12:38 p.m. February 26, 2010

Essays on the structure of reductive groups

From root datum to reductive group

Bill Casselman
University of British Columbia
cass@math.ubc.ca

A **Cartan matrix** is an integral matrix $C_{\alpha,\beta}$ indexed by a finite set Δ, satisfying the conditions

(a) $C_{\alpha,\alpha} = 2$;
(b) $C_{\alpha,\beta} \le 0$ for $\alpha \ne \beta$;
(c) there exists a diagonal matrix D with positive diagonal entries such that CD is symmetric and positive definite.

A **root datum** is a quadruple $\mathcal{L} = (L, \Delta, L^\vee, \Delta^\vee)$ in which (1) L is a free $\mathbb{Z}$-module of finite rank and $L^\vee$ is its dual $\mathrm{Hom}(L, \mathbb{Z})$; (2) Δ is a finite subset of L and $\Delta^\vee$ is one of $L^\vee$; (3) implicit is a map $\alpha \mapsto \alpha^\vee$ from Δ to $\Delta^\vee$ such that $(\langle \alpha, \beta^\vee \rangle)$ is a Cartan matrix. In the literature, this is usually called a **based** root datum. For my purposes it serves as a better beginning. I'll say later something about the distinction.

A reductive group defined over $\mathbb{C}$ is an affine algebraic group whose algebraic representations are all semi-simple—i.e. may be reduced into irreducible components. Examples identified in simple terms as matrix groups are the special and general linear groups, symplectic groups, and orthogonal groups. These groups are classified by root data. On the one hand, each such group determines a root datum, and on the other to every root datum is associated such a group. I'll begin with a few examples of how root data arise from some classical groups, in order to motivate the remainder of the paper, but this essay is concerned principally with the second step, assembling the group from the datum. This assembly takes place in several main steps: (1) constructing the root system and its Weyl group; (2) constructing the Lie algebra from the root system; (3) finally, constructing the reductive group itself. I'll not present complete proofs, especially in the later parts, but I hope to make the reader comfortable with important notions. Nearly all steps will be accompanied by explicit algorithms.

My original motivation in beginning this essay was to understand the ATLAS program written by the late Fokko du Cloux, but the narrative took an unexpected turn. I hope to take up this matter again sometime.

The principal reference I rely on for the parts on Lie algebra and group is [Cohen et al.:2004]. They rely in turn primarily on [Carter:1972]. The construction of a root datum from a reductive group, which I shall only mention briefly, can be found in [Borel:1963] and [Springer:1998]. The standard references on root systems are [Bourbaki:1968] and [Humphreys:1972].

Contents

I. Examples
1. Tori
2. The general linear groups
3. The symplectic groups
II. Root systems
4. Cartan matrices
5. The roots
6. Root data of rank two
7. Roots and the Weyl group
8. From root system to base

9. Positive roots and linear orders
10. Constructing roots
11. The Weyl group
12. Cosets

As for notation, if $\mathcal{L}$ is the root datum above, let

$$\begin{aligned} L_\Delta &= \text{sublattice of } L \text{ spanned by } \Delta \\ L_{\Delta^\vee} &= \text{sublattice of } L^\vee \text{ spanned by } \Delta^\vee \\ L_\mathbb{R} &= \text{the real vector space } L \otimes \mathbb{R}\,. \end{aligned}$$

To come: dominant roots (Prop. 25 on p. 165 in [B]), chains of roots, W_Θ and C_Θ, associates, general systems (not reduced) and relative roots, affine roots and group, Coxeter element and polynomial invariants (p. 169 etc. in [B]), Tits' extended Weyl group, construction of Lie algebra, structure constants (some mention of Tits' approach) the group, Tits system, Bruhat decomposition, covering by open cells.

Then general Coxeter groups, Kac-Moody algebras and groups

Part I. Examples

1. Tori

The simplest complex reductive groups are complex tori, which are copies of some $T = \mathbb{T}_n = (\mathbb{C}^\times)^n$. The group T can be embedded in $\mathbb{C}^{2n}$ as the closed affine variety $x_i y_i - 1 = 0$ for $1 \le i \le n$. The ring of affine functions on $\mathbb{T}_n$ is therefore that of all polynomials in the coordinates $x_i^{\pm 1}$. The algebraic characters of T are the monomial maps $(x_i) \mapsto \prod x_i^{m_i}$, which form a free module $X^*(T)$ over $\mathbb{Z}$ of rank n. This is its **character group**. The characters of T form a linear basis of the affine ring.

[tori-reductive] **Proposition 1.1.** *Every algebraic representation of a complex torus is a direct sum of characters.*

Proof. A representation of T is an algebraic homomorphism from T to $\mathrm{GL}_n(\mathbb{C})$. Because it is algebraic, we may write

$$\pi(t) = \sum_\chi c_\chi \, \chi(t) \, ,$$

where each c_χ is an $n \times n$ matrix. Since π is a homomorphism and complex characters of any group are linearly dependent (crucially used in Galois theory and sometimes called Dedekind's Theorem), we can deduce that

$$I = \sum c_\chi, \quad c_\chi c_\rho = \begin{cases} c_\chi & \text{if } \chi = \rho \\ 0 & \text{otherwise,} \end{cases}$$

which implies that $\mathbb{C}^n$ is the direct sum of the images of the idempotent operators c_χ, on which T acts as χ. ▯

One consequence of this is that tori are reductive. This is the only case in which this can be proven easily.

Elements of $X^*(T)$ are often expressed additively, in exponential notation. This is consistent with the identification of $X^*(\mathbb{C}^\times)$ with $\mathbb{Z}$—thus for λ in $X^*(T)$ I'll sometimes write $t \mapsto \lambda(t)$ but also sometimes $t \mapsto t^\lambda$. In additive notation the product of characters λ and μ is written $t \mapsto t^{\lambda+\mu}$.

The **cocharacter** group $X_*(T)$ of T is the group of all multiplicative homomorphisms from $\mathbb{T}_1$ to T. These are maps of the form $x \mapsto (x^{m_i})$, hence also make up a free $\mathbb{Z}$-module of rank n. They also will be often written as exponentials. This is canonically the dual of $X^*(T)$—for λ in $X^*(T)$ and $\mu^\vee$ the pairing $\langle \lambda, \mu^\vee \rangle$ with values in $\mathbb{Z}$ is defined by the formula

$$x^{\langle \lambda, \mu^\vee \rangle} = \big(x^{\mu^\vee}\big)^\lambda = \lambda\big(\mu^\vee(x)\big) \, .$$

A torus is determined by $X^*(T)$—on the one hand its affine ring is the group algebra of $X^*(T)$, and on the other

$$T(\mathbb{C}) = \mathrm{Hom}\big(X^*(T), \mathbb{C}^\times\big) = X_*(T) \otimes \mathbb{C}^\times \, .$$

The root datum corresponding to T is $\big(X^*(T), \emptyset, X_*(T), \emptyset\big)$. In general, the L in a root datum is $X^*(T)$ for some complex torus T, and $L^\vee$ is $X_*(T)$.

2. The general linear groups

Let $G = \mathrm{GL}_n(\mathbb{C})$ be the group of invertible $n \times n$ complex matrices, and let T be the subgroup of diagonal matrices, which is isomorphic to $\mathbb{T}_n$. The Lie algebra of G is the vector space $\mathfrak{g} = \mathfrak{gl}_n$ of all $n \times n$ matrices, and T acts on it by conjugation. It acts trivially on its own Lie algebra $\mathfrak{t}$. The other eigenspaces are parametrized by pairs $i \neq j$—the space $\mathfrak{g}_{i,j}$ is spanned by the elementary matrix $E_{i,j}$ with a single non-zero entry 1 in position (i, j). We have

$$t\, E_{i,j}\, t^{-1} = (t_i/t_j)\, E_{i,j} \;\text{ if }\; t = (t_i)\,,$$

so the non-trivial eigencharacters are the maps

$$\lambda_{i,j} \colon (t_k) \mapsto t_i/t_j\,,$$

which are called the **roots** of $\mathfrak{g}$ with respect to T.

Let ε_i be the character of T taking (t_k) to t_i. The ε_i form a basis of $X^*(T)$. In additive notation we have

$$t^{\lambda_{i,j}} = t^{\varepsilon_i}/t^{\varepsilon_j} = t^{\varepsilon_i - \varepsilon_j} \;\text{ so }\; \lambda_{i,j} = \varepsilon_i - \varepsilon_j\,.$$

The **positive roots** are those corresponding to the (i, j) with $i < j$ (upper right). The root spaces $\mathfrak{g}_{i,j}$ associated to positive roots are those in the Lie algebra $\mathfrak{n}$ of the subgroup N of upper triangular unipotent matrices. This is the unipotent radical of the subgroup B of all upper triangular matrices in G, whose Lie algebra is $\mathfrak{b} = \mathfrak{t} + \mathfrak{n}$.

In multiplicative notation we have

$$\frac{t_i}{t_j} = \frac{t_i}{t_{i+1}} \cdots \frac{t_{j-1}}{t_j}$$

and in additive

$$\lambda_{i,j} = \varepsilon_i - \varepsilon_j = \sum_{i \le k < j} \varepsilon_k - \varepsilon_{k+1} = \sum_{i \le k < j} \lambda_{k,k+1}\,.$$

The $n - 1$ roots $\alpha_k = \lambda_{k,k+1}$ make up the set Δ of **simple roots** of $\mathfrak{g}$. The equation above says that every positive root is a non-negative integral combination of simple roots.

Dual to the basis ε_i of $X^*(T)$ is the basis $\varepsilon_i^\vee$ of $X_*(T)$, taking x to the diagonal matrix t with $t_k = 1$ for $i \neq k$ and $t_i = 1$. Thus for $n = 4$

$$\varepsilon_2^\vee \colon\; x \mapsto \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & x & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}.$$

It is the basis dual to (ε_i) in the usual sense, since writing additively we have

$$\langle \varepsilon_i, \varepsilon_j^\vee \rangle = \delta_{i,j}$$

and writing multiplicatively we have

$$\varepsilon_i\big(\varepsilon_j^\vee(x)\big) = x^{\delta_{i,j}}\,.$$

For each pair $i \neq j$ define the **coroot**

$$\lambda_{i,j}^\vee = \varepsilon_i^\vee - \varepsilon_j^\vee\,.$$

and for $1 \le i < n$ set $\alpha_i^\vee = \lambda_{i,i+1}^\vee$. Thus $\lambda_{i,j}^\vee(x)$ is the diagonal matrix (t_k) for which

$$t_k = \begin{cases} x & \text{if } k = i \\ 1/x & \quad k = j \\ 1 & \text{otherwise.} \end{cases}$$

The Cartan matrix of this group is the matrix $(\langle \alpha_i, \alpha_j^\vee \rangle)$. It can be calculated explicitly by calculating conjugations:

$$\begin{aligned}\langle \alpha_i, \alpha_i^\vee \rangle &= 2 \\ \langle \alpha_i, \alpha_{i+1}^\vee \rangle &= -1 \\ \langle \alpha_{i+1}, \alpha_i^\vee \rangle &= -1 \\ \langle \alpha_i, \alpha_j^\vee \rangle &= 0 \ \text{ for } \ |j-i| > 1\,,\end{aligned}$$

so the Cartan matrix is

$$C = \begin{bmatrix} 2 & -1 & \cdot & \ldots & \cdot & \cdot \\ -1 & 2 & -1 & \ldots & \cdot & \cdot \\ \cdot & -1 & 2 & \ldots & \cdot & \cdot \\ & & \ldots & \ldots & & \\ \cdot & \cdot & \cdot & \ldots & 2 & -1 \\ \cdot & \cdot & \cdot & \ldots & -1 & 2 \end{bmatrix}.$$

It is manifestly symmetric. It is also positive definite, since it is the matrix $(\alpha_i \bullet \alpha_j)$ if the inner product is that inherited from the Euclidean norm

$$\left\| \sum x_i \varepsilon_i \right\|^2 = \sum x_i^2\,.$$

The root datum of G is the quadruple $\big(X^*(T), \Delta, X_*(T), \Delta^\vee\big)$,

Closely related to the group GL_n are the groups SL_n, matrices of determinant 1, and PGL_n, the quotient of GL_n by scalar matrices. The only essential difference for these is the specification of the torus T. In the first case it is the group of diagonal matrices (t_k) with $\prod t_k = 1$, and in the second the quotient of $\mathbb{T}_n$ by scalars. In both cases T is isomorphic to $\mathbb{T}_{n-1}$. The sets Δ and $\Delta^\vee$ are inherited naturally from GL_n. For SL_n, the lattice $X_*(T)$ has the $\alpha_i^\vee$ as basis, while for PGL_n the lattice $X^*(T)$ has the α_i as basis.

Motivation for the definition of the coroots $\lambda_{i,j}^\vee$ might seem at this point a bit deficient. The real point is that associated to every root $\lambda = \lambda_{i,j}$ is a unique homomorphism I'll call λ_* from $\mathrm{SL}_2(\mathbb{C})$ to G taking the diagonal matrices of SL_2 into T and mapping $E_{1,1}$ in SL_2 to $E_{i,j}$ in GL_n. Explicitly

$$\lambda_*\colon \begin{bmatrix} a & b \\ c & d \end{bmatrix} \longmapsto (g_{k,\ell})$$

where

$$g_{k.\ell} = \begin{cases} 1 & \text{if } k = \ell \text{ but } k \ne i, j \\ 0 & k \ne \ell, i, j \\ a & k = i, \ell = i \\ b & k = i, \ell = j \\ c & k = j, \ell = i \\ d & k = j, \ell = j \end{cases}.$$

Thus for $n = 4, i = 1, j = 3$ we have

$$\lambda_*\colon \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a & 0 & b & 0 \\ 0 & 1 & 0 & 0 \\ c & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The map $\lambda^\vee$ is the composition of λ_* with

$$x \longmapsto \begin{bmatrix} x & \cdot \\ \cdot & 1/x \end{bmatrix}.$$

Why is $G = \mathrm{GL}_n(\mathbb{C})$ reductive? Since its centre is a torus, we may express an algebraic representation as a direct sum of spaces onw hich the centre acts by a single character. Each of these is stable under G, so we must show that any algebraic representation of G on which the scalar matrices act as scalars decomposes into a direct sum of irreducible reresentations. This reduces to the same claim for representations of SL_n, and is a special case of a general theorem, for which I refer to §III.7 of [Jacobson:1962]. In structural terms, every affine algebraic group that does not contain a unipotent normal subgroup is reductive. Thus the symplectic group introduced in the next section is also reductive.

3. The symplectic groups

The symplectic group Sp_{2n} is that of all $2n \times 2n$ matrices X preserving a non-degenerate anti-symmetric form such as $\sum(x_i y_{i+n} - y_i x_{i+n})$ or, after a coordinate change, satisfying the equation

$$ {}^tX\,J\,X = J $$

where

$$ J = J_n = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix} \text{ with } \omega = \omega_n = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot \\ & & \dots & & \\ \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot \end{bmatrix} . $$

Sometimes I replaces ω, but there is a good reason for the choice here, as we'll see in a moment.

The Lie algebra of this group is the tangent space of this group at I. As for any algebraic group, it may be calculated by a very useful trick. An algebraic group defined over $\mathbb{C}$ determines also a group defined over any ring extension of $\mathbb{C}$, and in particular the **nil-ring** $\mathbb{C}[\varepsilon] = \mathbb{C} \oplus \mathbb{C}\cdot\varepsilon$ with $\varepsilon^2 = 0$. The tangent space may be identified with the linear space of all matrices X such that $I + \varepsilon X$ lies in $G\big(\mathbb{C}[\varepsilon]\big)$. Here this gives us the condition

$$ \begin{aligned} {}^t(I + \varepsilon X)J(I + \varepsilon X) &= J \\ J + \varepsilon({}^tXJ + JX) &= J \\ {}^tXJ + JX &= 0\,. \end{aligned} $$

This symplectic group contains a copy of GL_n made up of matrices

$$ \begin{bmatrix} X & \\ & \omega^{-1}\cdot{}^tX^{-1}\cdot\omega \end{bmatrix} $$

for arbitrary X in $\mathrm{GL}_n(k)$, and also unipotent matrices

$$ \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} $$

with ωX symmetric (that is to say, symmetric with respect to reflection in the SW-NE axis). The subgroup T of diagonal matrices in Sp_{2n}, which are those like

$$ \begin{bmatrix} a_1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & a_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & a_3 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & a_3^{-1} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_2^{-1} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & a_1^{-1} \end{bmatrix} $$

($n = 3$ here) is a maximal algebraic torus, of dimension n, with coordinates ε_i inherited from GL_n. The roots are in these terms the

$$\begin{gathered} \pm\varepsilon_i \pm \varepsilon_j \quad (i < j) \\ \pm 2\varepsilon_i \end{gathered}$$

and the co-roots

$$\begin{gathered} \pm\widehat{\varepsilon}_i \pm \widehat{\varepsilon}_j \quad (i < j) \\ \pm\widehat{\varepsilon}_i \,. \end{gathered}$$

It is significant that the two systems differ in the factor 2.

The simple roots are the $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$ for $i < n$ along with $\alpha_n = 2\varepsilon_n$. The first $n-1$ arise from the embedding of GL_n mentioned above, while the last comes about from a different embedding of SL_2:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \longmapsto \begin{bmatrix} I_{n-1} & \cdot & \cdot & \cdot \\ \cdot & a & b & \cdot \\ \cdot & c & d & \cdot \\ \cdot & \cdot & \cdot & I_{n-1} \end{bmatrix} .$$

The Cartan matrix is

$$C = \begin{bmatrix} 2 & -1 & \cdot & \ldots & \cdot & \cdot \\ -1 & 2 & -1 & \ldots & \cdot & \cdot \\ \cdot & -1 & 2 & \ldots & \cdot & \cdot \\ & & \ldots & \ldots & \ldots & \\ \cdot & \cdot & \cdot & \ldots & 2 & -1 \\ \cdot & \cdot & \cdot & \ldots & -2 & 2 \end{bmatrix} .$$

It is not symmetric, but multiplication on the right by the diagonal matrix (d_i) with

$$d_i = \begin{cases} 1 & \text{if } i < n \\ 2 & \text{if } i = n \end{cases}$$

makes it into the symmetric matrix

$$C = \begin{bmatrix} 2 & -1 & \cdot & \ldots & \cdot & \cdot \\ -1 & 2 & -1 & \ldots & \cdot & \cdot \\ \cdot & -1 & 1 & \ldots & \cdot & \cdot \\ & & \ldots & \ldots & \ldots & \\ \cdot & \cdot & \cdot & \ldots & 2 & -2 \\ \cdot & \cdot & \cdot & \ldots & -2 & 4 \end{bmatrix} ,$$

which is also $(\alpha_i \bullet \alpha_j)$ and hence positive definite. Here the inner product is, as for GL_n, inherited from the Euclidean inner product

$$\varepsilon_i \bullet \varepsilon_j = \delta_{i,j} \,.$$

The positive roots again correspond to upper nilpotent matrices, corresponding to the subgroup of unipotent upper triangular matrices. This is what motivates the choice of ω in the definition of J.

What has been shown in this section and the previous one holds very generally. If G is a connected complex reductive group, there exists in G a maximal complex torus T which is its own centralizer. The adjoint action of T on $\mathfrak{g}$ decomposes into the direct sum of the trivial action on $\mathfrak{t}$ and a direct sum of one-dimensional eigenspaces $\mathfrak{g}_\lambda$. The λ in $X^*(T)$ that appear are called the roots of the group. Every root together with an isomorphism $u\colon \mathbb{C} \cong \mathfrak{g}_\lambda$ gives rise to a unique homomorphism from $\mathrm{SL}_2(\mathbb{C})$ to G with the image of the diagonal matrices contained in T, compatible with the identification of $E_{1,1}$ in SL_2 and $u(1)$ in G. Restricted to the diagonal matrices, this defines $\lambda^\vee$. If we fix a Borel subgroup B of G containing T, we are in effect choosing positive roots as those occurring in the nilpotent radical of $\mathfrak{b}$. All positive roots may be expressed as integral non-negative combinations of simple roots. If Δ is the set of simple roots, the root datum of G is $(X^*(T), \Delta, X_*(T), \Delta^\vee)$.

Part II. Root systems

4. Cartan matrices

Recall that a **Cartan matrix** is an integral matrix $C_{\alpha,\beta}$ indexed by a finite set Δ, satisfying the conditions

(a) $C_{\alpha,\alpha} = 2$;
(b) $C_{\alpha,\beta} \le 0$ for $\alpha \ne \beta$;
(c) there exists a diagonal matrix D with positive diagonal entries such that CD is symmetric and positive definite.

The last condition implies that $C_{\alpha,\beta} = 0$ if and only if $C_{\beta,\alpha} = 0$. The first two conditions are trivial to verify, but the last is not at all trivial. If C is a Cartan matrix, then so also is every 2×2 diagonal submatrix

$$C = \begin{bmatrix} 2 & C_{\alpha,\beta} \\ C_{\beta,\alpha} & 2 \end{bmatrix} .$$

I'll look at these first. If C is a Cartan matrix then $C_{\alpha,\beta}$ and $C_{\beta,\alpha}$ are non-positive integers, and some

$$\begin{bmatrix} 2 & C_{\alpha,\beta} \\ C_{\beta,\alpha} & 2 \end{bmatrix} \begin{bmatrix} d_\alpha & 0 \\ 0 & d_\beta \end{bmatrix} = \begin{bmatrix} 2d_\alpha & C_{\alpha,\beta}d_\beta \\ C_{\beta,\alpha}d_\alpha & 2d_\beta \end{bmatrix}$$

with d_α, $d_\beta > 0$ will be symmetric and positive definite. If it is symmetric then it will be positive definite if and only if its determinant

$$d_\alpha d_\beta(4 - C_{\alpha,\beta}C_{\beta,\alpha}) > 0$$

which implies that, swapping α and β if necessary, either both $C_{\alpha,\beta} = C_{\beta,\alpha} = 0$ or $C_{\beta,\alpha} = -1$, $-3 \le C_{\alpha,\beta} \le -1$. Hence:

[twobytwoCartan] **Proposition 4.1.** *Up to transposition the only* 2×2 *Cartan matrices are*

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

and

$$\begin{bmatrix} 2 & -n \\ -1 & 2 \end{bmatrix}$$

with $1 \le n \le 3$.

The first case here is uninteresting and need not be considered further. In the second, the product CD is then

$$\begin{bmatrix} 2d_\alpha & -cd_\beta \\ -d_\alpha & 2d_\beta \end{bmatrix} .$$

The condition on D is invariant under positive scalar multiplication, so we may assume $d_\beta = 1/2$. Since CD is to be symmetric, this forces $d_\alpha = c/2$, making the product

$$\begin{bmatrix} c & -c/2 \\ -c/2 & 1 \end{bmatrix} ,$$

which is indeed positive definite. At this point, what we know about $n \times n$ Cartan matrices C is that each diagonal 2×2 submatrix of CD is proportional to one of four types lised above. The condition on diagonal 2×2 submatrices is thus simple to check. But the full condition on the signature is not easy to verify, and what one really does in practice is to consult a list of all possibilities, discovered a long time ago.

There is a convenient way to display this list. Every Cartan matrix corresponds to its **Dynkin diagram**. It is a partially oriented, labeled graph. The nodes of this graph are the elements in Δ. There is an edge between α and β if $C_{\alpha,\beta} \neq 0$. It is directed from α to β with label m if $C_{\alpha,\beta} = -m$. The labels are usually indicated by multiple links.

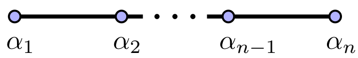
The possible Cartan matrices have all been classified. A Cartan matrix is called reducible if $\Delta = \Delta_1 \sqcup \Delta_2$ with $C_{\alpha,\beta} = 0$ for all $\alpha \in \Delta_1$, $\beta \in \Delta_2$, otherwise irreducible. A Cartan matrix is irreducible if and only if its Dynkin graph is connected. Relabeling if necessary, any Cartan matrix may be written as a direct sum of Cartan matrices that are irreducible. So it suffices to classify the irreducible Cartan matrices. Here is the list, following the conventions of [Bourbaki:1968].

• A_n

This is the root system corresponding to SL_n.

$$C_{i,j} = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } |i-j| = 1 \\ 0 & \text{if } |i-j| > 1 \end{cases}$$

$$C = \begin{bmatrix} 2 & -1 & \cdot & \ldots & \cdot & \cdot \\ -1 & 2 & -1 & \ldots & \cdot & \cdot \\ \cdot & -1 & 2 & \ldots & \cdot & \cdot \\ & & \ldots & \ldots & \ldots & \\ \cdot & \cdot & \cdot & \ldots & 2 & -1 \\ \cdot & \cdot & \cdot & \ldots & -1 & 2 \end{bmatrix}$$

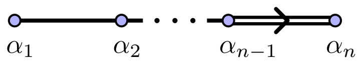


• B_n

This is the Cartan matrix of the orthogonal group of the quadratic form

$$\begin{bmatrix} 0 & 0 & \omega_n \\ 0 & 1 & 0 \\ \omega_n & 0 & 0 \end{bmatrix}.$$

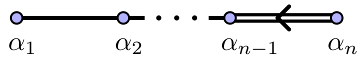
$$C = \begin{bmatrix} 2 & -1 & \cdot & \ldots & \cdot & \cdot \\ -1 & 2 & -1 & \ldots & \cdot & \cdot \\ \cdot & -1 & 2 & \ldots & \cdot & \cdot \\ & & \ldots & \ldots & \ldots & \\ \cdot & \cdot & \cdot & \ldots & 2 & -2 \\ \cdot & \cdot & \cdot & \ldots & -1 & 2 \end{bmatrix}$$



• C_n

This is the root system corresponding to $\mathrm{Sp}(2n)$.

$$C = \begin{bmatrix} 2 & -1 & \cdot & \ldots & \cdot & \cdot \\ -1 & 2 & -1 & \ldots & \cdot & \cdot \\ \cdot & -1 & 2 & \ldots & \cdot & \cdot \\ & & \ldots & \ldots & \ldots & \\ \cdot & \cdot & \cdot & \ldots & 2 & -1 \\ \cdot & \cdot & \cdot & \ldots & -2 & 2 \end{bmatrix}$$

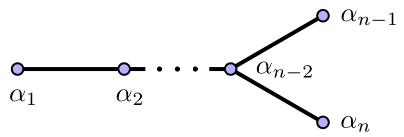


• D_n

This is the root system of the orthogonal group of the form

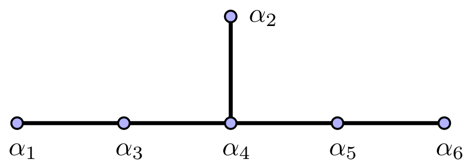
$$\begin{bmatrix} 0 & \omega_n \\ \omega_n & 0 \end{bmatrix} .$$

$$C = \begin{bmatrix} 2 & -1 & \ldots & \cdot & \cdot & \cdot & \cdot \\ -1 & 2 & \ldots & \cdot & \cdot & \cdot & \cdot \\ & & \ldots & & & & \\ \cdot & \cdot & \ldots & 2 & -1 & \cdot & \cdot \\ \cdot & \cdot & \ldots & -1 & 2 & -1 & -1 \\ \cdot & \cdot & \ldots & \cdot & -1 & 2 & \cdot \\ \cdot & \cdot & \ldots & \cdot & -1 & \cdot & 2 \end{bmatrix}$$



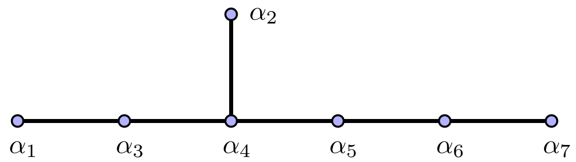
• E_6

$$C = \begin{bmatrix} 2 & \cdot & -1 & \cdot & \cdot & \cdot \\ \cdot & 2 & \cdot & -1 & \cdot & \cdot \\ -1 & \cdot & 2 & -1 & \cdot & \cdot \\ \cdot & -1 & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 & 2 & -1 \\ \cdot & \cdot & \cdot & \cdot & -1 & 2 \end{bmatrix}$$



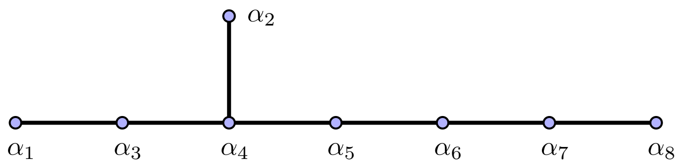
- E_7

$$C = \begin{bmatrix} 2 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 2 & \cdot & -1 & \cdot & \cdot & \cdot \\ -1 & \cdot & 2 & -1 & \cdot & \cdot & \cdot \\ \cdot & -1 & -1 & 2 & -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & -1 & 2 & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 2 \end{bmatrix}$$



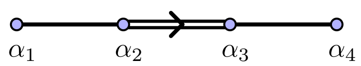
- E_8

$$C = \begin{bmatrix} 2 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 2 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot \\ -1 & \cdot & 2 & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & -1 & 2 & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 2 & -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 2 & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 2 \end{bmatrix}$$



- F_4

$$C = \begin{bmatrix} 2 & -1 & \cdot & \cdot \\ -1 & 2 & -2 & \cdot \\ \cdot & -1 & 2 & -1 \\ \cdot & \cdot & -1 & 2 \end{bmatrix}$$



• G_2

$$C = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$$

α_1 α_2

The Cartan matrices for A_n, B_n, C_n, D_n make up the infinite families, all of which have realizations as relatively simple matrix groups. The last 5 cases are isolated and called **exceptional**. They do not have simple matrix realizations.

It is useful to know that C is a Cartan matrix if and only if its transpose ${}^tC$ is, and that condition (c) in the definition of a Cartan matrix is equivalent to the condition that for some diagonal matrix D with positive entries DC is symmetric and positive definite. These assertions can both be verified easily by using D to define a positive definite inner product on $\mathbb{C}^\Delta$. As a consequence, $(L, \Delta, L^\vee, \Delta^\vee)$ is a root datum if and only if its dual $(L^\vee, \Delta^\vee, L, \Delta)$ is. This duality is at the heart of Langlands' conjectures regarding the interpretation of representations of reductive groups over local fields as well as conjectures about the occurrence of automorphic forms.

5. The roots

One might wonder why a root datum $\mathcal{L} = (L, \Delta, L^\vee, \Delta^\vee)$ does not incorporate the entire set of roots of a reductive Lie algebra. In fact, the usual definition of a root datum, as opposed to a based root datum, does exactly this. In this definition, a set of roots is characterized axiomatically. But this is somewhat awkward and also unecessary—one can construct the entire set of roots from $\mathcal{L}$. This is because the set of roots has a characteristic feature we haven't seen yet—it possesses a high degree of symmetry. The based root datum, in particular the Cartan matrix, is a more convenient starting point for computations.

We can see the symmetry easily in the example of GL_n. Here, the roots are the $\varepsilon_i - \varepsilon_j$ for $1 \le i, j \le n$. The symmetric group $\mathfrak{S}_n$, acting by permutations on the set of ε_i, permutes the roots as well. The roots for Sp_{2n} are also stable under a large group of linear transformations of L, the group generated by $\mathfrak{S}_n$ and the involutions $\varepsilon_i \mapsto \pm\varepsilon_i$.

There is a similar symmetry group in all cases. If $(L, \Delta, L^\vee, \Delta^\vee)$ is a root datum with associated Cartan matrix and α is in Δ then the linear transformation

$$s_\alpha \lambda = \lambda - \langle \lambda, \alpha^\vee \rangle \alpha$$

is a reflection, fixing the points λ of L in the hyperplane $\langle \lambda, \alpha^\vee \rangle = 0$ and taking α to $-\alpha$. Let W be the group generated by these reflections, the **Weyl group** of the datum. In the course of constructing the datum of a reductive group, it is proven that the reflections s_α take the set of roots to itself, and that every root is the transform of a root in Δ by some w in W. For us, this property is a matter of definition.

I can give at least some idea of the connection between the reflections defined purely in terms of the Cartan matrix and the roots. In the course of the construction of Δ and $\Delta^\vee$ are constructed homomorphisms λ_* from SL_2 to G. In SL_2 we have the element

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

Let T_1 be the group of diagonal matrices in SL_2. The element $\lambda_*(w)$ in G certainly conjugates $\lambda_*(T_1)$ to itself. It turns out that the centralizer of $\lambda_*(T_1)$ is T, so $\lambda_*(w)$ also conjugates T to itself, hence acts on $X^*(T)$. In fact, it acts on $X^*(T)$ exactly as s_α. In the case of GL_n, s_{α_i} swaps ε_i and ε_{i+1}.

As I have already mentioned, the Weyl group W of the root datum is the group generated by the reflections s_α. The elements of L in the W-orbit Σ of Δ are called the **roots** of the datum.

[rootsfinite] **Proposition 5.1.** *Both Σ and W are finite.*

Proof. Define an inner product on L by means of the positive definite matrix CD

$$\alpha \bullet \beta = \langle \alpha, \beta^\vee \rangle d_\beta \ \text{ in particular } \ \alpha \bullet \alpha = 2d_\alpha \,.$$

This inner product is non-degenerate on the subspace of L spanned by the simple roots. It follows immediately from the definition that

$$\langle \alpha, \beta^\vee \rangle = 2 \left(\frac{\alpha \bullet \beta}{\beta \bullet \beta} \right) ,$$

so that the formula for reflection becomes

$$s_\alpha \colon \lambda \mapsto \lambda - 2 \left(\frac{\lambda \bullet \alpha}{\alpha \bullet \alpha} \right) .$$

This means that s_α is orthogonal with respect to this inner product. Each reflection has integral coordinates, so the group W takes the lattice spanned by Δ into itself. Since it also preserves lengths, the orbit of Δ is contained in the lattice points of bounded length, hence finite. Since W embeds into the group of permutations of Σ, it too is finite. ▯

This Weyl group is $\mathfrak{S}_n$ in the case of GL_n because $\mathfrak{S}_n$ is generated by the elementary transpositions swapping i and $i+1$. (This can be proved most simply by induction.)

A **root system** in a real vector space V is a finite subset of non-zero vectors R with these properties:

(a) there exists a map $\rho \mapsto \rho^\vee$ from R to the linear dual $V^\vee$ such that $\langle \rho, \sigma^\vee \rangle \in \mathbb{Z}$ for all ρ, σ in R;
(b) the subspace of V annihilated by $R^\vee$ is complementary to the subspace spanned by R;
(c) the linear map

$$s_\rho \colon v \mapsto v - \langle v, \rho^\vee \rangle \rho$$

is a reflection that takes R to itself.

In the rest of this section I'll prove that $L_\mathbb{R}$ together with Σ form a root system.

First, the map $\alpha \mapsto \alpha^\vee$, which takes Δ to $L^\vee$, must be extended to a map defined on all of Σ.

[lambdavee] **Proposition 5.2.** *For every λ in Σ there exists a unique $\lambda^\vee$ in $L^\vee$ with these properties:*

(a) the linear map

$$s_\lambda \colon \mu \longmapsto \mu - \langle \mu, \lambda^\vee \rangle \lambda$$

is a reflection;
(b) $\lambda^\vee$ lies in the integral linear span $L_{\Delta^\vee}$ of $\Delta^\vee$ in $L^\vee$;
(c) the reflection s_λ takes Σ to itself.

As for (a), it just means that $\langle \lambda, \lambda^\vee \rangle = 2$, which guarantees that s_λ takes λ to $-\lambda$ in addition to fixing μ in the hyperplane $\langle \mu, \lambda^\vee \rangle = 0$.

Proof. The specification of a candidate for $\lambda^\vee$ is not difficult. Recall that if T is any invertible linear transformaion of $V = L_\mathbb{R}$, its contragredient $T^\vee$ is defined by the formula

$$\langle v, T^\vee u^\vee \rangle = \langle T^{-1} v, u^\vee \rangle \,.$$

If $\lambda = w\alpha$, the natural candidate for $\lambda^\vee$ is $w^\vee \alpha^\vee$. It is not at all obvious that this definition depends only on λ and not on the choices of α and w, but this will be shown momentarily.

It is relatively easy to show that, with this definition, s_λ *takes* Σ *to itself.* As before, let $V = L_{\mathbb{R}}$. Since w in W takes Σ to itself, as does s_α, the claim follows from:

[rootreflect] **Lemma 5.3.** *For* w *in* W*,* α *in* Δ*,* $\lambda = w\alpha$*,*

$$s_\lambda = w s_\alpha w^{-1} .$$

Proof. For v in V

$$\begin{aligned} s_\lambda v &= v - \langle v, w^\vee \alpha^\vee \rangle w\alpha \\ &= w(w^{-1}v - \langle w^{-1}v, \alpha^\vee \rangle \alpha) \\ &= w s_\alpha w^{-1} v . \end{aligned}$$ ∎

Since $(xy)^\vee = x^\vee y^\vee$, in order to show that $\lambda^\vee$ *lies in* $L_{\Delta^\vee}$ it suffices to show that $(s_\alpha)^\vee$ takes $L_{\Delta^\vee}$ to itself for all α in Δ. To see this, define the reflection

$$s_{\alpha^\vee} \colon u^\vee \longmapsto u^\vee - \langle \alpha, u^\vee \rangle \alpha^\vee$$

on $V^\vee$. The map $s_{\alpha^\vee}$ clearly takes $L_{\Delta^\vee}$ to itself. So the claim follows from:

[alphalambda] **Lemma 5.4.** *For* α *in* Δ

$$s_{\alpha^\vee} = (s_\alpha)^\vee .$$

Proof. For v in V, $u^\vee$ in $V^\vee$

$$\begin{aligned} \langle v, s_{\alpha^\vee} u^\vee \rangle &= \langle v, u^\vee - \langle \alpha, u^\vee \rangle \alpha^\vee \rangle \\ &= \langle v, u^\vee \rangle - \langle v, \alpha^\vee \rangle \langle \alpha, u^\vee \rangle \end{aligned}$$

while

$$\begin{aligned} \langle v, (s_\alpha)^\vee u^\vee \rangle &= \langle s_\alpha v, u^\vee \rangle \\ &= \langle v, u^\vee \rangle - \langle v, \alpha^\vee \rangle \langle \alpha, u^\vee \rangle . \end{aligned}$$ ∎

So now we know that the above definition of $\lambda^\vee$ satisfies the required conditions. We do not yet know the definition to be valid—i.e. that the definition of $\lambda^\vee$ as $w^\vee \alpha^\vee$ is independent of the choice of w and α. This follows from:

[bourbaki-uniqueness] **Lemma 5.5.** *Suppose* R *to be any finite subset spanning a real vector space* V*,* ρ *in* R*. There exists at most one reflection in* $\mathrm{GL}(V)$ *taking* ρ *to* $-\rho$*,* R *to itself.*

Proof. If there were two, their product T would take R to itself and act as the identity on V modulo $\mathbb{R}{\cdot}\rho$. Then

$$Tv = v + f(v)\rho$$

for some function f on L_R which is necessarily linear and vanishes on ρ. Thus

$$T^m v = v + m f(v)\rho$$

for all v. But since T preserves R, which spans V, it must be of finite order, and f must vanish everywhere. ∎

[lambdavee] **Lemma 5.6.** *For every* λ *in* Σ *there exists a unique* $\lambda^\vee$ *in* $L_{\Delta^\vee}$ *such that the reflection* s_λ *preserves* Σ*.*

Proof. If it lies in $L_{\Delta^\vee}$, it will vanish on the intersection of the kernels of the $\alpha^\vee$ in $L_{\mathbb{R}}$, which is complementary to the vector space V spanned by Σ. Since it preserves Σ it preserves V. The Lemma implies there is at most one reflection that acts properly, and there will be exactly one $\lambda^\vee$ such that s_λ is that reflection. ∎

♣ [lambdavee] This concludes as well the proof of Lemma 5.6, which implies in turn:

[rootsys] **Corollary 5.7.** *The set Σ, together with the map $\lambda \mapsto \lambda^\vee$, is a root system in $L_\mathbb{R}$.*

I'll call a **full root datum** a quadruple $(L, R, L^\vee, R^\vee)$ (together with an implicit map from R to $R^\vee$) in which (1) L is a free $\mathbb{Z}$-module and $L^\vee$ its dual; (2) the subset R in $L_\mathbb{R}$ together with $\rho \mapsto \rho^\vee$ is a root system. In the literature, this is what is usually called a root datum. Root systems, which have no lattice involved except that spanned by the roots, are designed to deal with Lie algebras. Several of these root data may correspond to the same root system, just as several Lie groups may have the same Lie algebra.

We'll see later how one can go back from a full root datum to a root datum—i.e. how to define $\Delta \subseteq \Sigma$.

6. Root data of rank two

The **rank** of a root datum is the cardinality of Δ. Many results about general systems reduce to results about systems of rank two, which I analyze in this section.

If the Cartan matrix for the pair α, β is

$$\begin{bmatrix} 2 & -1 \\ -n & 2 \end{bmatrix}$$

then

$$\begin{bmatrix} 2 & -1 \\ -n & 2 \end{bmatrix} \begin{bmatrix} d_\alpha & 0 \\ 0 & d_\beta \end{bmatrix} = \begin{bmatrix} 2d_\alpha & -d_\beta \\ -nd_\alpha & 2d_\beta \end{bmatrix} = \begin{bmatrix} \alpha \bullet \alpha & \alpha \bullet \beta \\ \alpha \bullet \beta & \beta \bullet \beta \end{bmatrix},$$

leading to

$$\begin{aligned} \alpha \bullet \beta &= -(n/2)(\alpha \bullet \alpha) \\ \beta \bullet \beta &= \quad n\,(\alpha \bullet \alpha)\,. \end{aligned}$$

Therefore if α and β are connected by an edge in the Dynkin diagram, the value of a single $\alpha \bullet \alpha$ determines $\alpha \bullet \beta$ and $\beta \bullet \beta$. If the Dynkin diagram is connected, the value of one $\alpha \bullet \alpha$ determines the whole matrix $(\beta \bullet \gamma)$.

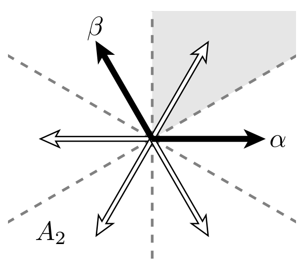
Let's look next in more detail at the rank two cases, with notation as above. If θ is the angle between α and β then

$$\begin{aligned} \cos\theta &= \frac{\alpha \bullet \beta}{\|\alpha\|\,\|\beta\|} \\ &= \frac{-(n/2)(\alpha \bullet \alpha)}{\sqrt{\alpha \bullet \alpha}\sqrt{\beta \bullet \beta}} \\ &= -\sqrt{n}/2\,. \end{aligned}$$

I'll now discuss each of the three cases in some detail. In each case, the Cartan matrix is given first, followed by a figure showing all the roots. The simple roots are in black; the other roots, which are obtained by repeatedly applying reflections in the simple ones, are outlined. Also in each figure is shown the acute cone $\{\alpha^\vee > 0\} \cap \{\beta^\vee > 0\}$, as well as the lines where roots vanish.

System A_2.

$$C = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

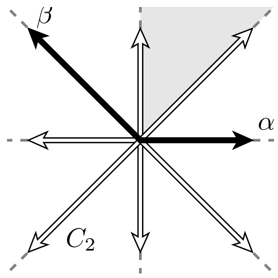


The product of the reflections s_α and s_β is a rotation through $120°$. The group W consists of

$$1,\ s_\alpha,\ s_\beta,\ s_\alpha s_\beta,\ s_\beta s_\alpha,\ s_\alpha s_\beta s_\alpha = s_\beta s_\alpha s_\beta\ .$$

System $C_2 = B_2$.

$$C = \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix}$$

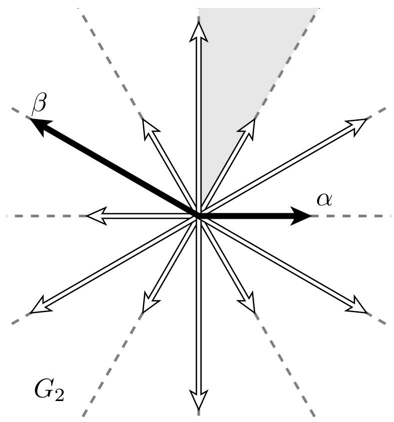


The product of the reflections s_α and s_β is a rotation through $90°$. The group W consists of

$$1,\ s_\alpha,\ s_\beta,\ s_\alpha s_\beta,\ s_\beta s_\alpha,\ s_\alpha s_\beta s_\alpha,\ s_\beta s_\alpha s_\beta,\ s_\alpha s_\beta s_\alpha s_\beta = s_\beta s_\alpha s_\beta s_\alpha\ .$$

System G_2.

$$C = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$$



The product of the reflections s_α and s_β is a rotation through $60°$. The group W consists of

$$\begin{gathered} 1,\ s_\alpha,\ s_\beta,\ s_\alpha s_\beta,\ s_\beta s_\alpha,\ s_\alpha s_\beta s_\alpha,\ s_\beta s_\alpha s_\beta,\ s_\alpha s_\beta s_\alpha s_\beta, \\ s_\beta s_\alpha s_\beta s_\alpha,\ s_\alpha s_\beta s_\alpha s_\beta s_\alpha,\ s_\beta s_\alpha s_\beta s_\alpha s_\beta,\ s_\alpha s_\beta s_\alpha s_\beta s_\alpha s_\beta = s_\beta s_\alpha s_\beta s_\alpha s_\beta s_\alpha\ . \end{gathered}$$

In each of the rank two cases, the product $s_\alpha s_\beta$ is a rotation through an angle $2\pi/m_{\alpha,\beta}$, where

$$m_{\alpha,\alpha} = 1, \quad 4\cos^2(\pi/m_{\alpha,\beta}) = C_{\alpha,\beta}C_{\beta,\alpha}\ ,$$

and more explicitly

$$m_{\alpha,\beta} = \begin{cases} 2 & \text{if } C_{\alpha,\beta}C_{\beta,\alpha} = 0 \\ 3 & \phantom{\text{if }} C_{\alpha,\beta}C_{\beta,\alpha} = 1 \\ 4 & \phantom{\text{if }} C_{\alpha,\beta}C_{\beta,\alpha} = 2 \\ 6 & \phantom{\text{if }} C_{\alpha,\beta}C_{\beta,\alpha} = 3. \end{cases}$$

For every Cartan matrix one may assign numbers according to this rule, defining what is called its **Coxeter matrix** $(m_{\alpha,\beta})$, and we always have relations

$$s_\alpha^2 = 1, \quad (s_\alpha s_\beta)^{m_{\alpha,\beta}} = 1\,.$$

It is easy to see that these are equivalent to

$$\begin{aligned} s_\gamma^2 &= 1 \quad \text{for } \gamma \in \Delta \\ s_\alpha s_\beta \ldots &= s_\beta s_\alpha \ldots \quad (m_{\alpha,\beta} \text{ terms on each side})\,. \end{aligned}$$

The second relation is called the **braid relation**.

[genrel] **Proposition 6.1.** *In rank two, the group W is defined by the generators s_α, s_β, and these relations.*

It is a **Coxeter group**.

Proof. I'll show that if $s_1 s_2 \ldots\ s_n = 1$ then the string $s_1 \diamond \ldots \diamond s_n$ is equivalent to the empty string by means of the given relations. This can be seen most easily by tracking the string around the unit circle. If it doubles back, a deletion of a pair $s_\gamma s_\gamma$ shortens it. Thus, one may assume the path is a combination of simple loops. But each loop is associated to either $(s_\alpha s_\beta)^{m_\alpha m_\beta}$ or its inverse, and may be eliminated. □

The following may be proven by inspection:

[fund2] **Proposition 6.2.** *In each case, the closure of the region*

$$C = \{\alpha > 0\} \cap \{\beta > 0\}$$

is a strict fundamental domain for W.

As may this:

[abneg] **Proposition 6.3.** *Suppose λ, μ to be two roots, and let $C = \{\lambda > 0\} \cap \{\mu > 0\}$. If C contains no part of any root line $\nu = 0$, then $\langle \lambda, \mu^\vee \rangle \le 0$.*

In fact, it can be seen that some $w\{\lambda, \mu\} = \{\alpha, \beta\}$. This Proposition holds even if $\langle \alpha, \beta^\vee \rangle = 0$.

These three cases all have a further feature in common. Let $\ell(w)$ be the length of the shortest string of reflections equal to w. For example, in the case of A_2 the length of $s_\alpha s_\beta s_\alpha$ is 3. The following may also be proved by inspection:

[wal] **Proposition 6.4.** *If γ is in Δ, then $\ell(s_\gamma w) > \ell(w)$ if and only if $\gamma > 0$ on wC.*

In other words, if and only if wC lies on the other side of the line $\gamma = 0$ from C.

Some of what I have said in this section applies to any pair of roots λ, μ. First of all, since $\mu^\vee$ lies in $L^\vee$, the pairing $\langle \lambda, \mu^\vee \rangle$ is always an integer. Since W preserves the metric and s_μ is in W, it must be an orthogonal reflection. Combining these two observations

$$2\left(\frac{\lambda \bullet \mu}{\mu \bullet \mu}\right) = \langle \lambda, \mu^\vee \rangle$$

must always be an integer. From this equation we get

$$\langle \lambda, \mu^\vee \rangle \langle \mu, \lambda^\vee \rangle = 4 \cdot \frac{(\lambda \bullet \mu)^2}{(\lambda \bullet \lambda)(\mu \bullet \mu)}\,.$$

But the cosine formula tells us that if θ is the angle θ between them then

$$\cos^2\theta = \frac{(\lambda \bullet \mu)^2}{(\lambda \bullet \lambda)(\mu \bullet \mu)}\,,$$

which implies that

$$\langle \lambda, \mu^\vee \rangle \langle \mu, \lambda^\vee \rangle$$

is 0, 1, 2, or 3 if λ and μ are linearly independent, and 4 otherwise.

[productC] **Proposition 6.5.** *For any two linearly independent roots λ, μ the product $\langle \lambda, \mu^\vee \rangle \langle \mu, \lambda^\vee \rangle$ is equal to 0, 1, 2, or 3.*

Furthermore:

[lengthratio] **Corollary 6.6.** *Suppose that λ, μ are distinct roots with $\lambda \bullet \mu \neq 0$ and $\|\mu\| \geq \|\lambda\|$. (a) If they are linearly indendent then*

$$\frac{\mu \bullet \mu}{\lambda \bullet \lambda} = 1,\ 2,\ \text{ or } 3\,.$$

(b) If they are proportional then $\mu = \pm 2\lambda$.

Proof. We have

$$(\mu \bullet \mu)\langle \lambda, \mu^\vee \rangle = (\lambda \bullet \lambda)\langle \mu, \lambda^\vee \rangle, \quad \frac{\mu \bullet \mu}{\lambda \bullet \lambda} = \frac{\langle \mu, \lambda^\vee \rangle}{\langle \lambda, \mu^\vee \rangle}\,.$$

But the right hand side can be only 1, 2, or 3.

If λ and μ are proportional, say $\mu = c\lambda$ with $c > 1$, then $\mu^\vee = \lambda^\vee/c$. Since $\langle \lambda, \mu^\vee \rangle = 2/c$ and $\langle \mu, \lambda^\vee \rangle = 2c$ are both integers, $c = 2$. ▯

[twolengths] **Proposition 6.7.** *If the Cartan matrix is irreducible, then among the roots there are at most two lengths. If λ and μ are two roots with $\|\lambda\| < \|\mu\|$ then the ratio $\|\mu\|/\|\lambda\|$ is either $\sqrt{2}$ or $\sqrt{3}$.*

In the proof, we'll need first:

[irrtrans] **Lemma 6.8.** *Assume the Cartan matrix to be irreducible. If λ and μ are two roots, there exists w in W such that $w\mu \bullet \lambda \neq 0$.*

Proof of the Lemma. Since every root is a W-transform of a simple root and W preserves lengths, we may assume μ in Δ. Let Δ_0 be the subset of α in Δ perpendicular to $W\lambda$, V_0 the subspace of V spanned by Δ_0, and V_1 its orthogonal complement, Δ_1 the intersection of Δ with V_1. By assumption, μ lies in Δ_1. Each of V_0, V_1 is W-stable, and $V = V_0 \oplus V_1$. If β is any element of Δ, the -1 eigenspace of s_β has dimension one, and hence must lie completely in either V_0 or V_1. Thus $\Delta = \Delta_0 \sqcup \Delta_1$, and the two pieces are orthogonal. Since Δ is irreducible and $\Delta_1 \neq \emptyset$, $\Delta_0 = \emptyset$. ▯

♣ [twolengths] *Proof* of the Proposition 6.7. Suppose λ, μ, ν three roots of different lengths. Since W preserves lengths, we may assume all lie in Δ, and also by permuting and scaling that

$$1 = \|\lambda\| < \|\mu\| < \|\nu\|\,.$$

♣ [lengthratio] By Lemma 6.8, all three ratios $\|\nu\|/\|\lambda\|$ etc. must satisfy the condition of Corollary 6.6 and the remark that follows it, which is possible only if

$$\|\lambda\| = 1,\ \|\mu\| = \sqrt{2},\ \|\nu\| = 2\,.$$

Since the Dynkin diagram is connected, this tells us that there exists in it a linear chain of roots in which the first node has length 1, the last length 2, and the intermediate ones length $\sqrt{2}$. So it remains to show that such a Dynkin diagram does not correspond to a Cartan matrix. Of course we can just check the list of possible Dynkin diagrams, but I offer it as an exercise that the associated Cartan matrix does not satisfy condition (c). (Hint: the choice of diagonal matrix D should be clear. Then show that CD is not invertible.) ▯

In fact the Proposition is true if and only if no such Dynkin diagram is acceptable.

[reduced] **Corollary 6.9.** *If λ and $c\lambda$ are both roots, then $c = \pm 1$.*

That is to say, the system is said to be **reduced**.

[samelength] **Proposition 6.10.** *If Δ is irreducible, two roots lie in the same W-orbit if and only if they have the same length.*

♣ [irrtrans] *Proof.* We may suppose $\|\lambda\| = \|\mu\| = 1$. By Lemma 6.8, we may suppose $\lambda \bullet \mu \neq 0$. Since the matrix

$$\begin{bmatrix} \lambda\bullet\lambda & \lambda\bullet\mu \\ \lambda\bullet\mu & \mu\bullet\mu \end{bmatrix}$$

is positive definite and $2\lambda\bullet\mu$ must be an integer, $\lambda\bullet\mu = \pm 1/2$. If $\lambda\bullet\mu = -1/2$, we may replace μ by $s_\lambda\mu$. But now λ and μ fit exactly into the diagram above for A_2 as α and β, and $s_\lambda s_\mu \lambda = \mu$. ▯

7. Roots and the Weyl group

Suppose $\mathcal{L}$ to be a root datum. *In this section, and only in this section, I shall use notation slightly different from that used elsewhere.* In the end the two systems of notation will be shown to be compatible. For each α in Δ, let $\overline{s}_\alpha$ be the simple reflection corresponding to α in Δ and let $\overline{S}$ be the set of all these simple reflections. We know from the analysis of the systems of rank 2 that the $\overline{s}_\alpha$ in $\overline{S}$ satisfy the relations

$$\overline{s}_\alpha^2 = 1, \quad (\overline{s}_\alpha \overline{s}_\beta)^{m_{\alpha,\beta}} = 1\,,$$

where $(m_{\alpha,\beta})$ is the associated Coxeter matrix. Define W to be the group defined by these generators s_α satisfying analogous relations, and let S be the set of s_α. The map $s_\alpha \mapsto \overline{s}_\alpha$ determines a homomorphism $w \mapsto \overline{w}$ from W to $\overline{W}$.

I recall the definition of W. It is the set of all strings $w = s_1 \circ \ldots \circ s_m$ (allowing the empty string) obtained by concatenating elements of S, modulo a certain equivalence. I say that $x \equiv y$ if either is obtained from the other by (a) an insertion or deletion of a duplication $s_\alpha \circ s_\alpha$ or (b) an insertion or deletion of a string $s_\alpha \circ s_\beta \circ s_\alpha \circ s_\beta \ldots$ in which there are $2m_{\alpha,\beta}$ terms all together. The product is defined by concatenation, the identity element is the empty string. It is immediately apparent that the inverse of a string is that string reversed. One easy consequence is that

$$s_\alpha \circ s_\beta \ldots = s_\beta \circ s_\alpha \ldots \quad (m_{\alpha,\beta} \text{ terms on each side}).$$

For w in W, let $\ell(w)$ be the length of the shortest string in its equivalence class. For s in S, $\ell(s\circ w)$ is either $\ell(w)+1$ or $\ell(w)-1$. Similarly $\ell(w\circ s) = \ell(w) \pm 1$. I write $s\circ w > w$ or $s\circ w < w$ depending on which case occurs. Let $V = L_\mathbb{R}$ and define

$$C = \{v \in V \mid v\bullet\alpha > 0 \text{ for all } \alpha \in \Delta\}\,.$$

The following is one of the fundamental facts about root systems. It is the main connection between the geometry of roots and the combinatorics of W.

[walext] **Theorem 7.1.** *Suppose w to be in W, $s = s_\alpha$ in S. Then*

(a) if $s\circ w > w$ then $\overline{w}C$ lies entirely in the region $\alpha > 0$;
(b) if $s\circ w < w$ then $\overline{w}C$ lies entirely in the region $\alpha < 0$.

♣ [wal] An important part of the proof is the case $|S| = 2$, which we have already seen (Proposition 6.4).

Proof. The proof begins with an observation:

[wtw] **Lemma 7.2.** *Suppose $\Theta \subset \Delta$. Let $T \subseteq S$ be the set of s_α for $\alpha \in \Theta$, and let W_T be the subgroup of W generated by T. Every w in W may be written as $x\circ y$ with (a) $x \in W_T$, $t\circ y > y$ for all t in T and (b) $\ell(w) = \ell(x) + \ell(y)$.*

I do not claim yet that x and y are unique, although that will be verified later on.

Proof of the Lemma. The following algorithm computes x and y:

$y := w$
$x := 1$
while $t{\circ}y < y$ for some t in T:
  $y := t{\circ}y$
  $x := x{\circ}t$

Since the length of y decreases in every iteration of the loop, the algorithm certainly stops. When it does so, $t{\circ}x > x$ for all t in T. In order to prove the Lemma, it suffices to verify that $w = x{\circ}y$ with $\ell(w) = \ell(x) + \ell(y)$ whenever entry into the loop is tested. These certainly hold at the first test, so it remains to see that they are not destroyed in the loop. Equality $w = x{\circ}y$ is certainly preserved since $t^2 = 1$. Since $\ell(w) = \ell(x) + \ell(y)$ to start and $\ell(w) = \ell(x{\circ}t{\circ}t{\circ}y)$, we also have

$$\begin{aligned} \ell(x) + \ell(y) &= \ell(xt) + \ell(ty) \\ &= \ell(x{\circ}t) + \ell(y) - 1 \\ \ell(x) &= \ell(x{\circ}t) - 1 \\ \ell(x{\circ}t) &= \ell(x) + 1 \\ \ell(w) &= \ell(x{\circ}t) + \ell(t{\circ}y) \end{aligned}$$

so this equality is preserved in the loop. ▯

This result will be made more precise later on, where we discuss the cosets $W_\Theta\backslash W$ in more detail.

♣ [walext] We now prove Theorem 7.1 by induction on $\ell(w)$. If $w = 1$ there is no problem. Suppose $\ell(w) > 1$. Suppose $s = s_\alpha$. If $y = s{\circ}w < w$ we must show that α is negative on wC. But then $s{\circ}y > y$, so by induction $\alpha > 0$ on yC. But then $\alpha < 0$ on $wC = s{\circ}yC$ since $s\alpha = -\alpha$.

Continue to suppose $s = s_\alpha$. and suppose $s{\circ}w > w$. It must be shown that $\alpha > 0$ on wC. Choose $t = s_\beta$ such that $t{\circ}w < w$. Find x in $W_{\alpha,\beta}$ and y in W satisfying the conditions of the Lemma. Since $t{\circ}w < w$, $\ell(y) < \ell(w)$. Since $s{\circ}y > y$ and $t{\circ}y > y$, induction lets us see that yC is contained in the region $C_{\alpha,\beta}$ where $\alpha > 0$, $\beta > 0$. Since $\ell(w) = \ell(x) + \ell(y)$, $\ell(s{\circ}x) = \ell(x) + 1$, and this is still valid if ℓ is the length
♣ [wal] in $W_{\alpha,\beta}$. From Proposition 6.4, we see that $\alpha > 0$ on the region $xC_{\alpha,\beta}$, hence on $wC = x{\circ}yC$ as well. ▯

[coxetergroup] **Corollary 7.3.** *The canonical map from W to $\overline{W}$ is an isomorphism.*

Proof. Suppose $\overline{w} = 1$. If $w \neq 1$ then $s{\circ}w < w$ for some $s = s_\alpha$ in S, which implies that $\alpha < 0$ on $wC = C$. Contradiction. ▯

Hence from now on I may revert to earlier notation and identify W with $\overline{W}$.

Given an irreducible root datum $\mathcal{L}$, the associated set of roots Σ is the W-orbit of Δ in L. Define the subset of **positive roots** Σ^+ to be those roots λ such that $\lambda \bullet C > 0$. In particular, $\Delta \subset \Sigma^+$. The negative roots are those < 0 on C.

[posneg] **Corollary 7.4.** *Every root is either positive or negative.*

Proof. If $\lambda = w\alpha$ with α in Δ, then $\lambda \bullet C = \alpha \bullet w^{-1}C$, and according to the Theorem either $\alpha > 0$ or $\alpha < 0$ on $w^{-1}C$. ▯

If λ is a non-negative combination of simple roots, it is certainly a positive root. Conversely:

[rootspos] **Proposition 7.5.** *Every positive root may be expressed as a non-negative integral combination of simple roots.*

Proof. Since the simple roots are linearly independent, the domain C is a simplicial cone, and a general result about convex cones asserts that any positive root is a non-negative linear combination of simple roots. But the coordinates of any root are integral. ▯

For every w in W let

$$L_w = \{\lambda \in \Sigma^+ \mid w^{-1}\lambda < 0\}\,.$$

♣ [walext] An equivalent condition is that $\lambda < 0$ on wC, or that $\lambda = 0$ separates C from wC. Thus Theorem 7.1 asserts neither more nor less than that $s_\alpha w > w$ if and only if α is not in L_w, or that $s_\alpha w < w$ if and only if α is in L_w.

[lw] **Proposition 7.6.** *If α lies in Δ then*

(a) $L_{s_\alpha} = \{\alpha\}$;
(b) if $s_\alpha w > w$ then $L_{s_\alpha w} = w^{-1}\{\alpha\} \sqcup L_w$.

Here as elsewhere $\sqcup$ means a disjoint union.

Proof. The hyperplane $\alpha = 0$ is a wall of C, so $\alpha = 0$ is the only root hyperplane that separates C from $s_\alpha C$. Any other root vanishing there must be an integral multiple of α, but that is excluded by ♣ [twolengths] Proposition 6.7. This proves (a).

I leave (b) as an exercise. ▯

[lwlength] **Corollary 7.7.** *For w in W, $|L_w| = \ell(w)$.*

Proof. By induction. ▯

The complement of the union of root hyperplanes $\lambda = 0$ is partitioned into connected components called **Weyl chambers**. One of these is the cone C. The following says that these chambers are a principal homogeneous set for W, and a bit more.

[funddom] **Proposition 7.8.** *The closure $\overline{C}$ of the cone C is a strict fundamental domain of W.*

That is to say, every v in $L_{\mathbb{R}}$ is the W-transform of a unique vector in $\overline{C}$.

Proof. The first step is to show that every v can be transformed by W to a point of $\overline{C}$. This is done by induction on the number of positive root hyperplanes separating v from C.

Suppose both v and $w(v)$ lie in C with $w \neq 1$. I'll prove by induction on $\ell(w)$ that $w(v) = v$. Say $x = s_\alpha w < w$. Then C and wC lie on opposite sides of the hyperplane $\alpha = 0$. But since v and $w(v)$ both belong to $\overline{C}$, it must lie on that hyperplane. But then $x(v) = v$. Apply induction. ▯

8. From root system to base

In this section and the next I'll explain two different ways to construct a base Δ from a root system Σ—i.e. to go backwards along the route followed so far. The next section introduces a completely new idea—new to this essay, that is to say, but not due to me.

Suppose that we are given a root system Σ in the real vector space V. I recall, this means Σ is a finite subset of non-zero vectors with these properties:

(a) there exists a map $\lambda \mapsto \lambda^\vee$ from Σ to the linear dual $V^\vee$ such that $\langle \lambda, \mu^\vee \rangle \in \mathbb{Z}$ for all λ, μ in Σ;
(b) the subspace of V annihilated by $\Sigma^\vee$ is complementary to the subspace spanned by Σ;
(c) the linear map

$$s_\lambda\colon\ v \mapsto v - \langle v, \lambda^\vee \rangle \lambda$$

is a reflection that takes Σ to itself.

♣ [reduced] I shall assume further that, following Corollary 6.9, (d) the root system is reduced. This means that if λ is a root then the only scalar multiples of λ that are also roots are $\pm\lambda$. The Weyl group W of the root system is that generated by all the s_λ. All of these preserve the subspace of V perpendicular to Σ. Hence if w in W acts trivially on Σ it acts trivially on all of V. Therefore W is finite. There thus exists at least one positive definite inner product preserved by W. Since each s_λ is orthogonal, we have

$$\langle \lambda, \mu^\vee \rangle = 2\left(\frac{\lambda \bullet \mu}{\mu \bullet \mu}\right)$$

for all roots λ, μ.

[rootlattice] **Proposition 8.1.** *The roots span a lattice in V.*

I recall that a lattice in a real vector space is a discrete subgroup, necessarily a finitely generated free $\mathbb{Z}$-module.

Proof. If the $\mathbb{Z}$ span of the roots is not discrete, there would exist vectors in it arbitrarily close to 0. This means that any linear function on V takes arbitrarily small values on the span. On the other hand, since $\langle \lambda, \mu^\vee \rangle$ lies in $\mathbb{Z}$, any coroot in $R^\vee$ defines a linear function on all of V taking only integral values. ∎

[subspaceroots] **Proposition 8.2.** *If U is a linear subspace of V then $\Sigma \cap U$,is a root system in U.*

Proof. Immediate. ∎

The **root hyperplanes** $\lambda = 0$ partition V, their complement is the disjoint union of convex conical sets called chambers. This configuration is invariant with respect to each of the reflections s_λ, hence preserved by W. Let C be one of these, and let Δ be the roots α that are > 0 on C such that $\alpha = 0$ is a wall of C (of codimension one).

[cartanfromroots] **Theorem 8.3.** *The set Δ is linearly independent, and $(\langle \alpha, \beta^{\vee>} \rangle)$ is a Cartan matrix. The associated set of roots is Σ.*

Proof. In several steps. The first step is to show that $\langle \alpha, \beta^\vee \rangle \le 0$ for $\alpha \ne \beta$ in Δ. The open region C has $\alpha = 0$ and $\beta = 0$ as walls. The region C is convex, so we can find points A on the first and B on the second such that the open line segment between them lies in C. It does not cross any root hyperplanes
♣ [subspaceroots] at all. By Proposition 8.2 the roots in the linear span of α and β make up a root system, so we can apply
♣ [abneg] Proposition 6.3 to deduce that $\langle \alpha, \beta^\vee \rangle \le 0$.

The next step is one in linear algebra.

[gaussab] **Proposition 8.4.** *Suppose $\mathfrak{A} = \{a_i\}$ to be a set of vectors in a Euclidean vector space such that (a) $a_i \bullet a_j \le 0$ for $i \ne j$; (b) there exists a single vector ρ such that $a_i \bullet \rho > 0$ for all i. Let M be the matrix $(a_i \bullet a_j)$. There exists a unipotent lower triangular matrix L with non-negative entries and an invertible diagonal matrix E wih positive entries such that*

$$L\, M\, {}^tL = E\,.$$

In particular, the set $\mathfrak{A}$ is linearly independent.

Proof. A careful application of Gauss elimination. Along the way we'll need a simple fact: *Suppose $u \bullet v \le 0$, $u \bullet \rho > 0$, and $v \bullet \rho > 0$. Let $u^\perp$ be the projection of u onto the hyperplane perpendicular to v. Then $u^\perp \bullet \rho > 0$ also.* For if $u = u_0 + u^\perp$, then the projection of u along v is

$$u_0 = \frac{u \bullet v}{v \bullet v}\, v = -cv$$

with $c \ge 0$. But then

$$u^\perp \bullet \rho = v \bullet \rho + c\, u \bullet \rho > 0\,.$$

Now to prove the Proposition. Let A be the matrix whose columns are the vectors a_i in $\mathfrak{A}$. The matrix $M = (a_i \bullet a_j)$ is then the same as ${}^tA\, QA$ if Q is the matrix defining the dot product. Let $a_1^\perp = a_1$, but for $i > 1$ let

$$a_i^\perp = a_i + \ell_i a_1 \quad \text{where} \quad \ell_i = -\frac{a_i \bullet a_1}{a_1 \bullet a_1}\,,$$

the projection of a_i onto the space perpendicular to a_1. By assumption, $\ell_i \ge 0$. If ℓ is the column matrix with entries ℓ_i and

$$L = \begin{bmatrix} 1 & {}^t\ell \\ \ell & I \end{bmatrix}$$

then

$$L\, M\, {}^tL = M^\perp\,.$$

where $M^{\perp}$ is the matrix $(a_i^{\perp} \bullet a_j^{\perp})$. The only non-zero entry in the first row or column of the matrix $M^{\perp}$ is $a_1 \bullet a_1 > 0$. According to the Lemma, the vectors $a = a_i^{\perp}$ still satisfy the condition $a \bullet \rho > 0$, so we may now continue by induction on the set $a_2^{\perp}$, etc. ∎

♣ [cartanfromroots] To prove Theorem 8.3, we take ρ to be any vector in C. ∎

♣ [gaussab] There is one other easy consequence of Proposition 8.4. Suppose Δ spans V, and let $\mathfrak{C}$ be the Cartan matrix determining it. For some positive diagonal matrix we know that $\mathfrak{C}D$ is positive definite and symmetric, ♣ [gaussab] satisfying the hypotheses of Proposition 8.4, so its inverse, which is $D^{-1}\mathfrak{C}^{-1}$, has non-negative entries, hence so does $\mathfrak{C}^{-1}$. But the rows of $\mathfrak{C}^{-1}$ are the coordinates with respect to Δ, of the basis (ϖ_α) dual to Δ. Therefore:

[cincpos] **Proposition 8.5.** *The closure $\overline{C}$ of the fundamental domain is contained in the closed cone spanned by Δ.*

Furthermore, a root is positive if and only if its dot-product with each ϖ_α is non-negative. Since this is dot-product is the coordinate of the root in the basis Δ, this proves very explicitly that every positive root is a non-negative combination of simple roots.

I have remarked that root systems, as opposed too root data, are designed to deal with Lie algebras, not groups. However, every root system gives rise to certain canonical root data. Define L_Δ to be the lattice in V spanned by Δ, which is contained in the lattice $(L_{\Delta^\vee})^\vee$ of v in V such that $\langle v, \lambda^\vee \rangle$ is integral. Both of these define root data. In some cases, for example the systems E_n, these are not distinct.

9. Positive roots and linear orders

The approach in the previous section is geometric, characterizing the set Δ in terms of the partition of V by root hyperplanes. This one will explain an alternate method, often seen in older literature.

Given a coordinate system on the real vector space V, we may define a total order on V by this condition:

$$(x_i) < (y_i) \text{ if for some } j \text{ we have } x_i = y_i \text{ for } i < j \text{ but } x_j = y_j\ .$$

In other words, this is lexicographic or dictionary order. The point (x_i) is positive if its first non-zero coordinate is positive. I'll call an order defined in this way a **linear order**. A linear order is translation invariant, also invariant with respect to positive scalar multiplication, and conversely every total order so invariant is a linear order.

If $(L, \Delta, L^\vee, \Delta^\vee)$ is a root datum such that Δ spans $V = L_{\mathbb{R}}$, then Δ is a basis of V, and according to ♣♣[r[potspeg] Corollary 7.4 and Proposition 7.5 the positive roots are those which are positive with respect to the linear order determined by this basis. The converse of this is the main result of the rest of this section:

[linearorder] **Proposition 9.1.** *Suppose Σ to determine a root system on V, on which a linear order has been given. The set of positive roots with respect to this order is the set of positive roots associated to some chamber of the partition determined by Σ.*

Proof. It will take several steps. The following lemma will be used more than once.

[differenceofroots] **Proposition 9.2.** *If λ, μ are roots and $\langle \lambda, \mu^\vee \rangle > 0$, then $\lambda - \mu$ is also a root.*

Proof. We may assume λ and μ to be linearly independent. Since $\lambda - \mu$ is a root if and only if $\mu - \lambda$ is one, we may swap λ and μ if necessary to make $\langle \lambda, \mu^\vee \rangle = -1$. But then $s_\mu \lambda = \lambda - \mu$ is a root. ∎

Suppose now a linear order to be given. For each j define α_j inductively to be the smallest root in V not in the subspace spanned by the α_i with $i < j$. Let Δ be the set of all α_i.

For $i \neq j$, we have $\alpha_i \bullet \alpha_j \leq 0$. Otherwise, say $i < j$. By the Lemma, $\alpha_j - \alpha_i$ is also a root, but it is also less than α_j, contradicting the definition.

Next I claim that every positive root is a non-negative linear combination of the α_i. I'll prove it by induction on the order of the roots. If λ is minimal, it is α_1 itself. Otherwise, suppose it is true for all roots less than λ. Suppose λ to be in the linear span of the α_i with $i \le j$, but not for $i < j$. If $\lambda = \alpha_i$ we are through. Otherwise $\lambda > \alpha_j$. I claim that for some α_i with $i \le j$ we must have $\lambda - \alpha_i$ equal to a root. If so, we are through, because $\lambda - \alpha_i$ will be less than λ. But if not, then by the Lemma $\lambda \bullet \alpha_i \le 0$ for all
♣ [cincpos] i, and hence by Proposition 8.5 it will be a non-positive linear combination of the α_i, which contradicts that it is a positive root. ▯

The definition of positive roots in terms of a linear order has one attractive feature—that the set of all roots is the disjoint union of positive and negative roots is immediate. On the other hand, an annoying feature is that the linear order is by no means determined by the choice of positive roots, so it is not intrinsic to a datum.

There is one useful application of the proof. Let Σ_j be the roots in the linear span of the α_i for $i \le j$.

[wtransformroots] **Suppose $\Lambda = \{\lambda_i\}$ to be a set of linearly independent roots, and for each j let Λ_j be the subset of λ_i with $i \le j$. There exists w in W such that $w\Lambda_i \subseteq \Sigma_i$. 9.3.**

10. Constructing roots

The set of roots is the orbit of Δ with respect to the group generated by the reflections s_α for α in Δ. How can we construct it, and more precisely construct it along with other relevant data?

The roots can be described explicitly for each of the known systems. This is done, for example, in [Bourbaki:1968]. So we could simply make up a list of roots for each system, or at least a way to reconstruct Bourbaki's list. But more interesting for our purposes is an algorithm that constructs the roots directly from the Cartan matrix. For one thing, this will turn out to take a negligible amount of time. And for another, building the roots will give us some additional structure we'll find useful subsequently.

There is one immediate simplification. If λ is a root, so is $-\lambda$. Therefore we only have to construct the positive roots. This is not obviously very hard. We know that every positive root is a non-negative integral combination

$$\lambda = \sum_{\alpha\in\Delta} \lambda_\alpha \alpha \,,$$

so we start with the elements of Δ itself, for which exactly one coefficient is non-zero, and keep applying reflections s_α to positive roots until we get nothing new. The formula for s_α is

$$s_\alpha\colon\ \lambda \longmapsto \lambda - \langle \lambda, \alpha^\vee\rangle \alpha \,,$$

so that

$$(s_\alpha\lambda)_\beta = \begin{cases} \lambda_\alpha - \sum_\beta \lambda_\beta \langle \beta, \alpha^\vee\rangle & \text{if } \beta = \alpha \\ \lambda_\beta & \text{otherwise.} \end{cases}$$

The first formula can be simplified by ignoring all β with $\langle \beta, \alpha^\vee\rangle = 0$. This gives us

$$(s_\alpha\lambda)_\alpha = -\lambda_\alpha - \sum_{\beta\sim\alpha} \lambda_\beta \langle \beta, \alpha^\vee\rangle$$

in which the sum is over those β linked to α in the Dynkin diagram.

This is simple, but for reasons that will soon appear, I prefer to keep track of the array $(\langle \lambda, \alpha^\vee\rangle)$ along with that of coefficients (λ_α). This makes calculating $(s_\alpha\lambda)_\alpha$ immediate, but it means that when we calculate $s_\alpha\lambda$ we must also calculate the numbers

$$\begin{aligned} \langle s_\alpha\lambda, \beta^\vee\rangle &= \langle \lambda - \langle \lambda, \alpha^\vee\rangle\alpha, \beta^\vee\rangle \\ &= \langle \lambda, \beta^\vee\rangle - \langle \lambda, \alpha^\vee\rangle\langle \alpha, \beta^\vee\rangle \,. \end{aligned}$$

This leads to

$$\langle s_\alpha\lambda, \beta^\vee\rangle = \begin{cases} -\langle\lambda, \alpha^\vee\rangle & \text{if } \beta = \alpha \\ \langle\lambda, \beta^\vee\rangle - \langle\lambda, \alpha^\vee\rangle\langle\alpha, \beta^\vee\rangle & \beta \sim \alpha \\ \langle\lambda, \beta^\vee\rangle & \text{otherwise.} \end{cases}$$

If

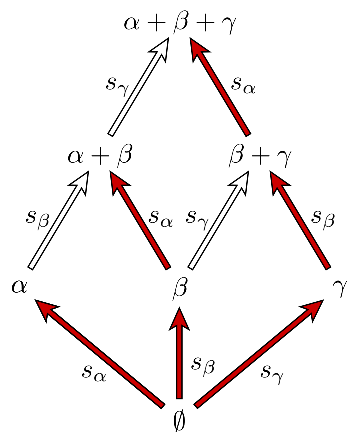
$$\lambda = \sum \lambda_\alpha \alpha$$

its height is $\sum \lambda_\alpha$. Thus the height of every simple root is 1. As we calculate roots, we are going to arrange things so we only move up in height. This is easy to do. In calculating $s_\alpha\lambda$ only the coefficient of α changes, and

$$(s_\alpha\lambda)_\alpha = \lambda_\alpha - \langle\lambda, \alpha^\vee\rangle \,.$$

Therefore $s_\alpha\lambda$ has height greater than that of λ if and only if $\langle\lambda, \alpha^\vee\rangle < 0$. I'll write $\lambda \prec s_\alpha\lambda$ in this situation.

The positive roots associated to a root datum are the nodes of a structure I call the **root graph**. It is an oriented, labeled graph. The obvious choice for the bottom layer of the graph consists of the nodes in Δ, but for mild convenience I create below these a dummy node $\emptyset$ of height 0, with an edge labeled by s_α directed from $\emptyset$ to each α in Δ. Then, there is an edge from λ to μ if $\mu = s_\alpha\lambda \succ \lambda$ for α in Δ, labeled by s_α. The following figure ilustrates this graph for A_3, with three simple roots α, β, γ.



Incidentally, the entire table of simple root reflections can be read off from this graph—if there is no edge leading from or to λ labeled by α, then $s_\alpha\lambda = \lambda$.

Paths from $\emptyset$ correspond to strings of simple reflections, the labels of the edges in the path. There will be in general several paths from $\emptyset$ to a given root, corresponding possibly to several strings. For example, from $\emptyset$ to $\alpha + \beta + \gamma$ in A_3 we have the paths

$$s_\alpha s_\beta s_\gamma$$
$$s_\beta s_\alpha s_\gamma$$
$$s_\gamma s_\beta s_\alpha$$
$$s_\beta s_\gamma s_\alpha \,.$$

I want to distinguish exactly one of the possible paths from $\emptyset$ to a root λ. First of all, assume that Δ is ordered or, equivalently, indexed by integers 1, 2, $\ldots$. To each root λ in the root graph pick out that path from $\emptyset$ to λ which, when read backwards, is lexicographically least. Thus among the four paths to $\alpha + \beta + \gamma$ the third is the distinguished one, since $\alpha < \gamma$ and $\beta < \gamma$ (recall, we are reading the string backwards).

In the figure above, distinguished paths are coloured dark.

We can therefore construct positive roots by constructing distinguished paths in the root graph. The advantage is that these paths can be constructed one edge at a time—given a path to λ, we calculate all the $\mu = s_\alpha \lambda \succ \lambda$ (verifying, in computation, $\langle \lambda, \alpha^\vee \rangle < 0$). If α is minimal in Δ such that $\langle \mu, \alpha^\vee \rangle > 0$, the the path continuing from λ to μ is distnguished, and we register μ as a root.

This process requires that we maintain two lists, one of all roots we have so far enountered, and another of those roots whose upward reflections have not been examined. The first list only grows, but the second grows and shrinks. I'll call the first the **root list**, the second the **process list**. To start, we put the simple roots on both lists. Then, as long as the process list is not empty, we remove an element λ from it. We then scan through the reflected roots $\mu = s_\alpha \lambda \succ \lambda$, and if the distinguished path continues to μ, we add it to both the root list and the process list. The process list is dynamic—items are both removed and added to it as time goes on. Programming deals with several different types of dynamic lists, and for technical reasons I use queues or FIFO (First In, First Out) lists. In this way, when I remove an item to be processed, all roots of smaller height have already been processed.

Incidentally, there is a single root of greatest height, called the **dominant root** $\widetilde{\alpha}$. For every β in Δ, $s_\beta \widetilde{\alpha} \prec \widetilde{\alpha}$. It is not the only root with this property, but it does dominate every other root in the sense that $\lambda_\beta \le \widetilde{\alpha}_\beta$ for all roots λ and all β.

11. The Weyl group

The **Weyl group** W of a root system is that generated by the reflections s_α with α in Δ. In this section I'll describe how to construct a list of all elements of W. The basic idea is rather similar to the method of listing all roots—we keep applying simple reflections to the identity element to get all elements of the Weyl group, taking care to construct each element only once.

As with roots, we'll represent each w in W by a string of simple reflections s_α, in this case by expressions for w as a product of simple roots of least length. Such a product is called a **reduced expression** for w. There may be several, but there is a simple way to pass from one to the other. Let S be the set of all simple reflections s_α. It turns out that the group W is defined by S together with relations

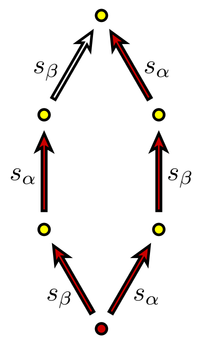
$$s_\alpha^2 = 1 \quad (s_\alpha s_\beta)^{m_{\alpha,\beta}} = 1$$

where $(m_{\alpha,\beta})$ is the Coxeter matrix of the system that we have seen above. These last relations may be replaced by the **braid relations**

$$s_\alpha s_\beta \ldots = s_\beta s_\alpha \ldots \quad (m_{\alpha,\beta} \text{ terms on each side}) .$$

Thus two product expressions give the same w in W if and only if one can be converted to the other by a sequence of these relations. For example, in the Weyl group of A_3 the expressions $s_\alpha s_\beta s_\alpha$ and $s_\beta s_\alpha s_\beta$ represent the same element. In general, in passing from one expression to another we might expect the length of the expression to go up, but [Tits:1968] has shown that two reduced expressions for the same w may be obtained from each other by braid relations alone.

As with roots, I define a graph whose nodes are elements of W, with an edge from w to ws_α if $\ell(ws_\alpha) > \ell(w)$, in which case I write $w \prec ws_\alpha$. For example, here is the graph for A_2:



A path in the graph from 1 to w is distinguished, as before, if when read backwards it is lexicographically least among the paths from 1 to w. In the figure, distinguished paths are dark. So we can list elements of W by listing distinguished paths.

How can we find edges in the graph of W? How tell whether or not a given path is distinguished? Answering both questions comes down to a fundamental property linking the geometry of the roots with the combinatorics of W: for any w in W, $\ell(ws_\alpha) > \ell(w)$ if and only if $w^{-1}C$ lies on the other side of the hyperplane $\alpha = 0$ from C. This happens if $\langle \alpha, x \rangle < 0$ for all x inside $w^{-1}C$. Since a root is either positive throughout the interior of C or negative throughout, this happens if and only if $\langle \alpha, w^{-1}x \rangle < 0$ for one x inside C. For this x I choose a vector ρ such that $\langle \alpha, \rho \rangle = 1$ for all α in Δ. So now we have this criterion: $ws_\alpha \succ w$ if and only if $\langle \alpha, w^{-1}\rho \rangle < 0$. Because ρ lies in the interior of the fundamental domain C, the map from w to $w^{-1}\rho$ is a bijection.

Therefore, as we construct our paths and elements w, we maintain the current value of $w^{-1}\rho$. When we ascend in the graph, we calculate $(ws_\alpha)^{-1}\rho = s_\alpha w^{-1}\rho$. Thus a distinguished path to w continues to $\mu = ws_\alpha$ if and only if $\langle \alpha, w^{-1}\rho \rangle < 0$ and α is least in Δ with $\langle \alpha, s_\alpha w^{-1}\rho \rangle < 0$.

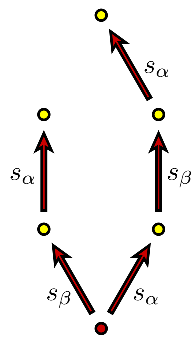
We build W by following paths in the Bruhat graph subject to this restriction, much as we constructed the roots. We use this to make a list of all of W, as we did for the list of roots.

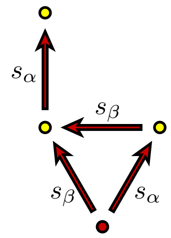
Then we can go through and calculate reflection tables for each w in W, to give the complete structure of W.

This process is relatively inefficient. For one thing, it becomes quickly apparent that the amount of work involved goes up rapidly with the rank of the root system, as you'd expect. For one thing, performing reflections involve doing arithmetic on vectors of dimension equal to the rank. There are faster methods. We'll see one in the next section.

12. Cosets

The process described in the previous section for making a list of elements in a Weyl group W depends on identifying admissible paths in a directed graph with elements of the group, and then constructing all such paths. (An admissible path is one that is compatible with orientation and starts at the base node.) The problem is, there are as many nodes in the graph as there are elements of the Weyl group—the graph is a tree. But there are other graphs for which admissible paths may be identified with elements of W, and they can be smaller. For example, in he figures below I display first of all the graph for A_2 constructed in the previous section, and next to it a smaller graph whose admissible paths also parametrize $\mathfrak{S}_3$.





This will turn out to be a special case of a general phenomenon. Here the saving is not great, but for arbitrary A_n we'll see the number of nodes in the graph shrink from $(n+1)!$ to $n(n+1)/2+1$.

There are some rather sophisticated ways to construct such a graph, but the simplest one uses an idea due to Fokko du Cloux, and depends on the cosets of a Weyl group with respect to smaller Weyl groups contained in it. Suppose $|\Delta| = n$. For $1 \le k \le n$ let Δ_k be the subset of i initial simple roots, let S_k be the set of simple reflections s_α for α in Δ_k, and let W_k be the subgroup of W generated by S_k. Let Σ be the set of all roots of the system, and let Σ_k be those which are linear combinations of the simple roots in Δ_k. The group W_k is the Weyl group of the root datum $(L, \Delta_k, L^\vee, \Delta_k^\vee)$, and Σ_k is the corresponding set of roots.

To each w in W is associated the subset R_w of Σ, the set of $\lambda > 0$ such that $w\lambda < 0$. Thus $R_w = L_{w^{-1}}$. If $w = s_\alpha$ then $R_w = \{\alpha\}$, and R_w may be calculated inductively by means of the formula

$$R_{xy} = R_y \sqcup y^{-1} R_x \ \text{ if } \ \ell(xy) = \ell(x) + \ell(y)\,.$$

Thus $|R_w| = \ell(w)$.

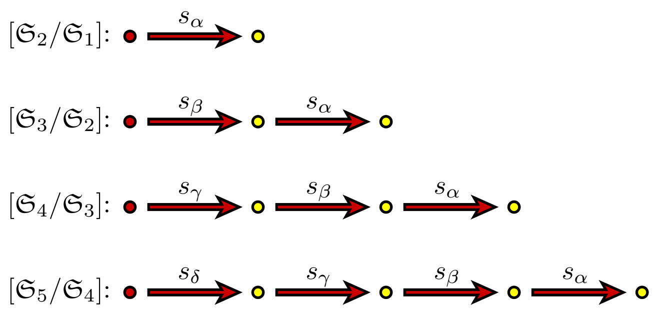
[coset] **Proposition 12.1.** *Suppose $\Theta \subseteq \Delta$. (a) An element w of W lies in W_Θ if and only if $R_w \subseteq \Sigma_\Theta^+$. (b) Every w in W may be factored uniquely as xy with y in W_Θ and $x\Theta > 0$.*

Thus if $[W/W_\Theta]$ is the subset of all w in W with $w\Theta > 0$, the projection from it to W/W_Θ is a bijection, and $[W/W_\Theta]$ is a set of distinguished representatives of W/W_Θ. Hence the product map is a bijection of $[W/W_\Theta] \times W_\Theta$ with W. In particular, s_m lies in in $[W_m/W_{m-1}]$, and w lies in $[W_m/W_{m-1}]$ if and only if the distinguished expression for w is of the form $s_{i_1} \ldots s_m$, with all $i_j \le m$. Arguing inductively, we deduce that if the size of S is n, the product map

$$[W_n/W_{n-1}] \times \ldots [W_1/W_0] \to W$$

is a bijection. The sizes of cosets are rather small, compared to the size of W. If $W = \mathfrak{S}_n$, for example, then $|W_m/W_{m-1}| = m$.

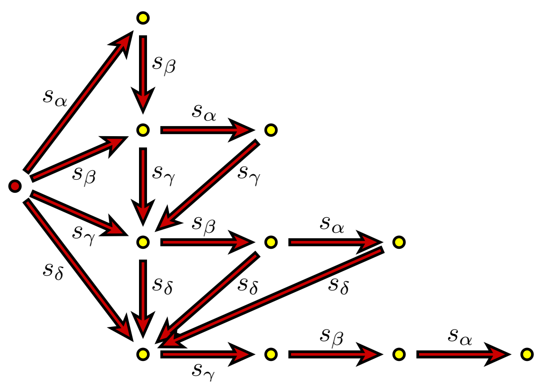
Associated to each coset is a graph analogous to the Bruhat graph for W itself. The nodes are elements of $[W_m/W_{m-1}]$, and there is an edge from w to $s_\alpha w$ if w and $s_\alpha w$ are both in $[W_m/W_{m-1}]$ with $\ell(s_\alpha w) > \ell(w)$. The base of the graph is the image of identity of W. There is still a notion of distinguished path, but with respect to left multiplication by elementary reflections: $w \prec x = s_\alpha w$ is distinguished if α is least with $s_\alpha x \prec x$. For example, here are the graphs of all the cosets of $\mathfrak{S}_5$, in which all edges are distinguished:



The graph of distinguished edges is a tree contained in the full graph of the coset.

It is relatively easy to multiply on the left by s in S_m on $[W_m/W_{m-1}]$: either sw lies again in $[W_m/W_{m-1}]$ or it is equal to wt for t in S_{m-1}. In practice, we can carry out this calculation by using the relationship between $w^{-1}\rho$ and product expressions. Thus we can list these cosets without much trouble. We can then make the tree of distinguished edges for each X_m, tacking on elements of S_m at the left by looking $w^{-1}\rho$. A fair amount of calculation will be involved, but cosets are small so this is not a serious problem.

We now make up a graph by combining all the graphs Γ_m of the cosets, by amalgamating all the identity nodes of the Γ_m into a single base node. In addition, for each node x of Γ_m other than the identity and for each $k > m$ we add an edge labeled by s_k from x to the node s_k of Γ_k. The number of nodes is $1 + \sum_{i \le n}(|\Gamma_i| - 1)$. The paths in this new graph startinga t the base is in bijection with elements of W, but the product expressions are in some reversed from the ones we got before—the paths now accepted are thoe where each s_i is least among the s in S with $sw < w$. That is to say, we have swapped left and right multiplication by elements of S. What follows is the graph for A_5:



Using this graph to traverse the Weyl group is very fast. Traversing E_8 (with $696, 729, 600$ elements!) on my laptop takes about 15 minutes, whereas doing this by the method described in the previous section wasn't obviously feasible at all.

We can also use the coset lists to multiply on the left by an element of S. We can build into the coset graph a table of left multiplications attached to each node. Given s in S and x in the coset, there are restricted possibilities for sx. Either $sx > x$ and sx is again in the coset, with an insertion of s to the left of s_m; or $sx < x$ in which case sx is definitely in the coset with a deletion to the left of s_m; or $sx = xt$ with $t < s_m$. If w is represented as $\prod w_i$ with w_i in $[W_i/W_{i-1}]$ then we can find sw by repeated multiplying in the cosets.