We defer the formal definition of the term *Coxeter group* until later, starting instead with a discussion of groups generated by reflections. Since it turns out that every Coxeter group can be realized as a reflection group, this discussion will in fact be about Coxeter groups anyway.

Let V be a vector space over $\mathbb{R}$, equipped with a scalar product (or bilinear form) $(u,v) \mapsto u \cdot v$ which satisfies

(a) $u \cdot v = v \cdot u$ for all u, $v \in V$, and

(b) $t \cdot (\lambda u + \mu v) = \lambda t \cdot u + \mu t \cdot v$ for all t, u, $v \in V$ and λ, $\mu \in \mathbb{R}$.

Note that we are leaving open the possibility that V may contain nonzero vectors v satisfying $v \cdot v \leq 0$. In the important special case that there are none of these, V is a *Euclidean space*, and $(u,v) \mapsto u \cdot v$ is an *inner product*.

The *radical* of V is the subspace

$$\mathrm{rad}(V) = \{\, v \in V \mid u \cdot v = 0 \text{ for all } u \in V \,\}.$$

Clearly $\mathrm{rad}(V) = \{0\}$ if V is Euclidean; note, however, that the converse of this statement is false.

We shall say that a linear transformation $V \to V$ is *orthogonal* if it is bijective and preserves the scalar product. The set of all orthogonal transformations of V is a group, $O(V)$, the *orthogonal group* of V (relative to the given bilinear form).

A *reflection* is an orthogonal transformation $\rho\colon V \to V$ of order 2 whose -1-eigenspace is 1-dimensional and not contained in the radical. Choose $a \neq 0$ in the -1-eigenspace, and let $v \in V$ be arbitrary. Since $\rho^2 = 1$ we find that

$$\rho(\rho v - v) = \rho^2 v - \rho v = -(\rho v - v),$$

and so $\rho v = v + \lambda a$ for some scalar λ. Since ρ is orthogonal,

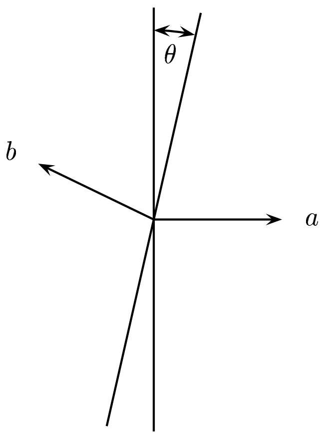
$$a \cdot v = (\rho a) \cdot (\rho v) = -a \cdot (v + \lambda a) = -a \cdot v - \lambda(a \cdot a),$$

and since there exists at least one $v \in V$ with $a \cdot v \neq 0$ (since $a \notin \mathrm{rad}(V)$) it follows that $a \cdot a \neq 0$. Substituting the resulting formula for λ into the formula for ρv yields

$$\rho v = v - 2\frac{(v \cdot a)}{(a \cdot a)}a \qquad \text{for all } v \in V. \tag{1}$$

Conversely, if a is any vector satisfying $a \cdot a \neq 0$ then it is easily checked that the transformation $\rho = \rho_a$ defined by eq. (1) is orthogonal and has order 2, and that a spans its -1-eigenspace. We call ρ_a the *reflection* corresponding to a (since in the familiar examples of two and three dimensional Euclidean spaces reflections are indeed of this form). The set of fixed points of ρ_a is a hyperplane (subspace of codimension 1 in V), commonly called the *mirror* of the reflection.

For the case of the Euclidean plane it is well known that the product of two reflections is a rotation through an angle equal to twice the angle between the mirrors. Thus, in the diagram shown, $\rho_a\rho_b$ is a



clockwise rotation through 2θ, and in particular if $\theta = \pi/n$, where n is a positive integer, then $\rho_a\rho_b$ has order n. In these circumstances, since a and b could be replaced by their negatives if we so wished, we can arrange either that the angle between a and b be π/n, or that it be $(n-1)\pi/n$. It happens to be convenient for the theory we shall develop to make the latter choice. The lengths of the vectors are not particularly important for us; so we will assume that they have unit length.

Proposition. *Suppose that $a, b \in V$ satisfy $a \cdot a = b \cdot b = 1$ and $a \cdot b = -\cos\theta$, where $\theta = \pi/n$ for some positive integer n. Then*

(i) $(\rho_a\rho_b)^i a = \frac{\sin(2i+1)\theta}{\sin\theta}a + \frac{\sin 2i\theta}{\sin\theta}b$ *(for all integers i), and*

(ii) $\rho_a\rho_b$ *has order n.*

Proof. Since $\rho_a a = -a$ and $\rho_a b = b + 2\cos\theta a$, and similar formulas apply for ρ_b, we conclude that the matrix of $\rho_a\rho_b$ in its action on the subspace with basis $\{a, b\}$ is

$$M = \begin{pmatrix} -1 & 2\cos\theta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2\cos\theta & -1 \end{pmatrix} = \frac{1}{\sin\theta} \begin{pmatrix} \sin 3\theta & -\sin 2\theta \\ \sin 2\theta & -\sin\theta \end{pmatrix}.$$

Induction can now be used to show that

$$M^i = \frac{1}{\sin\theta} \begin{pmatrix} \sin(2i+1)\theta & -\sin 2i\theta \\ \sin 2i\theta & -\sin(2i-1)\theta \end{pmatrix}$$

for all integers i, and this establishes part (i) of the proposition. It also shows that $\rho_a\rho_b$ has order n on the space spanned by a and b. But if $v \in V$ is arbitrary then $v = v' + \lambda a + \mu b$ for some v' such that $v' \cdot a = v' \cdot b = 0$, and some $\lambda, \mu \in \mathbb{R}$. Since $\rho_a\rho_b$ fixes v' it follows that $(\rho_a\rho_b)^n v = v$. □

It is valuable for our investigation of groups generated by reflections to have rather detailed information for the particular case of groups generated by two reflections. Hence our interest in the above example. It is fairly clear that the group W generated by ρ_a and ρ_b is the dihedral group of order $2n$ (the group of symmetries of a regular n-sided polygon). It is, indeed, a general fact that any group generated by two elements of order 2 is dihedral.

Proposition. *Let r, s be elements of order 2 (in any group) and suppose that rs has order n. Then the subgroup generated by r and s is dihedral of order $2n$.*

It is easy to check the truth of this by simply enumerating the elements of the subgroup generated by r and s. Since r and s are self-inverse, the only group elements obtainable as products of r's and s's are $1, r, s, rs, sr, rsr, srs, \ldots$, two products of each length greater than 0. But the relation $(rs)^n = 1$ gives $rsrs\ldots = srsr\ldots$, where there are n factors on each side; so the two products of length n give the same group element. And this is easily seen to be the element of maximal length in the subgroup, since pre- or post-multiplying it by either r or s will result in cancellation, giving one or other of the elements of length $n-1$.

Identifying r and s with ρ_a and ρ_b, observe that the elements of W of even length are powers of $\rho_a\rho_b$, and hence are rotations. There are n of these. The following lemma, whose proof we leave as an exercise, shows the elements of odd length are reflections.

Lemma. *Let $w \in O(V)$, and let $a \in V$ with $a \cdot a \neq 0$. Then $w\rho_a w^{-1} = \rho_{wa}$.*

If $w \in W$ then $\rho_{wa} = w\rho_a w^{-1}$ and $\rho_{wb} = w\rho_b w^{-1}$ will be in W (since $\rho_a, \rho_b \in W$). In particular, if we put

$$a_i = (\rho_a\rho_b)^i a = \frac{\sin(2i+1)\theta}{\sin\theta}a + \frac{\sin 2i\theta}{\sin\theta}b,$$

then $\rho_{a_i} \in W$. Note, for future reference, that if $0 \le i \le (n-1)/2$ then the coefficients of a and b in the expression for a_i are both nonnegative. Similarly, if we put

$$b_i = -(\rho_a\rho_b)^i b = \frac{\sin 2i\theta}{\sin\theta}a + \frac{\sin(2i-1)\theta}{\sin\theta}b$$

then ρ_{b_i} (which equals ρ_{-b_i}) is an element of W. The coefficients of a and b in b_i are both nonnegative for $1 \le i \le n/2$. The ρ_{a_i} for $0 \le i \le (n-1)/2$ and the ρ_{b_i} for $1 \le i \le n/2$, together with the n rotations, make up all the elements of W.

The larger the integer n in the above example, the closer $a \cdot b$ is to -1. Observe that $a \cdot b > -1$ whenever a and b are linearly independent unit vectors in a Euclidean space. But we do not wish to restrict ourselves to the Euclidean case.

Proposition. *Suppose that a, $b \in V$ are linearly independent and satisfy $a \cdot a = b \cdot b = 1$ and $a \cdot b = -t \le -1$. Put $\theta = \ln(t + \sqrt{t^2-1})$, so that $t = \cosh\theta$. Then*

(i) $(\rho_a\rho_b)^i a = \frac{\sinh(2i+1)\theta}{\sinh\theta} a + \frac{\sinh 2i\theta}{\sinh\theta} b$ *(for all integers i), and*

(ii) $\rho_a\rho_b$ *has infinite order.*

Proof. The matrix of $\rho_a\rho_b$ in its action on the subspace with basis $\{a, b\}$ is

$$M = \begin{pmatrix} -1 & 2t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2t & -1 \end{pmatrix} = \frac{1}{\sinh\theta} \begin{pmatrix} \sinh 3\theta & -\sinh 2\theta \\ \sinh 2\theta & -\sinh\theta \end{pmatrix}.$$

Induction can now be used to show that

$$M^i = \frac{1}{\sinh\theta} \begin{pmatrix} \sinh(2i+1)\theta & -\sinh 2i\theta \\ \sinh 2i\theta & -\sinh(2i-1)\theta \end{pmatrix},$$

and this yields the result. (The special case $t = 1$ gives $\theta = 0$, in which case $\sinh i\theta / \sinh\theta$ should be replaced by i.) $\square$

We are now ready to consider a more general situation: we investigate groups generated by a set of reflections $\{ \rho_a \mid a \in \Pi \}$, where every pair of elements a, $b \in \Pi$ satisfy the hypotheses of one or other of our propositions dealing with the two-generator cases. It will be convenient to use the following notation: if A is a set of vectors then

$$\mathrm{PLC}(A) = \{ \sum_{a \in A} \lambda_a a \mid \lambda_a \ge 0 \text{ for all } a, \text{ and } \lambda_a > 0 \text{ for some } a \}.$$

(The letters PLC stand for "positive linear combination".)

Definition. A subset Π of V is called a *root basis* if

(1) $a \cdot a = 1$ for all $a \in \Pi$, and if a, b are distinct elements of Π then either $a \cdot b = -\cos(\pi/n_{ab})$ for some integer $n_{ab} \ge 2$, or else $a \cdot b \le -1$ (in which case we define $n_{ab} = \infty$),

(2) $0 \notin \mathrm{PLC}(\Pi)$.

Notes.

(a) In the treatment of Coxeter groups given in Bourbaki's *Groupes et Algebres de Lie* Chap. IV, V et VI—the book which is the bible for the subject—the set Π is assumed to be linearly independent. However, our weaker assumption that $0 \notin \mathrm{PLC}(\Pi)$ suffices for the proofs, and turns out to be more convenient when it comes to looking at reflection subgroups of a Coxeter group.

If V is Euclidean then a root basis is necessarily linearly independent. This can be seen as follows. Suppose that $\sum_{a\in\Pi} \lambda_a a = 0$. Moving the terms with negative coefficients to the right-hand side we obtain $\sum_{a\in\Pi_1} |\lambda_a| a = \sum_{a\in\Pi_2} |\lambda_a| a = v$ (say), where $\Pi = \Pi_1 \dot\cup \Pi_2$. Now since V is Euclidean,

$$0 \le v \cdot v = \sum_{a\in\Pi_1} \sum_{b\in\Pi_2} |\lambda_a|\,|\lambda_b|\, a \cdot b \le 0$$

since $a \cdot b \le 0$ whenever a, $b \in \Pi$ with $a \ne b$. So $v \cdot v = 0$, and hence $v = 0$ (since V is Euclidean). It follows that $\sum_{a\in\Pi} |\lambda_a| a = 0$, and so (by (2)) $\lambda_a = 0$ for all a.

(b) Assuming that $0 \notin \mathrm{PLC}(\Pi)$ is equivalent to assuming that there exists an element of V^*, the space of all linear functionals $V \to \mathbb{R}$, which takes positive values on all elements of Π. It is easy to see that existence of such a functional implies that $0 \notin \mathrm{PLC}(\Pi)$. Conversely, assume that $0 \notin \mathrm{PLC}(\Pi)$, and let

$$X = \{ \sum_{a \in \Pi} \lambda_a a \mid \lambda_a \geq 0 \text{ for all } a, \text{ and } \sum_a \lambda_a = 1 \}.$$

Let $(\ ,\ )$ be a positive definite inner product on V. Since X is topologically closed it is easy to see that there exists a $v \in X$ with $||v||^2 = (v,v)$ minimal. Now let x be any element of X, and consider the function $f(t) = ||(1-t)v + tx||^2$, for $t \in [0,1]$. Since X is convex, f has a minimum at $t = 0$. So $f'(0) \geq 0$. But $f'(0) = 2(v,(x-v))$; so $(v,x) \geq (v,v) > 0$. It follows that the linear functional $x \mapsto (v,x)$ takes strictly positive values on X, and hence on all elements of $\mathrm{PLC}(\Pi)$, as required.

(c) The set $\mathrm{PLC}(\Pi)$ is *cone*: a set closed under addition and multiplication by positive scalars. If C is a cone in a finite-dimensional vector space V we define its *dual* C^* to be the subset of the dual space V^* consisting of those functionals α such that $\alpha v \geq 0$ for all $v \in C$. It can be shown that $C^{**} = \overline{C}$, the topological closure of C. Indeed, if $v \in C^{**}$ and $v \notin \overline{C}$, and if $u \in \overline{C}$ is chosen with $||u-v||$ minimal, then an argument similar to the one used in (b) above (considering $f(t) = ||(1-t)u + tx - v||^2$) can be used to show that $(x, u-v) \geq 0$ for all $x \in \overline{C}$. Applying this with $x = 2u$ and $x = \frac{1}{2}u$ enables us to conclude that $(u, u-v) = 0$, and hence the functional $x \mapsto (x, u-v)$ is in C^*. Since $v \in C^{**}$ by hypothesis, it follows that $(v, u-v) \geq 0$, whence $(u-v, u-v) \leq 0$, and $u = v$, contradicting $v \notin \overline{C}$.

Let Π be a root basis and let W be the subgroup of $\mathrm{GL}(V)$ generated by the reflections $\{ \rho_a \mid a \in \Pi \}$. These generators satisfy the following relations: $\rho_a^2 = 1$ for all $a \in \Pi$, and $(\rho_a \rho_b)^{n_{ab}} = 1$ for all $a, b \in \Pi$ such that $n_{ab} < \infty$. We aim to prove that these relations suffice to define W. So let $\widetilde{W}$ be a group generated by elements $\{ r_a \mid a \in \Pi \}$ satisfying $r_a^2 = 1$ (for all $a \in \Pi$) and $(r_a r_b)^{n_{ab}} = 1$ for all $a, b \in \Pi$ with $n_{ab} < \infty$, and suppose that there is a homomorphism $\phi\colon \widetilde{W} \to W$ with $\phi r_a = \rho_a$ for all $a \in \Pi$. We will show that ϕ is injective (which will show that the given relations are defining relations).

Since the group $\widetilde{W}$ is generated by $\{ r_a \mid a \in \Pi \}$, for each $w \in \widetilde{W}$ there exist $a_1, a_2, \ldots, a_l \in \Pi$ with $w = r_{a_1} r_{a_2} \cdots r_{a_l}$. The least such nonnegative integer l is called the *length*, $l(w)$, of w. The identity element is the unique $w \in \widetilde{W}$ of length 0. For $w \in \widetilde{W}$ and $x \in V$ define $wx \in V$ by $wx = (\phi w)x$ (remembering that $\phi w \in W \leq \mathrm{GL}(V)$).

The following two lemmas are trivial.

Lemma. *If $w \in \widetilde{W}$ and $a \in \Pi$ then $|l(w) - l(wr_a)| \leq 1$. Moreover, if $w \neq 1$ then there exists $b \in \Pi$ with $l(wr_b) = l(w) - 1$.*

Proof. If $w = r_{a_1} r_{a_2} \cdots r_{a_l}$ then $wr_a = r_{a_1} r_{a_2} \cdots r_{a_l} r_a$. Hence there is an expression for wr_a having length $l(w) + 1$, and so $l(wr_a) \leq l(w) + 1$. But since we also have $w = (wr_a)r_a$, a symmetrical argument gives $l(w) \leq l(wr_a) + 1$, whence the first assertion follows. For the other part, let $w = r_{a_1} r_{a_2} \cdots r_{a_l}$ with the a_i in Π and $l = l(w) > 0$, and let $b = a_l$. Then $wr_b = r_{a_1} r_{a_2} \cdots r_{a_{l-1}}$, and so $l(wr_b) \leq l(w) - 1$, and combined with the first part this gives $l(wr_b) = l(w) - 1$. □

Lemma. *If $w_1, w_2 \in \widetilde{W}$ then $l(w_1 w_2) \leq l(w_1) + l(w_2)$.*

Proof. Combining minimal length expressions for w_1 and w_2 gives an expression for $w_1 w_2$ whose length is $l(w_1) + l(w_2)$; so a minimal expression for $w_1 w_2$ is at most this long. □

Theorem. *Let $w \in \widetilde{W}$ and $a \in \Pi$. If $l(wr_a) \geq l(w)$ then $wa \in \mathrm{PLC}(\Pi)$.*

Proof. Choose $w \in \widetilde{W}$ of minimal length such that the assertion of the theorem statement fails for some $a \in \Pi$, and choose such an a. Certainly w is not the identity element 1, since $1a = a$ is trivially a positive linear combination of Π (since $a \in \Pi$). So $l(w) \geq 1$, and we may choose $b \in \Pi$ such that $w_1 = wr_b$ has length $l(w) - 1$. If $l(w_1 r_a) \geq l(w_1)$ then $l(w_1 r_c) \geq l(w_1)$ holds for both $c = a$ and $c = b$ (since $w_1 r_b = w$). Alternatively, if $l(w_1 r_a) < l(w_1)$, we define $w_2 = w_1 r_a$, and note that $l(w_2 r_c) \geq l(w_2)$ will hold for both

$c = a$ and $c = b$ if $l(w_2 r_b) \geq l(w_2)$. If this latter condition is not satisfied then we define $w_3 = w_2 r_b$. Continuing in this way we find, for some positive integer k, a sequence of elements $w_0 = w$, w_1, w_2, $\ldots$, w_k with $l(w_i) = l(w) - i$ for all $i = 0, 1, 2, \ldots, k$, and, when $i < k$,

$$w_{i+1} = \begin{cases} w_i r_a & \text{if } i \text{ is odd,} \\ w_i r_b & \text{if } i \text{ is even.} \end{cases}$$

Now since $0 \leq l(w_k) = l(w) - k$ we conclude that $l(w)$ is an upper bound for the possible values of k. Choosing k to be as large as possible, we deduce that $l(w_k r_c) \geq l(w_k)$ for both $c = a$ and $c = b$, for otherwise the process described above would allow a w_{k+1} to be found, contrary to the definition of k. By the minimality of our original counterexample it follows that $w_k a$ and $w_k b$ are both in $\mathrm{PLC}(\Pi)$.

We have that $w = w_k v$, where v is an alternating product of r_a's and r_b's, ending in r_b and with k factors altogether. Obviously this means that $l(v) \leq k$. But $w = w_k v$ gives $l(w) \leq l(w_k) + l(v)$; so $l(v) \geq l(w) - l(w_k) = k$, and hence $l(v) = k$. Furthermore, in view of the hypothesis that $l(wr_a) \geq l(w)$, and since $w_k v r_a = w r_a$, we have

$$l(w_k) + l(vr_a) \geq l(wr_a) \geq l(w) = l(w_k) + k = l(w_k) + l(v),$$

and hence $l(vr_a) \geq l(v)$. Note also that v cannot have a minimal length expression in which the final factor is r_a, for if so vr_a would have a strictly shorter expression.

Let $n = n_{ab}$. Since r_a and r_b satisfy the defining relations of the dihedral group of order $2n$, it follows that every element of the subgroup generated by r_a and r_b has an expression of length less than $n + 1$ as an alternating product of r_a's and r_b's. So, in particular, $l(v) \leq n$. Moreover if n is finite then the two alternating products of length n define the same element; so $l(v)$ cannot equal n, as v has no minimal expression ending with r_a. Now, as we have seen, the matrix of $(r_a r_b)^m$ on the space spanned by a and b (relative to the basis a, b) is

$$\begin{pmatrix} t_{2m+1} & -t_{2m} \\ t_{2m} & -t_{2m-1} \end{pmatrix}$$

where $t_i = \sin i\theta / \sin\theta$ (with $\theta = \pi/n$) if $n < \infty$, and $t_i = \sinh i\theta / \sinh\theta$ (with $\theta = \cosh^{-1}(-a \cdot b)$) if $n = \infty$. Observe that t_i is nonnegative for $0 \leq i \leq n$. If k is even, $k = 2m$ say, then $va = (r_a r_b)^m a = t_{2m+1} a + t_{2m} b$, while if k is odd, $k = 2m+1$, then $va = (r_b r_a)^m r_b a = -(r_b r_a)^{m+1} a = t_{2m+2} b + t_{2m+1} a$. Thus the coefficients of a and b are t_k and t_{k+1}, in one order or the other, and these numbers are both nonnegative since $k < n$.

We have now shown that $va = \lambda a + \mu b$, with $\lambda \geq 0$ and $\mu \geq 0$, and hence

$$wa = w_k va = w_k(\lambda a + \mu b) = \lambda w_k a + \mu w_k b \in \mathrm{PLC}(\Pi)$$

since $w_k a$, $w_k b \in \mathrm{PLC}(\Pi)$. This contradicts our original choice of w and a as a counterexample to the statement of the theorem; so the proof is completed. $\square$

Corollary. *If $w \in \widetilde{W}$ is such that $wa = a$ for all $a \in \Pi$, then $w = 1$. In particular, the kernel of the homomorphism $\rho\colon \widetilde{W} \to W \leq \mathrm{GL}(V)$ is trivial.*

Proof. If $w \neq 1$ then $l(w) > 0$, and so we can write $w = w' r_a$ with $a \in \Pi$ and $l(w') = l(w) - 1$. Since $l(w' r_a) > l(w')$ the theorem yields $w'a \in \mathrm{PLC}(\Pi)$; but

$$a = wa = (w' r_a)a = w'(r_a a) = w'(-a) = -w'a.$$

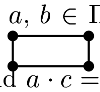
and hence $0 = a + w'a \in \mathrm{PLC}(\Pi)$, contradicting (2) of the definition of a root basis. $\square$

Definition. A *Coxeter group* is a group with a presentation of the following form:

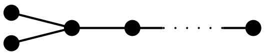
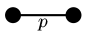
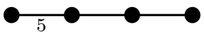
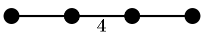
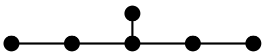
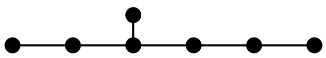
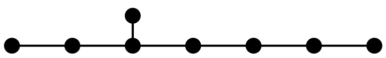
$$W = \mathrm{gp}\langle \{ r_a \mid a \in \Pi \} \mid r_a^2 = 1 \text{ for all } a \in \Pi,\ (r_a r_b)^{n_{ab}} = 1 \text{ for all } a,\, b \in \Pi \text{ with } a \neq b \rangle$$

where the set Π indexing the generators can be of any cardinality (although in these lectures it will always be assumed to be finite), and the parameters n_{ab} satisfy $n_{ab} = n_{ba} \in \mathbb{Z} \cup \{\infty\}$ and $n_{ab} \geq 2$ in all cases. A "relation" $(r_a r_b)^\infty = 1$ is to be interpreted as vacuous. (Note that allowing $n_{ab} = 1$ would not yield any extra groups, since the relation $(r_a r_b)^1 = 1$ would just give $r_a = r_b$, allowing one of the generators to be dropped.)

We have shown that the reflections corresponding to elements of a root basis generate a Coxeter group.

There is a convenient diagrammatic way of describing a root basis Π, as follows. Draw a graph that has one vertex for each $a \in \Pi$, and join the vertices corresponding to a, $b \in \Pi$ by an edge labelled n_{ab} if $n_{ab} > 2$. The label n_{ab} is usually omitted if $n_{ab} = 3$. Thus the graph  corresponds to a root basis $\Pi = \{a, b, c, d\}$ with $a \cdot b = b \cdot c = c \cdot d = d \cdot a = -\cos(\pi/3) = -1/2$, and $a \cdot c = b \cdot d = -\cos(\pi/2) = 0$. If a root basis Π can be split into two nonempty disjoint subsets Π_1 and Π_2 such that $n_{ab} = 2$ whenever $a \in \Pi_1$ and $b \in \Pi_2$ then the diagram of Π will be disconnected; it will consist of the diagram for Π_1 alongside the diagram for Π_2. Furthermore, if $a \in \Pi_1$ and $b \in \Pi_2$ then $a \cdot b = 0$, and the reflections ρ_a and ρ_b commute. In such cases we say that the root basis is *reducible*. If V and V' are vector spaces equipped with scalar products then we can define a scalar product on the external direct sum $V \oplus V'$ which extends the scalar products on V and V' and satisfies $v \cdot v' = 0$ whenever $v \in V$ and $v' \in V'$. It is easily seen that $V \oplus V'$ is Euclidean if V and V' both are. If Π and Π' are root bases in V and V' then $\Pi \mathbin{\dot\cup} \Pi'$ will be a reducible root basis in the direct sum.

The following table describes the classification of Euclidean root bases and the corresponding Coxeter groups:

Type	Diagram	Rank	Exponents	Order
A_n		n	$1, 2, 3, \ldots, n$	$(n+1)!$
$B_n = C_n$		n	$1, 3, 5, \ldots, 2n-1$	$2^n n!$
D_n		n	$1, 3, 5, \ldots, 2n-3, n-1$	$2^{n-1} n!$
$I_2(p)$		2	$1, p-1$	$2p$
H_3		3	$1, 5, 9$	120
H_4		4	$1, 11, 19, 29$	14400
F_4		4	$1, 5, 7, 11$	1152
E_6		6	$1, 4, 5, 7, 8, 11$	$2^7 \cdot 3^4 \cdot 5$
E_7		7	$1, 5, 7, 9, 11, 13, 17$	$2^{10} \cdot 3^4 \cdot 5 \cdot 7$
E_8		8	$1, 7, 11, 13, 17, 19, 23, 29$	$2^{14} \cdot 3^5 \cdot 5^2 \cdot 7$

and arbitrary direct products of groups from this list.

Notes.

1. The diagram can be regarded as an abbreviated notation for the presentation of the associated Coxeter group. Each vertex of the diagram represents a generator, and the parameters n_{ab} which appear in the relations $(r_a r_b)^{n_{ab}} = 1$ are determined by the edges and edge-labels of the diagram. The direct product of two Coxeter groups is the reducible Coxeter group whose diagram is the union of the diagrams of the two factors. The connected components of the diagram of a Coxeter group W are diagrams of Coxeter

groups which we will refer to as the *irreducible components* of W; clearly W is the direct product of these components.

2. The rank is the number of generators in the presentation—that is, the cardinality of Π—and is also the dimension of the Euclidean space on which the group acts as a group generated by reflections.

3. The classification almost coincides with the classification of complex semisimple Lie algebras. Indeed, the Weyl group of such a Lie algebra is a Euclidean reflection group, and all the groups on the above list can occur, with the exceptions of H_3, H_4 and $I_2(p)$ for $p = 5$ and all $p \geq 7$. The direct sum of two complex semisimple Lie algebras gives a complex semisimple Lie algebra, the Weyl group of which is the direct product of the Weyl groups of the factors. So a finite Coxeter group corresponds to a Lie algebra provided its irreducible components are all of types A_n, B_n, D_n, E_n, F_4 or $I_2(6)$ (this last case corresponding to the Lie algebra of type G_2).

4. For $n > 2$ the Lie algebras B_n and C_n are not isomorphic, but their Weyl groups are; hence the "$B_n = C_n$" on the above list.

5. All of the irreducible finite Coxeter groups can be described as the symmetry groups of various regular or nearly regular objects in n-dimensional space (where n is the rank). For instance, the group of type A_n, which is just $\mathrm{Sym}(n+1)$, is the symmetry group of a regular n-simplex (the n-dimensional generalization of the 2-dimensional equilateral triangle and the 3-dimensional regular tetrahedron). The group of type B_n is the symmetry group of the hypercube (measure polytope) or its dual, the cross polytope (the generalization of the octahedron: its $2n$ vertices lie on the n axes of a rectangular coordinate system at a fixed distance from the origin). The 2^{n-1} points with coordinates $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ where $\varepsilon_i = \pm 1$ for all i and $\varepsilon_1\varepsilon_2 \cdots \varepsilon_n = 1$ form the vertices of a polytope whose symmetry group is the Coxeter group of type D_n; this group is isomorphic to the group of all monomial matrices whose nonzero entries are all ± 1, with an even number of -1's. Types H_4 and F_4 are symmetry groups of regular 4-dimensional polytopes, while E_6, E_7 and E_8 correspond to certain semiregular polytopes. (See Coxeter's book "Regular Polytopes".)

6. The exponents of a finite Coxeter group W of rank n are n numbers $m_1, m_2, \ldots, m_n$, possessing many remarkable properties. Let $c = r_{a_1} r_{a_2} \cdots r_{a_n} \in W$, where $\Pi = \{a_1, a_2, \ldots, a_n\}$. That is, c is the product, in some order, of all the generators of W in the Coxeter presentation. Obviously c depends on the choice of the ordering of Π, but since the Coxeter diagram of a finite Coxeter group is a graph with no circuits, and since the generators corresponding to any pair of non-adjacent vertices commute with each other, a relatively easy graph-theoretic argument shows that different orderings give conjugate elements c. These elements are called the *Coxeter elements* of W. They have order $h = 2N/n$, where N is the number of reflections in W and n is the rank. The eigenvalues of c as a transformation of n-dimensional space are ω^{m_1}, ω^{m_2}, …, ω^{m_n}, where $\omega = e^{2\pi i/h}$; this is the reason for the name "exponents". It turns out that the numbers $m_1 + 1$, $m_2 + 1$, …, $m_n + 1$ are the degrees of certain homogeneous polynomials in n variables, known as the "basic polynomial invariants of W". The product $(m_1 + 1)(m_2 + 1) \cdots (m_n + 1)$ equals the order of W, and the sum $m_1 + m_2 + \cdots + m_n$ equals N, the number of reflections. (Proving these facts is beyond the scope of these lectures!)

Suppose that V is a Euclidean space, and G is a finite subgroup of $O(V)$ generated by reflections. We shall show that there exists a root basis such that the corresponding reflections generate G, whence it will follow that G is a Coxeter group. Thus the classification above describes all finite reflection groups that act on Euclidean space.

Let $\Phi = \{\, a \in V \mid \rho_a \in G \text{ and } ||a|| = 1 \,\}$ (the *root system* of G). Observe that if $a \in \Phi$ then $-a \in \Phi$ also. Choose a $v \in V$ such that $a \cdot v \neq 0$ for all $a \in \Phi$—this is possible since Φ is a finite set—and define $\Phi^+ = \{\, a \in \Phi \mid a \cdot v > 0 \,\}$. Let Π be a subset of Φ^+ that is minimal subject to $\Phi^+ \subseteq \mathrm{PLC}(\Pi)$. Note that $0 \notin \mathrm{PLC}(\Pi)$, since $a \cdot v > 0$ for all $a \in \Pi$ and hence for all $a \in \mathrm{PLC}(\Pi)$.

If $a \in \Pi$ then $a \notin \mathrm{PLC}(\Pi \backslash \{a\})$, for otherwise we would find that $\mathrm{PLC}(\Pi \backslash \{a\}) = \mathrm{PLC}(\Pi)$, contrary to minimality of Π. Nor is $-a \in \mathrm{PLC}(\Pi \backslash \{a\})$, for this would give $0 \in \mathrm{PLC}(\Pi)$. Now suppose, for a contradiction, that $a, b \in \Pi$ satisfy $a \cdot b > 0$ and $a \neq b$. Since $\rho_{(\rho_a b)} = \rho_a \rho_b \rho_a \in G$ we conclude that $\rho_a b \in \Phi$, and hence either $\rho_a b \in \Phi^+ \subseteq \mathrm{PLC}(\Pi)$ or $-\rho_a b \in \Phi^+ \subseteq \mathrm{PLC}(\Pi)$. But $\rho_a b = a - \lambda b$ where $\lambda > 0$, and so we obtain either $a = \lambda b + \sum_{c \in \Pi} \lambda_c c$ or $\lambda b = a + \sum_{c \in \Pi} \lambda_c c$ for some nonnegative scalars λ_c. Collecting coefficients we will either obtain $\pm a \in \mathrm{PLC}(\Pi\backslash\{a\})$ or $\pm b \in \mathrm{PLC}(\Pi\backslash\{b\})$, none of which are possible. So $a \cdot b \leq 0$ whenever $a, b \in \Pi$ with $a \neq b$.

It can now be seen, by means of the argument used in Note (a) to the definition of a root basis, that Π is linearly independent. So if $a, b \in \Pi$ with $a \neq b$ then whenever $\lambda a + \mu b \in \Phi$ the coefficients λ and μ cannot have opposite signs. But since $(\rho_a \rho_b)^i a = \frac{\sin(i+1)\theta}{\sin\theta} a + \frac{\sin i\theta}{\sin\theta} b$, where the angle between a and b is $-\theta$, it is easily seen that θ must be of the form π/n, and so it follows that Π is a root basis.

It remains to show that G is generated by $\{\, \rho_a \mid a \in \Pi\}$. Let W be the subgroup of G generated by $\{\, \rho_a \mid a \in \Pi\}$; since G is generated by the reflections it contains, it suffices to show that $\rho_b \in W$ for all $b \in \Phi^+$. This can be done by induction on the number of $c \in \Phi^+$ such that $\rho_b c \notin \Phi^+$. We leave it as an exercise (which will be easier after we have developed a little more of the theory), and turn instead to the classification of root bases in Euclidean space.

As we have seen, a root basis Π in a Euclidean space is necessarily linearly independent; moreover, if $a, b \in \Pi$ then $a \cdot b = -\cos(\pi/n_{ab})$ for some integer n_{ab}, since $a \cdot b \leq -1$ is impossible. So the problem can be rephrased as follows: if Π is a basis for a real vector space V, and a scalar product is defined on V by

$$\begin{aligned} &a \cdot a = 1 \quad \text{for all } a \in \Pi \\ &a \cdot b = -\cos(\pi/n_{ab}) \quad \text{for all } a,\, b \in \Pi \text{ with } a \neq b \end{aligned}$$

where the n_{ab} are integers greater than 1, under what circumstances is V Euclidean? The key lemma says that if V is not Euclidean, and the definition of the scalar product is modified by increasing some of the integers n_{ab}, then V remains non-Euclidean.

Lemma. Let V and the scalar product $(x, y) \mapsto x \cdot y$ on V be as defined above, and suppose a second scalar product $(x, y) \mapsto x * y$ on V is given by $a * a = 1$ for all $a \in \Pi$ and $a * b = -\cos(\pi/m_{ab})$ for $a, b \in \Pi$ with $a \neq b$. Suppose that $m_{ab} \geq n_{ab}$ in all cases. If there exists a nonzero $v \in V$ with $v \cdot v \leq 0$ then there exists a nonzero $u \in V$ with $u * u \leq 0$.

Proof. Suppose that $v = \sum_{a\in\Pi} \lambda_a a$ is nonzero and satisfies $v \cdot v \leq 0$. Put $u = \sum_{a\in\Pi} |\lambda_a| a$. Then $u \neq 0$. Now the inequality $m_{ab} \geq n_{ab}$ yields that $0 \geq -\cos(\pi/n_{ab}) \geq -\cos(\pi/m_{ab})$ for all $a, b \in \Pi$ with $a \neq b$, and so

$$\begin{aligned} u * u &= \sum_{a\in\Pi} |\lambda_a|^2 - \sum_{\substack{a,\, b\in\Pi \\ a\neq b}} |\lambda_a|\,|\lambda_b| \cos(\pi/m_{ab}) \\ &\leq \sum_{a\in\Pi} \lambda_a^2 - \sum_{\substack{a,\, b\in\Pi \\ a\neq b}} \lambda_a \lambda_b \cos(\pi/n_{ab}) \\ &= v \cdot v \leq 0, \end{aligned}$$

establishing the claim. □

If Π and Π' are root bases, let us say that $\Pi \leq \Pi'$ if there exists an injective map $a \mapsto a'$ from Π to Π' such that $n_{ab} \leq n_{a'b'}$ whenever $a, b \in \Pi$ with $a \neq b$. The lemma shows that if Π is non-Euclidean and $\Pi \leq \Pi'$ then Π' is non-Euclidean also. The classification of Euclidean root bases now proceeds by exhibiting a longish list of non-Euclidean ones, and then checking that if Π' is any root basis then either $\Pi' \geq \Pi$ for some Π on the list, whence Π' is non-Euclidean, or else the irreducible components of Π' appear in the table we gave above.

The 3-dimensional groups

The irreducible 3-dimensional Euclidean reflection groups, types A_3, B_3 and H_3, are (respectively) the groups of symmetries of a regular tetrahedron, a cube, and a regular dodecahedron. Note that a regular octahedron, being in some sense the dual of a cube, has the same group of symmetries as a cube. Similarly, a regular icosahedron has the same group of symmetries as a regular dodecahedron. It is not hard to investigate these groups directly without making use of any of the general theory of Coxeter groups.

Tetrahedron. Given a regular tetrahedron centred at the origin, there is a uniquely determined "dual" tetrahedron which is congruent to the given one, and has the property that each edge of the given tetrahedron is perpendicularly bisected by an edge of the dual. The vertices of the two tetrahedra together give the vertices of a cube. Let a and c be the position vectors relative to the origin of the midpoints of a pair of parallel (but not opposite) edges e_1 and e_2. Then a and c are perpendicular. The face opposite to that determined by e_1 and e_2 has two edges which are not parallel to e_1 and e_2; let b be the position vector of the midpoint of one of these. Then b makes an angle of $\frac{2}{3}\pi$ with a, since a rotation of the cube through a third of a revolution about one of its diagonals can be seen to take a to b. Similarly, the angle between b and c is also $\frac{2}{3}\pi$. Summarizing:

$$\begin{aligned} \text{angle}(a,b) &= \tfrac{2}{3}\pi, && \rho_a\rho_b \text{ has order } 3,\\ \text{angle}(b,c) &= \tfrac{2}{3}\pi, && \rho_b\rho_c \text{ has order } 3,\\ \text{angle}(a,c) &= \tfrac{1}{2}\pi, && \rho_a\rho_c \text{ has order } 2. \end{aligned}$$

The reflecting plane corresponding to ρ_a perpendicularly bisects one edge of the tetrahedron we started with, and the opposite edge lies in this plane. Thus ρ_a is a symmetry of the tetrahedron; the same is true for ρ_b and ρ_c. Furthermore, it is also easy to see that these three reflections generate the group of all symmetries of the tetrahedron (which is just Sym(4)).

Cube. Using the same cube as in the previous example, let b and c be as before, but now let a be the position vector of the midpoint of that face which is perpendicular to e_2 but does not include the edge corresponding to b. Then

$$\begin{aligned} \text{angle}(a,b) &= \tfrac{3}{4}\pi, && \rho_a\rho_b \text{ has order } 4,\\ \text{angle}(b,c) &= \tfrac{2}{3}\pi, && \rho_b\rho_c \text{ has order } 3,\\ \text{angle}(a,c) &= \tfrac{1}{2}\pi, && \rho_a\rho_c \text{ has order } 2. \end{aligned}$$

The reflecting plane corresponding to a is parallel to and midway between a pair of opposite faces of the cube, and there are two other similar reflections which together with ρ_a generate a group isomorphic to $C_2 \times C_2 \times C_2$. The group generated by ρ_b and ρ_c is isomorphic to Sym(3), and acts on the $C_2 \times C_2 \times C_2$ by permuting the three generating reflections. The group $\text{Sym}(3) \ltimes (C_2 \times C_2 \times C_2)$ generated by ρ_a, ρ_b and ρ_c is the full symmetry group of the cube. (This extends naturally to n dimensions: the symmetry group of the "hypercube" is a Euclidean reflection group isomorphic to $\text{Sym}(n) \ltimes (C_2)^n$.)

Dodecahedron. This time we can choose a, b and c to be the position vectors of A, B and C, midpoints of three suitable edges of the regular dodecahedron, satisfying

$$\begin{aligned} \text{angle}(a,b) &= \tfrac{4}{5}\pi, && \rho_a\rho_b \text{ has order } 5,\\ \text{angle}(b,c) &= \tfrac{2}{3}\pi, && \rho_b\rho_c \text{ has order } 3,\\ \text{angle}(a,c) &= \tfrac{1}{2}\pi, && \rho_a\rho_c \text{ has order } 2. \end{aligned}$$

Each of the twelve pentagonal faces has five lines of symmetry bisecting it, and each such line determines a plane through O, the reflection in which is a symmetry of the dodecahedron. Each of these planes bisects four faces, and so we obtain a set $\mathcal{H}$ of $5 \times 12/4 = 15$ planes corresponding to fifteen reflection symmetries. Each plane in $\mathcal{H}$ passes through a uniquely determined pair of opposite edges—note that there are thirty edges altogether—and perpendicularly bisects another pair of opposite edges. Furthermore, the line through O normal to the plane bisects a third pair of opposite edges. In particular the reflection in the plane orthogonal to the position vector of the midpoint of an edge is indeed a symmetry of the dodecahedron.

The points A, B and C can be chosen as follows. Let P be the centre of one of the faces, let Q be the midpoint of one of the edges of this face and let R be one of the vertices on this edge. The plane containing the triangle QRO is in the set $\mathcal{H}$, and so its normal through O passes through points A and A' which are midpoints of opposite edges. Choose A to be the one which is on the same side of the plane QRO as the point P. Similarly, choose B on the same side of PRO as Q such that BO normal to PRO, and C on the same side of PQO as R with CO normal to PQO. It can be shown that the reflections ρ_a, ρ_b and ρ_c thus determined—that is, the reflections in the planes QRO, PRO and PQO—generate the full group of symmetries of the dodecahedron.

(This group is in fact isomorphic to $\text{Alt}(5) \times C_2$. As we have seen above, each plane in $\mathcal{H}$ determines three pairs of opposite edges; these three pairs are mutually perpendicular to each other, and the fifteen pairs of opposite edges split into five such sets of three. The symmetry group permutes these five sets and also contains the transformation $-I$, which fixes each of the sets by taking each edge to its opposite.)

Returning to the example of the tetrahedron, recall that we found three elements $r = \rho_a$, $s = \rho_b$ and $t = \rho_c$ satisfying the relations

$$r^2 = s^2 = t^2 = 1,\ rsr = srs,\ sts = tst,\ rt = tr,$$

and generating a group isomorphic to Sym(4). We can easily show that the abstract group W defined by these generators and relations is in fact isomorphic to Sym(4). Note first that $r \mapsto (1,2)$, $s \mapsto (2,3)$ and $t \mapsto (3,4)$ certainly defines a surjective homomorphism $W \to \mathrm{Sym}(4)$, as these three permutations generate Sym(4) and satisfy the required relations. Thus it remains to show that the order of W is at most 24. Now the subgroup $P = \langle r, s\rangle$ contains at most six elements, namely 1, r, s, rs, sr and $rsr = srs$. We proceed to list the right cosets of P in W. Firstly there is P itself, which is fixed by right multiplication by r and s. Right multiplication by t takes P to Pt. Right multiplying by r fixes Pt, since $(Pt)r = Prt = Pt$, and right multiplication by t takes Pt back to P. Right multiplication by s takes Pt to Pts, which is fixed by t (since $Ptst = Psts = Pts$), and taken back to Pt by s. Right multiplying by r takes Pts to $Ptsr$, but no further cosets can be found, since s and t both fix $Ptsr$ (since $(Ptsr)s = Ptrsr = Prtsr = Ptsr$ and $(Ptsr)t = Ptstr = Pstsr = Ptsr$), while r takes $Ptsr$ back to Pts. These calculations have shown that the (nonempty) set $P \cup Pt \cup Pts \cup Ptsr$ is invariant under right multiplication by the generators of W, and therefore equals the whole of W.

Similar calculations can be used to show that the group of symmetries of the cube (of order 48) is defined abstractly by three generators r, s and t and the defining relations

$$r^2 = s^2 = t^2 = 1,\ rsrs = srsr,\ sts = tst,\ rt = tr,$$

while the group of symmetries of the dodecahedron (order 120) is defined by generators r, s and t with defining relations

$$r^2 = s^2 = t^2 = 1,\ rsrsr = srsrs,\ sts = tst,\ rt = tr.$$

Lecture 3

Let Π be an arbitrary root basis in a vector space V, and let W be the associated Coxeter group. The set $\Phi = \{\, wa \mid a \in \Pi \text{ and } w \in W \,\}$ is called the *root system* of W in V; elements of Φ are called *roots*, elements of Π *simple roots*. We define $\Phi^+ = \Phi \cap \mathrm{PLC}(\Pi)$ and $\Phi^- = \{\, -v \mid v \in \Phi^+ \,\} = \Phi \cap \mathrm{NLC}(\Pi)$.

Note that when an element of Φ is expressed as a linear combination of the elements of the root basis Π, the coefficients are not necessarily rational integers. Consequently the above definition of "root system" conflicts somewhat with the definition appropriate in the study of Lie algebras. Furthermore, for us all roots have length 1, which is not the case for root systems of Lie algebras. But a root system in the Lie algebra sense becomes a root system in our sense if each root is replaced by a unit vector in the same direction.† And our definition gives a root system to every Coxeter group, not merely those which are Weyl groups of Lie algebras.

The theorem from Lecture 1 has the following result as a corollary.

Proposition. $\Phi = \Phi^+ \mathbin{\dot\cup} \Phi^-$.

Proof. The fact that $\Phi^+$and Φ^- have empty intersection is immediate from Part (2) of the definition of a root basis.

Let $b \in \Phi$. By the definition of Φ there exist $w \in W$ and $a \in \Pi$ with $b = wa$. Let $w' = w\rho_a$, noting that the lengths of w and w' differ by at most 1. Now if $l(w') \geq l(w)$ then $l(w\rho_a) \geq l(w)$ and the theorem

† For Kac-Moody Lie algebras, only the real roots in the root system of the Lie algebra correspond to elements of the root system (in our sense) of the Weyl group.

yields that $b = wa \in \Phi^+$. On the other hand if $l(w) \geq l(w')$ then $l(w'\rho_a) \geq l(w')$ and the theorem yields that $w'a \in \Phi^+$, whence $b = w'\rho_a a = w'(-a) = -w'a \in \Phi^-$. $\square$

The preceding proof shows, incidentally, that $l(w\rho_a) = l(w)$ is not a possibility, since the root b cannot simultaneously be positive and negative. So we have

Corollary. If $w \in W$ and $a \in \Pi$ then

$$l(w\rho_a) = \begin{cases} l(w)+1 & \text{if } wa \in \Phi^+ \\ l(w)-1 & \text{if } wa \in \Phi^-. \end{cases}$$

Let $a \in \Pi$ and $b \in \Phi^+$, and write $b = \sum_{c\in\Pi} \lambda_c c$ with $\lambda_c \geq 0$ for all c. Suppose that $\rho_a b \in \Phi^-$, so that $b - 2(b\cdot a)a = -\sum_{c\in\Pi} \mu_c c$ for some $\mu_c \geq 0$. We find that

$$(\lambda_a + \mu_a - 2(b\cdot a))a + \sum_{c\in\Pi\backslash\{a\}} (\lambda_c + \mu_c)c = 0.$$

Now take the scalar product with a. Since $c \cdot a \leq 0$ for all $c \in \Pi\backslash\{a\}$, whereas $a \cdot a = 1$, we conclude that the coefficient of a is nonnegative. But this gives $0 \in \mathrm{PLC}(\Pi)$, unless all the coefficients are 0, which forces $\lambda_c = \mu_c = 0$ for all $c \in \Pi\backslash\{a\}$, which in turn forces b to be a scalar multiple of a. Since $b \cdot b = 1$ whenever b is a root, the only scalar multiples of a which are roots are a and $-a$. So the only positive root b with $\rho_a b$ negative is $b = a$; thus ρ_a permutes the positive roots other than a.

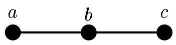
Definition. For $w \in W$ define $N(w) = \{\, b \in \Phi^+ \mid wb \in \Phi^- \,\}$

Our discussion above showed that if $a \in \Pi$ then $N(\rho_a) = \{a\}$. This combined with the corollary above yields that if $l(w\rho_a) > l(w)$ then $N(w\rho_a) = \rho_a N(w) \cup \{a\}$, and, by induction, if $w = \rho_{a_1}\rho_{a_2}\cdots\rho_{a_l}$ with $l = l(w)$ and $a_i \in \Pi$ for all i, then $N(w) = \{a_l, \rho_{a_l}a_{l-1}, \rho_{a_l}\rho_{a_{l-1}}a_{l-2}, \ldots, \rho_{a_l}\rho_{a_{l-1}}\cdots\rho_{a_2}a_1\}$. In particular, $|N(w)| = l(w)$.

If $N(w) \neq \Phi^+$ then there must be some $a \in \Pi$ for which $wa \in \Phi^+$, and then $l(w\rho_a) > l(w)$. If W is finite then it must have an element of maximal length, and hence we obtain the following result:

Proposition. *If W is finite then there exists an element $w_\Pi \in W$ with $N(w_\Pi) = \Phi^+$.*

Properties of the root system provide the most powerful tools for investigating W. Before proceeding with the theory, it may be useful to look at some specific examples of root systems.

Examples (1) Type A_3, 

.

We have $\Pi = \{a, b, c\}$ with $a \cdot b = -\frac{1}{2} = b \cdot c$ and $a \cdot c = 0$. Using the formula $\rho_a v = v - 2(v\cdot a)a$ we find that

$$\rho_a a = -a, \qquad \rho_a b = b + a, \qquad \rho_a c = c.$$

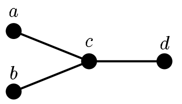
Similarly

$$\rho_b a = a + b, \qquad \rho_b b = -b, \qquad \rho_b c = c + b,$$
$$\rho_c a = a, \qquad \rho_c b = b + c, \qquad \rho_c c = -c.$$

We can quickly generate all the roots in the root system by applying ρ_a, ρ_b and ρ_c to all vectors which are already known to be roots, using the simple roots to start the process. For example, our calculations above have already shown that $a + b$ and $b + c$ are roots, and so $\rho_a(b + c) = \rho_a b + \rho_a c = a + b + c$ must also be a root. Furthermore, since the negative of every simple root is a root (for example, $-a = \rho_a a$), it follows that the negative of every root is a root. So the following twelve vectors are all roots:

$$a,\ b,\ c,\ a+b,\ b+c,\ a+b+c,\ -a,\ -b,\ -c,\ -a-b,\ -b-c \text{ and } -a-b-c.$$

It is easily shown that this set is closed under the action of ρ_a, ρ_b and ρ_c; hence it is the entire root system. Quite naturally the roots with the minus signs are called *negative roots*, the others *positive roots*.

(2) Type D_4, 

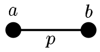
Here we have $\Pi = \{a, b, c, d\}$, with $a \cdot c = b \cdot c = d \cdot c = -\frac{1}{2}$ and $a \cdot b = b \cdot d = d \cdot a = 0$. Hence the actions of the simple reflections on the basis vectors are as follows:

$$\begin{array}{llll}
\rho_a(a) = -a & \rho_a(b) = b & \rho_a(c) = c + a & \rho_a(d) = d \\
\rho_b(a) = a & \rho_a(b) = -b & \rho_b(c) = c + b & \rho_b(d) = d \\
\rho_c(a) = a + c & \rho_c(b) = b + c & \rho_c(c) = -c & \rho_c(d) = d + c \\
\rho_d(a) = a & \rho_d(b) = b & \rho_d(c) = c + d & \rho_d(d) = -d.
\end{array}$$

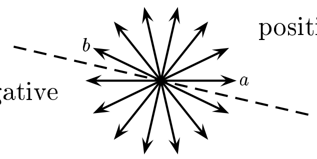
This already shows that $a + c$, $b + c$ and $d + c$ are roots. Now $\rho_a(b + c) = \rho_a(b) + \rho_a(c) = b + a + c$ must be a root, and similarly $b + d + c$ and $a + d + c$ are roots also. Continuing in this way we find that $\rho_a(b + c + d) = a + b + c + d$ is a root, and so is $\rho_c(a + b + c + d) = a + b + 2c + d$. And, as always, the negative of a root is necessarily a root. Applying the simple reflections to the roots found so far turns out to not yield any more. Hence we have found all the roots; there are twenty-four altogether. They are

$$a,\ b,\ c,\ d,\ a + c,\ b + c,\ d + c,\ a + b + c,\ b + d + c,\ d + a + c,\ a + b + c + d,\ a + b + 2c + d$$

and their negatives. Observe once again the natural division of the root system into a set of positive roots and a set of negative roots (as guaranteed by the proposition we proved above).

(3) Type $I_2(p)$, 

The group in this case is simply the dihedral group of order $2p$, and the vector space V can be identified with the Euclidean plane. The simple roots a and b are unit vectors, and the angle between them is $(p-1)\pi/p$. Since $\rho_a\rho_b$ acts as a rotation through $2\pi/p$, applying powers of $\rho_a\rho_b$ to a and $-b$ yields $2p$ distinct roots. Specifically, for each i from 0 to $2p - 1$, let v_i be the unit vector such that the anticlockwise angle from a to v_i is $i\pi/p$; these vectors v_i are the roots. The following diagram illustrates this when $p = 7$:



Observe that the roots on the positive side of the dashed line are positive combinations of a and b, since they lie "in between" a and b (in an angular sense). Those on the other side of the dashed line are the negatives of the positive roots, hence negative combinations of a and b.

If we let $\theta = \pi/p$ then the matrix of $\rho_a\rho_b$ on V relative to the basis a, b is the matrix M which appeared in the proof in Lecture 1. Using the formula for M^k given there we see that

$$(\rho_a\rho_b)^k a = \frac{\sin(2k + 1)\theta}{\sin\theta} a + \frac{\sin 2k\theta}{\sin\theta} b$$

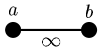
and

$$(\rho_a\rho_b)^k(-b) = \frac{\sin 2k\theta}{\sin\theta} a + \frac{\sin(2k - 1)\theta}{\sin\theta} b,$$

so that in fact the root system is

$$\Phi = \Big\{ \frac{\sin(m + 1)\theta}{\sin\theta} a + \frac{\sin m\theta}{\sin\theta} b \;\Big|\; 0 \le m < 2p \Big\}.$$

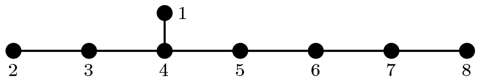
Observe that both coefficients are nonnegative when $0 \le m < p$, nonpositive when $p \le m < 2p$.

(4) The infinite dihedral group, 

In this case $a \cdot b$ can be -1 or anything smaller. If it is -1 then

$$\begin{array}{ll}
\rho_a a = -a & \rho_b a = a + 2b \\
\rho_a b = a + 2b & \rho_b b = -b.
\end{array}$$

A short calculation shows that the roots are precisely all vectors of the form $na + mb$ with $|n - m| = 1$ (as could have been predicted by letting $\theta \to 0$ in the $I_2(p)$ example).

(5) Type E_8,  .

In this case it turns out that there are 120 positive roots and 120 negative roots. They are all integral linear combinations of the simple roots; determining them as combinations of the simple roots is somewhat tedious, but not too bad really, and gives information which can be useful occasionally. Alternatively, the root system can be described as follows. Let $e_1, e_2, \ldots, e_8$ be an orthogonal basis of eight-dimensional Euclidean space, with each e_i having length $1/\sqrt{2}$. Define

$$\Phi = \{\, \pm e_i \pm e_j \mid i \neq j \,\} \cup \{\, \tfrac{1}{2} \sum_{i=1}^{8} \varepsilon_i e_i \mid \varepsilon_i = \pm 1 \text{ and } \prod_{i=1}^{8} \varepsilon_i = 1 \,\}.$$

It is not hard to check that $u \cdot u = 1$ and $u \cdot v \in \{-\frac{1}{2}, 0, \frac{1}{2}\}$ for all $u, v \in \Phi$ with $u \neq \pm v$. Thus the angle between two vectors in Φ is always either $\pi/3$, $\pi/2$ or $2\pi/3$. Define $a_2 = \frac{1}{2}(e_1 - e_2 - e_3 - e_4 - e_5 - e_6 - e_7 + e_8)$ and $a_1 = e_1 + e_2$, and, for $3 \leq i \leq 8$, define $a_i = -e_{i-2} + e_{i-1}$. Then if we take $\Pi = \{\, a_i \mid 1 \leq i \leq 8 \,\}$ it can be checked that the inner products $a_i \cdot a_j$ are as they should be for consistency with the numbering chosen above for the vertices of the Coxeter diagram. (That is, $a_i \cdot a_j = -\frac{1}{2}$ if the vertices i and j are adjacent, $a_i \cdot a_j = 0$ otherwise.) It is also not hard to check that the reflections ρ_{a_i} do indeed preserve Φ. A root $\sum_{i=1}^{8} t_i e_i$ is positive if the largest i with $t_i \neq 0$ has $t_i > 0$. The "highest" root (that is, the root $\sum n_i a_i$ with $\sum n_i$ maximal) is the positive root which is perpendicular to all the simple roots but a_8, namely $e_7 + e_8 = 3a_1 + 2a_2 + 4a_3 + 6a_4 + 5a_5 + 4a_6 + 3a_7 + 2a_8$.

It is possible to explicitly describe the linear transformations which comprise W, the Coxeter group of type E_8, as matrices relative to the basis $e_1, \ldots, e_8$. Firstly, there are the $8!$ permutation matrices, and 2^7 diagonal matrices with diagonal entries ± 1 and determinant 1. These generate a group of order $2^7 8!$ which is a subgroup S of W. The idea is to investigate the cosets of S. Now if $v_1, \ldots, v_8$ is any orthonormal basis of the space of eight component row vectors then there are $2^7 8!$ orthogonal matrices of the form xg where $x \in S$ and g is the matrix whose rows are $v_1, \ldots, v_8$. (These matrices are obtained from g by permuting the rows and multiplying an even number of rows by -1.) We proceed to describe a large number of orthonormal bases which give rise to elements of W.

Let $\eta_1, \eta_2, \ldots, \eta_8$ be signs $\eta_i = \pm 1$ with $\prod_{i=1}^{8} \eta_i = 1$, and define

$$v_{ij} = \begin{cases} \frac{3}{4} & \text{if } i = j \\ -\frac{1}{4}\eta_i \eta_j & \text{if } i \neq j. \end{cases}$$

Let v_i be the row whose jth entry is v_{ij}. Then $v_1, \ldots, v_8$ form an orthonormal basis and give rise to a coset of S in W. In fact this gives 64 cosets, corresponding to the 64 choices for the signs η_i, comprising a single (S, S) double coset in W.

Choose a division of the set $\{1, 2, \ldots, 8\}$ into two subsets J and J' of four elements each—there are thirty-five ways of doing this—and let $\varepsilon = \pm 1$. Let $J = \{j_1, j_2, j_3, j_4\}$ and $J' = \{j_5, j_6, j_7, j_8\}$, and let α_{ij} be the (i, j)-entry of the matrix

$$X = \begin{pmatrix} 1 & \varepsilon & \varepsilon & \varepsilon & 0 & 0 & 0 & 0 \\ \varepsilon & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ \varepsilon & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ \varepsilon & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \varepsilon & \varepsilon & \varepsilon \\ 0 & 0 & 0 & 0 & \varepsilon & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & \varepsilon & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & \varepsilon & -1 & -1 & 1 \end{pmatrix}$$

Let v_i be the vector whose kth entry is $\frac{1}{2}\alpha_{ij_k}$; then the matrix g whose rows are the v_i is in W. Since there were 35 possible partitions of $\{1, 2, \ldots, 8\}$ as $J \cup J'$ and two choices for ε, we have in fact obtained another 70 cosets of S, which again comprise one double coset. (The elements of this double coset are simply what you get by permuting the rows and columns of $\frac{1}{2}X$ and multiplying an even number of rows and/or columns by -1.)

We have now described $1+64+70=135$ cosets of S in W, and in view of the assertion in Lecture 2 that $|W|=2^{14}\cdot 3^5\cdot 5^2\cdot 7$, this must be all of them.

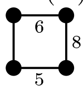
Let us return now to the general situation. For each $J\subseteq\Pi$ we define W_J to be the subgroup of W generated by $\{\,\rho_a \mid a\in J\,\}$. These subgroups are known as the *standard parabolic subgroups* of W. A *parabolic subgroup* is any conjugate of a W_J. It is immediate from the definition of root basis that any subset of a root basis is also a root basis, and so, since the the group generated by the reflections corresponding to a root basis is always a Coxeter group, parabolic subgroups of Coxeter groups are themselves Coxeter groups. In other words, we have the following result.

Proposition. *If $J\subseteq\Pi$ then W_J is isomorphic to the Coxeter group*

$$\widetilde{W}_J=\langle\,\{\,s_a\mid a\in J\,\}\mid s_a^2=1 \text{ and } (s_as_b)^{n_{ab}}=1 \text{ for all } a,\,b\in J\,\rangle.$$

To emphasize that this is not as trivial as it may appear, let us look at what it says in a particular case. Let

$$W=\langle\, r,s,t,u \mid 1=r^2=s^2=t^2=u^2=(rs)^3=(rt)^5=(ru)^2=(st)^2=(su)^6=(ut)^8\,\rangle,$$

(that is, W is the group corresponding to the diagram 

), and let $W_{\{r,s,t\}}$ be the parabolic subgroup of W generated by r, s and t. Let

$$\widetilde{W}=\langle\, r,s,t \mid 1=r^2=s^2=t^2=(rs)^3=(rt)^5=(st)^2\,\rangle.$$

It is clear that there is a homomorphism from $\widetilde{W}$ onto $W_{\{r,s,t\}}$, since the defining relations of $\widetilde{W}$ are satisfied in W. It is not clear—although this is what the proposition asserts—that the map is an isomorphism. The relations in W which involve u could have some consequences which do not involve u, since it is conceivable that u could be eliminated from some collection of equations, yielding some relation that holds in $W_{\{r,s,t\}}$ and is not a consequence of the defining relations of $\widetilde{W}$. In general, for an arbitrary presentation of a group G, one would not expect to obtain a presentation for the subgroup generated by a subset of the generators of G by simply taking those defining relations which involve only these generators. But, remarkably, it does work for Coxeter groups. The fact that it was so easy to prove demonstrates the power of the theorem we proved in Lecture 1.

If $J\subseteq\Pi$ then the root system Φ_J of W_J is the subset $\{\,wa\mid w\in W_J, a\in J\,\}$ of the root system of W. From our discussion of the sets $N(w)$ above it can be seen that $N(w)\subseteq\Phi_J^+$ whenever $w\in W_J$. And if W_J is finite then it has an element w_J of maximal length, with $N(w_J)=\Phi_J^+$.

Our next objective is to prove the following remarkable property of Coxeter groups:

Theorem. *Suppose that W is any Coxeter group of finite rank, and H is a finite subgroup of W. Then H is a subgroup of some finite parabolic subgroup of W.*

Notes.

1. The finite standard parabolic subgroups of W are easily found by inspection of the diagram of W and perusal of the list of finite Coxeter groups given in Lecture 2.
2. The theorem applies in particular to cyclic subgroups of W, and the conclusion is that the only elements of finite order are the elements of the finite parabolic subgroups.

We need a few preliminaries before starting the proof of the theorem. Firstly, we have

Lemma. *The group W is finite if and only if the root system Φ is finite.*

Proof. If $|W|<\infty$ then $|\Pi|<\infty$ and

$$|\Phi|=\bigl|\{\,wa\mid w\in W,\ a\in\Pi\,\}\bigr|\le|W|\,|\Pi|<\infty.$$

Conversely, assume that $|\Phi|=M<\infty$. For each $w\in W$ define $\sigma_w\colon\Phi\to\Phi$ by $\sigma_w(v)=wv$ for all $v\in\Phi$. Then σ_w is a permutation of Φ, and furthermore $w\mapsto\sigma_w$ is a homomorphism $\sigma\colon W\to\mathrm{Sym}(\Phi)$, the symmetric group on Φ. Now if $w \in \ker\sigma$ then $wb = b$ for all $b \in \Phi$, which in turn implies that $wa = a$ for all $a \in \Pi$. But as we saw in Lecture 1 this implies that $w = 1$, and so we conclude that σ is injective. Hence

$$|W| \le |\mathrm{Sym}(\Phi)| = M! < \infty$$

as required. □

There is a natural action of W on V^* as follows: if $w \in W$ and $\alpha \in V^*$ then $w\alpha \in V^*$ is defined by

$$(w\alpha)v = \alpha(w^{-1}v) \qquad \text{for all } v \in V.$$

(Note that if the scalar product on V were nondegenerate then the map $V \to V^*$ which takes each $v \in V$ to the linear functional $x \mapsto v \cdot x$ would be an isomorphism. This would enable us to identify V^* with V, and we would be able to avoid mention of the dual space. But the scalar product is not always nondegenerate.)

As pointed out in Note (b) to the definition of a root basis, the set

$$C = \{\, \alpha \in V^* \mid \alpha a > 0 \text{ for all } a \in \Pi \,\}$$

is nonempty. Clearly, C is a cone in V^*. Its closure is the dual of the cone PLC(Π) in V. For each $\alpha \in V^*$ let $\mathrm{Neg}(\alpha) = \{\, b \in \Phi^+ \mid \alpha b < 0 \,\}$, and observe that the closure of C is

$$\overline{C} = \{\, \alpha \in V^* \mid \mathrm{Neg}(\alpha) \cap \Pi = \emptyset \,\}.$$

If $\alpha \in \overline{C}$ then clearly $\alpha b \ge 0$ for all $b \in \Phi^+$, since $\Phi^+ \subseteq \mathrm{PLC}(\Pi)$. Conversely, if $\alpha b \ge 0$ for all $b \in \Phi^+$ then *a fortiori* $\alpha a \ge 0$ for all $a \in \Pi$. So we have that

$$\overline{C} = \{\, \alpha \in V^* \mid \alpha b \ge 0 \text{ for all } b \in \Phi^+ \,\} = \{\, \alpha \in V^* \mid \mathrm{Neg}(\alpha) = \emptyset \,\}.$$

Proposition. *Let $\alpha \in \overline{C}$ and let $J = \{\, a \in \Pi \mid \alpha a = 0 \,\}$. Then*

$$\{\, w \in W \mid w\alpha = \alpha \,\} = W_J = \{\, w \in W \mid w\alpha \in \overline{C} \,\}.$$

Proof. For all $a \in J$ and $v \in V$, we have

$$(\rho_a\alpha)v = \alpha(\rho_a v) = \alpha(v - 2(v \cdot a)a) = \alpha v$$

since $\alpha a = 0$, and so $\rho_a\alpha = \alpha$. By induction on the length of w we conclude that $w\alpha = \alpha$ for all $w \in W_J$. So $W_J \subseteq \{\, w \in W \mid w\alpha = \alpha \,\}$.

Since it is trivial that $\{\, w \in W \mid w\alpha = \alpha \,\} \subseteq \{\, w \in W \mid w\alpha \in \overline{C} \,\}$, it remains only to prove that $\{\, w \in W \mid w\alpha \in \overline{C} \,\} \subseteq W_J$. So let $w \in W$ and suppose that $w\alpha = \alpha' \in \overline{C}$. If $w \ne 1$ we may choose $a \in \Pi$ with $l(w\rho_a) = l(w) - 1$, and write $w\rho_a = w_1$. Then $w_1 a \in \Phi^+$, and

$$0 \le \alpha a = (w^{-1}\alpha')(a) = \alpha'(wa) = \alpha'(w_1(\rho_a a)) = -\alpha'(w_1 a) \le 0$$

since $\alpha' \in \overline{C}$. It follows that $\alpha a = 0$, and so $a \in J$. Hence, by the first part of this proof, $\rho_a\alpha = \alpha$. Therefore

$$w_1\alpha = w_1\rho_a\alpha = w\alpha = \alpha',$$

and, furthermore, $w_1 W_J = w_1\rho_a W_J = wW_J$. Now if $w_1 \ne 1$ we may repeat the argument and find $w_2 \in W$ with $l(w_2) = l(w_1) - 1$ such that $w_2\alpha = \alpha'$ and $w_2 W_J = w_1 W_J = wW_J$. And then if $w_2 \ne 1$ we can repeat the argument again, and so construct a sequence $w, w_1, w_2, w_3, \ldots$ of elements of W. Since the length decreases at each step the sequence must terminate; however, the process can be continued so long as $w_i \ne 1$. Thus $w_i = 1$ for some i, and

$$wW_J = w_1 W_J = w_2 W_J = \cdots = w_i W_J = W_J,$$

proving that $w \in W_J$, as required. $\square$

It is a consequence of this proposition that if $\alpha \in V^*$ can be mapped into $\overline{C}$ by an element of W then its stabilizer in W is a parabolic subgroup. In the case that $|W| < \infty$, every orbit of W on V^* contains an element of $\overline{C}$; however, in the infinite case, this is no longer true.

Proposition. $\{ w\alpha \mid \alpha \in \overline{C} \text{ and } w \in W \} = \{ \beta \in V^* \mid |\mathrm{Neg}(\beta)| < \infty \}$.

Proof. Let $\alpha \in \overline{C}$ and $w \in W$. If $b \in \Phi^+$ then $(w\alpha)b = \alpha(w^{-1}b)$ is nonnegative if $w^{-1}b \in \Phi^+$, and so

$$\mathrm{Neg}(w\alpha) = \{ b \in \Phi^+ \mid \alpha(w^{-1}b) < 0 \} \subseteq \{ b \in \Phi^+ \mid w^{-1}b \in \Phi^- \} = N(w^{-1}),$$

a finite set. So $\{ w\alpha \mid \alpha \in \overline{C} \text{ and } w \in W \} \subseteq \{ \beta \in V^* \mid |\mathrm{Neg}(\beta)| < \infty \}$.

Conversely, let $\beta \in V^*$ be such that set $\mathrm{Neg}(\beta)$ is finite; we show that $\beta = w\alpha$ for some $\alpha \in \overline{C}$ and $w \in W$. If $\mathrm{Neg}(\beta) = \emptyset$ then $\beta \in \overline{C}$, and so we may take $\alpha = \beta$ and $w = 1$. Otherwise $\beta \notin \overline{C}$, and so $\mathrm{Neg}(\beta) \cap \Pi \neq \emptyset$, and we may choose $a \in \Pi$ with $\beta a < 0$. Observe that $a \notin \mathrm{Neg}(\rho_a\beta)$, since $(\rho_a\beta)(a) = -\beta a > 0$. But ρ_a permutes $\Phi^+ \backslash \{a\}$, and if $b \in \Phi^+ \backslash \{a\}$ then $b \in \mathrm{Neg}(\rho_a\beta)$ if and only if $(\rho_a\beta)b = \beta(\rho_a b)$ is negative, which holds if and only $\rho_a b \in \mathrm{Neg}(\beta)$. So $\mathrm{Neg}(\rho_a\beta) = \rho_a(\mathrm{Neg}(\beta)\backslash\{a\})$ has fewer elements than $\mathrm{Neg}(\beta)$, and arguing by induction on the size of this set we deduce that $\rho_a\beta = w'\alpha$ for some $\alpha \in \overline{C}$ and $w' \in W$. Thus $\beta = w\alpha$ where $w = \rho_a w' \in W$, as required. $\square$

Define $U = \{ \beta \in V^* \mid |\mathrm{Neg}(\beta)| < \infty \}$. This set is a cone in V^*, known as the *Tits cone*. The proposition showed that $U = \bigcup_{w \in W}(w\overline{C})$; thus the action of W preserves U. Clearly W also preserves U°, the topological interior of U. Recall also that stabilizers of elements of $U$are parabolic subgroups.

Proposition. *The interior of U consists of those $\alpha \in U$ which have finite stabilizer in W.*

We defer the proof of this, and instead deduce from it the theorem about finite subgroups of W. Let H be a finite subgroup, and choose $\alpha \in C$. Put $\beta = \sum_{h \in H} h\alpha$. Since $\alpha \in U^\circ$, and U° is a W-invariant cone, it follows that $\beta \in U^\circ$. Hence the stabilizer of β is a finite parabolic subgroup of W. Since H is contained in the stabilizer of β, the theorem is proved. $\square$

Lecture 4

Much of today's lecture is based on a 1994 PhD thesis by D. Krammer (from the University of Utrecht). But first, let us attend to the proof that was deferred.

Lemma. *Let $\alpha \in \overline{C}$ and let $J = \{ a \in \Pi \mid \alpha a = 0 \}$. Then $\alpha \in U^\circ$ if and only if W_J is finite.*

Proof. Since $C \neq \emptyset$ we may choose a $\beta \in C$. Now every neighbourhood of α will contain a point γ on the line $\{ (1-t)\alpha + t\beta \mid t \in \mathbb{R} \}$ with $t < 0$. Then $\gamma a < 0$ for all $a \in J$, and hence $\gamma b < 0$ for all $b \in \Phi_J^+$, the root system of W_J. If W_J is infinite then so is Φ_J^+, whence $\mathrm{Neg}(\gamma)$ is infinite, and $\gamma \notin U$.

Suppose on the other hand that W_J is finite, and put

$$\delta = \min\{ \alpha a/(w\beta)a \mid a \in \Pi\backslash J \text{ and } w \in W_J \}.$$

This is a positive number, since $N(w^{-1}) \subseteq \Phi_J$ whenever $w \in W_J$, so that $w^{-1}a \in \Phi^+$ and $\beta(w^{-1}a) > 0$ for $a \in \Pi\backslash J$. We have

$$2|W_J|\alpha a > \delta \sum_{w \in W_J} (w\beta)a \qquad (\$)$$

for all $a \in \Pi\backslash J$. Furthermore, since $\sum_{w \in W_J}(w\beta)$ is fixed by all elements of W_J it must vanish on J. So the inequality ($) will hold whenever a is a positive linear combination of Π in which at least one element of $\Pi\backslash J$ appears with nonzero coefficient. So for all $a \in \Phi^+\backslash\Phi_J$,

$$2|W_J|\alpha a > \delta \sum_{w \in W_J} (w\beta)a > \delta(\beta a)$$

(since $\beta(w^{-1}a) > 0$ for all $w \in W_J$). In particular, $\beta' = 2|W_J|\alpha - \delta\beta$ has the property that $\beta' a > 0$ for all $a \in \Phi^+ \backslash \Phi_J$, and $\beta' a < 0$ for all $a \in \Phi_J^+$. It follows that $w_J\beta' \in C \subseteq U^\circ$, and hence β, $\beta' \in U^\circ$. But since U° is a cone it follows that $(1/2|W_J|)(\beta' + \delta\beta) \in U^\circ$; that is, $\alpha \in U^\circ$, as required. □

Corollary. *The interior of U consists of those elements of U which have finite stabilizer in W.*

Proof. Let $\alpha \in U$. Then there exists $w \in W$ such that $w\alpha \in \overline{C}$, and $\alpha \in U^\circ$ if and only if $w\alpha \in U^\circ$. Now the stabilizer of α is the parabolic subgroup $w^{-1}W_Jw$, where $J = \{\, a \in \Pi \mid (w\alpha)a = 0 \,\}$, and W_J is finite if and only if $w\alpha \in U^\circ$. □

A cone is called *solid* if it has nonempty interior, and *pointed* if there is no nonzero x such that x and $-x$ are both in the cone. A *proper cone* is a closed, solid, pointed cone. If W is an irreducible Coxeter group such that the scalar product is not positive semidefinite (so that there exists $v \in V$ with $v \cdot v < 0$) then it can be shown that the closure of the Tits cone is proper. For the so-called "affine Weyl groups" the scalar product is positive semidefinite but not positive definite; in this case the closure of Tits cone is a half-space. For the Euclidean (positive definite) case, the Tits cone is the whole of V^*.

There is a famous theorem of Frobenius which asserts that a matrix with nonnegative entries has an eigenvalue ρ such that

(1) ρ is real,

(2) if $\lambda \in \mathbb{C}$ is any other eigenvalue then $|\lambda| \leq \rho$,

(3) there exists a ρ-eigenvector whose entries are nonnegative.

(Property (2) says that ρ is the *spectral radius* of the matrix.)

There is a generalization of this theorem which states that if ϕ is a linear operator on a real vector space such that $\phi K \subseteq K$ for some proper cone K, then ϕ has an eigenvalue ρ satisfying (1) and (2), and there is a ρ-eigenvector in K. (See A. Berman and R. J. Plemmons "Nonnegative matrices in the mathematical sciences" (Academic Press, New York, 1979).)

When W is irreducible and V is not positive semidefinite then we have the interesting situation that a whole group of linear operators preserves a proper cone. In this case it can also be shown that if $w \in W$ is not contained in any proper parabolic subgroup then the spectral radius of w is strictly greater than 1. So there exists $\alpha \in \overline{U}$ such that $w\alpha = \lambda\alpha$ for some $\lambda > 1$. Since the same reasoning can be applied to w^{-1}, we find that there also exists $\beta \in \overline{U}$ with $w\beta = \lambda^{-1}\beta$. If α were actually in U, we would be able to find $g \in W$ such that $g\alpha \in \overline{C}$, and then we would have $(gwg^{-1})(g\alpha) = \lambda(g\alpha)$, showing that $g\alpha$ and $\lambda(g\alpha)$ are distinct points of $\overline{C}$ in the same W-orbit. This contradicts a proposition from last lecture. So α, and similarly β, lie in $\overline{U} \backslash U$.

Let us investigate the orbits on the root system of a cyclic subgroup $\langle w \rangle$ of W. If $a \in \Phi$ and there exists an integer $k > 0$ such that $w^k a = a$ then we say that a is *w-periodic*. If a is not w-periodic then the roots of the form $w^k a$ (where $k \in \mathbb{Z}$) are all distinct; but there are only finitely many roots in $N(w)$. So there exists an M such that $\pm w^k a \notin N(w)$ whenever $k > M$ or $k < -M$. Hence if $k > m$ and $w^k a \in \Phi^+$ then $w^{k+1}a \in \Phi^+$, and similarly if $w^k a \in \Phi^-$ then $w^{k+1}a \in \Phi^-$. Likewise, if $k < -m$ and $w^k a \in \Phi^+$ then $w^{k-1}a \in \Phi^+$, and if $w^k a \in \Phi^-$ then $w^{k-1}a \in \Phi^-$.

Definition. We say that a root a is *w-outward* if there exists an M such that $w^k a \in \Phi^-$ for all $k > M$ and $w^k a \in \Phi^+$ for all $k < -M$. We say that a is *w-inward* if $-a$ is w-outward, and *w-even* if it is neither inward nor outward, and not periodic.

Note that if a, a' lie in the same $\langle w \rangle$ orbit, then if a is w-outward so is a'. The same applies for inward, even and periodic. Furthermore, if $\mathcal{O}$ is a $\langle w \rangle$ orbit then $\{\, ga \mid a \in \mathcal{O} \,\}$ is a $\langle gwg^{-1} \rangle$ orbit, and since $N(g)$ is finite it is easily seen that if $\mathcal{O}$ consists of w-outward (resp. inward, even, periodic) roots then $gcalO$ consists of gwg^{-1}-outward (resp. inward, even, periodic) roots. In particular, the number of outward orbits is a conjugacy class invariant.

Suppose, as above, that α, $\beta \in \overline{U} \backslash U$ are eigenvectors for w with eigenvalues $\lambda > 1$ and λ^{-1} respectively. Suppose further that $\gamma = \alpha + \beta \in U^\circ$ (as seems geometrically reasonable, and can in fact be proved with some effort). Replacing w by gwg^{-1}, which replaces $\alpha+\beta$ by $g(\alpha+\beta)$, we can assume that $\gamma\in\overline{C}$. Let $J=\{\,a\in\Pi\mid\gamma a=0\,\}$, and note that W_J is a finite subgroup, since $\gamma\in U^\circ$. If $a\in\Phi$ then $\gamma(w^ia)=\lambda^{-i}(\alpha a)+\lambda^i(\beta a)$; if this is positive then $w^ia\in\Phi^+$, if negative then $w^ia\in\Phi^-$ (since $\gamma\in\overline{C}$). If $\gamma(w^ia)=0$ then $w^ia\in\Phi_J$ (a finite set). If $\alpha a>0$ and $\beta a<0$ then $\lambda^{-i}(\alpha a)+\lambda^i(\beta a)$ decreases monotonically as i increases, and there will be a unique value of i for which w^ia is positive and $w^{i+1}a$ is negative. In other words, the $\langle w\rangle$-orbit containing a is outward and contains a unique element of $N(w)$. If αa is negative and βa is positive then the orbit is inward and contains no element of $N(w)$. If αa and βa are both nonnegative, or both nonpositive, then the orbit is even and and contains no element of $N(w)$, unless $\alpha a=\beta a=0$, in which case the orbit is periodic and contained in Φ_J. So the length of w (the cardinality of $N(w)$) is the number of outward orbits plus $|N(w)\cap\Phi_J|$. Consider now a power w^k of w, where $k>0$. In each w-outward orbit there will be precisely k elements of $N(w^k)$. (If i is the unique integer such that $w^ia\in\Phi^+$ and $w^{i+1}a\in\Phi^-$ then the elements of the orbit which are in $N(w^k)$ are precisely w^ja for $i-k<j\le i$.) The length of w^k will thus be k times the number of outward orbits plus $|N(w^k)\cap\Phi_J$. Since Φ_J is finite, this gives us the following interesting result.

Proposition. *With w as above, $\lim_{k\to\infty}l(w^k)/k$ equals the number of w-outward orbits on Φ.*

Note also that since $l(w^k)$ and $l(gw^kg^{-1})$ can differ by at most $2l(g)$, the limit $\lim_{k\to\infty}l(w^k)/k$ is a conjugacy class invariant.

We close this lecture with some further facts about the Tits cone, and related things. First of all, let us determine the dual of the tits cone.

Proposition. *The dual of U is $\bigcap_{w\in W}w(\overline{\mathrm{PLC}(\Pi)})$, the set of all $v\in V$ such that wv is a nonnegative linear combination of Π for all $w\in W$.*

Proof. We have
$$\begin{aligned}U^* &= \{\,v\in V\mid \alpha v\ge 0 \text{ for all } \alpha\in U\,\}\\ &= \{\,v\in V\mid (w\alpha)v\ge 0 \text{ for all } \alpha\in\overline{C} \text{ and } w\in W\,\}\\ &= \{\,v\in V\mid \alpha(w^{-1}v)\ge 0 \text{ for all } \alpha\in\overline{C} \text{ and } w\in W\,\}\\ &= \{\,v\in V\mid w^{-1}v\in\overline{\mathrm{PLC}(\Pi)} \text{ for all } w\in W\,\}\end{aligned}$$
since $\overline{\mathrm{PLC}(\Pi)}=\mathrm{PLC}(\Pi)^{**}=\overline{C}^*$. Note also that $\overline{\mathrm{PLC}(\Pi)}=\mathrm{PLC}(\Pi)\cup\{0\}$ is the set of all nonnegative linear combinations of Π. □

Proposition. *Suppose that $v\in\mathrm{PLC}(\Pi)$ and $v\cdot a\le 0$ for all $a\in\Pi$. Then $v\in U^*$.*

Proof. We use induction on $l(w)$ to show that $wv\in\mathrm{PLC}(\Pi)$ for all $w\in W$. This holds by hypothesis when $l(w)=0$.

Assume then that $l(w)>0$, so that $N(w)\ne\emptyset$ and hence $N(w)\cap\Pi\ne\emptyset$, and choose $a\in N(w)\cap\Pi$. Then $w=w'\rho_a$ with $l(w')=l(w)-1$, and so the inductive hypothesis yields $w'v\in\mathrm{PLC}(\Pi)$. Now
$$\begin{aligned}wv &= (w'\rho_a)v\\ &= w'(v-2(v\cdot a)a)\\ &= w'v-2(v\cdot a)w'a\\ &\in \mathrm{PLC}(\Pi)\end{aligned}$$
since $w'a=-wa\in\Phi^+\subseteq\mathrm{PLC}(\Pi)$ and $-2(v\cdot a)\ge 0$. □

Proposition. *For all $v\in U^*$ we have $v\cdot v\le 0$.*

Proof. For simplicity we shall assume that Π is a basis of V, although with more care it is not necessary to assume this. Suppose that the result is false, and choose an $x\in U^*$ with $x\cdot x>0$. Scaling x we may assume that $x\cdot x=1$.

For each $v \in V$ let $S(v)$ be the coefficient sum for v; that is, if $v = \sum_{a\in\Pi} \lambda_a a$ then $S(v) = \sum_{a\in\Pi} \lambda_a$. Let $x = \sum_{a\in\Pi} \lambda_a a$, and choose M such that $\lambda_a \le M$ for all $a \in \Pi$. Now $1 = x \cdot x = \sum_{a\in\Pi} \lambda_a(a \cdot x)$, and so there exists $a \in \Pi$ with $\lambda_a(a \cdot x) \ge (1/n)$, where $n = |\Pi|$. This gives $a \cdot x \ge 1/n\lambda_a \ge 1/nM > 0$. Now

$$\rho_a x = x - 2(a \cdot x)a = \sum_{b\in\Pi} \lambda'_b b$$

where $\lambda'_b = \lambda_b$ for $b \ne a$ and $\lambda'_a = \lambda_a - 2(a \cdot x)$. So

$$S(\rho_a x) = S(x) - 2(a \cdot x) < S(x) - (2/nM).$$

Also, $0 < \lambda'_b \le M$ for all b, and $(\rho_a x) \cdot (\rho_a x) = 1$. So we may repeat the argument, each time decreasing the coefficient sum by at least $2/nM$. Eventually this gives a negative coefficient sum, contradiction. □

(The above argument is taken from a paper by G.Maxwell, J. Alg. 79 (1982), 78–97.)

Lecture 5

We have mentioned in passing that Coxeter groups arise in Lie theory. Today we shall attempt to say a little more about this.

A group G is said to have a BN-pair if it has subgroups B and N such that $B \cap N$ is a normal subgroup of N, and the quotient group $W = N/B \cap N$ is generated by a set R of involutions such that the following conditions hold:

(1) B and N together generate G,

(2) $rBr \ne B$ for all $r \in R$,

(3) $(BwB)(BrB) \subseteq BwB \cup BwrB$ for all $w \in W$ and $r \in R$.

The group W is called the *Weyl group* of the BN-pair.

Note that if $w \in W$ then $w = n(B \cap N)$ for some $n \in N$, and $wB = n(B \cap N)B = nB$.

Part (3) of the definition is of course the main axiom; it says that in some sense multiplication of (B, B) double cosets in G is not too complicated. (Indeed, it is hard to devise anything simpler which does not force B to be normal in G.) If the number of (B, B) double cosets in G is finite then it can be shown that the group W is necessarily a finite Coxeter group, with the generating set R corresponding to the simple reflections. (In this finite case it is also possible to weaken the axioms, so that the existence of the group N is not assumed, but derived as a consequence of (3).)

A fundamental property of groups with BN-pairs is the *Bruhat decomposition*:

Proposition. *If G has a BN-pair with Weyl group W then $G = \bigcup_{w\in W} BwB$ (disjoint union).*

It follows immediately from Part (3) of the definition that if W_J is a standard parabolic subgroup of the Weyl group (generated by the subset J of the generating set R) then $G_J = BW_JB$ is a subgroup of G. Not surprisingly, these are called *standard parabolic subgroups* of G. What is perhaps surprising is that they are the only subgroups of G that contain B.

As an example, consider the group $G = \mathrm{GL}_n(K)$ of $n \times n$ invertible matrices over the field K. Let B be the group of all upper triangular matrices in G, and N the group of all $n \times n$ permutation matrices. Since $B \cap N = \{1\}$ in this case, we have $W = N$.

The Bruhat decomposition can be checked as follows. Let $g \in G$ be arbitrary; we will show that pre- and post-multiplying g by suitable upper triangular matrices gives a permutation matrix w, whence $g \in BwB$. Now adding multiples of a row of g to higher rows is equivalent to premultiplying by a suitable upper triangular matrix, and adding multiples of a column to later columns is the same as postmultiplying by a suitable upper triangular matrix. So we start by finding the largest value of i such that the first entry of the ith row of g is nonzero, and then we add multiples of the ith row to higher rows to make all the other entries of the first column zero. Having done this, we add multiples of the first column to subsequent columns to make all entries of the ith row zero, except for the first entry. Now move on to the second column and repeat the process: find the lowest nonzero entry—in the jth row, say (and note that $j \neq i$)—and then add multiples of this row to higher rows so that there will only be one nonzero entry in the second column; then add multiples of the second column to subsequent columns so that the jth row is zero in all entries but the second. Then go on to the third column, and so on. Eventually we will get a monomial matrix (only one nonzero entry in each row and column), and then premultiplying by a suitable diagonal matrix will give the permutation matrix we seek.

The set R of involutory generators of W consists of the permutation matrices corresponding to the transpositions $r_i = (i, i+1)$ (for $1 \leq i \leq n-1$). For each ordered pair (i, j) with $i \neq j$ we define the *root subgroup* $X_{i,j}$ to consist of those matrices which have 1's on the diagonal and 0's in all other positions except the (i, j)th. It is easily seen that each $X_{i,j}$ is indeed a group, and isomorphic to the additive group of the field K.

Lemma. *For each i with $1 \leq i \leq n-1$ the group $B \cap r_i B r_i$ consists of all upper triangular matrices which are zero in the $(i, i+1)$th position. Furthermore, $B = X_{i,i+1}(B \cap r_i B r_i)$.*

From this we see that if $w \in W$ is arbitrary, and $1 \leq i \leq n-1$, then

$$(BwB)(Br_iB) = BwX_{i,i+1}(B \cap r_iBr_i)r_iB = BwX_{i,i+1}r_iB.$$

Now $wX_{i,i+1}w^{-1} = X_{wi,w(i+1)}$ (where $i \mapsto wi$ is the permutation corresponding to the permutation matrix w), and this is a subgroup of B if $wi < w(i+1)$. So in this case $BwX_{i,i+1}r_iB = Bwr_iB$. Suppose, on the other hand, that $wi > w(i+1)$, and note that every non-identity element of $X_{i,i+1}$ lies in $DX_{i+1,i}r_iX_{i+1,i}$, where D is the diagonal group, as follows from the 2×2 matrix formula

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} t & 0 \\ 0 & -t^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t^{-1} & 1 \end{pmatrix}.$$

We deduce that $BwX_{i,i+1}r_iB \subseteq Bwr_iB \cup BwX_{i+1,i}r_iX_{i+1,i}r_iB = Bwr_iB \cup BwB$, since $wX_{i+1,i}w^{-1}$ and $r_iX_{i+1,i}r_i$ are both contained in B. We deduce that Part (3) of the BN axioms holds.

In the case $n = 3$ the group G acts on a projective plane $\mathcal{P}$. The points of $\mathcal{P}$ can be identified with rays through the origin in K^3, and the lines of $\mathcal{P}$ with planes through the origin. Consider the following two subgroups of G:

$$P = \left\{ \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix} \right\}, \qquad Q = \left\{ \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix} \right\}.$$

Then P is the stabilizer of the ray which is the x-axis, and Q the stabilizer of the plane $z = 0$. The group G acts transitively on the rays through the origin and transitively also on the planes through the origin, and so the rays are in bijective correspondence with the cosets gP in G, and the planes in bijective correspondence with the cosets gQ in G. Thus we are able to reconstruct the projective plane from the group G, by calling the cosets of P "points" and the cosets of Q "lines". A point gP is incident with a line hQ if and only if the cosets have nonempty intersection; it is not hard to check that this condition is in agreement with the geometrical condition that the ray in question lies on the plane in question.

Motivated by this example, suppose now that G is any group with a BN-pair for which the Weyl group is the dihedral group of order 6

$$W = \mathrm{gp}\langle\, r, s \mid r^2 = s^2 = (rs)^3 = 1 \,\rangle$$

(that is, $W \cong S_3$), and let $P = B \cup BrB$ and $Q = B \cup BsB$. Then P and Q are parabolic subgroups of G, and it is not hard to show that if we take left cosets of P as points and the left cosets of Q as lines, with incidence defined as before, then we obtain a projective plane. This makes it natural to consider analogous coset geometries arising from BN-pairs with other Weyl groups.

If the Weyl group is dihedral of order $2n$,

$$W = \mathrm{gp}\langle\, r, s \mid r^2 = s^2 = (rs)^n = 1 \,\rangle,$$

then we again take the cosets of $P = B \cup BrB$ to be points and the cosets of $Q = B \cup BsB$ to be lines, but this time instead of a projective plane we obtain a geometry known as a *generalized n-gon*. A generalized triangle ($n = 3$) is just a projective plane: every two points are incident with a common line, and every two lines are incident with a common point. For a generalized quadrangle ($n = 4$), for each point p and each line l, either p is incident with l or else there is a unique line through p that intersects with l. For a generalized pentagon ($n = 5$) the condition is that for any two points p_1 and p_2, either they lie on a common line or else there is a unique line through p_1 which passes through a point p' such that there is a line through p' and p_2. In a generalized n-gon, any two objects (points or lines) should be part of at least one ordinary n-sided polygon, and there should be no subgeometries which are polygons with fewer than n sides. (So a point-line-point-line ... path from one object to another must be unique if it is of length less than n.)

When the Weyl group of a BN-pair is an arbitrary Coxeter group, the coset geometry obtained is known as a *building*. For each $J \subseteq R$ the left cosets of the standard parabolic subgroup G_J are "objects of type J" in the geometry. If the Weyl group is of type A_n (isomorphic to the symmetric group S_{n+1}) then the building is a projective geometry, the cosets of the maximal parabolic subgroups corresponding to the points, lines, projective planes, etc., which are subgeometries of the given geometry. (For example, if $G = \mathrm{GL}_{n+1}(K)$ then the (maximal) parabolic subgroup G_J corresponding to the set $J = R \backslash \{r_i\}$ is

$$G_J = \left\{ \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \,\middle|\, A \in \mathrm{GL}_i(K),\ C \in \mathrm{GL}_{n-i}(K) \right\}$$

and the cosets of G_J correspond to the i-dimensional linear subspaces of K^{n+1} (which are $(i-1)$-dimensional objects of projective n-space).)

A *chamber system* is a set C together with a collection of equivalence relations on C parametrized by elements of a set I. For each $i \in I$ the corresponding equivalence relation on C is called *i-adjacency*, and denoted by $\underset{i}{\sim}$. (So if c, $d \in C$ are i-adjacent we write $c \underset{i}{\sim} d$.) If $f = i_1 i_2 \dots i_l$ is a sequence of elements of I, an *f-gallery* from c to d is a sequence $c = c_0, c_1, \dots, c_l = d$ of elements of C with $c_{j-1} \underset{i_j}{\sim} c_j$ for all j from 1 to l.

A *building* is a chamber system Δ such that the equivalence relations are parametrized by the set R of simple reflections of a Coxeter group W, and having a *W-distance* function $\delta\colon \Delta \times \Delta \to W$ with the following property: whenever x, $y \in \Delta$ and f is a minimal length expression for an element $w \in W$, there exists an f-gallery from x to y if and only if $w = \delta(x, y)$.

Given a group with a BN-pair, let Δ be the set of left cosets of the subgroup B, and define the W-distance $\delta(gB, hB)$ to be the element $w \in W$ such that $(gB)^{-1}hB = BwB$. For each $r \in R$, chambers gB and hB are r-adjacent if and only if $gG_{\{r\}} = hG_{\{r\}}$.

Returning to the example of $\mathrm{GL}_{n+1}(K)$, note that B is the stabilizer of a *flag*: a sequence of subspaces $V_0 = \{0\} \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_{n+1} = K^{n+1}$, where $\dim V_i = i$. To be precise, V_i consists of those column vectors which are zero in their last $n + 1 - i$ entries. Since G acts transitively on the flags, the flags are in bijective correspondence with the left cosets of B, and so the flags can be regarded as the chambers of the building. Two flags are i-adjacent if and only if they coincide in all terms but the ith.

It should be emphasized that, just as there are projective planes that are not associated with $\mathrm{GL}_3(K)$ for any field K, so there are buildings that are not derived from groups with BN-pairs. One can hope to investigate buildings using geometrical concepts and methods that are not tied to standard group-theoretical concepts like subgroup and quotient group, and then possibly apply the results thus obtained to prove things about BN-pairs which cannot be proved by purely group-theoretic methods.