Chapter 5

Coxeter groups

Motivated by the examples of finite reflection groups (Chapter 1) and affine Weyl groups (Chapter 4), we embark on the general study of Coxeter groups. After introducing the basic notions in 5.1–5.3, we examine the 'root system' in 5.4–5.7, following Deodhar [4]. This leads to the 'Strong Exchange Condition' (5.8). Then we study the Bruhat ordering in 5.9–5.11. Finally, we look more closely at parabolic subgroups, deriving an inductive formula to express Poincaré series as rational functions in 5.12 and finding a fundamental domain for the action of our group in 5.13.

5.1 Coxeter systems

We define a **Coxeter system** to be a pair (W, S) consisting of a group W and a set of generators $S \subset W$, subject only to relations of the form

$$(ss')^{m(s,s')} = 1,$$

where $m(s,s) = 1$, $m(s,s') = m(s',s) \geq 2$ for $s \neq s'$ in S. In case no relation occurs for a pair s, s', we make the convention that $m(s,s') = \infty$. Formally, W is the quotient F/N, where F is a free group on the set S and N is the normal subgroup generated by all elements

$$(ss')^{m(s,s')}.$$

Call $|S|$ the **rank** of (W,S). The canonical image of S in W is a generating set which might conceivably be smaller than S, but in fact it will soon turn out to be in bijection with S (5.3). In the meantime, we may allow ourselves to write $s \in W$ for the image of $s \in S$, whenever this creates no real ambiguity in the arguments. Moreover, we may refer to W itself as a **Coxeter group**, when the presentation is understood.

Although a good part of the theory goes through for arbitrary S, we shall always assume that S is finite.

This definition is of course motivated by the two geometric examples studied earlier: finite groups generated by reflections (Chapter 1) and affine Weyl groups (Chapter 4). However, the subject becomes vastly more general when the choices of the $m(s, s')$ are essentially unrestricted. As a result, the reader may well be skeptical at this point about the depth or interest of such a generalization. It will be seen presently that Coxeter groups do admit a sort of geometric interpretation as groups generated by 'reflections' (in a weak sense), and that they share many interesting features. The special cases just mentioned are the ones most often encountered in applications, but there are further useful classes of Coxeter groups (e.g., the 'hyperbolic' ones, and the 'Weyl groups' associated with Kac–Moody Lie algebras). While the general theory may be regarded at first as mainly a nice unification of existing theories, it also suggests new viewpoints and problems.

To specify a Coxeter system (W, S) is to specify a finite set S and a symmetric matrix M indexed by S, with entries in $\mathbf{Z} \cup \{\infty\}$ subject to the conditions: $m(s, s) = 1$, $m(s, s') \geq 2$ if $s \neq s'$. Equivalently, one can draw an undirected graph Γ with S as vertex set, joining vertices s and s' by an edge labelled $m(s, s')$ whenever this number (∞ allowed) is at least 3. If distinct vertices s and s' are not joined, it is then understood that $m(s, s') = 2$. As a simplifying convention, the label $m(s, s') = 3$ may be omitted. As in 2.1, Γ is called a **Coxeter graph**.

Here are a couple of examples not previously encountered.

Example 1. In case all $m(s, s') = \infty$ when $s \neq s'$, we call W a **universal** Coxeter group (see Dyer [2]). If $|S| = 2$, W is just the infinite dihedral group $\mathcal{D}_\infty$, an affine Weyl group of type $\widetilde{A}_1$.

Example 2. Let $S = \{s_1, s_2, s_3\}$, with $m(s_1, s_2) = 3$, $m(s_1, s_3) = 2$, $m(s_2, s_3) = \infty$, so the Coxeter graph is

$$\circ\!-\!\!\!-\!\circ \overset{\infty}{-\!\!\!-} \circ$$

The resulting Coxeter group W turns out to be isomorphic to $\mathrm{PGL}(2, \mathbf{Z}) = \mathrm{GL}(2, \mathbf{Z})/\{\pm 1\}$. Denote the canonical map $\mathrm{GL}(2, \mathbf{Z}) \to \mathrm{PGL}(2, \mathbf{Z})$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then send the generators s_1, s_2, s_3 to the respective elements of order 2 in $\mathrm{PGL}(2, \mathbf{Z})$:

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

By checking the orders of the products, we see that this assignment induces a homomorphism $\varphi : W \to \mathrm{PGL}(2, \mathbf{Z})$. The image of φ includes the subgroup $\mathrm{PSL}(2, \mathbf{Z})$ of index 2, since $\varphi(s_1 s_3)$ and $\varphi(s_2 s_3)$ respectively come from elementary matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

which are well known to generate $\mathrm{SL}(2, \mathbf{Z})$. Because $\mathrm{PSL}(2, \mathbf{Z})$ does not contain the images of matrices of determinant -1 representing the s_i, we conclude that φ is surjective. To see that φ is injective, one can use the standard fact that $\mathrm{PSL}(2, \mathbf{Z})$ is the free product of the groups of orders 2 and 3 generated by $\varphi(s_1 s_3)$ and $\varphi(s_1 s_2)$. (W is an example of a 'hyperbolic' Coxeter group; see 6.8 below. It is discussed from several perspectives in Brown [1], pp. 40–46.)

It is notoriously difficult to say much about a group given only by generators and relations — for example, is the group trivial or not? In our case, we can see right away that W has order at least 2. Start with a homomorphism from the free group F onto the multiplicative group $\{1, -1\}$, defined by sending each element of S to -1. It is obvious that all elements $(ss')^{m(s,s')}$ lie in the kernel, so there is an induced epimorphism $\varepsilon : W \to \{1, -1\}$ sending the image of each $s \in S$ to -1. In particular, each of these generators of W does have order 2. The map ε is the generalization for an arbitrary Coxeter group of the sign character of the symmetric group.

Proposition *There is a unique epimorphism $\varepsilon : W \to \{1, -1\}$ sending each generator $s \in S$ to -1. In particular, each s has order 2 in W.* □

Note that when $|S| = 1$, W is just a group of order 2. When $|S| = 2$, W is dihedral, of order $2m(s, s') \leq \infty$ if $S = \{s, s'\}$. So we are already well acquainted with these types of Coxeter groups in the guise of reflection groups.

Exercise 1. Denote the kernel of ε by W^+. If $S = \{s_1, \ldots, s_n\}$, prove that W^+ is generated by the elements $s_i s_n$ $(1 \leq i \leq n-1)$.

Exercise 2. If W has rank n and all $m(s, s')$, $s \neq s'$, are even, then $|W| \geq 2^n$.

5.2 Length function

Since the generators $s \in S$ have order 2 in W, each $w \neq 1$ in W can be written in the form $w = s_1 s_2 \cdots s_r$ for some s_i (not necessarily distinct) in S. If r is as small as possible, call it the **length** of w, written $\ell(w)$, and call any expression of w as a product of r elements of S a **reduced expression**. By convention, $\ell(1) = 0$. More formally, a reduced expression should be viewed as an ordered r-tuple $(s_1, \ldots, s_r)$. Note that the lengths of partial products are predictable when $w = s_1 \cdots s_r$ is reduced: $\ell(s_1 \cdots s_{r-1}) = r - 1$, $\ell(s_2 \cdots s_{r-1}) = r - 2$, etc. However, the length function has its subtleties, because a typical element of W may have numerous reduced expressions.

Exercise. Prove that W is of 'universal' type (5.1) if and only if each element has a unique reduced expression.

Here are some elementary properties of the length function:

(L1) $\ell(w) = \ell(w^{-1})$. [If $w = s_1 \cdots s_r$, $w^{-1} = s_r \cdots s_1$, so $\ell(w^{-1}) \leq \ell(w)$, and similarly for w^{-1} in place of w.]

(L2) $\ell(w) = 1$ if and only if $w \in S$.

(L3) $\ell(ww') \leq \ell(w) + \ell(w')$. [If $w = s_1 \cdots s_p$ and $w' = s'_1 \cdots s'_q$, then the product $ww' = s_1 \cdots s_p s'_1 \cdots s'_q$ has length at most $p + q$.]

(L4) $\ell(ww') \geq \ell(w) - \ell(w')$. [Apply (L3) to the pair $ww', (w')^{-1}$, then use (L1).]

(L5) $\ell(w) - 1 \leq \ell(ws) \leq \ell(w) + 1$, for $s \in S$ and $w \in W$. [Use (L3) and (L4).]

Proposition *The homomorphism $\varepsilon : W \to \{1, -1\}$ of 5.1 is given by $\varepsilon(w) = (-1)^{\ell(w)}$. As a result, $\ell(ws) = \ell(w) \pm 1$, for all $s \in S, w \in W$, and similarly for $\ell(sw)$.*

Proof. Write a reduced expression $w = s_1 \cdots s_r$. Then

$$\varepsilon(w) = \varepsilon(s_1) \cdots \varepsilon(s_r) = (-1)^r = (-1)^{\ell(w)},$$

as required. Now $\varepsilon(ws) = -\varepsilon(w)$ implies that $\ell(ws) \neq \ell(w)$. By property (L5) above, the lengths must differ by precisely 1. □

In our study of Coxeter groups (as in the special cases treated earlier), we shall often prove theorems by induction on $\ell(w)$. It will therefore be essential to understand the precise relationship between $\ell(w)$ and $\ell(ws)$ (or $\ell(sw)$). For this we need a way to represent W concretely.

5.3 Geometric representation of W

Given a Coxeter system (W, S), it is too much to expect a faithful representation of W as a group generated by (orthogonal) reflections in a euclidean space. But we can get a reasonable substitute if we redefine a **reflection** to be merely a linear transformation which fixes a hyperplane pointwise and sends some nonzero vector to its negative. The idea is to begin with a vector space V over $\mathbf{R}$, having a basis $\{\alpha_s | s \in S\}$ in one-to-one correspondence with S, and then to impose a geometry on V in such a way that the 'angle' between α_s and $\alpha_{s'}$ will be compatible with the given $m(s, s')$, i.e., with the previously studied geometry of dihedral groups. Accordingly, we define a symmetric bilinear form B on V by requiring:

$$B(\alpha_s, \alpha_{s'}) = -\cos \frac{\pi}{m(s, s')}.$$

(This expression is interpreted to be -1 in case $m(s, s') = \infty$.) Evidently $B(\alpha_s, \alpha_s) = 1$, while $B(\alpha_s, \alpha_{s'}) \leq 0$ if $s \neq s'$. Since α_s is non-isotropic, the subspace H_s orthogonal to α_s relative to B is complementary to the line $\mathbf{R}\alpha_s$.

For each $s \in S$ we can now define a reflection $\sigma_s : V \to V$ by the rule:

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

Clearly $\sigma_s \alpha_s = -\alpha_s$, while σ_s fixes H_s pointwise. In particular, we see that σ_s has order 2 in $\mathrm{GL}(V)$.

A quick calculation (left to the reader) shows that σ_s preserves the form B, i.e., $B(\sigma_s\lambda, \sigma_s\mu) = B(\lambda, \mu)$ for all $\lambda, \mu \in V$. As a result, each element of the subgroup of $\mathrm{GL}(V)$ generated by the $\sigma_s (s \in S)$ will also preserve B.

Our first task is to show that there exists a homomorphism from W onto this linear group, sending s to σ_s. For this it is enough to check that

$$(\sigma_s \sigma_{s'})^{m(s,s')} = 1 \text{ whenever } s \neq s'.$$

Set $m := m(s, s')$ and consider first the two-dimensional subspace $V_{s,s'} := \mathbf{R}\alpha_s \oplus \mathbf{R}\alpha_{s'}$. We claim that *the restriction of B to $V_{s,s'}$ is positive semidefinite, and moreover is nondegenerate precisely when $m < \infty$.* To check the first part, just take any $\lambda = a\alpha_s + b\alpha_{s'}\,(a, b \in \mathbf{R})$, and compute

$$B(\lambda, \lambda) = a^2 - 2ab\cos(\pi/m) + b^2 = (a - b\cos(\pi/m))^2 + b^2\sin^2(\pi/m) \geq 0.$$

In turn, the form is positive definite on $V_{s,s'}$ if $\sin\,(\pi/m) \neq 0$, i.e., $m < \infty$ (whereas otherwise the nonzero vector $\alpha_s + \alpha_{s'}$ is isotropic).

Having seen precisely how the form B behaves on $V_{s,s'}$, we note further that σ_s and $\sigma_{s'}$ leave $V_{s,s'}$ stable: just look at the defining formula for each reflection. So it makes sense to calculate the order of $\sigma_s\sigma_{s'}$ viewed as an operator on $V_{s,s'}$. Two cases are possible:

(a) $m < \infty$. Here the form is positive definite, so we find ourselves in the familiar situation of the euclidean plane. Both σ_s and $\sigma_{s'}$ act as orthogonal reflections. Since $B(\alpha_s, \alpha_{s'}) = -\cos(\pi/m) = \cos(\pi - (\pi/m))$, the angle between the rays $\mathbf{R}^+\alpha_s$ and $\mathbf{R}^+\alpha_{s'}$ is $\pi - (\pi/m)$, forcing the angle between the two reflecting lines to be π/m. From our previous study of dihedral groups (1.1), we recognize $\sigma_s\sigma_{s'}$ as a rotation through the angle $2\pi/m$; it therefore has order m.

(b) $m = \infty$. Here $B(\alpha_s, \alpha_{s'}) = -1$. If $\lambda = \alpha_s + \alpha_{s'}$, $B(\lambda, \alpha_s) = 0 = B(\lambda, \alpha_{s'})$, so that both σ_s and $\sigma_{s'}$ fix λ. In turn, $\sigma_s\sigma_{s'}\alpha_s = \sigma_s(\alpha_s + 2\alpha_{s'}) = 3\alpha_s + 2\alpha_{s'} = 2\lambda + \alpha_s$, and by iteration, $(\sigma_s\sigma_{s'})^k\alpha_s = 2k\lambda + \alpha_s$ $(k \in \mathbf{Z})$. This implies that $\sigma_s\sigma_{s'}$ has infinite order on $V_{s,s'}$ (and therefore also on V).

In case (a), the fact that B is nondegenerate on $V_{s,s'}$ implies that V is the orthogonal direct sum of $V_{s,s'}$ and its orthogonal complement; evidently both σ_s and $\sigma_{s'}$ fix the latter subspace pointwise. Thus $\sigma_s\sigma_{s'}$ also has order m on V. To summarize:

Proposition *There is a unique homomorphism $\sigma : W \to \mathrm{GL}(V)$ sending s to σ_s, and the group $\sigma(W)$ preserves the form B on V. Moreover, for each pair $s, s' \in S$, the order of ss' in W is precisely $m(s, s')$.* □

This last observation removes any possible ambiguity in the status of the generators $s \in S$: if $s \neq s'$ in the subset S of the free group F, then also $s \neq s'$ in W, as promised in 5.1, and the subgroup of W generated by s, s' is dihedral of order $2m(s, s')$. Now we know that W is not 'too small'. It remains to be seen that W is not 'too big', i.e., that σ has trivial kernel (Corollary 5.4 below). This will require a closer study of the action on V.

For convenience we shall refer to the homomorphism σ as the **geometric representation** of W. (However, it should be emphasized that there may be other interesting ways to represent W as a group generated by 'reflections', e.g., acting in a hyperbolic space. See Vinberg [1]–[5].)

Question. If W is an affine Weyl group, how does the geometric representation compare with the action on euclidean space described in Chapter 4? (This will be discussed in 6.5.)

Exercise. Prove that $s, s' \in S$ are conjugate in W if and only if the following condition is satisfied: (*) There are elements $s = s_1, s_2, \ldots, s_k = s'$ in S for which every s_is_{i+1} has (finite) odd order.

($\Leftarrow$) In case $w = ss'$ itself has odd order $2p+1$, note that $w^psw^{-p} = s'$. Iterate!

($\Rightarrow$) Fix $s \in S$, and consider the set S' of all s' satisfying (*). It must be shown that no element of $S'' := S \setminus S'$ is conjugate to s. Define $f : S \to \{1, -1\}$ by $f(S') = 1$, $f(S'') = -1$. Show that f induces a homomorphism from W to $\{1, -1\}$. Then all conjugates of s must lie in $\mathrm{Ker}\, f$.

5.4 Positive and negative roots

In this section we obtain a precise criterion for $\ell(ws)$ to be greater or smaller than $\ell(w)$, in terms of the action of W on V. This will be the key to all further combinatorial properties of W relative to the generating set S. To avoid cumbersome notation, we may write $w(\alpha_s)$ in place of $\sigma(w)(\alpha_s)$.

First we introduce the **root system** Φ of W, consisting of a set of unit vectors in V permuted by W. Define Φ to be the collection of all vectors $w(\alpha_s)$, where $w \in W$ and $s \in S$. These are unit vectors, because W preserves the form B on V. Note that $\Phi = -\Phi$, since $s(\alpha_s) = -\alpha_s$. If α is any root, we can write it uniquely in the form

$$\alpha = \sum_{s \in S} c_s \alpha_s \ (c_s \in \mathbf{R}).$$

Call α **positive** (resp. **negative**) and write $\alpha > 0$ (resp. $\alpha < 0$) if all $c_s \geq 0$ (resp. all $c_s \leq 0$). For example, each α_s is positive. Write Φ^+ and Φ^- for the respective sets of positive and negative roots. It will be an immediate consequence of the theorem below that these sets exhaust Φ.

Note that, in contrast to the situation in Chapter 1, we have in effect specified once and for all a set of 'simple' roots.

We also have to introduce at this point the **parabolic subgroup** W_I of W, defined as in 1.10 to be the subgroup generated by a given subset $I \subset S$. (More generally, we refer to any conjugate of such a subgroup as a parabolic subgroup.) In the following section, W_I will be seen to be a Coxeter group in its own right. For the present, we just note that it has a length function ℓ_I relative to the generating set of involutions I. It is clear that $\ell(w) \leq \ell_I(w)$ for all $w \in W_I$. (It will be seen in 5.5 that equality holds.)

Theorem *Let $w \in W$ and $s \in S$. If $\ell(ws) > \ell(w)$, then $w(\alpha_s) > 0$. If $\ell(ws) < \ell(w)$, then $w(\alpha_s) < 0$.*

Proof. Observe that the second statement follows from the first, applied to ws in place of w: indeed, if $\ell(ws) < \ell(w)$, then $\ell((ws)s) > \ell(ws)$, forcing $ws(\alpha_s) > 0$, i.e., $w(-\alpha_s) > 0$, or $w(\alpha_s) < 0$.

To prove the first statement, we proceed by induction on $\ell(w)$. In case $\ell(w) = 0$, we have $w = 1$, and there is nothing to prove. If $\ell(w) > 0$, we can find an $s' \in S$ for which $\ell(ws') = \ell(w) - 1$, say by choosing s' to be the last factor in a reduced expression for w. Since $\ell(ws) > \ell(w)$ by assumption, we see that $s \neq s'$. Set $I := \{s, s'\}$, so that W_I is dihedral. Now we make a crucial choice within the coset wW_I. Consider the set

$$A := \{v \in W | v^{-1}w \in W_I \text{ and } \ell(v) + \ell_I(v^{-1}w) = \ell(w)\}.$$

Evidently $w \in A$. Choose $v \in A$ for which $\ell(v)$ is as small as possible, and write $v_I := v^{-1}w \in W_I$. Thus $w = vv_I$, with $\ell(w) = \ell(v) + \ell_I(v_I)$. The strategy now is to analyze how each of v and v_I acts on roots.

Observe that $ws' \in A$: Indeed, $(s'w^{-1})w = s'$ lies in W_I, while $\ell(ws') + \ell_I(s') = (\ell(w) - 1) + 1 = \ell(w)$. The choice of v therefore forces $\ell(v) \leq \ell(ws') = \ell(w) - 1$. This will allow us to apply the induction hypothesis to the pair v, s. But for this we need to compare the lengths of v and vs.

Suppose it were true that $\ell(vs) < \ell(v)$, i.e., $\ell(vs) = \ell(v) - 1$. Then we could calculate as follows:

$$\begin{array}{rcll} \ell(w) & \leq & \ell(vs) + \ell((sv^{-1})w) & \text{[use (L3) from 5.2]} \\ & \leq & \ell(vs) + \ell_I(sv^{-1}w) & [\text{since } sv^{-1}w \in W_I \text{ and } \ell \leq \ell_I] \\ & = & (\ell(v) - 1) + \ell_I(sv^{-1}w) & \\ & \leq & \ell(v) - 1 + \ell_I(v^{-1}w) + 1 & \\ & = & \ell(v) + \ell_I(v^{-1}w) & \\ & = & \ell(w). & \end{array}$$

So equality holds throughout, forcing $\ell(w) = \ell(vs) + \ell_I((sv^{-1})w)$ and therefore $vs \in A$, contrary to $\ell(vs) < \ell(v)$. This contradiction shows that we must instead have $\ell(vs) > \ell(v)$. By induction, we obtain: $v(\alpha_s) > 0$. An entirely similar argument shows that $\ell(vs') > \ell(v)$, whence $v(\alpha_{s'}) > 0$.

Since $w = vv_I$, we will be done if we can show that v_I maps α_s to a nonnegative linear combination of α_s and $\alpha_{s'}$.

We claim that $\ell_I(v_Is) \geq \ell_I(v_I)$. Otherwise we would have:

$$\ell(ws) = \ell(vv^{-1}ws) \leq \ell(v) + \ell(v^{-1}ws) = \ell(v) + \ell(v_Is)$$

$$\leq \ell(v) + \ell_I(v_Is) < \ell(v) + \ell_I(v_I) = \ell(w),$$

contrary to $\ell(ws) > \ell(w)$. In turn, it follows that any reduced expression for v_I in W_I (an alternating product of factors s and s') must end in s'. Consider the two possible cases:

(a) If $m(s, s') = \infty$, an easy direct calculation shows that $v_I(\alpha_s) = a\alpha_s + b\alpha_{s'}$, with $a, b \geq 0$ and $|a - b| = 1$. Indeed, $B(\alpha_s, \alpha_{s'}) = -1$, so that $s'(\alpha_s) = \alpha_s + 2\alpha_{s'}, ss'(\alpha_s) = 2\alpha_{s'} + 3\alpha_s, s'ss'(\alpha_s) = 3\alpha_s + 4\alpha_{s'}$, and so on.

(b) If $m := m(s, s') < \infty$, notice that $\ell_I(v_I) < m$. Indeed, m is clearly the maximum possible value of ℓ_I, and an element of length m in W_I has a reduced expression ending with s. So v_I can be written as a product of fewer than $m/2$ terms ss', possibly preceded by one factor s'. Direct calculation will now show that $v_I(\alpha_s)$ is a nonnegative linear combination of α_s and $\alpha_{s'}$. (A rough sketch should make the argument transparent.) Recall that we are now working in the euclidean plane, with unit vectors α_s and $\alpha_{s'}$ at an angle of $\pi - \pi/m$, and ss' rotates

α_s through an angle of $2\pi/m$ toward $\alpha_{s'}$. So the rotations involved in v_I move α_s through at most an angle of $\pi - 2\pi/m$, still within the positive cone defined by α_s and $\alpha_{s'}$. If v_I further involves a reflection corresponding to s', the resulting vector still lies within this positive cone, because the angle between α_s and the reflecting line is $(\pi/2) - (\pi/m)$. □

Corollary *The representation $\sigma : W \to \mathrm{GL}(V)$ is faithful.*

Proof. Let $w \in \operatorname{Ker} \sigma$. If $w \neq 1$, there exists $s \in S$ for which $\ell(ws) < \ell(w)$. The theorem says that $w(\alpha_s) < 0$. But $w(\alpha_s) = \alpha_s > 0$, which is a contradiction. □

5.5 Parabolic subgroups

With Theorem 5.4 in hand, we can get more precise information about the internal structure of W. First we want to clarify (as promised) the nature of the parabolic subgroups W_I $(I \subset S)$. The set I and the corresponding values $m(s, s')$ give rise to an abstractly defined Coxeter group $\overline{W_I}$, to which our previous results apply. In particular, $\overline{W_I}$ has a geometric representation of its own. This can obviously be identified with the action of the group generated by all σ_s $(s \in I)$ on the subspace V_I of V spanned by all α_s $(s \in I)$, since the bilinear form B restricted to V_I agrees with the form B_I defined by $\overline{W_I}$. The group generated by these σ_s is just the restriction to V_I of the group $\sigma(W_I)$. On the other hand, $\overline{W_I}$ maps canonically onto W_I, yielding a commutative triangle:

$$\begin{tikzcd} \overline{W_I} \arrow[rr] \arrow[dr] & & \mathrm{GL}(V_I) \\ & W_I \arrow[ur] & \end{tikzcd}$$

Since the map $\overline{W_I} \to \mathrm{GL}(V_I)$ is injective by 5.4, we conclude that W_I is isomorphic to $\overline{W_I}$ and is therefore itself a Coxeter group.

Theorem (a) *For each subset I of S, the pair (W_I, I) with the given values $m(s, s')$ is a Coxeter system.*

(b) *Let $I \subset S$. If $w = s_1 \cdots s_r$ $(s_i \in S)$ is a reduced expression, and $w \in W_I$, then all $s_i \in I$. In particular, the function ℓ agrees with ℓ_I on W_I, and $W_I \cap S = I$.*

(c) *The assignment $I \mapsto W_I$ defines a lattice isomorphism between the collection of subsets of S and the collection of subgroups W_I of W.*

(d) *S is a minimal generating set for W.*

Proof. We have just verified (a). For (b), use induction on $\ell(w)$, noting that $\ell(1) = 0 = \ell_I(1)$. Suppose $w \neq 1$, and set $s = s_r$. According to

Theorem 5.4, $w(\alpha_s) < 0$. Since $w \in W_I$, we can also write $w = t_1 \cdots t_q$ with all $t_i \in I$. Therefore

$$w(\alpha_s) = \alpha_s + \sum_{i=1}^{q} c_i \alpha_{t_i} \ (c_i \in \mathbf{R}).$$

Because $w(\alpha_s) < 0$, we must have $s = t_i$ for some i, forcing $s \in I$. In turn, $ws = s_1 \cdots s_{r-1} \in W_I$, and the expression is reduced. By induction, all $s_i \in I$. The remaining assertions of (b) are clear.

To prove (c), suppose $I, J \subset S$. If $W_I \subset W_J$, then $I = W_I \cap S \subset W_J \cap S = J$, thanks to (b). Thus $I \subset J$ (resp. $I = J$) if and only if $W_I \subset W_J$ (resp. $W_I = W_J$). It is clear that $W_{I \cup J}$ is the subgroup of W generated by W_I and W_J. On the other hand, (b) implies that $W_{I \cap J} = W_I \cap W_J$. This yields the desired lattice isomorphism. To prove (d), suppose that a subset I of S generates W, so $W_I = W = W_S$. According to (c), $I = S$. □

Example. When the Coxeter group in question is an affine Weyl group W_a associated with a Weyl group W (Chapter 4), W itself is a parabolic subgroup of W_a: its Coxeter graph is obtained from that of W_a by removing a single vertex. In particular, the length functions of these groups are compatible.

5.6 Geometric interpretation of the length function

Our next goal is to extract from Theorem 5.4 a more precise description of the way in which W permutes Φ. Once we have this information in hand, we can explore more deeply the internal structure of W itself. Recall that Φ is the disjoint union of the sets Φ^+ and Φ^- of positive and negative roots. For brevity, write $\Pi = \Phi^+$.

Proposition (a) *If $s \in S$, then s sends α_s to its negative, but permutes the remaining positive roots.*

(b) *For any $w \in W$, $\ell(w)$ equals the number of positive roots sent by w to negative roots.*

Proof. Note that part (a) is a special case of part (b); but it is needed in the proof of (b).

(a) Suppose $\alpha > 0$, but $\alpha \neq \alpha_s$. Since all roots are unit vectors, α cannot be a multiple of α_s. We can therefore write

$$\alpha = \sum_{t \in S} c_t \alpha_t,$$

where all coefficients are nonnegative and some $c_t > 0$, $t \neq s$. Applying s to α only modifies this sum by adding some constant multiple of α_s, so the coefficient of α_t remains strictly positive. It follows that $s(\alpha)$ cannot be a negative root, so it lies in Π and is obviously distinct from α_s. Thus $s(\Pi \setminus \{\alpha_s\}) \subset \Pi \setminus \{\alpha_s\}$. Apply s to both sides to get the reverse inclusion.

(b) If $w \in W$, define $n(w)$ to be the number of positive roots sent by w to negative roots, so

$$n(w) = \text{Card}\ \Pi(w), \quad \text{where } \Pi(w) := \Pi \cap w^{-1}(-\Pi).$$

(It is not instantly obvious that $n(w)$ is finite, but this will follow from the proof that $n(w) = \ell(w)$.) Notice that part (a) implies that $n(s) = 1$ for $s \in S$.

To see that $n(w)$ behaves like the length function, we first verify that, for $s \in S, w \in W$, the condition $w(\alpha_s) > 0$ implies $n(ws) = n(w) + 1$, whereas $w(\alpha_s) < 0$ implies $n(ws) = n(w) - 1$. Indeed, if $w(\alpha_s) > 0$, part (a) implies that $\Pi(ws)$ is the disjoint union of $s(\Pi(w))$ and $\{\alpha_s\}$. Similarly, if $w(\alpha_s) < 0$, we get $\Pi(ws) = s(\Pi(w) \setminus \{\alpha_s\})$, with $\alpha_s \in \Pi(w)$.

Now we proceed by induction on $\ell(w)$ to prove that $n(w) = \ell(w)$ for all $w \in W$. This is clear if $\ell(w) = 0$, and also (by part (a)) if $\ell(w) = 1$. Theorem 5.4 says that $\ell(ws) = \ell(w) + 1$ (resp. $\ell(w) - 1$) just when $w(\alpha_s) > 0$ (resp. < 0). Combining this with the preceding paragraph and the induction hypothesis completes the proof. $\square$

As in the case of finite reflection groups, part (a) of the proposition is invoked frequently, usually as a device for recognizing that a positive root obtained in the course of an argument is none other than α_s (because s sends it to a negative root).

Exercise 1. Given a reduced expression $w = s_1 \cdots s_r$ ($s_i \in S$), set $\alpha_i := \alpha_{s_i}$ and $\beta_i := s_r s_{r-1} \cdots s_{i+1}(\alpha_i)$, interpreting β_r to be α_r. Prove that $\Pi(w)$ consists of the r distinct positive roots $\beta_1, \ldots, \beta_r$.

Exercise 2. If W is infinite, prove that the length function takes arbitrarily large values, hence that Φ is infinite. (Therefore the scalar $-1 \in \text{GL}(V)$ does not lie in $\sigma(W)$.) If W is finite, prove that there is one and only one element $w_\circ \in W$ of maximum length, and that $w_\circ$ maps Π onto $-\Pi$.

Exercise 3. Use the fact that $\ell(w) = n(w)$ to give another proof of part (b) of Theorem 5.5. [Note that for $w \in W_I$, $n(w) \geq n_I(w)$ is clear, if n_I has the obvious meaning.]

5.7 Roots and reflections

By the way $\sigma : W \to \mathrm{GL}(V)$ was defined, each $s \in S$ acts on V as a reflection. More generally, we can associate a reflection in $\mathrm{GL}(V)$ with each root $\alpha \in \Phi$, as follows. Say $\alpha = w(\alpha_s)$ for some $w \in W, s \in S$. Consider how wsw^{-1} acts on V:

$$\begin{aligned} wsw^{-1}(\lambda) &= w[w^{-1}(\lambda) - 2B(w^{-1}(\lambda), \alpha_s)\,\alpha_s] \\ &= \lambda - 2B(w^{-1}(\lambda), \alpha_s)\,w(\alpha_s) \\ &= \lambda - 2B(\lambda, w(\alpha_s))\,w(\alpha_s) \\ &= \lambda - 2B(\lambda, \alpha)\alpha. \end{aligned}$$

It follows that wsw^{-1} depends only on α, not on the choice of w and s. So we may denote it by s_α. Moreover, s_α acts on V as a reflection, sending α to $-\alpha$ and fixing pointwise the hyperplane orthogonal to α. Of course, both α and $-\alpha$ determine the same reflection $s_\alpha = s_{-\alpha}$. Denote by T the set of all reflections s_α, $\alpha \in \Phi$. Thus

$$T = \bigcup_{w \in W} wSw^{-1}.$$

In order to pass back and forth freely between roots and reflections, we should observe that the correspondence $\alpha \mapsto s_\alpha$ is bijective (for $\alpha \in \Pi$). Indeed, suppose that $s_\alpha = s_\beta$. From the above formula for s_α (with $\lambda = \beta$) we get $\beta = B(\beta, \alpha)\alpha$, forcing $\alpha = \beta$ since both are unit vectors in Π.

One other observation is useful:

Lemma *If $\alpha, \beta \in \Phi$ and $\beta = w(\alpha)$ for some $w \in W$, then $ws_\alpha w^{-1} = s_\beta$.*

Proof. This is immediate from the above formula for a reflection and the fact that B is W-invariant. □

The following proposition generalizes Theorem 5.4 to arbitrary reflections.

Proposition *Let $w \in W, \alpha \in \Pi$. Then $\ell(ws_\alpha) > \ell(w)$ if and only if $w(\alpha) > 0$.*

Proof. As in the proof of Theorem 5.4, it will be enough to verify the 'only if' part. Proceed by induction on $\ell(w)$, the case $\ell(w) = 0$ being trivial. If $\ell(w) > 0$, there exists $s \in S$ such that $\ell(sw) < \ell(w)$. Then $\ell((sw)s_\alpha) = \ell(s(ws_\alpha)) \geq \ell(ws_\alpha) - 1 > \ell(w) - 1 = \ell(sw)$. By induction, $sw(\alpha) > 0$. Suppose $w(\alpha) < 0$. The only negative root made positive by s is $-\alpha_s$ (5.6), so $w(\alpha) = -\alpha_s$. But then $sw(\alpha) = \alpha_s$ would imply $(sw)s_\alpha(sw)^{-1} = s$ (by the above lemma), whence $ws_\alpha = sw$. This contradicts $\ell(ws_\alpha) > \ell(w) > \ell(sw)$. As a result, $w(\alpha)$ must be positive. □

5.8 Strong Exchange Condition

We are now able to prove a key fact about the nature of reduced expressions in W, which is at the heart of what it means to be a Coxeter group.

Theorem *Let $w = s_1 \cdots s_r$ ($s_i \in S$), not necessarily a reduced expression. Suppose a reflection $t \in T$ satisfies $\ell(wt) < \ell(w)$. Then there is an index i for which $wt = s_1 \cdots \widehat{s_i} \cdots s_r$ (omitting s_i). If the expression for w is reduced, then i is unique.*

Proof. Write $t = s_\alpha$ (say $\alpha > 0$). Since $\ell(wt) < \ell(w)$, Proposition 5.7 forces $w(\alpha) < 0$. Because $\alpha > 0$, there exists an index $i \leq r$ for which $s_{i+1} \cdots s_r(\alpha) > 0$ but $s_i s_{i+1} \cdots s_r(\alpha) < 0$. According to part (a) of Proposition 5.6, the only positive root which s_i sends to a negative root is α_{s_i}, so $s_{i+1} \cdots s_r(\alpha) = \alpha_{s_i}$. Now Lemma 5.7 implies $(s_{i+1} \cdots s_r)t(s_r \cdots s_{i+1}) = s_i$, or $wt = s_1 \cdots \widehat{s_i} \cdots s_r$ as required.

In case $\ell(w) = r$, consider what would happen if there were distinct indices $i < j$ such that $wt = s_1 \cdots \widehat{s_i} \cdots s_j \cdots s_r = s_1 \cdots s_i \cdots \widehat{s_j} \cdots s_r$. After cancelling, this gives $s_{i+1} \cdots s_j = s_i \cdots s_{j-1}$, or $s_i \cdots s_j = s_{i+1} \cdots s_{j-1}$, allowing us to write $w = s_1 \cdots \widehat{s_i} \cdots \widehat{s_j} \cdots s_r$. This contradicts the assumption that $\ell(w) = r$. □

Exercise 1. Prove a version of the theorem in which the hypothesis reads: $\ell(tw) < \ell(w)$.

We shall refer to the main assertion of the theorem as the **Strong Exchange Condition**. If t is required to lie in S, the resulting weaker statement is called the **Exchange Condition**, generalizing what we proved in the case of finite reflection groups (1.7) and affine Weyl groups (4.6):

Corollary (a) *Suppose $w = s_1 \cdots s_r$ ($s_i \in S$), with $\ell(w) < r$. Then there exist indices $i < j$ for which $w = s_1 \cdots \widehat{s_i} \cdots \widehat{s_j} \cdots s_r$. (This is called the* **Deletion Condition**.)

(b) *If $w = s_1 \cdots s_r$ ($s_i \in S$), then a reduced expression for w may be obtained by omitting certain s_i (an even number, in fact).*

Proof. (a) The hypothesis implies that there exists an index j for which $\ell(w's_j) < \ell(w')$, where $w' := s_1 \cdots s_{j-1}$. Applying the Exchange Condition to the pair w', s_j, we get $w's_j = s_1 \cdots \widehat{s_i} \cdots s_{j-1}$, or $w = s_1 \cdots \widehat{s_i} \cdots \widehat{s_j} \cdots s_r$.

(b) This follows inductively from (a). □

This brings us full circle: recall that the proof in 1.9 shows that any group generated by a set S of involutions and satisfying the Deletion Condition must be a Coxeter group. The theory developed so far in this chapter should, in principle, allow us to answer any reasonable question about Coxeter groups. In practice, some ingenuity is often required. For example, it turns out to be true that the subset of S involved in writing a reduced expression for an element $w \in W$ is independent of the particular reduced expression chosen. A related fact is the equality $W_I \cap W_J = W_{I \cap J}$. The reader might think about how to prove these using the Exchange Condition (see 5.10 below for a less direct approach).

Exercise 2. Let $I \subset S$. Prove that W_I is normal in W if and only if all $s \in S \setminus I$ commute with all $s' \in I$. In terms of the Coxeter graph, this means that I corresponds to a union of some connected components. [Use the Exchange Condition to analyze the length of $ss's$ in W_I.]

Exercise 3. Suppose $w \in W$ acts on V as a reflection, in the sense that there exists a unit vector $\alpha \in V$ for which $w(\lambda) = \lambda - 2B(\lambda, \alpha)\alpha$ for all $\lambda \in V$. Prove that α is a root and $w = s_\alpha$. [First show that, if $s \in S$ and $\ell(ws) < \ell(w)$, then either $\ell(sws) = \ell(w) - 2$ or else $w(\alpha_s) = -\alpha_s$, using just the fact that $w^2 = 1$: find a reduced expression $w = s_1 \cdots s_r$ with $s_r = s$, so $w = s_r \cdots s_1$ is also reduced, and use the Exchange Condition together with 5.6. Now proceed by induction on $\ell(w)$, to show that $w(\beta) = -\beta$ for some root β, whence $\beta = \alpha$ or $-\alpha$, and w is the reflection belonging to α.]

Exercise 4. If $I \subset S$, set $T_I := \bigcup_{w \in W_I} wIw^{-1}$ (the set of reflections in the Coxeter group W_I). Prove that $T \cap W_I = T_I$. [If $t \in T \cap W_I$, write $t = wsw^{-1} = s_1 \cdots s_r$, with $s \in S$, $w \in W$, $s_i \in I$ for all i, and $\ell(ws) > \ell(w)$. Use the Exchange Condition to show that $t = (w')^{-1}s'w'$ for some $s' = s_i, w' = s_{i+1} \cdots s_r$.]

5.9 Bruhat ordering

Among the possible ways to partially order W in a way compatible with the length function, the most useful has proven to be the Bruhat ordering, defined as follows.

As before, T is the set of reflections in W with respect to roots. Write $w' \to w$ if $w = w't$ for some $t \in T$ with $\ell(w) > \ell(w')$. Then define $w' < w$ if there is a sequence $w' = w_0 \to w_1 \to \ldots \to w_m = w$. It is clear that the resulting relation $w' \leq w$ is a partial ordering of W (reflexive, antisymmetric, transitive), with 1 as the unique minimal element. Following Verma [2], we call it the **Bruhat ordering**. The terminology is motivated by the way this ordering arises for Weyl groups in connection with inclusions among closures of Bruhat cells for a corresponding semisimple algebraic group. In view of the way the ordering is defined, it should not be surprising to find the Strong Exchange Condition used below in investigating its properties.

The definition has a one-sided appearance, since we have written t on the right in defining the arrow relation. But this version could just as well be replaced by a left-sided version. Say $w = w's_\alpha$, with $\ell(w) > \ell(w')$. Setting $\beta = w'(\alpha)$, we get $(w')^{-1}s_\beta w' = s_\alpha$, hence $w = s_\beta w'$. (On the other hand, if we had insisted that t belong to S, the resulting partial ordering, sometimes called the **weak ordering**, would actually have a one-sided nature, as the reader can check for dihedral groups. We won't pursue this possibility here, but see Björner [2].)

One other remark about the definition: when $w' \to w$, the precise length difference is not specified; it must be odd but need not be 1 (as seen already in dihedral groups). So it is not clear at first whether two immediately adjacent elements in the Bruhat ordering must differ in length by just 1. This turns out to be true, but requires some delicate arguments (5.11).

Another natural question about the ordering will also be deferred. If $I \subset S$, the Coxeter group W_I has a Bruhat ordering of its own; does this agree with the restriction to W_I of the Bruhat ordering of W? The answer will be given in 5.10.

Exercise. Prove that $v < w$ if and only if $v^{-1} < w^{-1}$.

Example 1. If W is a dihedral group $\mathcal{D}_m$, $m \leq \infty$, all elements of distinct lengths are comparable in the Bruhat ordering (but not in the weak ordering): $v < w$ if and only if $\ell(v) < \ell(w)$.

Example 2. If W is the symmetric group $\mathcal{S}_n$, each element π can be represented by the string of n integers $(\pi(1), \ldots, \pi(n))$. Then $\pi \leq \sigma$ if and only if σ is obtainable from π by a sequence of transpositions (ij), where $i < j$ and i occurs to the left of j in π. For example, when $n = 5$, we have $24153 \to 42153 \to 45123 \to 54123$, or more directly, $24153 \to 54123$. Another criterion, due to Deodhar, goes as follows. Given a sequence of integers $(a_1, \ldots, a_k)$, denote by $[a_1, \ldots, a_k]$ the sequence rewritten in increasing order. Order $\mathbf{Z}^k$ by $(a_1, \ldots, a_k) \leq (b_1, \ldots, b_k)$ iff $a_i \leq b_i$ for all i. Then $\pi \leq \sigma$ iff $[\pi(1), \ldots, \pi(k)] \leq [\sigma(1), \ldots, \sigma(k)]$ for $1 \leq k \leq n$.

Example 3. There is added symmetry in case W is finite, with longest element $w_\circ$ (see Exercise 2 in 5.6). One sees easily that $v \leq w$ if and only if $w_\circ w \leq w_\circ v$. (This will be used in 7.6.)

One rather subtle property of the Bruhat ordering is needed in 5.10:

Proposition *Let $w' \leq w$ and $s \in S$. Then either $w's \leq w$ or else $w's \leq ws$ (or both).*

Proof. The proof reduces quickly (as the reader should check) to the case $w' \to w$, where $w = w't$ ($t \in T$) and $\ell(w) > \ell(w')$. If $s = t$, there is nothing to prove, so we assume $s \neq t$. Two cases have to be analyzed:

(a) If $\ell(w's) = \ell(w') - 1$, then $w's \to w' \to w$, forcing $w's \leq w$.

(b) If $\ell(w's) = \ell(w') + 1$, we shall argue that $w's < ws$. Since $(w's)t' = ws$ for the reflection $t' = sts$, it is enough to show that $\ell(w's) < \ell(ws)$. Suppose the contrary, i.e., $\ell(ws) < \ell(w's)$. Then the Strong Exchange Condition (5.8) can be applied to the pair $t', w's$ as follows. For any reduced expression $w' = s_1 \cdots s_r$, $w's = s_1 \cdots s_r s$ is also reduced, since $\ell(w's) > \ell(w')$ by assumption. Then $ws = (w's)t'$ is obtained from $w's$ by omitting one factor in this reduced decomposition. This factor cannot be s, since $s \neq t$. Thus $ws = s_1 \cdots \hat{s}_i \cdots s_r s$ for some i, or $w = s_1 \cdots \hat{s}_i \cdots s_r$, contradicting $\ell(w) > \ell(w')$. □

5.10 Subexpressions

There is a very simple and useful characterization of the Bruhat ordering in terms of **subexpressions** of a given reduced expression $w = s_1 s_2 \cdots s_r$, by which we mean products (not necessarily reduced, and possibly empty) of the form $s_{i_1} \cdots s_{i_q}$ $(1 \leq i_1 < i_2 < \ldots < i_q \leq r)$. Formally, the given reduced expression is an ordered r-tuple of elements of S, and a subexpression is a q-tuple obtained by discarding some or all of these elements.

Theorem *Let $w = s_1 \cdots s_r$ be a fixed, but arbitrary, reduced expression for w. Then $w' \leq w$ if and only if w' can be obtained as a subexpression of this reduced expression.*

Proof. Let us first show that any $w' < w$ occurs as a subexpression of the given reduced expression for w. Start with the case $w' \to w$, say $w = w't$. Since $\ell(w') < \ell(w)$, the Strong Exchange Condition can be applied to the pair t, w to yield $w' = wt = s_1 \cdots \hat{s}_i \cdots s_r$ for some i. This argument can be iterated. If in turn $w' \to w'$, with $w' = w't'$, apply the Strong Exchange Condition to the pair $t', w' = s_1 \cdots \hat{s}_i \cdots s_r$ (which is not required to be a reduced expression!) to obtain

$$w' = w't' = s_1 \cdots \hat{s}_i \cdots \hat{s}_j \cdots s_r$$

or else

$$w' = s_1 \cdots \hat{s}_j \cdots \hat{s}_i \cdots s_r.$$

In the other direction, we are given a subexpression $s_{i_1} \cdots s_{i_q}$ and must show it to be $\leq w$. Here we can use induction on $r = \ell(w)$, the case $r = 0$ being trivial. If $i_q < r$, the induction hypothesis can be applied to the reduced expression $s_1 \cdots s_{r-1}$ to yield:

$$s_{i_1} \cdots s_{i_q} \leq s_1 \cdots s_{r-1} = ws_r < w.$$