h_1, \dots, h_m$ is a minimal generating set for the ideal in A^W that they generate. Since h has positive degree, there exists j such that $\frac{\partial h}{\partial y_j} \neq 0$. By minimality of $\deg h$, this derivative is not in $\ker \psi$, so $m > 0$. In particular, for $m+1 \leq j \leq k$, we can write $h_j = \sum_{a=1}^m \alpha_{j,a} h_a$ for homogeneous $\alpha_{j,a} \in A^W$ with $\deg \alpha_{j,a} = \deg h_j - \deg h_a = \deg f_a - \deg f_j$ if it is nonzero. So we can rearrange the above equation to get an expression of the form

$$\sum_{j=1}^m h_j p_{i,j} = 0, \qquad p_{i,j} = \frac{\partial f_j}{\partial x_i} + \sum_{b=m+1}^k \alpha_{b,j} \frac{\partial f_b}{\partial x_i}.$$

By Lemma 3.7, we have $p_{i,1} \in I$ for each i. In other words, $p_{i,1}$ is a homogeneous linear combination of the f_i. Then the following element is also in I (we use Lemma 3.8):

$$\sum_{i=1}^n x_i p_{i,1} = \sum_{i=1}^n x_i \frac{\partial f_1}{\partial x_i} + \sum_{b=m+1}^k \alpha_{b,1} \sum_{i=1}^n x_i \frac{\partial f_b}{\partial x_i} = (\deg f_1) f_1 + \sum_{b=m+1}^k (\deg f_b) \alpha_{b,1} f_b,$$

and is a homogeneous linear combination of the f_i:

$$(\deg f_1) f_1 + \sum_{b=m+1}^k (\deg f_b) \alpha_{b,1} f_b = \sum_{j=1}^k \beta_j f_j.$$

All of the terms on the left side are homogeneous of degree f_1. On the right side, if $\beta_1 \neq 0$, then $\deg \beta_1 > 0$ (since the expression is also equal to $\sum_{i=1}^n x_i p_{i,1}$), so $\deg(\beta_1 f_1) > \deg f_1$, so it must cancel with some other terms in the sum. Once we remove those terms, we can rearrange the above equation to write f_1 as a linear combination of $f_2, \dots, f_k$, which contradicts that they are a minimal set of generators. This shows that $f_1, \dots, f_k$ are algebraically independent.

Since the transcendence degree of A^W is n (Proposition 3.6), we must have $n = k$. ☐

The generators need not be unique (for example, W could be trivial, in which case taking any basis for the linear polynomials works), but their degrees are since they are encoded into the Hilbert series: if $\deg f_i = d_i$, then $\mathrm{H}_{A^W}(t) = \prod_i (1 - t^{d_i})^{-1}$.

Theorem 3.10. *Let $G \subset \mathbf{GL}_n(\mathbf{C})$ be a finite group and assume that A^G is generated by an algebraically independent set of homogeneous elements $f_1, \dots, f_n$ and let $d_i = \deg f_i$. Let T be the set of reflections in G.*

Then $d_1 \cdots d_n = |G|$ and $d_1 + \cdots + d_n = |T| + n$.

Proof. Since A^G is generated by algebraically independent elements of degrees $d_1, \dots, d_n$, its Hilbert series is

$$\prod_{i=1}^n \frac{1}{1 - t^{d_i}} = \mathrm{H}_{A^G}(t) = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(1 - \rho(g) t)},$$

where the second equality follows from Molien's formula (Theorem 3.3). For the right side, since $\det(1 - \rho(g)t)$ factors as $\prod (1 - \zeta_i t)$ where ζ_i are the eigenvalues of $\rho(g)$, this is $(1-t)^n$ for $g = 1$, it is $(1-t)^{n-1}(1 - \omega_g t)$ for a root of unity $\omega_g \neq 1$ if $g \in T$, and otherwise is not divisible by $(1-t)^{n-1}$. Hence, we multiply both expressions for $\mathrm{H}_{A^G}(t)$ by $(1-t)^n$, we get

$$\prod_{i=1}^n \frac{1}{1 + t + \cdots + t^{d_i - 1}} = \frac{1}{|G|}(1 + \sum_{g \in T} \frac{1 - t}{1 - \omega_g t} + (1-t)^2 F(t))$$

where $F(t)$ is a rational function whose denominator is nonzero at $t = 1$. Now set $t = 1$ to get $\prod_{i=1}^n d_i^{-1} = 1/|G|$, or $d_1 \cdots d_n = |G|$.

Next, take the derivative with respect to t of the above identity to get

$$-\left(\prod_{i=1}^n \frac{1}{1+t+\cdots+t^{d_i-1}}\right) \sum_{i=1}^n \frac{1+2t+\cdots+(d_i-1)t^{d_i-2}}{1+t+\cdots+t^{d_i-1}} = \frac{1}{|G|}\sum_{g\in T} \frac{\omega_g - 1}{(1-\omega_g t)^2} + H(t).$$

where $H(t) = \frac{d}{dt}(1-t)^2 F(t)$ satisfies $H(1) = 0$. Next, set $t = 1$ to get

$$-\frac{1}{d_1\cdots d_n}\sum_{i=1}^n \frac{d_i(d_i-1)/2}{d_i} = -\frac{1}{|G|}\sum_{g\in T}\frac{1}{1-\omega_g}.$$

If $g \in T$ has order r, then $g, g^2, \dots, g^{r-1} \in T$ and $\omega_{g^i} = \omega_g^i$. By Lemma 3.11, their contribution to the sum is $\frac{r-1}{2}$. In particular, by grouping together terms based on the cyclic group that they generate, we see that the sum is $|T|/2$. Using $|G| = d_1 \cdots d_n$, we conclude that $\sum_{i=1}^n (d_i - 1) = |T|$. □

Lemma 3.11. *Let ω be a primitive rth root of unity. Then*

$$\sum_{i=1}^{r-1} \frac{1}{1-\omega^i} = \frac{r-1}{2}.$$

Proof. Define $f(x) = \prod_{i=1}^{r-1}(x - \omega^i)$. Using the product rule for derivatives,

$$\frac{f'(x)}{f(x)} = \sum_{i=1}^{r-1} \frac{1}{x-\omega^i}.$$

Now substitute $x = 1$. Since $f(x) = 1 + x + \cdots + x^{r-1}$, we have $f(1) = r$ and $f'(1) = \sum_{i=1}^{r-1} i = r(r-1)/2$, so the result follows. □

Theorem 3.12 (Shephard–Todd)**.** *Let $G \subset \mathbf{GL}_n(\mathbf{C})$ be a finite group and suppose that A^G is generated by algebraically independent elements. Then G is generated by reflections.*

Proof. Let $T \subset G$ be the set of reflections and let $H \subset G$ be the subgroup generated by T. By Theorem 3.9, A^H has a set of algebraically independent generators. Let $d_1 \le \cdots \le d_n$ be the degrees of the generators for A^G and let $e_1 \le \cdots \le e_n$ be the degrees of the generators for A^H. Since $A^G \subseteq A^H$, we must have $e_i \le d_i$ for all i (the first i generators of A^G are elements of A^H and hence require at least i of the generators of A^H to be generated).

By Theorem 3.10, we have

$$d_1 + \cdots + d_n = |T| + r = e_1 + \cdots + e_n,$$

so $e_i \le d_i$ cannot be strict for any i. But then $|H| = e_1 \cdots e_n = d_1 \cdots d_n = |G|$, so $G = H$ is generated by T. □

3.5. **Coinvariant ring.** As before, W is a reflection group and we let I be the ideal in A generated by A^W. The coinvariant ring is the quotient A/I. Pick a homogeneous basis $\{\overline{v}_i\}$ for A/I as a vector space and pick homogeneous representatives v_i for $\overline{v}_i + I$.

Lemma 3.13. *The v_i form a basis for A as an A^W-module.*

Proof. Let $f \in A$ be an element. We claim that it is in the A^W-module spanned by the v_i. It suffices to handle the case when f is homogeneous, and then we will prove the claim by induction on $\deg f$. The base case $\deg f = 0$ is clear, so assume $\deg(f) > 0$.

Let $\overline{f}$ be its image in A/I. Then by definition, we can find scalars such that $\overline{f} = \sum_i c_i \overline{v}_i$. Hence $f - \sum_i c_i v_i \in I$, so we can write it as $\sum_j h_j f_j$ where f_j are the basic invariants and the h_j are homogeneous. The expression $\sum_i c_i v_i$ is homogeneous of the same degree as f, so $\deg(h_j) + \deg(f_j) = \deg(f)$ for all j, and in particular, $\deg(h_j) < \deg(f)$ for all j. By induction, each h_j is spanned by the v_i with coefficients in A^W. Substituting these expressions shows that f is also spanned by the v_i with coefficients in A^W.

Now we claim that if $u_1, \dots, u_m \in A$ are homogeneous such that their images in A/I are linearly independent, then $u_1, \dots, u_m$ are linearly independent over A^W. We prove this by induction on m, with the case $m = 1$ being clear. If $u_1, \dots, u_m$ are dependent over A^W, then we have an expression of the form $h_1 u_1 + \cdots + h_m u_m = 0$ for $h_i \in A^W$. By Lemma 3.7, since $u_1 \notin I$, it must be that h_1 is in the A^W ideal generated by $h_2, \dots, h_m$, i.e., $h_1 = \alpha_2 h_2 + \cdots + \alpha_m h_m$ with $\alpha_i \in A^W$. But then we have

$$h_2(u_2 + \alpha_2 u_1) + \cdots + h_m(u_m + \alpha_m u_1) = 0.$$

Since $\alpha_i u_1 \in I$, the images of $u_2 + \alpha_2 u_1, \dots, u_m + \alpha_m u_1$ are linearly independent, so by induction $h_2 = \cdots = h_m = 0$, and hence $h_1 = 0$. □

Corollary 3.14. *A/I is a vector space of dimension $|W|$ and A is a free A^W-module of rank $|W|$. Furthermore, A/I is isomorphic to the regular representation of W.*

Proof. From the last result, $\dim(A/I) = \operatorname{rank}_{A^W} A$. If A is a free A^W-module of rank N, then by tensoring with $\operatorname{Frac}(A^W)$, we see that $A \otimes \operatorname{Frac}(A^W)$ is a dimension N vector space over $\operatorname{Frac}(A^W)$. We claim that $A \otimes \operatorname{Frac}(A^W) = \operatorname{Frac}(A)$. First, by Proposition 3.6, $\operatorname{Frac}(A)$ is a degree $|W|$ extension over $\operatorname{Frac}(A^W)$. Hence if $f \in A$, then $1/f$ satisfies a (monic) polynomial equation with coefficients in $\operatorname{Frac}(A^W)$, say

$$(1/f)^n + a_{n-1}(1/f)^{n-1} + \cdots + a_0 = 0.$$

Multiplying by f^{n-1}, we see $1/f = -(a_{n-1} + \cdots + a_0 f^{n-1})$. Since the a_i and f belong to $A \otimes \operatorname{Frac}(A^W)$, we conclude that $1/f$ does too, which shows that $A \otimes \operatorname{Frac}(A^W) = \operatorname{Frac}(A)$. Again using Proposition 3.6, we see that $N = |W|$.

For the last statement, pick a basis $\overline{v}_1, \dots, \overline{v}_N$ for A/I and lift them to representatives $v_1, \dots, v_N \in A$. After inverting A^W, this gives a basis for $\operatorname{Frac}(A)$ as a $\operatorname{Frac}(A^W)$-vector space. Now we appeal to the normal basis theorem in Galois theory [L, §VI.13] which says that $\operatorname{Frac}(A)$ is a regular representation of W as a vector space over $\operatorname{Frac}(A)^W = \operatorname{Frac}(A^W)$. This tells us that the trace of any non-identity element of W is 0, and so the same is true if we reduce modulo I. So we see that A/I is a regular representation (Proposition 2.3). □

Corollary 3.15. *Let $d_i = \deg(f_i)$. Then*

$$\mathrm{H}_{A/I^W}(t) = \frac{\mathrm{H}_A(t)}{\mathrm{H}_{A^W}(t)} = \prod_{i=1}^n \frac{1 - t^{d_i}}{1 - t} = \prod_{i=1}^n (1 + t + \cdots + t^{d_i - 1}).$$

Proof. Corollary 3.14 implies that

$$\mathrm{H}_{A/I^W}(t)\mathrm{H}_{A^W}(t) = \mathrm{H}_A(t),$$

so the result follows from $\mathrm{H}_{A^W}(t) = \prod_{i=1}^n (1 - t^{d_i})^{-1}$ and $\mathrm{H}_A(t) = \prod_{i=1}^n (1 - t)^{-1}$. □

Remark 3.16. In commutative algebra terminology, the fact that A/I is finite-dimensional means that $f_1, \dots, f_n$ form a system of parameters. Any system of parameters in a polynomial ring (more generally, for any Cohen–Macaulay ring) is automatically a regular sequence, which implies the Hilbert series formula for A/I that we just obtained. □

3.6. **Jacobian.** For any polynomials $h_1, \dots, h_n$, we define the **Jacobian** by

$$J = J(h_1, \dots, h_n) = \det \left(\frac{\partial h_i}{\partial x_j} \right)_{i,j=1,\dots,n} = \det \left(\frac{\partial h_j}{\partial x_i} \right)_{i,j=1,\dots,n}.$$

While this depends on the choice of basis $x_1, \dots, x_n$, a change of coordinates only changes J by a scalar (the determinant of the change of coordinates).

Lemma 3.17. *$h_1, \dots, h_n$ are algebraically independent if and only if $J(h_1, \dots, h_n) \neq 0$.*

Proof. First suppose that the h_i are algebraically dependent. Hence we can find a nonzero polynomial in new variables $F(y_1, \dots, y_n)$ such that $F(h_1, \dots, h_n) = 0$. Pick F so that it has smallest possible degree with this property. For each j, using the chain rule, we have

$$0 = \frac{\partial}{\partial x_j} F(h_1, \dots, h_n) = \sum_{i=1}^n \frac{\partial F}{\partial y_i}(h_1, \dots, h_n) \frac{\partial h_i}{\partial x_j}.$$

In particular, we have

$$\begin{pmatrix} \frac{\partial F}{\partial y_1}(h_1, \dots, h_n) & \cdots & \frac{\partial F}{\partial y_n}(h_1, \dots, h_n) \end{pmatrix} \left(\frac{\partial h_i}{\partial x_j} \right)_{i,j=1,\dots,n} = 0$$

Since F is not constant, at least one of its partial derivatives is nonzero. By minimality of $\deg F$, for each of these nonzero derivatives, we have $\frac{\partial F}{\partial x_j}(h_1, \dots, h_n) \neq 0$. Hence we see that $J = 0$.

Now suppose that $h_1, \dots, h_n$ are algebraically independent. By considering transcendence degree, the maximal number of algebraically independent polynomials is n. So for each i, the set $x_i, h_1, \dots, h_n$ is algebraically dependent. For each i, we let F_i be a polynomial in new variables $y_0, \dots, y_n$ of minimal degree such that $F_i(x_i, h_1, \dots, h_n) = 0$. For each j, we have

$$0 = \frac{\partial}{\partial x_k} F_i(x_i, h_1, \dots, h_n) = \frac{\partial F_i}{\partial y_0}(x_i, h_1, \dots, h_n)\delta_{i,j} + \sum_{k=1}^n \frac{\partial F_i}{\partial y_k}(x_i, h_1, \dots, h_n) \frac{\partial h_k}{\partial x_j}.$$

Combining these gives a matrix identity

$$\left(\frac{\partial F_i}{\partial y_j}(x_i, h_1, \dots, h_n) \right)_{i,j=1,\dots,n} \left(\frac{\partial h_i}{\partial x_j} \right)_{i,j=1,\dots,n} = -\left(\frac{\partial F_i}{\partial y_0}(x_i, h_1, \dots, h_n)\delta_{i,j} \right)_{i,j=1,\dots,n}.$$

Since the h_i are algebraically independent, F_i is a positive degree polynomial with respect to y_0 for each i. In particular, $\frac{\partial F_i}{\partial y_0} \neq 0$ and hence $\frac{\partial F_i}{\partial y_0}(x_i, h_1, \dots, h_n) \neq 0$ by minimality of $\deg F_i$, so the matrix on the right side is a diagonal matrix with nonzero entries. In particular, by taking determinants of both sides, we see that $J \neq 0$. □

We continue the same notation, so $W \subset \mathbf{GL}_n(\mathbf{C})$ is a finite reflection group and $f_1, \dots, f_n \in A = \mathbf{C}[x_1, \dots, x_n]$ are generators for A^W.

We let $\mathcal{J} = J(f_1, \dots, f_n)$, which is nonzero by the previous lemma and is homogeneous of degree $\sum_{i=1}^n (\deg(f_i) - 1)$.

Example 3.18. Consider $W = \mathfrak{S}_n$ acting on $\{x \in \mathbf{C}^n \mid x_1 + \cdots + x_n = 0\}$. Then the basic invariants can be taken to be the (normalized) power sums $\frac{1}{d}\sum_{i=1}^n x_i^d$. In that case, the Jacobian is the usual Vandermonde determinant

$$\mathcal{J} = \det\left(x_j^{i-1}\right)_{i,j=1,\dots,n} = \prod_{i<j}(x_j - x_i).$$

The solutions to $\mathcal{J}(x) = 0$ consist of the union of the reflection hyperplanes. This is a general fact to be proven next. □

3.7. **Solomon's theorem.** We now consider the algebra of differential forms in $x_1, \dots, x_n$ with coefficients in A. Algebraically, if V is the span of $x_1, \dots, x_n$, this is the ring $\mathrm{Sym}(V) \otimes \bigwedge(V)$ where $E = \bigwedge(V)$ denotes the exterior algebra of V, so we denote this algebra by $A \otimes E$. Concretely, this is an algebra over A, with multiplication denoted $\wedge$, generated by symbols $\mathrm{d}x_i$ subject to the relations $\mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i$. In particular, $A \otimes E$ is a free A-module of rank 2^n with basis given by $\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_r}$ with $1 \le i_1 < \cdots < i_r \le n$. We will denote this element by $\mathrm{d}x_I$ where $I = \{i_1, \dots, i_r\}$.

There is a bigrading on $A \otimes E$ where an element $h\mathrm{d}x_I$, with h homogeneous, has bidegree $(\deg h, |I|)$. Given a bigraded vector space M, we denote the bidegree (d, e)-component by $M_{d,e}$ and define a bigraded Hilbert series by

$$\mathrm{H}_M(t,u) = \sum_{d,e} (\dim M_{d,e}) t^d u^e.$$

In particular,

$$\mathrm{H}_{A\otimes E}(t,u) = \mathrm{H}_A(t,u)\mathrm{H}_E(t,u) = \left(\frac{1+u}{1-t}\right)^n.$$

We will extend the differential notation to make it linear, i.e., $\mathrm{d}(u+v) = \mathrm{d}u + \mathrm{d}v$. The action of $g \in W$ is given by

$$g \sum_I h_I \mathrm{d}x_I = \sum_I g h_I \mathrm{d}gx_{i_1} \wedge \cdots \wedge \mathrm{d}gx_{i_r}.$$

Our goal in this section is to understand $(A \otimes E)^W$.

A basic, but important, observation is that for polynomials $h_{i,j}$ with $i, j = 1, \dots, n$, we have

$$(\sum_i h_{i,1}\mathrm{d}x_1) \wedge \cdots \wedge (\sum_i h_{i,n}\mathrm{d}x_n) = \det(h_{i,j})\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n. \tag{3.19}$$

In particular, for $g \in \mathbf{GL}_n(\mathbf{C})$, we have $g(\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n) = (\det g)\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n$.

For a general polynomial h, we define

$$\mathrm{d}h = \sum_{i=1}^n \frac{\partial h}{\partial x_i}\mathrm{d}x_i.$$

This satisfies the product rule:

$$\mathrm{d}(h_1 h_2) = h_1 \mathrm{d}h_2 + h_2 \mathrm{d}h_1.$$

Lemma 3.20. *For any $g \in \mathbf{GL}_n(\mathbf{C})$, we have $g\mathrm{d}h = \mathrm{d}(gh)$.*

Proof. Next, if it holds for h_1 and h_2, then by the product rule it holds for h_1h_2:

$$g\mathrm{d}(h_1h_2) = g(h_1\mathrm{d}h_2+h_2\mathrm{d}h_1) = gh_1g(\mathrm{d}h_2)+gh_2(g\mathrm{d}h_1) = gh_1\mathrm{d}(gh_2)+gh_2\mathrm{d}(gh_1) = \mathrm{d}(g(h_1h_2)).$$

Since the statement holds for $h = x_i$ by definition, by induction on degree, and what we just said, it then holds when h is a monomial. Finally, if it holds for h_1 and h_2 it also holds for $h_1 + h_2$, so it holds for all polynomials. □

We say that $h \in A$ is a **skew-invariant** if $gh = (\det g)^{-1}h$ (where det is defined using that G is a subgroup of $\mathbf{GL}_n(\mathbf{C})$) for all $g \in W$.

For each reflection hyperplane H, consider the subgroup W_H of reflections that fix H. Let r_H be its order and let ℓ_H be a linear equation whose solution set is H.

Proposition 3.21. *(1) $\mathcal{J}$ is a skew-invariant.*
(2) $\mathcal{J}$ is a nonzero scalar multiple of $\prod_H \ell_H^{r_H-1}$.
(3) Every skew-invariant is of the form $h\mathcal{J}$ for some $h \in A^W$.

Proof. Use (3.19) with $h_{i,j} = \frac{\partial f_j}{\partial x_i}$ to get

$$\mathrm{d}f_1 \wedge \cdots \wedge \mathrm{d}f_n = \mathcal{J}\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n.$$

For each i, $\mathrm{d}f_i$ is a W-invariant by Lemma 3.20. Hence both sides are W-invariant. For the right side, we have for $g \in W$:

$$g(\mathcal{J}\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n) = (g\mathcal{J})(\det g)\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n.$$

Hence $(\det g)(g\mathcal{J}) = \mathcal{J}$ which implies that $\mathcal{J}$ is a skew-invariant, and proves (1).

Now suppose that F is an arbitrary skew-invariant. Given H, do a linear change of coordinates to variables $y_1, \dots, y_n$ so that $y_1 = \ell_H$ and $y_2, \dots, y_n$ are fixed by W_H. Then for a generator $g \in W_H$, we have $gF = (\det g)^{-1}F$ but also $gy_1^{p_1} \cdots y_n^{p_n} = (\det g)^{p_1}y_1^{p_1} \cdots y_n^{p_n}$ where $\det g$ is a primitive r_Hth root of unity. This implies that F is divisible by $\ell_H^{r_H-1}$. In particular, every skew-invariant is divisible by $\prod_H \ell_H^{r_H-1}$.

Now note that $\deg \mathcal{J} = \sum_{i=1}^n (d_i - 1)$, which is the number of reflections by Theorem 3.10, which can also be written $\sum_H (r_H - 1)$, so that $\mathcal{J}$ and $\prod_H \ell_H^{r_H-1}$ have the same degree. So, using the previous paragraph, we conclude that (2) holds. This also implies (3). □

Theorem 3.22 (Solomon)**.** *The ring of invariants $(A \otimes E)^W$ is freely generated as an exterior algebra over A^W by $\mathrm{d}f_1, \dots, \mathrm{d}f_n$. In other words, every element is uniquely of the form $\sum_I h_I\mathrm{d}f_I$ where $h_I \in A^W$ and $\mathrm{d}f_I = \mathrm{d}f_{i_1} \wedge \cdots \wedge \mathrm{d}f_{i_r}$ and $i_1 < \cdots < i_r$ are the elements of I.*

Proof. We prove the statement in the second form. First, we prove uniqueness, i.e., if $\sum_I h_I\mathrm{d}f_I = 0$, then $h_I = 0$ for all subsets I. We can break this relation into its homogeneous components with respect to the $\mathrm{d}x_I$, and assume that all I have size p for some fixed p. For a given subset I, let $I^c = [n] \setminus I$ and multiply the expression by $\mathrm{d}f_{I^c}$. Note that $\mathrm{d}f_I \wedge \mathrm{d}f_J = 0$ if $I \cap J \neq \varnothing$ and is $\pm\mathcal{J}\mathrm{d}x_{[n]}$ otherwise. Since $\mathcal{J} \neq 0$, we conclude that $h_I = 0$ for each I.

Next, we need to show that every element $\omega \in (A \otimes E)^W$ is of the form that we specified. Again, we may assume that ω is homogeneous in the $\mathrm{d}x_i$. We have $(A \otimes E)^W \subset \mathrm{Frac}(A) \otimes E$, and the argument above also shows that the elements $\mathrm{d}f_I$ are linearly independent over $\mathrm{Frac}(A)$. Since the exterior algebra has dimension 2^n, this also shows that the $\mathrm{d}f_I$ form a basis. So we can write $\omega = \sum_I c_I\mathrm{d}f_I$ with $c_I \in \mathrm{Frac}(A)$. Now we average over the elements of W:

$$\omega = \frac{1}{|W|}\sum_I (\sum_{g \in W} gc_I)\mathrm{d}f_I = \sum_I \frac{s_I}{t_I}\mathrm{d}f_I$$

where $s_I, t_I \in A$ and $s_I/t_I = \frac{1}{|W|}\sum_{g\in W} gc_I \in \mathrm{Frac}(A)^W$. Multiply by $\mathrm{d}f_{I^c}$ to get

$$\pm c_I \mathfrak{J} \mathrm{d}x_{[n]} = \omega \wedge \mathrm{d}f_{I^c} = \pm \frac{s_I}{t_I}\mathfrak{J}\mathrm{d}x_{[n]}.$$

Since $\omega \in A \otimes E$, we must have $c_I \mathfrak{J} \in A$. The right side is W-invariant, and hence so is the left side. This implies that $c_I\mathfrak{J}$ is a skew-invariant, and hence is divisible by $\mathfrak{J}$ by Proposition 3.21. In particular, $c_I \in A^W$, so we are done. □

Corollary 3.23. *The bigraded Hilbert series of $(A \otimes E)^W$ is*

$$\mathrm{H}_{(A\otimes E)^W}(t,u) = \prod_{i=1}^n \frac{1+ut^{d_i-1}}{1-t^{d_i}}.$$

Let $e_p(x_1,\dots,x_n) = \sum_{1\le i_1<\cdots<i_p\le n} x_{i_1}\cdots x_{i_p}$.

Theorem 3.24. *$e_p(d_1-1,\dots,d_n-1)$ is the number of $g \in W$ such that the eigenvalue 1 has multiplicity exactly $n-p$.*

Proof. By adapting the proof of Molien's formula, we can show that

$$\mathrm{H}_{(A\otimes E)^W}(t,u) = \frac{1}{|W|}\sum_{g\in W} \frac{\det(1+\rho_V(g)u)}{\det(1-\rho_V(g)t)}.$$

Combining this with Corollary 3.23 gives

$$\prod_{i=1}^n \frac{1+ut^{d_i-1}}{1-t^{d_i}} = \frac{1}{|W|}\sum_{g\in W} \frac{\det(1+\rho_V(g)u)}{\det(1-\rho_V(g)t)}.$$

For each p, equating the coefficients of u^p on both sides gives (let $\omega_1(g),\dots,\omega_n(g)$ be the eigenvalues of g)

$$\frac{e_p(t^{d_1-1},\dots,t^{d_n-1})}{\prod_{i=1}^n(1-t^{d_i})} = \frac{1}{|W|}\sum_{g\in W}\frac{e_p(\omega_1(g),\dots,\omega_n(g))}{(1-\omega_1 t)\cdots(1-\omega_n t))}.$$

Next, let $e'_p(x_1,\dots,x_n) = e_p(1-x_1,\dots,1-x_n)$; by expanding its terms, we see that it is a linear combination of $e_0(x),\dots,e_p(x)$. In particular, since the above formula is valid for all p, we can replace the e_p in the numerator with e'_p and get another identity

$$\frac{e_p(1-t^{d_1-1},\dots,1-t^{d_n-1})}{\prod_{i=1}^n(1-t^{d_i})} = \frac{1}{|W|}\sum_{g\in W}\frac{e_p(1-\omega_1(g),\dots,1-\omega_n(g))}{(1-\omega_1 t)\cdots(1-\omega_n t)}.$$

Now multiply both sides by $(1-t)^{n-p}$ and set $t=1$. For the left side (set $[d] = (1-t^d)/(1-t)$), by pulling out $(1-t)^p$ from each term in the numerator, we get:

$$(1-t)^n \frac{e_p([d_1-1],\dots,[d_n-1])}{\prod_{i=1}^n(1-t^{d_i})}\bigg|_{t=1} = \frac{e_p([d_1-1],\dots,[d_n-1])}{\prod_{i=1}^n[d_i]}\bigg|_{t=1} = \frac{e_p(d_1-1,\dots,d_n-1)}{|W|}$$

where in the last equality, we used $d_1\cdots d_n = |W|$ (Theorem 3.10).

For the right side, consider each summand. If the multiplicity of the eigenvalue 1 for g is greater than $n-p$, then $e_p(1-\omega_1(g),\dots,1-\omega_n(g)) = 0$. If it is less than $n-p$, then $(1-t)^{n-p}/((1-\omega_1 t)\cdots(1-\omega_n t))$ evaluated at $t=1$ is 0. If it is exactly $n-p$, then $e_p(1-\omega_1(g),\dots,1-\omega_n(g)) = (1-\omega_1(g))\cdots(1-\omega_p(g))$ where $\omega_1(g),\dots,\omega_p(g)$ are the eigenvalues different from 1. Hence the summand is 1, and we see that the right hand side is simply $1/|W|$ times the number of g such that the multiplicity of 1 as an eigenvalue is $n-p$. ∎

Remark 3.25. When $p = 1$ in the theorem, $e_1(d_1-1, \dots, d_n-1) = \sum_i (d_i-1)$ is the number of reflections, i.e., the number of elements that have 1 as an eigenvalue with multiplicity $n-1$, so this includes our previous formula. Also, note that

$$d_1 \cdots d_n = ((d_1-1)+1)\cdots((d_n-1)+1) = \sum_{p=0}^{n} e_p(d_1-1, \dots, d_n-1)$$

which the theorem says is $|W|$, so it also recovers that formula. ∎

3.8. **Spherical Coxeter complex.** We now restrict to the case that W is a Coxeter group and show that the Hilbert series of A/I gives the Poincaré series $W(t)$.

First we identify the fundamental domain from §1.8. Recall that if V is the geometric representation of W, then this is defined by

$$D = \{f \in V^* \mid f(\alpha_s) \geq 0 \text{ for all } s \in S\}.$$

By Proposition 2.11, the bilinear form B_W is nondegenerate and hence defines an isomorphism $V \to V^*$ via $v \mapsto B_W(v, -)$. The inverse image of D is then (we also call it D)

$$D = \{v \in V \mid B_W(v, \alpha_s) \geq 0 \text{ for all } s \in S\}.$$

We call this the **fundamental chamber**.

Proposition 3.26. *The union of the W-translates of D is all of V.*

Proof. Let $I \subseteq S$ be a subset. Let $w_{I,0} \in W_I$ be the maximal length element, as in Lemma 1.33. Then $w_{I,0}\alpha_s$ is negative for all $s \in I$ and otherwise $w_{I,0}\alpha_s = \alpha_s$. In particular,

$$w_{I,0}D = \{v \in V \mid B_W(v, \alpha_s) \leq 0 \text{ for all } s \in I \text{ and } B_W(v, \alpha_s) \geq 0 \text{ for all } s \notin I\}.$$

and the union over all subsets I is all of V. ∎

These translates will be called **chambers**. Each one is homeomorphic to the non-negative orthant in $\mathbf{R}^n$. Each subset $I \subseteq S$ gives a face D_I of D, given by

$$D_I = \{v \in D \mid B_W(v, \alpha_s) = 0 \text{ for all } s \in I\}.$$

Now consider the $(n-1)$-dimensional sphere $\mathrm{S}^{n-1} = \{v \in V \mid B_W(v, v) = 1\}$ and intersect with the W translates of the D_I. This gives a triangulation of S^{n-1} by simplices, which is the **spherical Coxeter complex**. We will now make use of some basic facts about (reduced) homology. (The conclusion below can be obtained without using any topology, see [H1, §1.16]. However, I think it is more natural to just appeal to some topology.) The triangulation gives us a chain complex $\mathbf{F}_\bullet$ that computes the reduced homology of S^{n-1} where $\mathbf{F}_i$ (for $i = -1, \dots, n-1$) is the $\mathbf{Q}$-vector space with basis given by W translates of D_I with $|I| = n - i - 1$ (i.e., the i-dimensional faces). The (rational) reduced homology of S^{n-1} is given by

$$\widetilde{\mathrm{H}}_i(\mathrm{S}^{n-1}; \mathbf{Q}) = \begin{cases} \mathbf{Q} & \text{if } i = n-1 \\ 0 & \text{else} \end{cases}.$$

Proposition 3.27. *As class functions on W, we have*

$$\det = \sum_{I \subseteq S} (-1)^{|I|} 1 \uparrow_{W_I}^{W}.$$

Proof. Since the triangulation is W-invariant, each w gives a chain map $w \colon \mathbf{F}_\bullet \to \mathbf{F}_\bullet$, and the Hopf trace formula tells us that

$$(-1)^{n-1} \operatorname{Tr}(w \mid \widetilde{\mathrm{H}}_{n-1}(\mathrm{S}^{n-1}; \mathbf{Q})) = \sum_{i=-1}^{n-1} (-1)^i \operatorname{Tr}(w \mid \mathbf{F}_i).$$

Since w is an orthogonal matrix, it acts trivially on $\widetilde{\mathrm{H}}_{n-1}(\mathrm{S}^{n-1}; \mathbf{Q})$ if and only if it is orientation-preserving and by -1 otherwise. In other words, the trace is simply $\det w$.

Next, $\mathbf{F}_i$ is a permutation representation on the set of i-dimensional faces. There is one orbit for each subset $I \subseteq S$ with $|I| = n - i - 1$, and the D_I are representatives. By Theorem 1.31, the stabilizer of D_I is W_I. So by Example 2.4, we see that

$$\operatorname{Tr}(w \mid \mathbf{F}_i) = \sum_{I,\ |I|=n-i-1} 1 \uparrow_{W_I}^{W} (w).$$

Combining all of this gives the desired formula. □

Recall the definition of the Poincaré series from §1.9

$$W(t) = \sum_{w \in W} t^{\ell(w)}.$$

As before, we let $A = \mathbf{C}[x_1, \dots, x_n]$ with the action of W by linear change of coordinates. The action of W on $\mathbf{C}^n$ is isomorphic to the complexification of $V^* \cong V$.

Let $A^{W,\varepsilon}$ denote the submodule of skew-invariants of A. By Proposition 3.21, this is a free A^W-module of rank 1 with its generator in degree $\sum_{i=1}^n (d_i - 1)$. By Theorem 3.10, this is the number of reflections of W, which is also the number of positive roots, and hence $\ell(w_0)$. So

$$\mathrm{H}_{A^{W,\varepsilon}}(t) = \frac{t^{\ell(w_0)}}{\prod_{i=1}^n (1 - t^{d_i})}.$$

Theorem 3.28. *If W is a finite Coxeter group, then*

$$W(t) = \mathrm{H}_{A/I^W}(t) = \prod_{i=1}^n \frac{1 - t^{d_i}}{1 - t}.$$

Proof. Define $Q_W(t) = \mathrm{H}_{A/I^W}(t) = \prod_{i=1}^n \frac{1-t^{d_i}}{1-t}$.

By Frobenius reciprocity (Theorem 2.5), we have $\dim A_d^{W_I} = \langle 1 \uparrow_{W_I}^{W}, \operatorname{char}(A_d) \rangle_W$. In particular, using Proposition 3.27, we get

$$\begin{aligned}
\sum_{I \subseteq S} (-1)^{|I|} \mathrm{H}_{A^{W_I}}(t) &= \sum_{I \subseteq S} (-1)^{|I|} \sum_{d \geq 0} \langle 1 \uparrow_{W_I}^{W}, \operatorname{char}(A_d) \rangle_W t^d \\
&= \sum_{d \geq 0} \langle \det, \operatorname{char}(A_d) \rangle_W t^d = \mathrm{H}_{A^{W,\varepsilon}}(t) \\
&= \frac{t^{\ell(w_0)}}{\prod_{i=1}^n (1 - t^{d_i})} = \frac{t^{\ell(w_0)}}{(1-t)^n Q_W(t)}
\end{aligned}$$

If $|I| = p$, then the restriction of the geometric representation of W to W_I is isomorphic to the geometric representation of W_I direct sum a copy of a trivial representation of dimension $n-p$. Hence if $d'_1, \dots, d'_p$ are the degrees of the basic invariants of W_I on its geometric representation, then $1, \dots, 1, d'_1, \dots, d'_p$ are the degrees of the basic invariants of W_I on $\mathbf{C}^n$, and so

$$\mathrm{H}_{A^{W_I}}(t) = \frac{1}{(1-t)^{n-p}\prod_{i=1}^p (1-t^{d'_i})} = \frac{1}{(1-t)^n Q_{W_I}(t)}.$$

Hence, after multiplying both sides by $(1-t)^n Q_W(t)$, the first equation above becomes

$$\sum_{I \subseteq S} (-1)^{|I|} \frac{Q_W(t)}{Q_{W_I}(t)} = t^{\ell(w_0)}.$$

Now we prove the result by induction on $|S|$. The case $|S| = 0$ is clear: $W(t) = 1 = Q_W(t)$. If $|S| > 0$, we have

$$\frac{t^{\ell(w_0)} - (-1)^{|S|}}{W(t)} = \sum_{I \subsetneq S} \frac{(-1)^{|I|}}{W_I(t)} = \sum_{I \subsetneq S} \frac{(-1)^{|I|}}{Q_{W_I}(t)} = \frac{t^{\ell(w_0)} - (-1)^{|S|}}{Q_W(t)},$$

the first equality follows from (1.35), the second equality follows by induction since $|I| < |S|$, and the third follows from the identity we just proved. Hence $W(t) = Q_W(t)$. ☐

Remark 3.29. We previously showed by direct calculation that the Poincaré series of $W = \mathfrak{S}_n$ is $\prod_{i=2}^n \frac{1-t^i}{1-t}$. This is consistent with the above result since the degrees of the invariants of $\mathfrak{S}_n$ on $\mathbf{C}^n$ are $1, \dots, n$ (we could also use $\{x \in \mathbf{C}^n \mid x_1 + \cdots + x_n = 0\}$ in which case the degrees are $2, \dots, n$). ☐

We will see later how to get a closed form formula for W when its Coxeter graph is positive semidefinite (the affine Weyl groups). This will make use of Solomon's theorem and the "toroidal Coxeter complex", in which we get a triangulation of a torus rather than a sphere.

3.9. **Examples.**

We start with a useful fact which gives an easy way to check that polynomials are algebraically independent.

Lemma 3.30. *Let $h_1, \dots, h_n$ be homogeneous polynomials such that $h_1(\alpha) = \cdots = h_n(\alpha) = 0$ implies that $\alpha = 0$, i.e., there is no nontrivial solution to the h_i. Then the h_i are algebraically independent.*

Proof. We just provide a sketch. Let I be the ideal generated by the h_i. Using the Hilbert Nullstellensatz, the condition on the h_i implies that the radical of I is the ideal generated by the variables. This implies that A/I is a finite-dimensional $\mathbf{C}$-vector space. If we pick a basis and consider their preimages in A, then they give generators for A as a module over $\mathbf{C}[h_1, \dots, h_n]$. Hence the fraction field of A is a finite extension of the fraction field of $\mathbf{C}[h_1, \dots, h_n]$, so the latter has transcendence degree n over $\mathbf{C}$ and hence the h_i are algebraically independent. ☐

The condition on the h_i is strictly stronger than being algebraically independent. For example, x^2, xy are algebraically independent but have a nontrivial solution $(0, 1)$. However, it will be easier to check than showing that the Jacobian is nonzero for our examples.

We will use one other fact, essentially proven in Theorem 3.12: if $h_1, \dots, h_n \in A^W$ are homogeneous and algebraically independent and $(\deg f_1)\cdots(\deg f_n) = |W|$, then the h_i generate A^W.

Example 3.31. A large source of reflection groups is the family called $G(m,p,n)$ where m,p,n are positive integers such that p divides m. First, for $p = 1$, this is the subgroup of $n \times n$ matrices such that each row and entry has exactly one nonzero entry, that each nonzero entry is an mth root of unity. Then $|G(m,1,n)| = n!m^n$. For general p, $G(m,p,n)$ the subgroup where the product of the nonzero entries is an (m/p)th root of unity. Then $|G(m,1,n)| = n!m^n/p$. We claim this is generated by reflections (let ω be a primitive mth root of unity). When $p = 1$: we can take as a set of generators the adjacent transpositions in S_n together with the diagonal matrix whose entries are $\omega, 1, \dots, 1$. If $p = m$, we replace that last matrix by the block sum of $\begin{pmatrix} 0 & \omega \\ \omega^{-1} & 0 \end{pmatrix}$ with an identity matrix of size $n-2$. For general p, $G(m,p,n)$ is generated by $G(m,m,n)$ together with the diagonal matrix whose entries are $\omega^p, 1, \dots, 1$. To see that this is correct, note that $G(m,p,n)$ is generated by permutation matrices together with diagonal matrices whose entries are $\omega^{a_1}, \dots, \omega^{a_n}$ such that $a_i \in \mathbf{Z}/m$ and $\sum_i a_i = 0 \pmod p$. The group $G(m,m,n)$ already contains the diagonal matrices satisfying $\sum_i a_i = 0 \pmod m$, so any diagonal matrix in $G(m,p,n)$ differs by a diagonal matrix in $G(m,m,n)$ by a power of the new generator we added.

These are usually irreducible, but not always. For example, $G(1,1,n) \cong S_n$ acting on $\mathbf{C}^n$, which is not its geometric representation since there is an extra trivial factor. Almost all of the irreducible reflection groups are one of the $G(m,p,n)$ (there are 34 exceptions, which were classified by Shephard–Todd [ST]).

This group is a Coxeter group in a few cases:

- When $m = 1$, we get the symmetric group S_n.
- When $m = 2$ and $p = 1$, we get the Coxeter groups of type B_n.
- When $m = p = 2$, we get the Coxeter groups of type D_n.
- When $m = p$ and $n = 2$, we get the dihedral groups $\mathrm{I}_2(m)$.

Now we describe generators for the invariant ring. We start with $S_n \cong G(1,1,n)$.

In that case, define $e_p(x_1, \dots, x_n) = \sum_{i_1 < \cdots < i_p} x_{i_1} \cdots x_{i_p}$ as before. They are clearly S_n-invariant. Note that

$$\sum_{i=0}^{n} (-1)^p e_p(x) t^{n-p} = (t - x_1)\cdots(t - x_n),$$

so that $e_p(\alpha) = 0$ for all p if and only if $\alpha = 0$. So $e_1, \dots, e_n$ are algebraically independent and the product of their degrees is $n! = |S_n|$, which means they generate.

Now consider $G(m,1,n)$ in general. We now consider the polynomials $e_p(x_1^m, \dots, x_n^m)$, which are clearly $G(m,1,n)$-invariant. If we had a nonzero solution for these polynomials, by extracting mth roots we would get a nonzero solution for the original $e_1, \dots, e_n$. Hence they are algebraically independent. The product of their degrees is now $m^n n! = |G(m,1,n)|$, so they generate the ring of invariants.

Finally, for $G(m,p,n)$, we take $e_p(x_1^m, \dots, x_n^m)$ for $p = 1, \dots, n-1$ and $e_n(x_1^{m/p}, \dots, x_n^{m/p}) = (x_1 \cdots x_n)^{m/p}$. Again, these are $G(m,p,n)$-invariant, and algebraically independent and the product of their degrees is $|G(m,p,n)|$ so they generate the ring of invariants. □

There are 34 additional irreducible complex reflection groups, numbered 4 through 37 in the Shephard–Todd classification. The ones of rank 3 have interesting connections to projective plane geometry. We'll describe a couple, see [D] for some more about reflection groups and algebraic geometry.

Example 3.32. Group number 24 is closely related to the Klein quartic: this is the degree 4 equation

$$x^3y + y^3z + z^3x = 0,$$

whose (projective) solution set is a Riemann surface of genus 3 in the projective plane. Its automorphism group is of order 168 (the maximum possible size allowed by Hurwitz's bound) and is isomorphic to $\mathbf{PSL}_2(\mathbf{F}_7)$, which is the group of determinant 1 2×2 matrices over $\mathbf{Z}/7$ modulo the subgroup generated by $-I$. Then, precisely, group 24 is isomorphic to $\mathbf{Z}/2 \times \mathbf{PSL}_2(\mathbf{F}_7)$, and its invariants have degrees $4, 6, 14$. The degree 4 invariant is the Klein quartic and the degree 6 invariant is its Hessian (determinant of the second partial derivatives). □

Example 3.33. For group 25, we consider the Hesse pencil. This is the family of cubic curves (λ, μ are projective parameters):

$$\lambda(x_0^3 + x_1^3 + x_2^3) + \mu x_0x_1x_2 = 0.$$

Think of the indices as elements of $\mathbf{Z}/3$ and let ω be a primitive 3rd root of unity. We have two linear operators $\sigma(x_i) = x_{i+1}$ and $\tau(x_i) = \omega^i x_i$. They generate a nonabelian group H of order 27 (they don't commute, but their commutator is multiplication by ω), and group 25 N normalizes H and $N/H \cong \mathbf{SL}_2(\mathbf{F}_3)$. The size of the group is 648. (This is a more general construction in the theory of level structures on abelian varieties.) Furthermore, there are 12 reflection planes, which are projectively lines. To obtain them: there are 4 values of $[\lambda : \mu]$ for which the cubic curve above is singular, and in that case it is the union of 3 lines. This gives the 12 lines we want. □

4. Affine Weyl groups

4.1. **Affine representation.** Consider a Coxeter group (W_a, S_a) associated with one of the graphs in Proposition 2.17 and let V_a be its geometric representation. We let $n+1 = \dim V_a$, and pick an element $s_0 \in S_a$ so that the graph on $S = S_a \setminus s_0$ is of type X_n if (W_a, S_a) is of type $\widetilde{\mathrm{X}}_n$. We know that the kernel of the bilinear form B_{W_a} is 1-dimensional, denote it by $V_a^\perp = \ker B_{W_a}$. We let $\delta \in V_a^\perp$ be the unique vector such that the coefficient of α_0 is 1, which is possible by Theorem 2.13.

Define

$$Z = \{f \in V_a^* \mid f(\delta) = 0\},$$
$$E = \{f \in V_a^* \mid f(\delta) = 1\}.$$

Then Z is a vector space and E is an affine space over Z, i.e., we have a simply transitive action $Z \times E \to E$.

A map $\varphi\colon E \to E$ is an **affine transformation** if there exists a linear map $\psi\colon Z \to Z$ (necessarily unique) such that $\varphi(e + z) = \varphi(e) + \psi(z)$ for all $e \in E$ and $z \in Z$. The set of affine transformations forms a group under composition denoted $\mathbf{Aff}(E)$.

Lemma 4.1. $\mathbf{Aff}(E) \cong \{g \in \mathbf{GL}(V_a^*) \mid g(E) = E\}$.

Proof. First, suppose that $g \in \mathbf{GL}(V_a^*)$ satisfies $g(E) = E$. Then for any $e \in E$ and $z \in Z$, we have $g(e + z) = g(e) + g(z)$ by linearity, so $g\colon E \to E$ is affine. Conversely, suppose we have an affine transformation $\varphi\colon E \to E$ with corresponding linear map $\psi\colon Z \to Z$. Pick a basis $v_0, \dots, v_n$ such that $v_0 = \delta$ and let $v_0^*, \dots, v_n^*$ be the dual basis for V_a^*. Then $v_1^*, \dots, v_n^*$ is a basis for Z and we define $g \in \mathbf{GL}(V_a^*)$ by

$$g(c_0 v_0^* + \cdots + c_n v_n^*) = c_0 \varphi(v_0^*) + \psi(c_1 v_1^* + \cdots + c_n v_n^*).$$

Then $g(E) = E$ and its restriction to E is φ: any element of E is of the form $v_0^* + v'$ where v' is a span of $v_1^*, \dots, v_n^*$, and so $g(v_0^* + v') = \varphi(v_0^*) + \psi(v') = \varphi(v_0^* + v')$. □

Since $w\delta = \delta$ for all w by the formula for how each $s \in S_a$ acts, w preserves both Z and E and hence we get a homomorphism

$$W_a \to \mathbf{Aff}(E).$$

Via the identification of $\mathbf{Aff}(E)$ with a subgroup of $\mathbf{GL}(V_a^*)$, we see that this homomorphism is injective (since the same holds for the geometric representation).

Next, B_{W_a} descends to a positive definite form on $V_a/V_a^\perp$ (which we call B_W) and we have an identification $Z \cong (V_a/V_a^\perp)^*$. Using B_W, we also get an identification $(V_a/V_a^\perp)^* \cong V_a/V_a^\perp$ where $f \in (V_a/V_a^\perp)^*$ is identified with the unique vector v_f such that $f(x) = B_W(v_f, x)$ for all $x \in V_a/V_a^\perp$. In particular, we get a W_a-invariant positive definite form on Z.

For $s \in S_a$, define

$$Z_s = \{f \in V_a^* \mid f(\alpha_s) = 0\}, \qquad E_s = E \cap Z_s.$$

Since the coefficient of α_0 in δ is nonzero, $\{\alpha_s \mid s \in S\} \cup \{\delta\}$ is linearly independent. Hence $E \cap \bigcap_{s \in S} Z_s$ is a single point, call it e_0. Explicitly, we have

$$e_0(\alpha_{s_0}) = 1, \qquad e_0(\alpha_s) = 0 \text{ for } s \neq s_0.$$

We can identify Z with E via $z \mapsto e_0 + z$. We denote the inverse by $e' \mapsto e' - e_0$. This identification allows us to transfer the form to E, call it B_E.

Lemma 4.2. *B_E is invariant under $(W_a)_{e_0}$.*

Proof. If $w \in (W_a)_{e_0}$, i.e., $we_0 = e_0$, then for $e', e'' \in E$, we have

$$\begin{aligned} B_E(we', we'') &= B_W(we' - e_0, we'' - e_0) \\ &= B_W(w(e' - e_0), w(e'' - e_0)) \\ &= B_W(e' - e_0, e'' - e_0) \\ &= B_E(e', e''). \end{aligned}$$

□

Proposition 4.3. *W_a is isomorphic to a subgroup of $\mathbf{Aff}(E)$ generated by affine reflections.*

Proof. For each $s \in S$, E_s is an affine hyperplane passing through e_0. The action of s on E fixes E_s, and hence e_0, so that it preserves B_E. Since it has order 2, it must be the reflection with respect to E_s.

On the other hand, E_{s_0} does not contain e_0. Let $\alpha_0' \in Z$ be the element corresponding to the linear functional $v \mapsto B_{W_a}(\alpha_{s_0}, v)$. Given $v \in V$ and $z \in Z$, we have

$$\begin{aligned} (s_0(e_0 + z))(v) &= (e_0 + z)(s_0 v) \\ &= (e_0 + z)(v - 2B_{W_a}(v, \alpha_{s_0})\alpha_{s_0}) \\ &= (e_0 + z)(v) - 2(e_0 + z)(\alpha_{s_0})\alpha_0'(v), \\ &= (e_0 + z)(v) - 2(1 + B_W(z, \alpha_0'))\alpha_0'(v) \end{aligned}$$

In other words, s_0 is an affine reflection. □

4.2. **Semidirect product structure.** Let W be an irreducible finite Weyl group of rank n with geometric representation V as in §2.4. Let Φ be the set of roots constructed there with simple roots $\alpha_1, \dots, \alpha_n$. This differs from our general theory in that they generally do not have the same length 1. We use $(,)$ in place of B_W to distinguish this fact. However, the $\mathbf{Z}$-span L of Φ is a lattice preserved by the action of W. We let $\widetilde{\alpha}$ denote the highest root.

For each root $\alpha \in \Phi$, we define $\alpha^\vee = \frac{2\alpha}{(\alpha,\alpha)}$. We let $L^\vee$ denote the $\mathbf{Z}$-span of $\alpha^\vee$ as α ranges over the roots and call it the **coroot lattice**.

If Φ is of type X_n, let Γ_a be the Coxeter graph of type $\widetilde{\mathrm{X}}_n$. (From our discussion, Φ is determined by W except if $W = W(\mathrm{B}_n)$ in which case we have two root systems, type B and type C.) This is obtained by adding a single node to the Coxeter graph of W; let α_0 be the new simple root.

Lemma 4.4. *The kernel of the bilinear form for Γ_a is spanned by $\delta = \alpha_0 + \widetilde{\alpha}$.*

Proof. This can be checked in each case directly, so I won't elaborate on it. □

For $k \in \mathbf{Z}$, we define $s_{\alpha,k} \in \mathbf{Aff}(V)$ by

$$s_{\alpha,k}(v) = v - ((\alpha, v) - k)\alpha^\vee.$$

We omit the check that $s_{\alpha,0}$ preserves the form $(,)$, and call this just s_α. This is a reflection with respect to the affine hyperplane

$$H_{\alpha,k} = \{v \in V \mid (\alpha, v) = k\}.$$

For $v \in V$, let $t_v \in \mathbf{Aff}(V)$ denote translation by v, i.e., $t_v(x) = x + v$.

Lemma 4.5. *We have the following properties.*

(1) If $w \in W$, then $wH_{\alpha,k} = H_{w\alpha,k}$ and $ws_{\alpha,k}w^{-1} = s_{w\alpha,k}$.
(2) We have $s_{\alpha,1}s_\alpha = t_{\alpha^\vee}$.
(3) If $v \in V$ satisfies $(v, \alpha) \in \mathbf{Z}$, then $t_v H_{\alpha,k} = H_{\alpha,k+(v,\alpha)}$ and $t_v s_{\alpha,k} t_v^{-1} = s_{\alpha,k+(v,\alpha)}$.

Proof. (1) For $w \in W$ and $x \in H_{\alpha,k}$, we have $(w\alpha, wx) = (\alpha, x) = k$, so $wH_{\alpha,k} = H_{w\alpha,k}$. Also, we have for $x \in V$,

$$ws_{\alpha,k}w^{-1}(x) = w(w^{-1}x - ((\alpha, w^{-1}x) - k)\alpha^\vee) = x - ((w\alpha, x) - k)w(\alpha^\vee) = s_{w\alpha,k}(x)$$

since $w(\alpha^\vee) = (w\alpha)^\vee$.

(2) It's clear from the formula that $s_{\alpha,k} = t_{k\alpha^\vee}s_\alpha$. Take $k = 1$ and multiply both sides on the right by s_α to get $s_{\alpha,1}s_\alpha = t_{\alpha^\vee}$.

(3) The first equality is clear since if $x \in H_{\alpha,k}$, then $(\alpha, x+v) = k+(\alpha, v)$. For the second, we have for $x \in V$,

$$\begin{aligned} t_v s_{\alpha,k} t_v^{-1}(x) &= t_v s_{\alpha,k}(x - v) \\ &= t_v(x - v - ((\alpha, x - v) - k)\alpha^\vee) \\ &= x - ((\alpha, x) - (k + (\alpha, v)))\alpha^\vee \\ &= s_{\alpha,k+(\alpha,v)}(x). \end{aligned}$$

□

We have identifications $V \cong Z \cong E$ where the first map sends v to the linear functional $B_{W_a}(v, -)$ and the second map is adding the vector e_0. For the following, note that we are using different conventions now for what the simple roots are compared to the last section, which affects the definition of δ (and potentially some scalars). We take $\delta = \alpha_0 + \widetilde{\alpha}$ to be consistent with the current section.

Lemma 4.6. *Via the identification $V \cong E$, $s_{\widetilde{\alpha},1}$ coincides with $s_0 \in W_a$.*

Proof. We have previously shown that s_0 is an affine reflection on E, so it suffices to compute the affine hyperplane that it fixes. For $v \in V$ and $z \in Z$, we have

$$(s_0(e_0 + z))(v) = (e_0 + z)(v - 2B_{W_a}(v, \alpha_0)\alpha_0),$$

and hence $s_0(e_0 + z) = e_0 + z$ if and only if $(e_0 + z)(\alpha_0) = 0$. Since $e_0(\alpha_0) = 1$ and $\alpha_0 \equiv -\widetilde{\alpha} \pmod{V_a^\perp}$, this is equivalent to $z(\widetilde{\alpha}) = 1$. Hence, for $u \in V$, this condition translates, via our identification of V with E, to $B_{W_a}(u, \widetilde{\alpha}) = 1$, and B_{W_a} restricts to $(,)$ on V, so we see the fixed affine hyperplane of s_0 is $H_{\widetilde{\alpha},1}$. □

Corollary 4.7. *The group generated by $s_{\alpha_1}, \dots, s_{\alpha_n}, s_{\widetilde{\alpha},1}$ is isomorphic to W_a.*

Corollary 4.8.
(1) W_a contains t_v for every $v \in L^\vee$.
(2) For every root α and $k \in \mathbf{Z}$, we have $s_{\alpha,k} \in W_a$.

Proof. We refer to the 3 items in Lemma 4.5. From Corollary 4.7, W_a is generated by $s_{\alpha_1}, \dots, s_{\alpha_n}, s_{\widetilde{\alpha},1}$. Since W acts transitively on the set of roots, by (1) we have $s_{\alpha,1} \in W_a$ for every root α. Hence by (2), we get $t_{\alpha^\vee} \in W_a$ for every root α. Since $(\alpha, \alpha^\vee) = 2$, by (3) we have $t^k_{\alpha^\vee} s_\alpha t^{-k}_{\alpha^\vee} = s_{\alpha,2k}$ and $t^k_{\alpha^\vee} s_{\alpha,1} t^{-k}_{\alpha^\vee} = s_{\alpha,2k+1}$ for all k. □

We let $T_{L^\vee} = \{t_v \mid v \in L^\vee\}$. We have a usual action of W on $T_{L^\vee}$ via $w \cdot t_v = t_{wv}$. The semidirect product $T_{L^\vee} \rtimes W$ is defined to be $T_{L^\vee} \times W$ as a set, with the product $(t_v, w)(t_{v'}, w') = (t_v t_{wv'}, ww')$.

Corollary 4.9. *$T_{L^\vee}$ is a normal subgroup of W_a and $W_a \cong T_{L^\vee} \rtimes W$.*

Proof. We have $wt_vw^{-1} = t_{wv}$, so $T_{L^\vee}$ is a normal subgroup. Note that $W \cap T_{L^\vee}$ is just the identity element, since W preserves the zero vector in V, and the only element in $T_{L^\vee}$ that does that is t_0. Hence the composition $\varphi\colon W \to W_a \to W_a/T_{L^\vee}$ is injective. Next, $s_{\widetilde{\alpha}} = s_{\widetilde{\alpha},1} t_{\widetilde{\alpha}^\vee}$ and $s_{\widetilde{\alpha}} \in W$, so the coset $s_{\widetilde{\alpha},1} T_{L^\vee}$ is in the image of φ. In other words, $\varphi\colon W \to W_a/T_{L^\vee}$ is an isomorphism.

In particular, we have distinguished right coset representatives $\{T_{L^\vee} w \mid w \in W\}$ for $W_a/T_{L^\vee}$, and so every element is of the form $t_v w$ for a unique $w \in W$ and $v \in L^\vee$. Since $wt_{v'} = t_{wv'}w$, we have $t_v w t_{v'} w' = t_v t_{wv'} ww'$. □

4.3. **Alcoves.** In §1.8, we constructed a fundamental domain D for the action of W_a using its geometric realization. Using the notation from §4.1 again, we have

$$D = \{f \in V_a^* \mid f(\alpha_s) \geq 0 \text{ for all } s \in S_a\}.$$

As before, define $E = \{f \in V_a^* \mid f(\delta) = 1\}$.

As in the last section, V is the geometric representation of the finite Weyl group and we let $\alpha_1, \dots, \alpha_n \in V$ be the roots and let $\widetilde{\alpha}$ denote the highest root.

Lemma 4.10. *We have*

$$D \cap E = \{e_0 + z \mid (z, \alpha_i) \geq 0 \text{ for } 1 \leq i \leq n \text{ and } (z, \widetilde{\alpha}) \leq 1\}.$$

Proof. Suppose $e_0 + z \in D \cap E$. Then $(e_0 + z)(\alpha_s) \geq 0$ for all $s \in S_a$ and $(e_0 + z)(\delta) = 1$. Since $e_0(\alpha_s) = 0$ for $s \in S$, we get $z(\alpha_s) \geq 0$ for $s \in S$. Since $e_0(\alpha_{s_0}) = 1$, we get that $z(\alpha_{s_0}) \geq -1$. When we identify $(V_a/V_a^\perp)^*$ with V, for $s \in S$, α_s becomes some α_i for $1 \leq i \leq n$, and α_{s_0} becomes $-\widetilde{\alpha}$, so the conditions we just wrote are equivalent to $(z, \alpha_i) \geq 0$ for $1 \leq i \leq n$ and $B(z, \widetilde{\alpha}) \leq 1$.

Conversely, if these conditions hold, then $(e_0 + z)(\alpha_s) \geq 0$, and

$$1 = (e_0 + z)(\delta) = (e_0 + z)(\alpha_{s_0}) + z(\delta - \alpha_{s_0}).$$

The latter quantity is $(z, \widetilde{\alpha})$, so we see that $(e_0 + z)(\alpha_{s_0}) \geq 0$. □

We define

$$A_\circ = \{v \in V \mid (v, \alpha_i) > 0 \text{ for } 1 \leq i \leq n \text{ and } (v, \widetilde{\alpha}) < 1\},$$
$$A = \{v \in V \mid (v, \alpha_i) \geq 0 \text{ for } 1 \leq i \leq n \text{ and } (v, \widetilde{\alpha}) \leq 1\}.$$

We call $A_\circ$ the **fundamental alcove**. Note that $A_\circ$ is an open simplex and A is its closure. By Theorem 1.31, each W_a-orbit in E intersects $D \cap E$ in at most one point. This translates to saying that every W_a-orbit in V intersects A in at most one point.

Proposition 4.11. *The W_a-orbit of A is all of V.*

Proof. Fix $\lambda \in A_\circ$. Pick any $\mu \in V$. Then $T_{L^\vee}\mu$ is a discrete set, and since W is finite, this implies that $W_a\mu$ is a discrete set. In particular, we can pick $\nu \in W_a\mu$ whose distance from λ is minimized. If $\nu \in A$, we're done. Otherwise, there is a hyperplane that bounds $A_\circ$ and separates λ and ν. Let w be the corresponding reflection. Then $\|w\nu - \lambda\| < \|\nu - \lambda\|$.[1] Since $w\nu \in W_a\mu$, this contradicts the choice of ν, so we conclude that $\nu \in A$. □

Example 4.12. Consider the A_1 root system. Our model has been to take the line in $\mathbf{R}^2$ spanned by $(1, -1)$. We can also think of this as $V \cong \mathbf{R}$ and $\Phi = \{2, -2\}$ so $\alpha_1^\vee = 1$. In that case, $A_\circ = (0, 1)$ and our affine Weyl group is generated by the negation operator and reflection with respect to the point 1. □

Example 4.13. The irreducible rank 2 root systems $\widetilde{\mathrm{A}}_2$, $\widetilde{\mathrm{B}}_2$, $\widetilde{\mathrm{C}}_2$, and $\widetilde{\mathrm{G}}_2$ give tilings of the plane by triangles. The pictures are omitted here. □

4.4. **Length function.** Fix $w \in W_a$. We say that a hyperplane $H = H_{\alpha,k}$ separates $A_\circ$ and $wA_\circ$ if they lie on opposite sides of H. Alternatively, if we join some point of $A_\circ$ with some point of $wA_\circ$ with a line segment, then H separates if and only if it intersects this line segment. For each α, the hyperplanes $H_{\alpha,k}$ as k varies are parallel, so at most finitely many of them can intersect a finite line segment. As there are only finitely many choices for α, the number of separating hyperplanes is finite. We define

$$\mathcal{L}(w) = \{H_{\alpha,k} \mid H_{\alpha,k} \text{ separates } A_\circ \text{ and } wA_\circ\}.$$

We define $H_{s_0} = H_{\widetilde{\alpha},1}$ and for $s \in S$, we define $H_s = H_{\alpha_s,0}$.

Lemma 4.14. *For $s \in S_a$, $\mathcal{L}(s) = \{H_s\}$.*

Proof. It is clear that H_s separates $A_\circ$ and $sA_\circ$ so we have to show that it is the only separating hyperplane. Pick $x \in A_\circ$. Then $0 < (x, \alpha) < 1$ for all positive roots α. Then for $t \in S_a \setminus \{s\}$, we have $(sx, \alpha_t) = (x, s\alpha_t)$ and $s\alpha_t$ is a positive root and hence $0 < (sx, \alpha_t) < 1$, which means that $A_\circ$ and $sA_\circ$ are on the same side of $H_{\alpha,k}$ for any $\alpha \neq \alpha_s$. Since $(xs, \alpha_s) = -(x, \alpha_s)$ and $0 < (x, \alpha_s) < 1$, we see that $H_{\alpha_s,k}$ only separates if $k = 0$. □

Lemma 4.15. *Pick $w \in W_a$ and $s \in S_a$.*

[1] Picking an appropriate orthonormal basis for V, we may assume that w is negation of the last coordinate. Then $\|\nu - \lambda\|^2 - \|w\nu - \lambda\|^2 = -4ab$ where a is the last coordinate of ν and b is the last coordinate of λ. Since λ and ν are on opposite sides of the hyperplane, we have $ab < 0$.

(1) H_s is in exactly one of $\mathcal{L}(w^{-1})$ and $\mathcal{L}(sw^{-1})$.
(2) $s(\mathcal{L}(w^{-1}) \setminus \{H_s\}) = \mathcal{L}(sw^{-1}) \setminus \{H_s\}$.

Proof. (1) If $x \in w^{-1}A_\circ$, then $(\alpha_s, sx) = -(\alpha_s, x)$, so $w^{-1}A_\circ$ and $sw^{-1}A_\circ$ are on different sides of H_s.

(2) Suppose $H \in \mathcal{L}(w^{-1}) \setminus \{H_s\}$. Since $sH_s = H_s$, we see that $sH \neq H_s$. We claim that $sH \in \mathcal{L}(sw^{-1})$. If not, then $sw^{-1}A_\circ$ and $A_\circ$ are on the same side as sH, which implies that $w^{-1}A_\circ$ and $sA_\circ$ are on the same side as H. Since H separates $w^{-1}A_\circ$ and $A_\circ$, we see that it also separates $sA_\circ$ and $A_\circ$. But this forces $H = H_s$ by Lemma 4.15, which is a contradiction, so our claim is proven.

This shows that $s(\mathcal{L}(w^{-1}) \setminus \{H_s\}) \subseteq \mathcal{L}(sw^{-1}) \setminus \{H_s\}$. Replacing w by sw^{-1} shows that $s(\mathcal{L}(sw^{-1}) \setminus \{H_s\}) \subseteq \mathcal{L}(w^{-1}) \setminus \{H_s\}$. Applying s to both sides gives the reverse inclusion that we want. ∎

Proposition 4.16. *Let $w = s_{i_1} \cdots s_{i_r}$ be a reduced expression for $w \in W_a$.*

(1) The hyperplanes

$$H_{s_{i_1}}, s_{i_1}H_{s_{i_2}}, \dots, s_{i_1} \cdots s_{i_{r-1}} H_{s_{i_r}}$$

are all distinct.

(2) $\mathcal{L}(w) = \{H_{s_{i_1}}, s_{i_1}H_{s_{i_2}}, \dots, s_{i_1} \cdots s_{i_{r-1}} H_{s_{i_r}}\}$.

Proof. Suppose that (1) is false, so that there exist $p < q$ such that $s_{i_1} \cdots s_{i_{p-1}} H_{s_{i_p}} = s_{i_1} \cdots s_{i_{q-1}} H_{s_{i_q}}$, which implies that $H_{s_{i_p}} = s_{i_p} \cdots s_{i_{q-1}} H_{s_{i_q}}$. By Corollary 4.9, there is a unique $x \in L^\vee$ and $u \in W$ such that $s_{i_p} \cdots s_{i_{q-1}} = t_x u$. So by Lemma 4.5, we see that $(s_{i_p} \cdots s_{i_{q-1}}) s_{i_q} (s_{i_p} \cdots s_{i_{q-1}})^{-1} = s_{i_p}$. In particular, $s_{i_p} \cdots s_{i_{q-1}} s_{i_q} = s_{i_{p+1}} \cdots s_{i_{q-1}}$, which can be used to shorten the reduced expression for w, and hence is a contradiction.

Now we prove (2) by induction on $\ell(w)$. The base case $\ell(w) = 0$ consists of showing that $\mathcal{L}(1) = \varnothing$ which is clear. So assume $\ell(w) > 0$. By induction,

$$\mathcal{L}(s_{i_1} w) = \{H_{s_{i_2}}, s_{i_2} H_{s_{i_3}}, \dots, s_{i_2} \cdots s_{i_{r-1}} H_{s_{i_r}}\}$$

and by (1), the set has size $r - 1$. In particular, $\{s_{i_1} H_{s_{i_2}}, s_{i_1} s_{i_2} H_{s_{i_3}}, \dots, s_{i_1} s_{i_2} \cdots s_{i_{r-1}} H_{s_{i_r}}\}$ also has size $r-1$, and does not contain $H_{s_{i_1}}$ by (1). In particular, $\mathcal{L}(s_{i_1} w)$ does not contain $s_{i_1} H_{s_{i_1}} = H_{s_{i_1}}$. By Lemma 4.15, we have $H_{s_{i_1}} \in \mathcal{L}(w)$ and $s_{i_1}(\mathcal{L}(w) \setminus \{H_{s_{i_1}}\}) = \mathcal{L}(s_{i_1} w)$, which proves our claim. ∎

Corollary 4.17. *For all $w \in W_a$, we have $\ell(w) = |\mathcal{L}(w)|$.*

4.5. **Toroidal Coxeter complex.** We can modify the approach of §3.8 to get a nice factorization of the Poincaré polynomial of an affine Weyl group. As before, write $W_a \cong T_{L^\vee} \rtimes W$ for a finite Weyl group W and coroot lattice $L^\vee$.

This time, we have a triangulation of V by the W_a-translates of A. Each $s \in S_a$ corresponds to a facet of A, namely $\{v \mid (v, \alpha_s) = 0\}$ for $s \in S$ and $\{v \mid (v, \widetilde{\alpha}) = 1\}$ for s_0. Given a subset $I \subseteq S_a$, define A_I to be the intersection of the corresponding facets. Since A is a simplex, this is nonempty if and only if $I \neq S_a$.

Rather than work with this infinite triangulation, we consider the quotient $V/L^\vee$, which is an n-dimensional torus which now has a finite triangulation. This is the **toroidal Coxeter complex**. It has an action of the finite Weyl group $W_a/T_{L^\vee} \cong W$.

The toroidal Coxeter complex gives us a chain complex $\mathbf{F}_\bullet$ that computes the (real) homology of $V/L^\vee$ where $\mathbf{F}_i$ (for $i = 0, \dots, n$) is the $\mathbf{R}$-vector space with basis given by W

translates of the A_I with $i = n - |I|$. The homology of an n-dimensional torus is given by

$$\mathrm{H}_i(V/L^\vee; \mathbf{R}) = \bigwedge^i V$$

and this identification is compatible with the action of W on $V/L^\vee$ (the basic idea is that $V/L^\vee$ is a product of n circles and we can use the Künneth formula).

For $I \subsetneq S_a$, the parabolic subgroup $(W_a)_I$ is finite. Let $W(I)$ be the image of $(W_a)_I$ under the quotient map $W_a \to W$. Since $T_{L^\vee} \cong \mathbf{Z}^n$ has no nontrivial finite order elements and $(W_a)_I$ is finite, their intersection must be trivial, so that $(W_a)_I \to W(I)$ is an isomorphism. Note that $W(I)$ is *not* a parabolic subgroup of W if $s_0 \in I$. However, the image of s_0 in W is the reflection with respect to the highest root $\widetilde{\alpha}$, so $W(I)$ is a subgroup generated by reflections in the geometric representation of W.

Proposition 4.18. *As class functions on W, we have*

$$\sum_{i=0}^n (-1)^i \mathrm{char}(\bigwedge^i V) = (-1)^n \sum_{i=0}^n (-1)^{|I|} 1 \uparrow_{W(I)}^W .$$

Proof. Since the triangulation is W-invariant, each w gives a chain map $w\colon \mathbf{F}_\bullet \to \mathbf{F}_\bullet$, and the Hopf trace formula tells us that

$$\sum_{i=0}^n (-1)^i \operatorname{Tr}(w \mid \mathrm{H}_i(V/L^\vee; \mathbf{R})) = \sum_{i=0}^n (-1)^i \operatorname{Tr}(w \mid \mathbf{F}_i).$$

Next, $\mathbf{F}_i$ is a permutation representation on the set of i-dimensional faces. In V, there is one orbit for each subset $I \subseteq S_a$ with $|I| = n - i$, and the A_I are representatives. By Theorem 1.31, the stabilizer of A_I is $(W_a)_I$. Then the stabilizer of $A_I/L^\vee$ in W is $W(I)$. So by Example 2.4, we see that

$$\operatorname{Tr}(w \mid \mathbf{F}_i) = \sum_{I,\ |I|=n-i} 1 \uparrow_{W(I)}^W (w).$$

Combining all of this gives the desired formula. ∎

Theorem 4.19 (Bott)**.** *Let $d_1, \dots, d_n$ be the degrees of the basic invariants for W. Then*

$$W_a(t) = W(t) \prod_{i=1}^n \frac{1}{1 - t^{d_i - 1}} = \frac{1}{(1-t)^n} \prod_{i=1}^n \frac{1 - t^{d_i}}{1 - t^{d_i - 1}}.$$

Proof. We have

$$\sum_{I\subsetneq S_a}(-1)^{|I|}\mathrm{H}_{A^{W(I)}}(t) = \sum_{I\subsetneq S_a}(-1)^{|I|}\sum_{d\ge 0}\langle 1\uparrow_{W(I)}^{W}, \mathrm{char}(A_d)\rangle_W t^d \qquad \text{(Frobenius reciprocity, Theorem 2.5)}$$

$$= \sum_{i=0}^{n}(-1)^{n+i}\sum_{d\ge 0}\langle \mathrm{char}(\bigwedge^i V), \mathrm{char}(A_d)\rangle_W t^d \qquad \text{(Proposition 4.18)}$$

$$= \sum_{i=0}^{n}(-1)^{n+i}\sum_{d\ge 0}\dim(\bigwedge^i V^*\otimes A_d)^W t^d \qquad \text{(Proposition 2.3)}$$

$$= (-1)^n \mathrm{H}_{(A\otimes E)^W}(t,-1) \qquad \text{(V is self-dual, Proposition 2.3)}$$

$$= (-1)^n\prod_{i=1}^{n}\frac{1-t^{d_i-1}}{1-t^{d_i}} \qquad \text{(Corollary 3.23)}$$

Next, each group $W(I)$ in the original sum is a finite generated by reflections. So from Theorem 3.28 and Corollary 3.15, we have

$$\mathrm{H}_{A^{W(I)}}(t)\cdot(1-t)^n = \frac{1}{\mathrm{H}_{A/I^{W(I)}}(t)} = \frac{1}{W(I)(t)}.$$

Finally, using (1.35), and our first derivation, we have

$$\frac{(-1)^n}{W_a(t)} = \sum_{I\subsetneq S_a}\frac{(-1)^{|I|}}{W(I)(t)} = (-1)^n(1-t)^n\prod_{i=1}^{n}\frac{1-t^{d_i-1}}{1-t^{d_i}},$$

so in particular

$$W_a(t) = \frac{1}{(1-t)^n}\prod_{i=1}^{n}\frac{1-t^{d_i}}{1-t^{d_i-1}}.$$

The equality involving $W(t)$ follows from Theorem 3.28. □

Remark 4.20. This proof is due to Steinberg [St]. □

4.6. **Affine permutation groups.** We now give combinatorial descriptions for each of the 4 infinite series of affine Weyl groups. Some of this material is taken from [BB, Chapter 8].

4.6.1. *Type* $\widetilde{\mathrm{A}}_{n-1}$. For $n\ge 2$, let $\widetilde{\mathfrak{S}}_n$ be the set of bijections $\sigma\colon \mathbf{Z}\to\mathbf{Z}$ such that $\sigma(i+n) = \sigma(i)+n$ for all $i\in\mathbf{Z}$ and that such $\sum_{i=1}^n \sigma(i) = \binom{n+1}{2}$. We represent an element in "window notation" by $[\sigma(1),\dots,\sigma(n)]$.

The first condition implies that we have an induced bijection $\sigma\colon \mathbf{Z}/n\to\mathbf{Z}/n$, so that $\sigma(1),\dots,\sigma(n)$ represent different cosets of $\mathbf{Z}/n$.

Lemma 4.21. $\widetilde{\mathfrak{S}}_n$ *is a group under composition.*

This is the **affine symmetric group**.

Proof. First, the identity belongs to $\widetilde{\mathfrak{S}}_n$ since $1+\cdots+n=\binom{n+1}{2}$. Second, suppose $\sigma,\tau\in\widetilde{\mathfrak{S}}_n$. Then $\sigma(\tau(i+n)) = \sigma(\tau(i)+n) = \sigma(\tau(i))+n$. We can find unique integers $k_1,\dots,k_n$ so that $1 \le \tau(i) + k_i n \le n$. In particular, $\{\tau(1) + k_1 n, \dots, \tau(n) + k_n n\} = \{1, \dots, n\}$, so $\sum_{i=1}^n (\tau(i) + k_i n) = \binom{n+1}{2}$, which forces $\sum_{i=1}^n k_i = 0$. Finally,

$$\binom{n+1}{2} = \sum_{i=1}^n \sigma(\tau(i) + k_i n) = \sum_{i=1}^n (\sigma(\tau(i)) + k_i n) = \sum_{i=1}^n \sigma(\tau(i)),$$

so $\sigma\tau \in \widetilde{\mathfrak{S}}_n$.

Third, consider $\tau^{-1}(i+n) - n$. If we apply τ to it, we get i, so since it is a bijection, the original expression is $\tau^{-1}(i)$. We have

$$\sum_{i=1}^n \tau^{-1}(i) = \sum_{i=1}^n (\tau^{-1}(\tau(i) + k_i n)) = \sum_{i=1}^n (\tau^{-1}(\tau(i)) + k_i n) = \binom{n+1}{2},$$

so $\tau^{-1} \in \widetilde{\mathfrak{S}}_n$. □

Let $L = \{x \in \mathbf{Z}^n \mid x_1 + \cdots + x_n = 0\}$. Then $L = L^\vee$. For $x \in L$, we let $t_x \in \widetilde{\mathfrak{S}}_n$ be defined by $t_x(i) = i + x_i n$ for all $i = 1, \dots, n$. If we extend the notation x_i to mean $x_{i'}$ where $i' \in [n]$ is the coset representative of i, then $t_x(i) = i + x_i n$ for all i and hence $t_x(i+n) = t_x(i) + n$ for all i. Then $t_x t_y = t_{x+y}$ and $t_x^{-1} = t_{-x}$. We let $T_L = \{t_x \mid x \in L\}$.

As said before, we have a surjective map $\pi\colon \widetilde{\mathfrak{S}}_n \to \mathfrak{S}_n$ by considering the induced bijection on $\mathbf{Z}/n$. An element $\sigma \in \ker \pi$ satisfies $\sigma(i) \equiv i \pmod{n}$ for all i. Then $\sigma(i) = i + x_i n$ for some integers x_i such that $x_1 + \cdots + x_n = 0$, so $\sigma = t_x$. Furthermore it is clear that $T_L \subseteq \ker \pi$ and so $\ker \pi = T_L$.

Also, we have an injective homomorphism $\mathfrak{S}_n \to \widetilde{\mathfrak{S}}_n$ where the window notation of $\sigma \in \mathfrak{S}_n$ is simply $[\sigma(1), \dots, \sigma(n)]$.

In particular, we have distinguished right coset representatives $\{T_L w \mid w \in \mathfrak{S}_n\}$ for $\widetilde{\mathfrak{S}}_n / T_L$, and so every element is of the form $t_x w$ for unique choices of $x \in L$ and $w \in \mathfrak{S}_n$. Next, $w t_x w^{-1}(i) = i + x_{w^{-1}(i)} n = i + (wx)_i n$, so $w t_x w^{-1} = t_{wx}$. Hence $t_x w t_{x'} w' = t_x t_{wx'} w w'$.

Now let W_a be the affine Coxeter group of type $\widetilde{\mathrm{A}}_{n-1}$. The coroot lattice of $\mathfrak{S}_n$ is identified with L, so by Corollary 4.9 we get the following result.

Proposition 4.22. *We have an isomorphism $W_a \to \widetilde{\mathfrak{S}}_n$ given by $t_x w \mapsto t_x w$ where $x \in L$ and $w \in \mathfrak{S}_n$.*

We can make this more explicit. Pick $w \in \widetilde{\mathfrak{S}}_n$. For $x \in \mathbf{R}^n$, we define $x_{i+kn} = x_i - k$ for any integer k. Then we define $wx \in \mathbf{R}^n$ by $(wx)_i = x_{w^{-1}(i)}$. Then the action of $w \in \mathfrak{S}_n$ is by permutations as usual and $(t_y x)_i = x_{i - y_i n} = x_i + y_i$, so T_L acts by translations as usual.

Tracing through the isomorphism, the Coxeter generators are $s_i = [1, \dots, i+1, i, \dots, n]$, the $(i, i+1)$ transposition in the copy of $\mathfrak{S}_n$, for $i = 1, \dots, n$, and $s_0 = [0, 2, \dots, n-1, n+1]$.

Proposition 4.23. *For all $w \in \widetilde{\mathfrak{S}}_n$, we have*

$$\ell(w) = \sum_{1 \le i < j \le n} \left| \left\lfloor \frac{w(j) - w(i)}{n} \right\rfloor \right|.$$

Proof. By Corollary 4.17, $\ell(w)$ is the number of hyperplanes $H_{\alpha,k}$ that separate $A_\circ$ and $wA_\circ$. The relevant hyperplanes in question are of the form $x_i - x_j = k$ for $i < j$ and $k \in \mathbf{Z}$. Note that $x \in A_\circ$ if and only if $0 < x_i - x_j < 1$ for $i < j$. Hence if $y \in wA_\circ$, the number of hyperplanes separating y from $A_\circ$ is $\sum_{i<j} |\lfloor y_i - y_j \rfloor|$. Our goal is to rewrite this expression in terms of w.

Fix a point $x \in A_\circ$ and let $y = wx$. There is a unique permutation $\sigma \in \mathfrak{S}_n$ and integers $k_1, \dots, k_n$ so that $w^{-1}(i) = \sigma(i) + k_i n$ for $i = 1, \dots, n$. Then

$$\begin{aligned} |\lfloor y_i - y_j \rfloor| &= |\lfloor x_{w^{-1}(i)} - x_{w^{-1}(j)} \rfloor| \\ &= |\lfloor x_{\sigma(i)} - x_{\sigma(j)} + k_j - k_i \rfloor| \\ &= \left| \left\lfloor \frac{\sigma(j) - \sigma(i)}{n} + k_j - k_i \right\rfloor \right| \\ &= \left| \left\lfloor \frac{w^{-1}(j) - w^{-1}(i)}{n} \right\rfloor \right|, \end{aligned}$$

where in the third equality, we use that the sign of $x_{\sigma(i)} - x_{\sigma(j)}$ is the same as the sign of $(\sigma(j) - \sigma(i))/n$, and that both have absolute value between 0 and 1. Finally, we conclude by noting that $\ell(w) = \ell(w^{-1})$. □

4.6.2. *Type* $\widetilde{\mathrm{C}}_n$. For $n \geq 2$, set $N = 2n + 1$. We let $\widetilde{\mathfrak{S}}^C_n$ be the set of bijections $\sigma \colon \mathbf{Z} \to \mathbf{Z}$ such that $\sigma(i + N) = \sigma(i) + N$ and $\sigma(-i) = -\sigma(i)$ for all i. Via composition, these are exactly the $\sigma \in \widetilde{\mathfrak{S}}_N$ that commute with the bijections $i \mapsto i + N$ and $i \mapsto -i$, so $\widetilde{\mathfrak{S}}^C_n$ is a subgroup of $\widetilde{\mathfrak{S}}_N$.

Let $L^\vee = \{(x_{-n}, \dots, x_n) \in \mathbf{Z}^N \mid x_{-i} = -x_i\} \cong \mathbf{Z}^n$. For $x \in L^\vee$, we define $t_x(i) = i + x_i N$ for $i = -n, \dots, n$. Then for general i, we have $t_x(i) = i + x_{i'} N$ where $-n \leq i' \leq n$ is the representative of i modulo N. Then $t_x t_y = t_{x+y}$ and $t_x^{-1} = t_{-x}$, and we set $T_{L^\vee} = \{t_x \mid x \in L\}$.

Let $W = W(\mathrm{B}_n)$ be the Weyl group of type B_n. We realize this as the permutations σ of $[-n, n]$ such that $\sigma(-i) = -\sigma(i)$ for all i. We have a surjective map $\pi \colon \widetilde{\mathfrak{S}}^C_n \to W(\mathrm{B}_n)$ by considering the induced bijection modulo N (and using $[-n, n]$ as representatives). As before, $\ker \pi = T_{L^\vee}$ and we can show that $\widetilde{\mathfrak{S}}^C_n \cong T_{L^\vee} \rtimes W(\mathrm{B}_n)$.

Recall that the root system of type C_n consists of the vectors $e_i \pm e_j$ and $\pm 2e_i$ for $i, j \leq n$. Hence the coroot lattice is $\mathbf{Z}^n$ with the usual action of W by signed permutation matrices. So the action of W on $L^\vee$ can be identified with the action of W on its coroot lattice. Let W_a be the affine Weyl group of type $\widetilde{\mathrm{C}}_n$; Corollary 4.9 gives the following identification.

Proposition 4.24. *We have an isomorphism* $W_a \cong \widetilde{\mathfrak{S}}^C_n$.

4.6.3. *Type* $\widetilde{\mathrm{B}}_n$. The type $\widetilde{\mathrm{B}}_n$ affine Weyl group is very closely related to type $\widetilde{\mathrm{C}}_n$ since the corresponding finite Weyl group is the same. The main difference is the coroot lattice. Recall that the roots of the B_n root system are the vectors $e_i \pm e_j$ and e_i for $i, j \leq n$. Hence the coroot lattice is $\{(x_1, \dots, x_n) \in \mathbf{Z}^n \mid x_1 + \cdots + x_n \text{ is even}\}$ which is closed under the action of signed permutation matrices.

For $\sigma \in \widetilde{\mathfrak{S}}^C_n$ and integers i, j, define

$$\sigma[i, j] = |\{a \in \mathbf{Z} \mid a \leq i,\ \sigma(a) \geq j\}|.$$

Define

$$\widetilde{\mathfrak{S}}^B_n = \{\sigma \in \widetilde{\mathfrak{S}}^C_n \mid \sigma[n, n+1] \text{ is even}\}.$$

Lemma 4.25. *For* $\sigma \in \widetilde{\mathfrak{S}}^C_n$, *we have*

$$\sigma[n, n+1] = \sum_{i=1}^{n} \left\lfloor \frac{|\sigma(i)| + n}{N} \right\rfloor.$$

Proof. More specifically,

$$\left\lfloor \frac{|\sigma(i)|+n}{N} \right\rfloor = |\{k \in \mathbf{Z}_{\geq 0} \mid \sigma(i-kN) > n\}| + |\{k \in \mathbf{Z}_{\geq 0} \mid \sigma(-i-kN) > n\}|.$$

Summing the right side over $i = 1, \dots, n$ gives $\sigma[n, n+1]$. □

If $\sigma \in W(\mathrm{B}_n)$, then $\sigma[n, n+1] = 0$, so $W(\mathrm{B}_n) \subseteq \widetilde{\mathfrak{S}}_n^B$. If $\sigma \in \widetilde{\mathfrak{S}}_n^C$ is translation by $(x_{-n}, \dots, x_n)$, then $\sigma[n, n+1] = x_1 + \cdots + x_n$, so the translation subgroup of $\widetilde{\mathfrak{S}}_n^C$ intersected with $\widetilde{\mathfrak{S}}_n^B$ is translation by the type B_n coroot lattice. As before, we use Corollary 4.9 to conclude the following result.

Proposition 4.26. *$\widetilde{\mathfrak{S}}_n^B$ is isomorphic to the type $\widetilde{\mathrm{B}}_n$ affine Weyl group.*

4.6.4. *Type $\widetilde{\mathrm{D}}_n$.* The type $\widetilde{\mathrm{D}}_n$ affine Weyl group is closely related to the type $\widetilde{\mathrm{B}}_n$ affine Weyl group. The coroot lattice is the same since the type D_n root system consists of the vectors $e_i \pm e_j$. However, the finite Weyl group changes: $W(\mathrm{D}_n)$ consists of signed permutation matrices with an even number of signs.

Define

$$\widetilde{\mathfrak{S}}_n^D = \{\sigma \in \widetilde{\mathfrak{S}}_n^B \mid \sigma[0,1] \text{ is even}\}.$$

We still maintain the notation $N = 2n+1$.

Lemma 4.27. *For $\sigma \in \widetilde{\mathfrak{S}}_n^C$, we have*

$$\sigma[0,1] = \sum_{i=1}^{n} \left\lfloor \frac{|\sigma(i)|}{N} \right\rfloor + |\{i \mid 1 \leq i \leq n,\ \sigma(i) < 0\}|.$$

Proof. More specifically, we have

$$\left\lfloor \frac{|\sigma(i)|}{N} \right\rfloor = |\{k \in \mathbf{Z}_{>0} \mid \sigma(i-kN) > 0\}| + |\{k \in \mathbf{Z}_{>0} \mid \sigma(-i-kN) > 0\}|.$$

The sum of the right hand side is $\sigma[0,1]$ minus the number of i between 1 and n such that $\sigma(i) < 0$. □

In particular, if $\sigma \in W(\mathrm{B}_n)$, then $\sigma[0,1] = |\{i \mid 1 \leq i \leq n, \quad \sigma(i) < 0\}$, so we see that $W(\mathrm{B}_n) \cap \widetilde{\mathfrak{S}}_n^D = W(\mathrm{D}_n)$. Next, consider the translation t_x element which satisfies $t_x(i) = i + x_i N$ for $i = 1, \dots, n$. Then $\sigma[0,1] = x_1 + \cdots + x_n$ which we already know is even from the condition of being in $\widetilde{\mathfrak{S}}_n^B$. As before, we use Corollary 4.9 to conclude the following result.

Proposition 4.28. *$\widetilde{\mathfrak{S}}_n^D$ is isomorphic to the type $\widetilde{\mathrm{D}}_n$ affine Weyl group.*

5. Kazhdan–Lusztig polynomials

Throughout, (W, S) is a fixed Coxeter group.

5.1. **Hecke algebras.** Let A be a commutative ring. Let $\mathcal{E}$ be the free A-module with basis $\{T_w \mid w \in W\}$.

Theorem 5.1. *Let $a, b \in A$. There is a unique associative algebra structure on $\mathcal{E}$ such that for all $s \in S$ and $w \in W$:*

$$T_s T_w = \begin{cases} T_{sw} & \text{if } \ell(sw) > \ell(w) \\ aT_w + bT_{sw} & \text{if } \ell(sw) < \ell(w) \end{cases}.$$

This algebra will be denoted $\mathcal{E}_A(a, b)$.

Uniqueness is clear: the relations say that $T_w = T_{s_1} \cdots T_{s_r}$ whenever $s_1 \cdots s_r$ is a reduced expression for w (and T_1 is the multiplicative identity). Hence, the product $T_v T_w$ for any $v, w \in W$ can be deduced from the relations above.

We remark that if $\ell(sw) > \ell(w)$, then $T_w T_s = T_{s_1} \cdots T_{s_r} T_s = T_{ws}$ since all suffixes of the reduced expression $s_1 \cdots s_r s$ are also reduced expressions.

We will sketch the proof of existence, deferring to [H1, §§7.2, 7.3] for missing details.

The key is to construct a subalgebra of $\mathrm{End}(\mathcal{E})$ which is isomorphic to $\mathcal{E}$ as an A-module and that satisfies the multiplication above. To that end, for each $s \in S$, define $\lambda_s, \rho_s \in \mathrm{End}(\mathcal{E})$ by

$$\lambda_s(T_w) = \begin{cases} T_{sw} & \text{if } \ell(sw) > \ell(w) \\ aT_w + bT_{sw} & \text{if } \ell(sw) < \ell(w) \end{cases}, \qquad \rho_s(T_w) = \begin{cases} T_{ws} & \text{if } \ell(ws) > \ell(w) \\ aT_w + bT_{ws} & \text{if } \ell(ws) < \ell(w) \end{cases}.$$

We let $\mathcal{L}$ denote the subalgebra of $\mathrm{End}(\mathcal{E})$ generated by the λ_s for $s \in S$.

Lemma 5.2. *For all $s, t \in S$, $\lambda_s \rho_t = \rho_t \lambda_s$. In particular, ρ_t commutes with $\mathcal{L}$.*

Since λ_s, ρ_t are defined piecewise, there are several cases to consider, but we omit the proof.

Lemma 5.3. *The map $\varphi\colon \mathcal{L} \to \mathcal{E}$ defined by $\varphi(\lambda) = \lambda(T_1)$ is an isomorphism of A-modules.*

Proof. The fact that φ is A-linear is clear. For $w \in W$, pick a reduced expression $w = s_1 \cdots s_r$. Then $T_w = \varphi(T_{s_1} \cdots T_{s_r})$, so φ is surjective.

Finally, suppose that $\varphi(\lambda) = 0$. We claim that $\lambda = 0$. It suffices to show that $\lambda(T_w) = 0$ for all $w \in W$, and prove this by induction on $\ell(w)$. If $\ell(w) = 0$, then $w = 1$, and then $\lambda(T_1) = \varphi(\lambda) = 0$. In general, if $\ell(w) > 0$, write $w = vs$ for $s \in S$ and $\ell(v) < \ell(w)$. Then $\lambda(T_v) = 0$ by induction, and so

$$\lambda(T_w) = \lambda(\rho_s(T_v)) = \rho_s(\lambda(T_v)) = 0. \qquad \square$$

In particular, for $w \in W$, we can define $\lambda_w = \lambda_{s_1} \cdots \lambda_{s_r}$ for any reduced expression $s_1 \cdots s_r$ of w, and this is independent of the choice of reduced expression.

Lemma 5.4. *We have*

$$\lambda_s \lambda_w = \begin{cases} \lambda_{sw} & \text{if } \ell(sw) > \ell(w) \\ a\lambda_w + b\lambda_{sw} & \text{if } \ell(sw) < \ell(w) \end{cases}.$$

Proof. Let $s_1 \cdots s_r$ be a reduced expression for w. If $\ell(sw) > \ell(w)$, then $ss_1 \cdots s_r$ is a reduced expression for sw, so $\lambda_{sw} = \lambda_s \lambda_w$.

Otherwise, we have $\ell(sw) < \ell(w)$. Using the first case, we have $\lambda_s \lambda_{sw} = \lambda_w$, and hence $\lambda_s \lambda_w = \lambda_s^2 \lambda_{sw}$.

So it suffices to show that $\lambda_s^2 = a\lambda_s + b$. We apply both sides to a basis element T_v. If $\ell(sv) > \ell(v)$, then

$$\lambda_s^2(T_v) = \lambda_s(T_{sv}) = aT_{sv} + bT_v = (a\lambda_s + b)(T_v).$$

If $\ell(sv) < \ell(v)$, then

$$\lambda_s^2(T_v) = \lambda_s(aT_v + bT_{sv}) = a\lambda_s(T_v) + bT_v. \qquad \square$$

Let $A = \mathbf{Z}[q^{\pm 1/2}]$ (this is the ring of Laurent polynomials in q with a square root of q adjoined). The **Hecke algebra** of (W, S) is $\mathcal{H} = \mathcal{E}_A(q-1, q)$.

Remark 5.5. If we specialize $q = 1$, then $\mathcal{H}$ becomes the group algebra of W. $\square$

5.2. **R-polynomials.** For $s \in S$, we have $T_s^2 = (q-1)T_s + q$ in $\mathcal{H}$, which we rewrite as $T_s(T_s + 1 - q) = q$. Hence T_s is invertible with

$$T_s^{-1} = q^{-1}(T_s - (q-1)). \tag{5.6}$$

This implies that T_w is invertible in general. Our first goal is to prove that the change of basis between the inverses and the T_v is lower-triangular with respect to the Bruhat order. For $w \in W$, we write $\varepsilon_w = (-1)^{\ell(w)}$.

Theorem 5.7. *For $x \le w$, there exist polynomials $R_{x,w}(q)$ of degree $\ell(w) - \ell(x)$ such that $R_{w,w}(q) = 1$ and*

$$T_{w^{-1}}^{-1} = \varepsilon_w q^{-\ell(w)} \sum_{x \le w} \varepsilon_x R_{x,w}(q) T_x.$$

Furthermore, these polynomials are nonzero.

Proof. We prove this by induction on $\ell(w)$, the case $\ell(w) = 0$ being obvious since $T_1^{-1} = T_1$. If $\ell(w) > 0$, write $w = sv$ for $s \in S$ and $\ell(v) = \ell(w) - 1$. Then using induction and (5.6), we have

$$\begin{aligned}
T_{w^{-1}}^{-1} &= T_s^{-1} T_{v^{-1}}^{-1} \\
&= q^{-1}(T_s - (q-1))\varepsilon_v q^{-\ell(v)} \sum_{y \le v} \varepsilon_y R_{y,v}(q) T_y \\
&= \varepsilon_v q^{-\ell(w)} \sum_{y \le v} \varepsilon_y R_{y,v}(q) T_s T_y - \varepsilon_v q^{-\ell(w)} \sum_{y \le v} \varepsilon_y (q-1) R_{y,v}(q) T_y \\
&= \varepsilon_v q^{-\ell(w)} \sum_{\substack{y \le v \\ \ell(sy) > \ell(y)}} \varepsilon_y R_{y,v}(q) T_{sy} + \varepsilon_v q^{-\ell(w)} \sum_{\substack{y \le v \\ \ell(sy) < \ell(y)}} \varepsilon_y R_{y,v}(q)((q-1)T_y + qT_{sy}) \\
&\qquad - \varepsilon_v q^{-\ell(w)} \sum_{y \le v} \varepsilon_y (q-1) R_{y,v}(q) T_y
\end{aligned}$$

We can cancel the first term in the second sum with some of the summands in the third sum:

$$\begin{aligned}
&= \varepsilon_v q^{-\ell(w)} \sum_{\substack{y \le v \\ \ell(sy) > \ell(y)}} \varepsilon_y R_{y,v}(q) T_{sy} + \varepsilon_v q^{-\ell(w)} \sum_{\substack{y \le v \\ \ell(sy) < \ell(y)}} \varepsilon_y R_{y,v}(q) q T_{sy} \\
&\qquad - \varepsilon_v q^{-\ell(w)} \sum_{\substack{y \le v \\ \ell(sy) > \ell(y)}} \varepsilon_y (q-1) R_{y,v}(q) T_y
\end{aligned}$$

In the first sum, we see T_{sy} with $sy > y$ and $y \le v$. Then there is a reduced expression for v which contains a reduced expression for y as a subword. Multiplying this on the left by s, we see that a reduced expression for sy is a subword of a reduced expression for $sv = w$, so $sy \le w$. In the second sum, we see T_{sy} with $sy \le y \le v \le w$. So all of the terms T_x above satisfy $x \le w$.

Now pick $x \le w$ and consider the coefficient of T_x. We break this up into two cases depending on whether $x > sx$ or not. In case 1, suppose $x > sx$. Then T_x appears only in the first sum with $y = sx$, so we set $R_{x,w}(q) = R_{sx,sw}(q)$ which is nonzero by induction. Then

$$\deg R_{x,w}(q) = \ell(sw) - \ell(sx) = \ell(w) - 1 - (\ell(x) - 1) = \ell(w) - \ell(x).$$

Furthermore, for $x = w$, we have $R_{w,w}(q) = R_{sw,sw}(q) = 1$.

For case 2, suppose that $x < sx$. Then T_x does not appear in the first sum, but potentially appears in the second sum with $y = sx$ and appears in the third sum with $y = x$, so we set $R_{x,w}(q) = qR_{sx,sw}(q) + (q-1)R_{x,sw}(q)$. Here we take the convention that $R_{a,b} = 0$ if $a \not\le b$. Then

$$\deg(qR_{sx,sw}(q)) = 1 + \ell(sw) - \ell(sx) = 1 + (\ell(w) - 1) - (\ell(x) + 1) = \ell(w) - \ell(x) - 1$$

while

$$\deg((q-1)R_{x,sw}(q)) = 1 + \ell(sw) - \ell(x) = \ell(w) - \ell(x),$$

so $\deg R_{x,w}(q) = \ell(w) - \ell(x)$. Since $R_{x,sw}(q)$ is nonzero by induction, the same is true for $R_{x,w}(q)$. ☐

We call the $R_{x,w}(q)$ the **R-polynomials**. The proof gives the following recursion for computing them.

Corollary 5.8. *With the convention that $R_{a,b}(q) = 0$ if $a \not\le b$, pick $s \in S$ so that $w > sw$. Then for $x \le w$, we have*

$$R_{x,w}(q) = \begin{cases} R_{sx,sw}(q) & \text{if } x > sx \\ qR_{sx,sw}(q) + (q-1)R_{x,sw}(q) & \text{if } x < sx \end{cases}.$$

Corollary 5.9. *For all $x \le w$, we have*

$$R_{x,w}(q^{-1}) = \varepsilon_x \varepsilon_w q^{\ell(x)-\ell(w)} R_{x,w}(q).$$

Proof. We check this by induction on $\ell(w)$. If $\ell(w) = 0$, then $R_{1,1}(q) = 1$, so the formula holds. In general, pick $s \in S$ such that $sw < w$. We have two cases to check.

If $x > sx$, then

$$R_{x,w}(q^{-1}) = R_{sx,sw}(q^{-1}) = \varepsilon_{sx}\varepsilon_{sw} q^{\ell(sx)-\ell(sw)} R_{sx,sw}(q) = \varepsilon_x \varepsilon_w q^{\ell(x)-\ell(w)} R_{x,w}(q).$$

Otherwise, if $x < sx$, then

$$\begin{aligned} R_{x,w}(q^{-1}) &= q^{-1} R_{sx,sw}(q^{-1}) + (q^{-1} - 1) R_{x,sw}(q^{-1}) \\ &= q^{-1} \varepsilon_{sx}\varepsilon_{sw} q^{\ell(sx)-\ell(sw)} R_{sx,sw}(q) + (q^{-1} - 1)\varepsilon_x \varepsilon_{sw} q^{\ell(x)-\ell(sw)} R_{x,sw}(q) \\ &= \varepsilon_x \varepsilon_w q^{\ell(x)-\ell(w)+1} R_{sx,sw}(q) + (q-1)\varepsilon_x\varepsilon_w q^{\ell(x)-\ell(w)} R_{x,sw}(q) \\ &= \varepsilon_x\varepsilon_w q^{\ell(x)-\ell(w)} R_{x,w}(q). \end{aligned}$$ ☐

This immediately implies the following formula.

Corollary 5.10. *For all $w \in W$, we have*

$$T_{w^{-1}}^{-1} = \sum_{x \le w} q^{-\ell(x)} R_{x,w}(q^{-1}) T_x.$$

5.3. **Kazhdan–Lusztig polynomials.** Define $\iota\colon \mathcal{H} \to \mathcal{H}$ on $\mathbf{Z}[q^{\pm 1/2}]$ by $\iota(q^{1/2}) = q^{-1/2}$ and extend it to $\mathcal{H}$ on basis elements by $\iota(T_w) = T_{w^{-1}}^{-1}$.

Lemma 5.11. *ι is a ring homomorphism, so in particular, $\iota^2 = 1$.*

Proof. It's enough to show that $\iota(T_s)\iota(T_w) = \iota(T_s T_w)$ for all $s \in S$ and $w \in W$, since then we get that $\iota(T_v)\iota(T_w) = \iota(T_v T_w)$ for all $v, w \in W$ since we have an expression $T_v = T_{s_1} \cdots T_{s_r}$.

We have two cases depending on whether $sw > w$ or not. If $sw > w$, then

$$\iota(T_s T_w) = \iota(T_{sw}) = T_{w^{-1}s}^{-1} = (T_{w^{-1}} T_s)^{-1} = T_s^{-1} T_{w^{-1}}^{-1} = \iota(T_s)\iota(T_w).$$

Otherwise, if $sw < w$, then

$$\iota(T_s T_w) = \iota((q-1)T_w + qT_{sw}) = (q^{-1}-1)T_{w^{-1}}^{-1} + q^{-1} T_{w^{-1}s}^{-1}$$

and by (5.6), we have $\iota(T_s) = q^{-1}(T_s - (q-1))$, so

$$\iota(T_s)\iota(T_w) = q^{-1} T_s T_{w^{-1}}^{-1} + (q^{-1}-1) T_{w^{-1}}^{-1}.$$

Since Bruhat order is invariant under inversion, we have $w^{-1} > w^{-1}s$, and hence

$$T_{w^{-1}s} T_s T_{w^{-1}}^{-1} = T_{w^{-1}} T_{w^{-1}}^{-1}.$$

This shows that $T_s T_{w^{-1}}^{-1} = T_{w^{-1}s}^{-1}$, so we're done showing that ι is a ring homomorphism.

For the last statement, $\iota^2(T_w) = \iota(T_{w^{-1}}^{-1}) = \iota(T_{w^{-1}})^{-1} = (T_w^{-1})^{-1} = T_w$ where in the middle we used that ι is a ring homomorphism. ☐

Corollary 5.12. *For all $x \le w$ we have*

$$\sum_{x \le y \le w} \varepsilon_x \varepsilon_y R_{x,y}(q) R_{y,w}(q) = \delta_{x,w}.$$

Proof. Applying ι to Corollary 5.10 gives

$$\begin{aligned} T_w = \iota(T_{w^{-1}}^{-1}) &= \sum_{y \le w} q^{\ell(y)} R_{y,w}(q) T_{y^{-1}}^{-1} \\ &= \sum_{y \le w} q^{\ell(y)} R_{y,w}(q) \varepsilon_y q^{-\ell(y)} \sum_{x \le y} \varepsilon_x R_{x,y}(q) T_x \\ &= \sum_{x \le y \le w} \varepsilon_x \varepsilon_y R_{x,y}(q) R_{y,w}(q) T_x. \end{aligned}$$

Now compare the coefficient of T_x between the first and last expression: in the first, it is $\delta_{x,w}$ and in the last it is the sum that we claim. ☐

Theorem 5.13. *For each $w \in W$, there exists a unique $C_w \in \mathcal{H}$ such that*

(1) $\iota(C_w) = C_w$,

(2) there exist polynomials $P_{x,w}(q)$ for $x \le w$ such that $P_{w,w}(q) = 1$ and $\deg P_{x,w}(q) \le \frac{1}{2}(\ell(w) - \ell(x) - 1)$ for $x < w$, and

$$C_w = \varepsilon_w q^{\ell(w)/2} \sum_{x \le w} \varepsilon_x q^{-\ell(x)} P_{x,w}(q^{-1}) T_x.$$

The $P_{x,w}(q)$ are the **Kazhdan–Lusztig polynomials**. We will indicate some of their applications in the next section, though they go beyond the scope of this course.

Proof. Pick $w \in W$. Consider an element of the form

$$C'_w = \varepsilon_w q^{\ell(w)/2} \sum_{y \le w} \varepsilon_y q^{-\ell(y)} P'_{y,w}(q^{-1}) T_y$$

where for the moment the P' are arbitrary polynomials. We will show that the conditions the P' need to satisfy force them to be unique and then show that they can be satisfied. We have

$$\begin{aligned} \iota(C'_w) &= \varepsilon_w q^{-\ell(w)/2} \sum_{y \le w} \varepsilon_y q^{\ell(y)} P'_{y,w}(q) \sum_{x \le y} \varepsilon_y q^{-\ell(y)} \varepsilon_x R_{x,y}(q) T_x \\ &= \varepsilon_w q^{-\ell(w)/2} \sum_{y \le w} \sum_{x \le y} \varepsilon_x P'_{y,w}(q) R_{x,y}(q) T_x \end{aligned}$$

For $x \le w$, the coefficient of T_x in $\varepsilon_w \varepsilon_x q^{\ell(x)/2}(C'_w - \iota(C'_w))$ is

$$q^{\ell(w)/2-\ell(x)/2} P'_{x,w}(q^{-1}) - q^{-\ell(w)/2+\ell(x)/2} P'_{x,w}(q) - \sum_{x<y\le w} q^{-\ell(w)/2+\ell(x)/2} P'_{y,w}(q) R_{x,y}(q). \tag{5.14}$$

The theorem is equivalent to saying that there exist unique polynomials $P'_{x,w}(q)$ of degree $\le \frac{1}{2}(\ell(w) - \ell(x) - 1)$ and such that $P'_{w,w}(q) = 1$ that make all of the expressions (5.14) equal to 0. We prove that we can choose such polynomials that make (5.14) by induction on $\ell(w) - \ell(x)$. For the base case, we need to set $P'_{w,w}(q) = 1$.

Now pick $x < w$. With the degree assumption, the first quantity is a polynomial in $q^{1/2}$ while the second quantity is a polynomial in $q^{-1/2}$. In particular, no terms get cancelled, so at most one solution $P'_{x,w}(q)$ exists.

Hence we are forced to set the coefficients of $P'_{x,w}(q)$ so that the first term matches up with the sum of the positive powers of $q^{1/2}$ in the sum. So a solution exists if and only if the sum of the negative powers of $q^{1/2}$ in the sum agree with the second term. We see that this is equivalent to the condition $\iota(\alpha) = -\alpha$, where α is the sum in the above expression. So our final task is to verify this (in the second equality, we use Corollary 5.9 and (5.14) with z in place of x, and in the fourth equality, we use Corollary 5.12):

$$\begin{aligned} \iota(\alpha) &= \sum_{x<z\le w} R_{x,z}(q^{-1}) q^{\ell(w)/2-\ell(x)/2} P'_{z,w}(q^{-1}) \\ &= \sum_{x<z\le w} \varepsilon_x \varepsilon_z q^{\ell(x)-\ell(z)} R_{x,z}(q) q^{\ell(z)/2-\ell(x)/2} \sum_{z\le y\le w} q^{-\ell(w)/2+\ell(z)/2} P'_{y,w}(q) R_{z,y}(q) \\ &= \sum_{x<y\le w} q^{\ell(x)/2-\ell(w)/2} P'_{y,w}(q) \sum_{x<z\le y} \varepsilon_x \varepsilon_z R_{x,z}(q) R_{z,y}(q) \\ &= \sum_{x<y\le w} q^{\ell(x)/2-\ell(w)/2} P'_{y,w}(q)(-R_{x,y}(q)) = -\alpha. \end{aligned}$$

□

5.4. **Further remarks.** The applications of Kazhdan–Lusztig polynomials goes beyond the scope of the course, but we list a few references:

- For representations of Hecke algebras (and at $q = 1$, the corresponding Coxeter groups), see [BB, Chapter 6] for an introduction.

- The finite Weyl groups play a prominent role in the representation theory of semisimple Lie algebras. See [H2, Chapter 8] for the role of Kazhdan–Lusztig polynomials in describing characters of irreducible representations.
- See [H2, Chapter 8] also for a connection between Kazhdan–Lusztig polynomials and Schubert varieties. In fact, this was one of the early proofs that they have non-negative coefficients in the case when the Coxeter group is the Weyl group of a semisimple Lie algebra (or more generally, a Kac–Moody algebra). The general situation was resolved in [EW].

Some other things:

- I believe that it is still an open problem (interval conjecture) to determine if $P_{x,y}(q)$ is a combinatorial invariant of the interval $[x, y]$ in Bruhat order. That is, $[x, y]$ carries a poset structure and if this is isomorphic to $[x', y']$ for some other elements in a Bruhat order, does this force $P_{x,y}(q) = P_{x',y'}(q)$?
- There are also parabolic analogues of Kazhdan–Lusztig polynomials which are indexed by elements of W^P rather than W. These are well-suited for representation theory problems as above in relative settings (parabolic category $\mathcal{O}$ or partial flag varieties, etc.) The case where P comes from a Hermitian symmetric space is particularly well understood, see [EHP].

References

[BB] Anders Björner, Francesco Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Math. **231**, Springer, New York, 2005.

[D] Igor V. Dolgachev, Reflection groups in algebraic geometry, *Bull. Amer. Math. Soc.* **45** (2008), 1–60.

[EW] Ben Elias, Geordie Williamson, The Hodge theory of Soergel bimodules, *Ann. of Math. (2)* **180** (2014), no. 3, 1089–1136, arXiv:1212.0791v2.

[EHP] Thomas J. Enright, Markus Hunziker, W. Andrew Pruett, Diagrams of Hermitian type, highest weight modules, and syzygies of determinantal varieties, *Symmetry: Representation Theory and Its Applications*, Progress in Mathematics **257**, Springer (2014).

[H1] James Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Mathematics **29**, Cambridge University Press, Cambridge, 1990.

[H2] James Humphreys, *Representations of Semisimple Lie Algebras in the BGG Category* $\mathcal{O}$, Graduate Studies in Mathematic **94**, American Mathematical Society, Providence, 2008.

[L] Serge Lang, *Algebra*, 3rd edition, Graduate Texts in Math. **211**, Springer-Verlag, 2002.

[ST] G. C. Shephard, J. A. Todd, Finite unitary reflection groups, *Can. J. Math.* **6** (1954), 274–304.

[So] Louis Solomon, Invariants of finite reflection groups, *Nagoya Math. J.* **22** (1963), 57–64.

[St] Robert Steinberg, *Endomorphisms of linear algebraic groups*, Memoirs of the AMS **80**, American Mathematical Society, Providence, 1968.