Pinning-Sympy: Theoretical Underpinnings

Joshua Ruiter

July 22, 2026

Contents

1 Big picture **3**
- 1.1 What is an algebraic group? . . . 3
- 1.2 What is rigidity? . . . 3
- 1.3 What is a pinning? . . . 4

2 Special linear groups **7**
- 2.1 Torus subgroup and character group . . . 7
- 2.2 Pinning the special linear group . . . 9
  - 2.2.1 Conjugation by torus . . . 10
  - 2.2.2 Conjugation by Weyl group . . . 11
  - 2.2.3 Commutator formula . . . 13
- 2.3 "*Nonlinearity of matrix groups*" . . . 13

3 Forms **20**
- 3.1 Bilinear forms . . . 20
- 3.2 Hermitian and skew-hermitian forms . . . 23
- 3.3 Quadratic forms . . . 26
- 3.4 Isotropy and Witt index . . . 28

4 Pinnings **30**
- 4.1 Algebraic groups . . . 30
- 4.2 Lie algebras . . . 31
- 4.3 Lie algebras associated to classical groups . . . 32
- 4.4 Roots and root spaces . . . 33
- 4.5 Root systems . . . 36
- 4.6 Pinnings . . . 38

5 Special orthogonal groups **40**
- 5.1 Split, quasi-split, and otherwise . . . 40
- 5.2 Lie algebra . . . 42
- 5.3 Torus and characters . . . 43
- 5.4 Quasi-split groups . . . 45

5.4.1 The minimal example ($n = 4, q = 1$) . . . 45
5.4.2 Another example ($n = 6, q = 2$) . . . 48
5.4.3 The general quasi-split case ($n = 2q + 2$) . . . 49
5.5 General non-split case ($n \geq 2q + 2$) . . . 49

6 Special unitary groups **50**
6.1 Quasi-split case ($n = 2q$) . . . 50
6.2 Non-quasi-split case ($n > 2q$) . . . 53

1 Big picture

Algebraic groups have two structures. An abstract homomorphism between algebraic groups is a map that preserves one of the structures, but not necessarily the other. We want to prove that abstract homomorphisms are more rigid than they initially appear, by proving that they "kind of" preserve the other structure.

1.1 What is an algebraic group?

For the purposes of these notes, an algebraic group is a set which is both an algebraic variety (over a base field k), and a group. The algebraic variety structure is a "geometric" thing. A variety is a set of solutions to some polynomial equations. Polynomial equations are geometric in the sense that the set of solutions is something we can graph.

In this project, we are more focused on the "group" aspect of algebraic groups. The group side is something we can study from an abstract algebra and group theory perspective. We can analyze subgroups, conjugation, group homomorphisms, kernels, etc.

Group homomorphisms are an effective tool for extracting information about groups. The analogous tool for algebraic varieties is a regular map, which is roughly speaking something that behaves like a polynomial function. Usually, people studying algebraic groups study morphisms of algebraic groups, which are group homomorphisms that are also regular maps. However, you can also just look at "abstract" homomorphisms between algebraic groups, which is to say, just group homomorphisms that ignore the geometric structure.

1.2 What is rigidity?

To illustrate the idea of rigidity, we go on an aside about continuity and differentiability. In general, a continuous function $f : \mathbb{R} \to \mathbb{R}$ is not infinitely differentiable, and may even fail to be differentiable a single time. (By "differentiable," we mean "differentiable everywhere.") A continuous function may have corners ($y = |x|$) or cusps ($y = x^{2/3}$). On the other hand, every differentiable function is continuous. In the world of continuous functions, at one extreme are the "very differentiable" and "very continuous" polynomial functions. At the other end, there are continuous functions which are not differentiable anywhere, such as the Weierstrass function[1]. In the diagram below, we depict a "rigidity spectrum" for continuous functions, with rigidity increasing to the right.

Continuous functions

Not differentiable	Differentiable exactly once	Differentiable exactly twice	Infinitely differentiable
$\longleftarrow$			$\longrightarrow$
Weierstrass function, $\lvert x\rvert, x^{2/3}$	$x\lvert x\rvert$	$x^2\lvert x\rvert$	Polynomials $e^x, \sin x, \cos x$

The main issue with the graphic above is that actually, differentiable functions are extremely rare among all continuous functions. A better visualization would have 99% of the spectrum devoted to

[1] https://en.wikipedia.org/wiki/Weierstrass_function

functions like the Weierstrass function which are continuous but differentiable at no points or just a finite set of points.

In parallel to the discussion above, morphisms of algebraic groups are much more rigid than abstract homomorphisms. So we do not, in general, expect an abstract homomorphism to be regular. However, sometimes we can prove that an abstract homomorphism is "almost" regular, in a particular sense. These are called "rigidity theorems."

We want to prove rigidity theorems of the following form: for a particular fixed algebraic group $G(k)$, any abstract homomorphism $\rho : G(k) \to \mathrm{GL}_m(K)$ must be "almost regular" (K is some field extension of k). In this context, "almost regular" means a factorization of ρ, a way of writing ρ as a composition of two maps, where one of those is regular. The usual goal is to factor ρ into the form $\rho = \sigma \circ F$ where σ is regular.

$$\begin{tikzcd} G(k) \arrow[rr, rightsquigarrow, "\rho"] \arrow[dr, rightsquigarrow, "F"'] & & \mathrm{GL}_m(K) \\ & G(A) \arrow[ur, "\sigma"'] & \end{tikzcd}$$

Here A is some algebra. To emphasize the rigidity of σ, the diagram above uses squiggly arrows for abstract group homomorphisms, and a straight arrow for the more rigid regular group homomorphism. Philosophically speaking, we think of σ as the "regular part" of ρ.

1.3 What is a pinning?

Algebraic groups have a lot of structure. We can "pin" them to understand and see the structure better. The word "pinning" here comes from the metaphor of pinning the wings of a butterfly specimen, as you might see in a museum. A pinning of an algebraic group $G(k)$ is a collection of functions

$$X_\alpha : V_\alpha(k) \to G(k)$$

where α ranges over the set of "roots" Φ of G, and $V_\alpha(k)$ is just a vector space over k (often the dimension of $V_\alpha(k)$ is just one, so in that case $V_\alpha(k)$ is essentially just k with the group operation of addition). These pinning maps X_α should satisfy a variety of properties.

(P1) They are *usually* group homomorphisms, meaning that

$$X_\alpha(v) \cdot X_\alpha(w) = X_\alpha(v + w)$$

for all $v, w \in V_\alpha(k)$. More generally, this equation might fail and there will need to be an additional factor on one side to account for this defect.

$$X_\alpha(v) \cdot X_\alpha(w) = X_\alpha(v + w) \cdot \Big(\text{correction term}\Big)$$

(P2) The group $G(k)$ is generated by the images $X_\alpha\big(v\big)$, where α ranges over Φ and v ranges over $V_\alpha(k)$.

(P3) $G(k)$ contains a subgroup $T(k)$ called a "maximal k-split torus" and for any $t \in T(k)$, conjugating $X_\alpha(v)$ by t corresponds to scalar multiplication of the input vector v by a scalar called $\alpha(t)$.

$$t \cdot X_\alpha(v) \cdot t^{-1} = X_\alpha\Big(\alpha(t) \cdot v\Big)$$

(P4) There is a "Weyl group" W and for any root $\alpha \in \Phi$, there is a corresponding Weyl group element $w_\alpha \in G(k)$. Also, for any root $\beta \in \Phi$, conjugating $X_\beta(v)$ by w_α corresponds to changing the root β by some kind of reflection process to get a new root $\sigma_\alpha(\beta)$, and doing some invertible linear transformation ϕ to v.

$$w_\alpha \cdot X_\beta(v) \cdot w_\alpha^{-1} = X_{\sigma_\alpha(\beta)}\Big(\varphi(v)\Big)$$

Ideally, φ is not very complicated, something like $\phi(v) = -v$, but sometimes it might be something like: thinking of v as a vector with a 3 or 4 components, $\varphi(v)$ permutes these components and adds some negative signs to some of them.

(P5) Given two roots α and β with $\alpha + \beta \neq 0$, the commutator

$$\Big[X_\alpha(v), X_\beta(w)\Big] = \prod_{?} X_\gamma\Big(N(u,v)\Big)$$

is some product of other pinning map outputs X_γ and the inputs of those are determined by some function N (a polynomial function of u and v). This product is determined by the root system Φ. Often $N(u,v)$ is something like

$$N(u,v) = \pm uv$$

or

$$N(u,v) = c_1 u_1 v_1 + c_2 u_2 v_2 + \cdots + c_n u_n v_n$$

where $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ and the c_i are coefficients coming from the definition of the group G.

Next up are some exercises to get a better hands-on understanding of some aspects of the pinning definition. For all of the exercises below, fix a field k.

Exercise 1. Let G be a group of matrices, and let $X : k \to G$ be a group homomorphism from $(k,+)$ to G. That is,

$$X(u+v) = X(u) \cdot X(v)$$

where $+$ on the left means addition in k, and $\cdot$ on the right means matrix multplication. Show that $X(0)$ is the identity matrix. (Combined with property (P1), what does this say about pinning maps?)

Exercise 2. Show that any matrix of the form

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

with $a, b, c \in k$ can be written as a product of matrices of the form

$$\begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & & v \\ & 1 & \\ & & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & & \\ & 1 & w \\ & & 1 \end{pmatrix}$$

Be careful: It's not just as simple as the product

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

This exercise is illustrating how $\mathrm{SL}_3(k)$ is generated by the images of its pinning maps, a la property (P2). We will study this pinning in detail in section 2. The set of all matrices of this form also has a name: the group of **unipotent** matrices.

Exercise 3. Let $t_1, t_2, t_3 \in k^\times$ be nonzero elements of k. Let $t = \mathrm{diag}(t_1, t_2, t_3)$ be an arbitrary 3×3 diagonal matrix with entries in $k^\times$. Let

$$X = X_{\alpha_{12}}(u) = \begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix}$$

Verify by direct calculation that

$$t \cdot X \cdot t^{-1} = X_{\alpha_{12}}(t_1 t_2^{-1} u) = \begin{pmatrix} 1 & t_1 t_2^{-1} u & \\ & 1 & \\ & & 1 \end{pmatrix}$$

(This is just one example of (P3) for $\mathrm{SL}_3(k)$.)

Exercise 4. Let X and Y be the matrices below.

$$X = X_{\alpha_{12}}(u) = \begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix} \qquad Y = X_{\alpha_{23}}(v) = \begin{pmatrix} 1 & & \\ & 1 & v \\ & & 1 \end{pmatrix}$$

Verify that the commutator $[X, Y]$ has the form

$$X_{\alpha_{13}}(N(u,v)) = \begin{pmatrix} 1 & & N(u,v) \\ & 1 & \\ & & 1 \end{pmatrix}$$

and determine a formula for $N(u, v)$, thinking of N as a function $N : k \times k \to k$. (This is an example of (P5) for $\mathrm{SL}_3(k)$.)

2 Special linear groups

Fix a field k. Whenever convenient, we can assume that the characteristic of k is zero. This assumption is mostly not necessary, but it simplifies some things. The group of units of k is denoted $k^\times$ (in other words, $k^\times$ is the set of nonzero elements in k).

Definition 5. Fix a field k and a positive integer n. $\mathrm{M}_n(k)$ is the set of $n \times n$ matrices with entries in k. For $x \in \mathrm{M}_n(k)$, the **determinant** of x is an element of k. For 2×2 matrices, the usual formula applies.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

For 3×3 matrices, we can use the 2×2 formula and a cofactor expansion to calculate the determinant.

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \cdot \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \cdot \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \cdot \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

This process generalizes to larger $n \times n$ matrices. To calculate the determinant, we expand along the top row. We should not actually have to do this by hand at any point. The point being made here is just that the determinant of an $n \times n$ matrix is defined over any field k, because calculating a determinant only involves addition and multiplication operations in the field k.

Definition 6. Fix a positive integer n. The **special linear group** with coefficients in k is the group of $n \times n$ matrices with determinant 1.

$$\mathrm{SL}_n(k) = \{X \in \mathrm{M}_n(k) : \det X = 1\}$$

$\mathrm{SL}_n(k)$ is an algebraic group under matrix multiplication. The condition $\det(X) = 1$ may not initially appear to be a polynomial equation, but if you expand out the determinant, it is just a big polynomial in terms of the coefficients of the matrix X.

2.1 Torus subgroup and character group

Definition 7. The **diagonal subgroup** or **torus subgroup** in $\mathrm{SL}_n(k)$ is the subgroup of diagonal matrices. We will call it $T(k)$.

$$T(k) = \left\{ \begin{pmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{pmatrix} : t_i \in k^\times, \det t = t_1 t_2 \cdots t_n = 1 \right\}$$

Notice that the equation $t_1 t_2 \cdots t_n = 1$ means that none of the entries t_i can be zero for an element of $T(k)$. For an element of $T(k)$, the inverse is the matrix whose diagonal entries are the term-by-term inverses.

$$\begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix}^{-1} = \begin{pmatrix} t_1^{-1} & & \\ & \ddots & \\ & & t_n^{-1} \end{pmatrix}$$

Definition 8. Let $T(k)$ be the diagonal subgroup of $\mathrm{SL}_n(k)$ as above. For $i = 1, \ldots, n$, define a map $\alpha_i : T(k) \to k^\times$ by

$$\alpha_i \begin{pmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{pmatrix} = t_i$$

Since the determinant of an element of $T(k)$ is one, none of the individual diagonal entries t_i can be zero, so $t_i \in k^\times$. More importantly, α_i is a group homomorphism, meaning $\alpha_i(ts) = \alpha_i(t)\alpha_i(s)$ for any $s, t \in T(k)$.

Definition 9. The **character group** of $T(k)$ is the set of group homomorphisms from $T(k)$ to $k^\times$. It is called $X(T(k))$ or just $X(T)$.

$$X(T) = \mathrm{Hom}(T(k), k^\times) = \big\{\alpha : T(k) \to k^\times \mid \alpha \text{ is a group homomorphism}\big\}$$

Group homomorphisms $T(k) \to k^\times$ are called **characters** (of T). The **sum** of two characters α and β is defined to be

$$\alpha + \beta : T(k) \to k^\times \qquad (\alpha + \beta)(t) = \alpha(t)\beta(t)$$

where the multiplication on the right side is just multiplication in k. The set of characters $X(T)$ is an abelian group under this character sum operation (you checked this on the application last fall). In particular, the identity of $X(T)$ is the constant function

$$0_{X(T)} : T(k) \to k^\times \qquad t \mapsto 1$$

We call this function the **zero character**, and usually just denote it by 0. (You can check that $\alpha + 0_{X(T)} = \alpha$ for any $\alpha \in X(T)$.)

Definition 10. Let $T(k)$ and α_i be as above. For $1 \leq i, j \leq n$ and $i \neq j$, define $\alpha_{ij} = \alpha_i - \alpha_j$. In other words, α_{ij} is a group homomorphism $T(k) \to k^\times$ (that is, α_{ij} is a character of T), and it is described by

$$\alpha_{ij}(t) = (\alpha_i - \alpha_j)(t) = \alpha_i(t)\alpha_j(t)^{-1} = \alpha_{ij} \begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix} = t_i t_j^{-1}$$

Remark 11. In $T(k)$, we have the equation $\det(t) = 1$. Since the determinant of a diagonal matrix is the product of the diagonal entries, we get the equation

$$1 = \det(t) = t_1 \cdots t_n = \alpha_1(t) \cdots \alpha_n(t)$$

That is,

$$\alpha_1(t) \cdots \alpha_n(t) = 1 = 0_{X(T)}(t)$$

where $0_{X(T)} \in X(T)$ is the zero character. In other words, the sum

$$\alpha_1 + \cdots + \alpha_n \in X(T)$$

is the zero character. Translating this into an equation in $X(T)$, we get the following.

$$\alpha_1 + \cdots + \alpha_n = 0_{X(T)} = 0$$

We emphasize: this is not an equation in k, it is an equation in the abelian group $X(T)$.

Definition 12. For $n \geq 1$, define $\mathbb{Z}^n = \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}$ (n copies of $\mathbb{Z}$).

Lemma 13. *Let $\mathbb{Z}\vec{1}$ be the set of integer multiples of the vector $\vec{1} = (1, \ldots, 1)$.*

$$\mathbb{Z}\vec{1} = \{m(1, \ldots, 1) = (m, \ldots, m) : m \in \mathbb{Z}\}$$

Then $X(T) \cong \mathbb{Z}^n/\mathbb{Z}\vec{1}$ (isomorphism of abelian groups).

Proof. Let $e_i \in \mathbb{Z}^n$ be the n-tuple $(0, \ldots, 0, 1, 0, \ldots, 0)$ which is all zeros except for a one in the ith position. Then consider the map

$$\phi : \mathbb{Z}^n \to X(T) \qquad \phi\left(\sum_{i=1}^{n} m_i e_i\right) = \sum_{i=1}^{n} m_i \alpha_i$$

where the m_i are integers. Then ϕ is a group homomorphism, $\ker\phi = \mathbb{Z}\vec{1}$, and ϕ is onto, so by the first isomorphism theorem, $\mathbb{Z}^n/\ker\phi = \mathbb{Z}^n/\mathbb{Z}\vec{1} \cong X(T)$. $\square$

The following elements correspond under the above isomorphism.

$$\begin{aligned} e_i &\longleftrightarrow \alpha_i \\ e_i - e_j &\longleftrightarrow \alpha_{ij} = \alpha_i - \alpha_j \\ (1, -1, 0, 0, \ldots, 0) &\longleftrightarrow \alpha_{12} = \alpha_1 - \alpha_2 \\ (0, 1, -1, 0, \ldots, 0) &\longleftrightarrow \alpha_{23} = \alpha_2 - \alpha_3 \\ (0, 0, 1, -1, \ldots, 0) &\longleftrightarrow \alpha_{34} = \alpha_3 - \alpha_4 \end{aligned}$$

This is somewhat misleading – more literally, on the left side of this correspondence I mean that the coset $\overline{e}_i = e_i + \mathbb{Z}\vec{1}$ corresponds to the character α_i, and so on. This becomes apparent when we consider the corresponding elements

$$\vec{1} = (1, \ldots, 1) \longleftrightarrow \alpha_1 + \cdots + \alpha_n$$

because $\alpha_1 + \cdots + \alpha_n = 0$ in $X(T)$, but $\vec{1}$ is not the zero element in $\mathbb{Z}^n$. However, $\vec{1}$ does represent the zero coset in the quotient $\mathbb{Z}^n/\mathbb{Z}\vec{1}$; that is, $\vec{1} + \mathbb{Z}\vec{1} = \mathbb{Z}\vec{1} = 0$ in $\mathbb{Z}^n/\mathbb{Z}\vec{1}$.

2.2 Pinning the special linear group

Let k be a field of characteristic zero, and consider the special linear group $\mathrm{SL}_n(k)$.

Definition 14. Fix i, j with $1 \leq i, j \leq n$ and $i \neq j$. For $x \in k$, let $E_{ij}(x)$ be the matrix with x in the ij entry and 0 elsewhere. Let $e_{ij}(x) = I + E_{ij}(x)$ be the matrix with x in the ij entry, 1's on the diagonal, and 0 elsewhere. For example, in $\mathrm{SL}_2(k)$ we have

$$E_{12}(x) = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \qquad e_{12}(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$$

The determinant of $e_{ij}(x)$ is one, so $e_{ij}(x) \in \mathrm{SL}_n(k)$. Now, consider e_{ij} as a group homomorphism from k (viewed as an abelian group under addition) to $\mathrm{SL}_n(k)$.

$$e_{ij} : k \to \mathrm{SL}_n(k) \qquad x \mapsto e_{ij}(x)$$

This being a group homomorphism means that

$$e_{ij}(x+y) = e_{ij}(x) \cdot e_{ij}(y)$$

Since e_{ij} is a homomorphism, it must take the identity to the identity.

$$e_{ij}(0) = I = 1 \in \mathrm{SL}_n(k)$$

Also notice that $\ker(e_{ij}) = \{0\}$, so e_{ij} is injective.

Definition 15. Instead of e_{ij}, we will use the notation $X_{\alpha_{ij}}$ for the same thing. That is, $X_{\alpha_{ij}}(u) = e_{ij}(u)$ for $u \in k$.

Remark 16. It is useful to remember that $X_{\alpha_{ij}}(u)^{-1} = X_{\alpha_{ij}}(-u)$. This is a consequence of the fact that $X_{\alpha_{ij}}$ is a group homomorphism.

Definition 17. The **root system** of $\mathrm{SL}_n(k)$ is the set of characters α_{ij} for $i \neq j$. We denote it by $\Phi(\mathrm{SL}_n(k), T(k))$ or $\Phi(\mathrm{SL}_n, T)$ or just Φ.

$$\Phi = \{\alpha_{ij} : 1 \leq i, j \leq n, i \neq j\}$$

Elements of Φ are called **roots**. It is important not to forget that elements of Φ are characters, i.e. functions from $T(k)$ to $k^\times$. Specifically, α_{ij} is the function

$$\alpha_{ij} : T(k) \to k^\times \qquad \alpha_{ij}(t) = \alpha_i(t)\alpha_j(t)^{-1} = t_i t_j^{-1}$$

where $t = \mathrm{diag}(t_1, \ldots, t_n)$.

2.2.1 Conjugation by torus

$T(k)$ acts on $\mathrm{SL}_n(k)$ by conjugation. That is, there is a group action

$$T(k) \times \mathrm{SL}_n(k) \to \mathrm{SL}_n(k) \qquad t * x = txt^{-1}$$

Exercise 18. Verify that this is a group action, by showing that

- For any $t \in T(k)$ and $x \in \mathrm{SL}_n(k)$, the conjugation $t * x = txt^{-1}$ is in $\mathrm{SL}_n(k)$.
- If I is the identity matrix and $x \in \mathrm{SL}_n(k)$, then $I * x = IxI^{-1} = x$.
- For any two elements $s, t \in T(k)$, we have $(ts) * x = t * (s * x)$.

Lemma 19. *Let $1 \leq i, j \leq n$ and $i \neq j$. Let $t \in T(k)$ be a diagonal matrix, and let $u \in k$. Then*

$$t * X_{\alpha_{ij}}(u) = t \cdot X_{\alpha_{ij}(u)} \cdot t^{-1} = X_{\alpha_{ij}}\Big(\alpha_{ij}(t)u\Big)$$

In other words, for $\alpha \in \Phi(\mathrm{SL}_n(k), T(k))$, we have

$$t * X_\alpha(u) = t \cdot X_\alpha(u) \cdot t^{-1} = X_\alpha\Big(\alpha(t)u\Big)$$

Exercise 20. In the group $\mathrm{SL}_2(k)$, confirm by direct matrix calculation that for any $t = \mathrm{diag}(t_1, t_2)$ and $u \in k$, that we have

$$t \cdot X_{\alpha_{12}}(u) \cdot t^{-1} = X_{\alpha_{12}}\Big(\alpha_{12}(t)u\Big) = X_{\alpha_{12}}(t_1 t_2^{-1} u)$$

That is,

$$\begin{pmatrix} t_1 & \\ & t_2 \end{pmatrix} \begin{pmatrix} 1 & u \\ & 1 \end{pmatrix} \begin{pmatrix} t_1 & \\ & t_2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & t_1 t_2^{-1} u \\ & 1 \end{pmatrix}$$

Then convince yourself that nothing really changes when you do the "same" calculation in $\mathrm{SL}_n(k)$, replacing t with $\mathrm{diag}(t_1, t_2, \ldots, t_n)$.

2.2.2 Conjugation by Weyl group

Definition 21. Let $\alpha \in \Phi(\mathrm{SL}_n(k), T(k))$ be a root, so $\alpha = \alpha_{ij}$ for some $i \neq j$. Let $u \in k$, and assume $u \neq 0$. Then define

$$w_\alpha(u) = X_\alpha(u) \cdot X_{-\alpha}(-u^{-1}) \cdot X_\alpha(u)$$

Here, $-\alpha$ means the negative of α in the abelian group $X(T)$. Concretely, if $\alpha = \alpha_{ij} = \alpha_i - \alpha_j$, then $-\alpha = \alpha_{ji} = \alpha_j - \alpha_i$. The elements $w_\alpha(u)$ are called **Weyl group elements**. (There is a larger "Weyl group" including the $w_\alpha(u)$, but we omit the details for now.)

Example 22. In the group $\mathrm{SL}_3(k)$, with $\alpha = \alpha_{12} = \alpha_1 - \alpha_2$, where $X_\alpha(u) = e_{12}(u)$, that we get

$$w_\alpha(1) = w_{\alpha_{12}}(1) = \begin{pmatrix} & 1 & \\ -1 & & \\ & & 1 \end{pmatrix}$$

Exercise 23. Work out $w_{\alpha_{12}}(u)$ in more generality (still in $\mathrm{SL}_3(k)$), by doing concrete matrix multiplication with $e_{12}(u)$ matrices.

Definition 24. A **monomial matrix** is a matrix with exactly one nonzero entry in each row and each column. A **permutation matrix** is a monomial matrix where all nonzero entries are 1. A **signed permutation matrix** is a monomial matrix where all nonzero entries are 1 or -1.

Remark 25. In general, for $\alpha \in \Phi(\mathrm{SL}_n(k), T(k))$, the Weyl group elements $w_\alpha(u)$ are monomial matrices, and $w_\alpha(1)$ is a signed permutation matrix.

Exercise 26. Show that for $\alpha \in \Phi(\mathrm{SL}_n(k), T(k))$ and $u \in k^\times$, $w_\alpha(u)^{-1} = w_\alpha(-u)$.

Definition 27. Consider the space $\mathbb{R}^n$, and denote the dot product of two vectors $v, w \in \mathbb{R}^n$ by (v, w). Given $v \in \mathbb{R}^n$, the **hyperplane perpendicular to** v is the set of all vectors perpendicular to v.

$$P_v = \{w \in \mathbb{R}^n : (v, w) = 0\}$$

The **reflection across** P_v is the (linear) map

$$\sigma_v : \mathbb{R}^n \to \mathbb{R}^n \qquad \sigma_v(w) = w - \frac{2(v,w)}{(v,v)} v$$

Geometrically, this is actually a mirror reflection across P_v. Now consider $\Phi = \Phi(\mathrm{SL}_n(k), T(k))$. Identify Φ with a subset of $\mathbb{R}^n$ by identifying α_i with the standard basis vector $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 in the ith position. Then $\alpha_{ij} = \alpha_i - \alpha_j$ is identified with

$$\alpha_{ij} \leftrightarrow (0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots, 0)$$

with the 1 in the ith position and -1 in the jth position. Under this identification, the dot product operation on $\mathbb{R}^n$ becomes

$$(\alpha_i, \alpha_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

and then

$$\begin{aligned} (\alpha_{ij}, \alpha_{k\ell}) &= (\alpha_i - \alpha_j, \alpha_k - \alpha_\ell) \\ &= (\alpha_i, \alpha_k) - (\alpha_i, \alpha_\ell) - (\alpha_j, \alpha_k) + (\alpha_j, \alpha_\ell) \\ &= \delta_{ik} - \delta_{i\ell} - \delta_{jk} + \delta_{j\ell} \\ &\in \{-2, -1, 0, 1, 2\} \end{aligned}$$

Now, for $\alpha \in \Phi$, consider α as an element of $\mathbb{R}^n$ under the identification above. Then we have a hyperplane P_α of vectors perpendicular to α, and we can again use the reflection σ_α across P_α.

$$\sigma_\alpha : \mathbb{R}^n \to \mathbb{R}^n \qquad \sigma_\alpha(w) = w - \frac{2(\alpha, w)}{(\alpha, \alpha)}\alpha$$

Lemma 28. *Let $\alpha \in \Phi = \Phi(\mathrm{SL}_n(k), T(k))$, and let σ_α be the reflection above. If $\beta \in \Phi$, then $\sigma_\alpha(\beta) \in \Phi$. In other words, σ_α can be thought of as a map $\Phi \to \Phi$.*

$$\sigma_\alpha : \Phi \to \Phi \qquad \sigma_\alpha(\beta) = \beta - \frac{2(\alpha, \beta)}{(\alpha, \alpha)}\alpha$$

Example 29. When $n = 3$, we have

$$\Phi = \Phi(\mathrm{SL}_n, T) = \{\alpha_{12}, \alpha_{21}, \alpha_{13}, \alpha_{31}, \alpha_{23}, \alpha_{32}\}$$

We identify Φ with a subset of $\mathbb{R}^3$, the set of vectors

$$\widetilde{\Phi} = \{(1, -1, 0), (-1, 1, 0), (1, 0, -1), (-1, 0, 1), (0, 1, -1), (0, -1, 1)\}$$

These vectors are all perpendicular to the vector $(1, 1, 1)$. In other words, $\widetilde{\Phi}$ sits inside a 2D plane, so we can visualize $\widetilde{\Phi}$ in two dimensions. When we do, it looks like six vectors pointing at the six vertices of a regular hexagon.

Try to convince yourself that if you take any two of those vectors v and w, take P_v to be the line perpendicular to v, and reflect w across P_v, then that reflection $\sigma_v(w)$ is another one of the six vectors in Φ.

Proposition 30. *Let $\alpha, \beta \in \Phi(\mathrm{SL}_n(k), T(k))$ be roots. Then*

$$w_\alpha(1) \cdot X_\beta(u) \cdot w_\alpha(1)^{-1} = X_{\sigma_\alpha(\beta)}(\pm u)$$

2.2.3 Commutator formula

Definition 31. Let G be any group, and let $x, y \in G$. The **commutator** of x and y is

$$[x, y] = xyx^{-1}y^{-1}$$

Exercise 32. Show that x and y commute (i.e. $xy = yx$) if and only if $[x, y] = 1$ (here 1 is the identity element of the group G).

Exercise 33. Let $u, v \in k$. In the group $\mathrm{SL}_n(k)$, compute the commutator

$$\Big[X_{\alpha_{12}}(u), X_{\alpha_{23}}(v)\Big]$$

Proposition 34. *Let $\Phi = \Phi(\mathrm{SL}_n(k), T(k))$, and let $\alpha, \beta \in \Phi$ so that $\alpha \neq \pm\beta$. Let $u, v \in k$. Then*

$$\Big[X_\alpha(u), X_\beta(v)\Big] = \begin{cases} I & \alpha + \beta \notin \Phi \\ X_{\alpha+\beta}(\pm uv) & \alpha + \beta \in \Phi \end{cases}$$

That is, $X_\alpha(u)$ and $X_\beta(v)$ commute if and only if $\alpha + \beta \notin \Phi$ (or if $uv = 0$).

Remark 35. The signs in the previous proposition can be described more precisely as follows. Let $c_{\alpha\beta}$ be the sign, so

$$\Big[X_\alpha(u), X_\beta(v)\Big] = \begin{cases} I & \alpha + \beta \notin \Phi \\ X_{\alpha+\beta}(c_{\alpha\beta}uv) & \alpha + \beta \in \Phi \end{cases}$$

If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$, write $\alpha = \alpha_{ij}$ and $\beta = \alpha_{k\ell}$. For $\alpha + \beta$ to be a valid root, the set $\{i, j, k, \ell\}$ must contain exactly three distinct elements, meaning an index must match to allow cancellation. Out of the four possible index alignments, $i = k$ and $j = \ell$ yield sums that are not roots. The remaining two valid cases determine the sign:

Case	Sum $\alpha + \beta$	$c_{\alpha\beta}$
$j = k$	$\alpha_{ij} + \alpha_{jk} = \alpha_{ik}$	1
$i = \ell$	$\alpha_{ij} + \alpha_{ki} = \alpha_{kj}$	-1

You can check cases of this by hand by calculating a few commutators. For example,

$$\Big[X_{\alpha_{12}}(u), X_{\alpha_{31}}(v)\Big] = X_{\alpha_{32}}(-uv)$$

falls into the $i = \ell$ case.

2.3 "*Nonlinearity of matrix groups*"

In this section, we discuss a key construction from the 2009 paper "*Nonlinearity of matrix groups*" by Martin Kassabov and Mark Sapir, which utilizes a pinning of $\mathrm{SL}_n(k)$ to prove a theorem about rigidity. Technically, they study the elementary linear group but we discuss how this is the same as the special linear group under some assumptions. In particular we focus on the ideas of Theorem 3 as they apply to rigidity statements for abstract homomorphisms of algebraic groups.

Definition 36. A **unital ring** is a set R with two operations $+$ and $\times$, and these operations satisfy:

1. $(R,+)$ is an abelian group with identity element 0. That is, $+$ is associative, commutative, and every element a has a unique additive inverse $-a$.

$$\begin{aligned} a+(b+c) &= (a+b)+c \\ a+b &= b+a \\ a+0 &= a \\ a+(-a) &= 0 \end{aligned}$$

2. $(R,\times)$ is not a group, but is associative and has an identity element 1. Note that we do not assume $\times$ is commutative.

$$a(bc) = (ab)c \qquad a1 = 1a = a$$

3. Addition and multiplication play well together via the distributive property.

$$a(b+c) = ab+ac \qquad (a+b)c = ac+bc$$

Let R be a unital ring, and let K be a field of characteristic zero. Let n be a positive integer.

Definition 37. The **general linear group** of R is $\mathrm{GL}_n(R)$, the group of $n \times n$ invertible matrices with entries in R. Equivalently, $\mathrm{GL}_n(R)$ is the set of $n \times n$ matrices whose determinant is a unit (multiplicatively invertible) element of R. $\mathrm{SL}_n(R)$ is the subgroup of $\mathrm{GL}_n(R)$ of matrices of determinant one.

Definition 38. For $1 \leq i,j \leq n$ and $i \neq j$, and $r \in R$, let $x_{ij}(r)$ be the matrix with 1's on the diagonal and r in the (i,j) off-diagonal entry. Note that $\det x_{ij}(r) = 1$, so $x_{ij}(r) \in \mathrm{SL}_n(R)$.

In other places in this document, $x_{ij}(r)$ is called $X_{\alpha_{ij}}(r)$. We have previously observed that $X_{\alpha_{ij}} : R \to \mathrm{SL}_n(R)$ is a group homomorphism, which means in this context that for any $r_1, r_2 \in R$, we have

$$x_{ij}(r_1+r_2) = x_{ij}(r_1) \cdot x_{ij}(r_2)$$

Definition 39. Let G be a group, and let $S \subset G$ be a subset. The **subgroup generated by** S is the subset of G consisting of all elements that can be written as a product of elements in S and their inverses. Equivalently, it is the smallest subgroup of G containing S.

Definition 40. Let S be the following subset of $\mathrm{GL}_n(R)$.

$$S = \{x_{ij}(r) : 1 \leq i,j \leq n \text{ and } i \neq j \text{ and } r \in R\}$$

We define $\mathrm{EL}_n(R)$ to be the subgroup of $\mathrm{GL}_n(R)$ generated by S.

Lemma 41. $\mathrm{EL}_n(R) \subset \mathrm{SL}_n(R)$

Proof. Since every $x_{ij}(r)$ has determinant one, any product of them (or their inverses) will also have determinant one. ☐

Lemma 42. *If R is a field, then* $\mathrm{EL}_n(R) = \mathrm{SL}_n(R)$.

Proof. If R is a field, then Gaussian elimination works. What does Gaussian elimination do? It uses elementary row operations to rewrite a square matrix as the identity matrix. But each time you do an elementary row operation, that's like multiplying by some $x_{ij}(r)$ or a permutation matrix. ☐

The subgroups A and U

Let n and k be positive integers, and assume $k \geq 3$. For all of the following calculations, we can assume $k = 3$ without loss of generality. Let

$$\pi : \mathrm{EL}_k(R) \to \mathrm{GL}_n(K)$$

be a group homomorphism. R is any unital associative commutative ring, and K is a field of characteristic zero. The paper refers to π as a "linear representation," but this is just another way of saying group homomorphism into some general linear group. In particular, π is what I call an "abstract homomorphism," because we do not assume it is algebraic at all, even though $\mathrm{EL}_k(R)$ and $\mathrm{GL}_n(K)$ can be thought of as algebraic groups.

Definition 43. This notation is not in the paper, but I think it helps clarify things. Let

$$A = \{x_{13}(r) : r \in R\}$$

That is, A is the subset of $\mathrm{EL}_k(R)$ of matrices of the form

$$\begin{pmatrix} 1 & 0 & r & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix}$$

In other words, A is the image of the map $x_{13} : R \to \mathrm{EL}_k(R)$.

$$A = x_{13}(R)$$

Since x_{13} is a group homomorphism and R is a group (under addition), the image must be a subgroup. That is, A is a subgroup of $\mathrm{EL}_k(R)$. The group operation is matrix multiplication.

$$\begin{pmatrix} 1 & 0 & r_1 & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix} \begin{pmatrix} 1 & 0 & r_2 & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} 1 & 0 & r_1 + r_2 & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix}$$

Even though matrix multiplication is usually not commutative, in this context the matrix multiplication within A is commutative, because addition in R is commutative.

$$\begin{pmatrix} 1 & 0 & r_1 + r_2 & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} 1 & 0 & r_2 + r_1 & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix}$$

In fact, if we just view $(R, +)$ as an abelian group, we can think of x_{13} as a group isomorphism between $(R, +)$ and $(A, \times)$.

$$x_{13} : R \to A \qquad x_{13}(r) = \begin{pmatrix} 1 & 0 & r & \\ & 1 & 0 & \\ & & 1 & \\ & & & \ddots \end{pmatrix}$$

The inverse map is just sending that matrix to $x_{13}(r)$.

Definition 44. Let A be as above, and recall that we have a group homomorphism $\pi : \mathrm{EL}_k(R) \to \mathrm{GL}_n(K)$. Let

$$U = \pi(A) = \{\pi(a) : a \in A\} = \{\pi(x_{13}(r)) : r \in R\}$$

So we have A sitting inside $\mathrm{EL}_k(R)$, and U sitting inside $\mathrm{GL}_n(K)$. The diagram below depicts the situation. The upward arrows are just the inclusion maps, and the diagram commutes, meaning following a path of arrows along one side of the square is the same as going the other way.

$$\begin{tikzcd} \mathrm{EL}_k(R) \arrow[r, "\pi"] & \mathrm{GL}_n(K) \\ A \arrow[u, hook] \arrow[r, "\pi|_A"] & U \arrow[u, hook] \end{tikzcd}$$

Also note that $\pi|_A(A) = U$, or in other words, $\pi|_A$ is a surjective map from A to U.

Remark 45. R is a ring, so $(R, +)$ is an abelian group. Then $x_{13} : R \to A$ is a group isomorphism, so A is an abelian group. Then $\pi : A \to U$ is a surjective group homomorphism, so U is also an abelian group.

Definition 46. Let V be the "Zariski closure" of U inside $\mathrm{GL}_n(K)$. This is not something we need to understand in a very thorough way. Basically, V is still a subset of $\mathrm{GL}_n(K)$, and it should be "just a little bit bigger" than U. As an analogy, if U was an open interval on the real line, then V would just be the closed version of that interval, just adding in the endpoints. Or as a second analogy, if U was an open disk in the plane, not including its boundary, then V would be the closed disk, just U plus the boundary.

We defined V, but we will mostly ignore it and focus on U.

Theorem 3 - multiplication in U

The big idea of Theorem 3 is to make U into a ring by adding a second operation. In order to make an abelian group into a unital ring, we need to:

- Define a second operation
- Verify that the second operation is associative and has a unity element
- Verify that the distributive property holds

In general, there is not a particularly "natural" way to make an abelian group into a ring. Even if you find a way, it may not be the only way. Technically, Theorem 3 is about making V into a ring, but you can mostly ignore this and pretend that it's happening to U.

Definition 47. Next we describe the Weyl group elements of $\mathrm{SL}_3(R)$ used in the proof of Theorem 3. Set

$$w_{12} = \begin{pmatrix} & -1 & \\ 1 & & \\ & & 1 \end{pmatrix} \qquad w_{23} = \begin{pmatrix} 1 & & \\ & & 1 \\ & -1 & \end{pmatrix}$$

The paper does not call these "Weyl group elements," but they are elements of the Weyl group of $\mathrm{SL}_3(R)$.

Lemma 48. *For any $r \in R$,*

$$w_{12} \cdot x_{13}(r) \cdot w_{12}^{-1} = x_{23}(r) \qquad w_{23} \cdot x_{13}(r) \cdot w_{23}^{-1} = x_{12}(r)$$

Remark 49. I will not discuss the proof; you can just do some matrix calculations to verify it yourself. However, I want to make a note here on how these fit into the theory of root systems and our general Weyl group reflection formula. In the notation of this document, the first equation is

$$w_{\alpha_{12}}(-1) \cdot X_{\alpha_{13}}(r) \cdot w_{\alpha_{12}}(-1)^{-1} = X_{\alpha_{23}}(r)$$

Hopefully you remember that in the general version of the formula, the right hand side subscript on X is described in terms of a reflection. That general version is

$$w_\alpha(u) \cdot X_\beta(r) \cdot w_\alpha(u)^{-1} = X_{\sigma_\alpha(\beta)}(\pm r)$$

The reason α_{23} shows up on the right side of the formula above is that $\alpha_{23} = \sigma_{\alpha_{12}}(\alpha_{13})$.

Remark 50. We have talked about the general commutator formula in SL_n. In the context of Theorem 3, the only version of this that we need is

$$\Big[x_{12}(r_1), x_{23}(r_2)\Big] = x_{13}(r_1 r_2)$$

Also note that in general, group homomorphisms can "go inside" commutators, as follows. If $\pi : G \to H$ is a group homomorphism and $x, y \in G$, then

$$\pi\big([x,y]\big) = \pi(xyx^{-1}y^{-1}) = \pi(x)\pi(y)\pi(x)^{-1}\pi(y)^{-1} = [\pi(x), \pi(y)]$$

The version of this to remember is

$$\pi\big([x,y]\big) = [\pi(x), \pi(y)]$$

Definition of multiplication in U. Finally, we are ready to define the multiplication operation in U, which is the big accomplishment of Theorem 3. Our goal is to define a second operation $U \times U \to U$ (and then extend this to V) to make U into a ring, and furthermore to do this in a way which makes the map

$$\rho = \pi \circ x_{13} : R \to \mathrm{GL}_n(K)$$

into a ring homomorphism. That is, we want to have

$$\rho\left(r_1 \underset{R}{+} r_2\right) = \rho(r_1) \underset{U}{+} \rho(r_2)$$
$$\rho\left(r_1 \underset{R}{\times} r_2\right) = \rho(r_1) \underset{U}{\times} \rho(r_2)$$

On the left side, the operations $\underset{R}{+}$ and $\underset{R}{\times}$ are in the ring R. On the right side, $\underset{U}{+}$ actually means matrix multiplication and $\underset{U}{\times}$ means some other weird operation, which we haven't defined yet. The way to define it is

$$U \times U \to \mathrm{GL}_n(K) \qquad u_1 \times u_2 = \Big[\pi(w_{23}) \cdot u_1 \cdot \pi(w_{23})^{-1}, \pi(w_{12}) \cdot u_2 \cdot \pi(w_{12})^{-1}\Big]$$

All of the dots here are matrix multiplication inside $\mathrm{GL}_n(K)$. Note that in Kassabov and Sapir, they define it without some of these π's. Their version looks like

$$u_1 \times u_2 = \left[w_{23} \cdot u_1 \cdot w_{23}^{-1}, w_{12} \cdot u_2 \cdot w_{12}^{-1}\right]$$

I think that the π is meant to be implied, because without it this just doesn't make sense because because matrices from different groups are being multiplied together.

We want this weird multiplication to be a map into U, not just a map into $\mathrm{GL}_n(K)$, so let's confirm that. Let $u_1, u_2 \in U$. Then we can always write them in the form $u_1 = \rho(r_1) = \pi(x_{13}(r_1))$ and $u_2 = \rho(r_2) = \pi(x_{13}(r_2))$ for some $r_1, r_2 \in R$.

$$\begin{aligned}
\rho(r_1) \times \rho(r_2) &= \left[\pi(w_{23}) \cdot \rho(r_1) \cdot \pi(w_{23})^{-1}, \pi(w_{12}) \cdot \rho(r_2) \cdot \pi(w_{12})^{-1}\right] \\
&\text{rewrite } \rho(r_1) \text{ as } \pi(x_{13}(r_1)) \\
&= \left[\pi(w_{23}) \cdot \pi(x_{13}(r_1)) \cdot \pi(w_{23})^{-1}, \pi(w_{12}) \cdot \pi(x_{13}(r_2)) \cdot \pi(w_{12})^{-1}\right] \\
&\text{use the fact that } \pi \text{ is a group homomorphism to ``factor it out''} \\
&= \left[\pi\left(w_{23} \cdot x_{13}(r_1) \cdot w_{23}^{-1}\right), \pi\left(w_{12} \cdot x_{13}(r_2) \cdot w_{12}^{-1}\right)\right] \\
&\text{move } \pi \text{ outside the commutator using the fact that it is a group homomorphism} \\
&= \pi\Big(\left[w_{23} \cdot x_{13}(r_1) \cdot w_{23}^{-1}, w_{12} \cdot x_{13}(r_2) \cdot w_{12}^{-1}\right]\Big) \\
&\text{now use Weyl group conjugation formulas} \\
&= \pi\Big(\left[x_{12}(r_1), x_{23}(r_2)\right]\Big) \\
&\text{now use commutator formula} \\
&= \pi\Big(x_{13}(r_1 r_2)\Big) \\
&= \rho(r_1 r_2)
\end{aligned}$$

Ok, what was the point of all of this calculating? We defined a map $\times : U \times U \to \mathrm{GL}_n(K)$ by this weird conjugation formula, and what we just showed is that for any $r_1, r_2 \in R$, we have

$$\rho(r_1 r_2) = \rho(r_1) \times \rho(r_2)$$

Since ρ is a surjective map from R to U, every element of U can be written in the form $\rho(r)$ for some $r \in R$. Thus, for any $u_1, u_2 \in U$ we can write them as

$$u_1 = \rho(r_1) \qquad u_2 = \rho(r_2)$$

and then

$$u_1 \times u_2 = \rho(r_1) \times \rho(r_2) = \rho(r_1 r_2)$$

In particular, the product $u_1 \times u_2$ is in U. So the map we defined as a map $U \times U \to \mathrm{GL}_n(K)$ is actually a map $U \times U \to U$.

At this point, there is a bit of technical work to show that the two operations we have defined actually satisfy the axioms for a ring. This involves some details beyond the scope of our project, but it's not actually super interesting.

Summary. In "*Nonlinearity of matrix groups*," particularly in Definition 2 and Theorem 3, Kassabov and Sapir showed how using some of the formulas involved in a pinning of a matrix group can be useful to prove something about rigidity. In particular, they used pinning formulas related to the group SL_n, and used the formulas to assist in defining a weird kind of multiplication operation on the set of matrices of the form

$$\begin{pmatrix} 1 & & x \\ & 1 & \\ & & 1 \end{pmatrix}$$

It's a few steps more from there to proving something about rigidity, but that's beyond our current scope.

3 Forms

In this section, we build up the necessary background on bilinear forms, hermitian forms, and quadratic forms needed to describe orthogonal and unitary groups. Our objective is just to study them enough to understand what the Witt index of a form is. An excellent alternate reference on bilinear and quadratic forms is the following notes by Keith Conrad `https://kconrad.math.uconn.edu/blurbs/linmultialg/bilinearform.pdf`.

3.1 Bilinear forms

Before discussing bilinear forms in general, we look at the classical example of a bilinear form: the dot product on $\mathbb{R}^n$.

Example 51. Let $n \geq 1$, and consider the vector space $V = \mathbb{R}^n$ (as a vector space of dimension n over $\mathbb{R}$, with elements thought of as column vectors). The usual dot product operation is a map

$$\begin{aligned} b : V \times V &\to \mathbb{R} \\ b(x, y) &= x \cdot y \\ &= b\Big((x_1, \dots, x_n), (y_1, \dots, y_n)\Big) \\ &= (x_1, \dots, x_n) \cdot (y_1, \dots, y_n) \\ &= x_1 y_1 + \cdots + x_n y_n \\ &= \sum_{i=1}^{n} x_i y_i \end{aligned}$$

The map b has the following properties: for any $x, y, z \in V$, and any $s, t \in \mathbb{R}$,

$$\begin{aligned} b(x, y + z) &= b(x, y) + b(x, z) \\ b(x + y, z) &= b(x, z) + b(y, z) \\ b(sx, ty) &= (st)b(x, y) \end{aligned}$$

These properties are summarized by saying that b is a **bilinear map**, or by saying that b is **linear in each entry**. Another way to think about b is the following. Given $x, y \in \mathbb{R}^n$, view them as column vectors, and then

$$b(x, y) = x^T \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} y$$

where x^T means the transpose (so turning x from a column vector into a row vector).

A bilinear form (on a vector space over a field k) is a generalization of these properties of the dot product.

Definition 52. Fix a field k. A **vector space over** k is an abelian group V (where we write the operation as $+$) together with a map

$$k \times V \to V$$

with the properties that for any $a, b \in k$ and $u, v \in V$ we have

$$\begin{aligned} 0v &= 0 \\ 1v &= v \\ a(bv) &= (ab)v \\ a(v+w) &= av + aw \\ (a+b)v &= av + bv \end{aligned}$$

Example 53. Fix a positive integer n. Then $V = k^n$ is a vector space over k, where vector addition is component-wise,

$$(v_1, v_2, \ldots, v_n) + (u_1, u_2, \ldots, u_n) = (v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n)$$

and scalar multiplication is also component-wise.

$$a(v_1, \ldots, v_n) = (av_1, \ldots, av_n)$$

Definition 54. The **dimension** of a vector space V over k is the size of any basis (a linearly independent spanning set). We will only encounter finite dimensional vector spaces. It is a theorem that the size of any two bases are the same, so the dimension is well-defined.

The dimension of k^n is n. Essentially, any two vector spaces of the same dimension are "the same," meaning that they are isomorphic as vector spaces, so every finite dimensional vector space over k is isomorphic to k^n for some n.

Definition 55. Let k be a field, and let V be a vector space of (finite) dimension n over k (so essentially you can think of V as k^n). A **bilinear form** on V is a map

$$b : V \times V \to k$$

such that for any $u, v, w \in V$ and $\lambda, \mu \in k$, we have

$$\begin{aligned} b(u, v+w) &= b(u,v) + b(u,w) \\ b(u+v, w) &= b(u,w) + b(v,w) \\ b(\lambda v, \mu w) &= \lambda\mu\, b(v,w) \end{aligned}$$

Definition 56. A bilinear form is **symmetric** if $b(v, w) = b(w, v)$ for all $v, w \in V$.

Example 57. The dot product on $\mathbb{R}^n$ is a symmetric bilinear form.

Definition 58. Let $b : V \times V \to k$ be a bilinear form, and fix an ordered basis $\beta = \{e_1, \ldots, e_n\}$ of V. The **matrix of b with respect to the basis** β is the matrix $B = \mathrm{Mat}_\beta(b)$ whose ijth entry is $b(e_i, e_j)$.

$$B = \mathrm{Mat}_\beta(b) = \Big(b(e_i, e_j)\Big)$$

Definition 59. The **Kronecker delta function** is defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Example 60. Let $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the usual dot product, and let $e_i \in \mathbb{R}^n$ be the usual basis vector with 1 in the ith entry and zeros elsewhere. Then $\beta = \{e_1, \ldots, e_n\}$ is called the **standard basis** of $\mathbb{R}^n$. Note that

$$b(e_i, e_j) = \delta_{ij}$$

and the matrix of b with respect to β is

$$B = \mathrm{Mat}_\beta(b) = \Big(b(e_i, e_j)\Big) = (\delta_{ij})$$

which is just the identity matrix.

Definition 61. Let $B \in \mathrm{M}_n(k)$ be an $n \times n$ matrix. Let V be a vector space over k of dimension n, and fix an ordered basis $\beta = \{e_1, \ldots, e_n\}$ of V. For any $v \in V$, we can write it uniquely in terms of this basis.

$$v = \sum_{i=1}^{n} v_i e_i$$

Let $[v]_\beta$ be the column vector $(v_1, \ldots, v_n) \in k^n$. In particular, $[e_i]_\beta$ is the vector with zeros except for a 1 in the ith entry. The **bilinear form associated to B (with respect to the basis β)** is the bilinear form $b_{\beta,B} : V \times V \to k$ defined by

$$b_{\beta,B}(v, w) = [v]_\beta^T B [w]_\beta$$

Notice that if $v = e_i$ and $w = e_j$ then $b_{\beta,B}(e_i, e_j)$ is the ijth entry of the matrix B.

$$b_{\beta,B}(e_i, e_j) = [e_i]_\beta^T B [e_j]_\beta = B_{ij}$$

Lemma 62. *Let k be a field, let V be a vector space of dimension n over k, and fix an ordered basis β of V. Then there is a one-to-one correspondence between bilinear forms and elements of $\mathrm{M}_n(k)$.*

$$\begin{aligned} \{\text{bilinear forms}\} &\leftrightarrow \mathrm{M}_n(k) \\ b &\leftrightarrow \mathrm{Mat}_\beta(b) \\ b_{\beta,B} &\leftrightarrow B \end{aligned}$$

Exercise 63. Show that the matrix of a symmetric bilinear form is a symmetric matrix (satisfies $B^T = B$).

Definition 64. A matrix $B \in \mathrm{M}_n(k)$ is **nondegenerate** if it has nonzero determinant (aka is an element of $\mathrm{GL}_n(k)$). A bilinear form b is **nondegenerate** if the associated matrix is nondegenerate. (It is not immediately obvious from the definition, but it is true, that the associated matrix being nondegenerate does not depend on the choice of basis in which the matrix is written down.)

Definition 65. Let b, b' be bilinear forms on a k-vector space V with respective associated matrices B and B' (with respect to a fixed basis). Then b, b' are **equivalent** if there exists an invertible matrix P such that

$$B' = P^T B P$$

Exercise 66. Prove that equivalence of bilinear forms is an equivalence relation.

Definition 67. Let b be a bilinear form on V and fix a basis of V. A bilinear form is **diagonal** if its associated matrix B is diagonal. A bilinear form is **diagonalizable** if it is equivalent to a diagonal bilinear form.

Theorem 68. *As long as the characteristic is not 2, every nondegenerate symmetric bilinear form is diagonalizable. That is, there always exists a basis in which the matrix of b is a diagonal matrix.*

Proof. See `https://kconrad.math.uconn.edu/blurbs/linmultialg/bilinearform.pdf`. □

3.2 Hermitian and skew-hermitian forms

Example 69. Let n be a positive integer. We view $\mathbb{C}^n$ as a vector space over $\mathbb{C}$, thinking of elements as column vectors $(z_1, \ldots, z_n)$ with $z_i \in \mathbb{C}$. The **complex dot product** is the following operation.

$$\begin{aligned} &h : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \\ &h(w, z) = h\Big((w_1, \ldots, w_n), (z_1, \ldots, z_n)\Big) = w_1\overline{z}_1 + \cdots + w_n\overline{z}_n = \sum_{i=1}^{n} w_i\overline{z}_i \end{aligned}$$

Exercise 70. Verify that the complex dot product h has the following properties. For any $u, v, w \in \mathbb{C}^n$ and $\lambda, \mu \in \mathbb{C}$,

$$\begin{aligned} h(u, v + w) &= h(u, v) + h(u, w) \\ h(u + v, w) &= h(u, w) + h(v, w) \\ h(\lambda u, v) &= \lambda h(u, v) \\ h(u, \mu v) &= \overline{\mu} h(u, v) \\ h(u, v) &= \overline{h(v, u)} \end{aligned}$$

In other words, h is almost a symmetric bilinear form, except for two differences. First, when you pull a scalar out of the second entry, you have to take the complex conjugate. Second, h is not symmetric, but "conjugate-symmetric."

Definition 71. Let V be a vector space over $\mathbb{C}$. A **hermitian form** or **sesquilinear form** on V is a map

$$h : V \times V \to \mathbb{C}$$

such that for any $u, v, w \in V$ and $\lambda, \mu \in \mathbb{C}$,

$$\begin{aligned} h(u, v + w) &= h(u, v) + h(u, w) \\ h(u + v, w) &= h(u, w) + h(v, w) \\ h(\lambda u, v) &= \lambda h(u, v) \\ h(u, \mu v) &= \overline{\mu} h(u, v) \\ h(u, v) &= \overline{h(v, u)} \end{aligned}$$

The second-to-last equation is often phrased by saying that "h is conjugate-linear in the second entry." The last equation can be thought of as saying that h is "conjugate-symmetric."

Example 72. Complex dot product is a hermitian form on $\mathbb{C}^n$.

Definition 73. A **skew-hermitian form** or **anti-hermitian form** is nearly the same as a hermitian form, except that $h(u,v) = \overline{h(v,u)}$ is replaced by the condition $h(u,v) = -\overline{h(v,u)}$.

In order to generalize hermitian forms to arbitrary fields, we need to discuss quadratic field extensions.

Definition 74. A **field extension** L/k is a pair of fields L and k where k is a subfield of L.

Example 75. The following are examples of field extensions.

$$\mathbb{C}/\mathbb{R} \qquad \mathbb{R}/\mathbb{Q} \qquad \mathbb{Q}(\sqrt{2})/\mathbb{Q} \qquad \mathbb{Q}(\sqrt{3})/\mathbb{Q}$$

Definition 76. Let L/k be a field extension. We can view L as a vector space over k. The **degree** of the field extension is the dimension of L as a k-vector space.

Example 77. The degree of the field extension $\mathbb{C}/\mathbb{R}$ is 2, because the dimension is 2. A basis is $\{1, i\}$.

Definition 78. Let k be a field. Let $d \in k$ be a nonzero element, such that there is no element $a \in k$ so that $a^2 = d$ (that is, d is does not already have a "square root" in k). Then we can define a new field by

$$k\left(\sqrt{d}\right) = \left\{a + b\sqrt{d} : a, b \in k\right\}$$

Addition in $k\left(\sqrt{d}\right)$ is described by

$$(a_1 + b_1\sqrt{d}) + (a_2 + b_2\sqrt{d}) = (a_1 + a_2) + (b_1 + b_2)\sqrt{d}$$

and multiplication is

$$(a_1 + b_1\sqrt{d})(a_2 + b_2\sqrt{d}) = (a_1a_2 + b_1b_2d) + (a_1b_2 + a_2b_1)\sqrt{d}$$

One way to think about $k\left(\sqrt{d}\right)$ is that we are taking k and adding in solutions to the quadratic equation $x^2 - d = 0$. The way to make this precise is to quotient the polynomial ring $k[x]$ by the principal ideal generated by $x^2 - d$.

$$k\left(\sqrt{d}\right) \cong k[x]/\left\langle x^2 - d\right\rangle$$

Example 79. $\mathbb{C} = \mathbb{R}\left(\sqrt{-1}\right)$

Example 80. For any integer d that is not the square of another integer, we can construct the field $\mathbb{Q}\left(\sqrt{d}\right)$ and get a field which is strictly bigger than $\mathbb{Q}$. If $d > 0$, then we can think of $\mathbb{Q}\left(\sqrt{d}\right)$ as a subfield of $\mathbb{R}$, but if $d < 0$ then we can just think of $\mathbb{Q}\left(\sqrt{d}\right)$ as a subfield of $\mathbb{C}$.

Exercise 81. Show that $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}(\sqrt{18})$.

Lemma 82. *For any field k and any $d \in k$ such that $d \neq a^2$ for any $a \in k$, the field extension $k\left(\sqrt{d}\right)/k$ has degree 2.*

Proof. The set $\left\{1, \sqrt{d}\right\}$ is a basis, because every element of $k\left(\sqrt{d}\right)$ can be uniquely written in the form $a + b\sqrt{d}$ where $a, b \in k$. $\square$

Definition 83. A **quadratic field extension** is a field extension of the form $k\left(\sqrt{d}\right)/k$ for some field k, where d is some element of k that is not already a square in k.

Lemma 84. *If* $\operatorname{char}(k) \neq 2$, *then every field extension* L/k *of degree 2 is a quadratic extension.*

Since we will generally assume that $\operatorname{char} k = 0$ or at least $\operatorname{char} k \neq 2$, for us "quadratic extension" and "extension of degree 2" are effectively synonymous.

Definition 85. For a matrix X with entries in L, the **conjugate matrix**, denoted $\overline{X}$, is the matrix formed by applying conjugation to each individual entry of X. The transpose matrix is denoted X^t or X^T. The **conjugate transpose** matrix $\overline{X^T} = \overline{X}^T$ is denoted X^* for short. (Other common notations include $X^\dagger$ and X^H.)

Definition 86. Let L be a field. An **involution** on L is a field automorphism $\sigma : L \to L$ such that $\sigma^2 = \mathrm{Id}_L$ (that is, $\sigma(\sigma(x)) = x$ for every $x \in L$). Being a field automorphism is the same as being a ring isomorphism from L to itself. In other words, σ is a bijection and

$$\sigma(x + y) = \sigma(x) + \sigma(y) \qquad \sigma(xy) = \sigma(x)\sigma(y)$$

for all $x, y \in L$. Usually when we have an involution on a field, we don't write it as $\sigma(x)$, instead we just write $\overline{x}$ (because the traditional example is complex conjugation on $\mathbb{C}$).

Exercise 87. Let L/k be a quadratic field extension, so $L = k(\sqrt{d})$ for some non-square $d \in k$. Verify that

$$L \to L \qquad a + b\sqrt{d} \mapsto \overline{a + b\sqrt{d}} = a - b\sqrt{d} \qquad a, b \in k$$

is an involution on L, and that $\overline{x} = x$ for any $x \in k$.

The next definition generalizes hermitian forms to vector spaces over fields other than $\mathbb{C}$.

Definition 88. Let L be a field with an involution σ, which we write as $\sigma(x) = \overline{x}$. Let V be a vector space over L. A σ**-hermitian form** or just **hermitian form** on V is a map

$$h : V \times V \to L$$

such that for any $u, v, w \in V$ and $\lambda, \mu \in L$,

$$\begin{aligned}
h(u, v + w) &= h(u, v) + h(u, w) \\
h(u + v, w) &= h(u, w) + h(v, w) \\
h(\lambda u, v) &= \lambda h(u, v) \\
h(u, \mu v) &= \overline{\mu} h(u, v) \\
h(u, v) &= \overline{h(v, u)}
\end{aligned}$$

Essentially, a hermitian form is a symmetric bilinear form except for two changes: bringing out scalars from the second entry requires conjugation, and instead of being symmetric it is "conjugate-symmetric."

Definition 89. As in the complex case, a **skew-σ-hermitian form** is the same as a σ-hermitian form, except that instead of $h(u,v) = \overline{h(v,u)}$ we have $h(u,v) = -\overline{h(v,u)}$.

Definition 90. Let h be a hermitian or skew-hermitian form on an n-dimensional L-vector space V. Fix a basis $\beta = \{v_1, \ldots, v_n\}$ of V. The **matrix of h with respect to the basis β** is a matrix $H = \mathrm{Mat}_\beta(h)$ whose (i,j) entry is $h(v_i, v_j)$.

$$H = \big(h(v_i, v_j)\big)$$

Conversely, given any hermitian matrix H there is an associated hermitian form h defined on basis elements by

$$h(v_i, v_j) = H_{ij}$$

For a hermitian form, the equation $h(u,v) = \overline{h(v,u)}$ is equivalent to the condition $H = \overline{H}^T$, i.e. H is a hermitian matrix. For a skew-hermitian form, $h(u,v) = -\overline{h(v,u)}$ is equivalent to $H = -\overline{H}^T$, i.e. H is a skew-hermitian matrix. We will denote the conjugate transpose of a matrix by H^*.

Lemma 91. *There is a one-to-one correspondence between σ-hermitian forms and σ-hermitian matrices. (This is also true if you add the modifier "skew" to both.)*

Definition 92. A bilinear or hermitian form is **nondegenerate** if its associated matrix is invertible.

3.3 Quadratic forms

Definition 93. Let V be an n-dimensional k-vector space. A **quadratic form** on V is a map $q : V \to k$ described by a quadratic polynomial. That is, q is of the form

$$q(x_1, \ldots, x_n) = \sum_{i,j} a_{ij} x_i x_j$$

with $a_{ij} \in k$ (where $(x_1, \ldots, x_n)$ are the coordinates in some chosen basis).

Example 94. Let $k = \mathbb{R}$ and $V = \mathbb{R}^3$. Then the following are quadratic forms on V.

$$q_1(x,y,z) = x^2 + y^2 + z^2 \qquad q_2(x,y,z) = 3x^2 - y^2 + xy \qquad q_3(x,y,z) = xy + 3xz$$

Remark 95. A quadratic form always satisfies

$$q(\lambda v) = \lambda^2 q(v)$$

for all $\lambda \in k$ and $v \in V$. Slightly more subtly, given any quadratic form v, the map

$$b_q : V \times V \to k \qquad b_q(u,v) = \frac{1}{2}\Big(q(u+v) - q(u) - q(v)\Big)$$

is a bilinear form on V.

Example 96. Let $q\begin{pmatrix}x\\y\end{pmatrix} = x^2 + y^2$ on $\mathbb{R}^2$. Then the associated bilinear form b_q is

$$\begin{aligned}
b_q\left(\begin{pmatrix}x_1\\y_1\end{pmatrix}, \begin{pmatrix}x_2\\y_2\end{pmatrix}\right) &= \frac{1}{2}\left(q\begin{pmatrix}x_1+x_2\\y_1+y_2\end{pmatrix} - q\begin{pmatrix}x_1\\y_1\end{pmatrix} - q\begin{pmatrix}x_2\\y_2\end{pmatrix}\right)\\
&= \frac{1}{2}\Big((x_1+x_2)^2 + (y_1+y_2)^2 - (x_1^2+y_1^2) - (x_2^2+y_2^2)\Big)\\
&= \frac{1}{2}\Big(2x_1x_2 + 2y_1y_2\Big)\\
&= x_1x_2 + y_1y_2
\end{aligned}$$

which is the usual dot product operation on $\mathbb{R}^2$. If we use the standard basis $\{e_1, e_2\}$ on $\mathbb{R}^2$, the matrix of this bilinear form is the identity matrix.

$$B = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

Example 97. Let $q\begin{pmatrix}x\\y\end{pmatrix} = x^2 - y^2$ on $\mathbb{R}^2$. The associated bilinear form is

$$\begin{aligned}
b_q\left(\begin{pmatrix}x_1\\y_1\end{pmatrix}, \begin{pmatrix}x_2\\y_2\end{pmatrix}\right) &= \frac{1}{2}\left(q\begin{pmatrix}x_1+x_2\\y_1+y_2\end{pmatrix} - q\begin{pmatrix}x_1\\y_1\end{pmatrix} - q\begin{pmatrix}x_2\\y_2\end{pmatrix}\right)\\
&= \frac{1}{2}\Big((x_1+x_2)^2 - (y_1+y_2)^2 - (x_1^2-y_1^2) - (x_2^2-y_2^2)\Big)\\
&= \frac{1}{2}\Big(2x_1x_2 - 2y_1y_2\Big)\\
&= x_1x_2 - y_1y_2
\end{aligned}$$

If we use the standard basis $\{e_1, e_2\}$ on $\mathbb{R}^2$, the matrix of this bilinear form is

$$B = \begin{pmatrix}b(e_1,e_1) & b(e_1,e_2)\\ b(e_2,e_1) & b(e_2,e_2)\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$$

Lemma 98. *When* $\operatorname{char} k \neq 2$, *there is a one-to-one correspondence between symmetric bilinear forms and quadratic forms. Given a quadratic form* q, *the associated bilinear form is described by*

$$b_q(u,v) = \frac{1}{2}\Big(q(u+v) - q(u) - q(v)\Big)$$

Given a symmetric bilinear form b, *the associated quadratic form is*

$$q_b(v) = b(v,v)$$

Remark 99. If you want to associate a matrix to a quadratic form, just turn it into the equivalent bilinear form first and use that matrix.

Theorem 100. *When* $\operatorname{char} k \neq 2$, *any quadratic form on a* k*-vector space* V *of dimension* n *is equivalent to a quadratic form that looks like*

$$q(x_1, \ldots, x_n) = a_1x_1^2 + \cdots + a_nx_n^2$$

for some $a_1, \ldots, a_n \in k$.

Definition 101. A **quadratic space** is a pair (V, q) where V is a vector space and q is a quadratic form on V.

Example 102. Let $q(x_1, \dots, x_n) = x_1^2 + \cdots + x_n^2$. Then $(\mathbb{R}^n, q)$ is a quadratic space. (This is the quadratic form associated to the traditional dot product on $\mathbb{R}^n$.)

Definition 103. Let (V, q) be a quadratic space. A **quadratic subspace** or just **subspace** of (V, q) is just a vector subspace with the restriction of q to that subspace.

3.4 Isotropy and Witt index

Definition 104. Let b be a symmetric bilinear form or hermitian form or skew-hermitian form on a vector space V. A vector $v \in V$ is **isotropic** (with respect to b) if $v \neq 0$ and $b(v, v) = 0$. The form b is **isotropic** if there exists at least one isotropic vector. (Note that if v is isotropic, then any nonzero scalar multiple of v is also isotropic, so if there is any isotropic vector then there are many.) The form b is **anisotropic** if there are no isotropic vectors. The form b is **totally isotropic** if all nonzero vectors are isotropic.

Let q be a quadratic form on a vector space V. A vector $v \in V$ is **isotropic** (with respect to q) if $v \neq 0$ and $q(v) = 0$. As with bilinear and hermitian forms, the form q is called **isotropic** if there exists at least one isotropic vector, **anisotropic** if there are no isotropic vectors, and **totally isotropic** if all nonzero vectors are isotropic.

Example 105. The traditional dot product on $\mathbb{R}^n$ is an anisotropic, nondegenerate, symmetric bilinear form.

Example 106. The traditional complex dot product on $\mathbb{C}^n$ is an anisotropic, nondegenerate, hermitian form.

Example 107. Return to the quadratic form $q(x, y) = x^2 - y^2$ on $\mathbb{R}^2$, with associated bilinear form $b((x_1, y_1), (x_2, y_2)) = x_1x_2 - y_1y_2$. This q is isotropic, because $(1, 1)$ is an isotropic vector. However, it is not totally isotropic because $q(v)$ isn't always zero.

Example 108. Let h be the skew-hermitian form represented by the matrix

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

Then h is isotropic, but not totally isotropic.

Lemma 109. *Let b be a symmetric bilinear form on a k-vector space V where $\operatorname{char} k \neq 2$. Then b is totally isotropic if and only if $b(v, w) = 0$ for all $v, w \in V$.*

Proof. Clearly if $b(v, w) = 0$ for all v, w then b is totally isotropic, so we just need to prove the other direction. Suppose b is totally isotropic, that is, $b(v, v) = 0$ for all $v \in V$. Let $v, w \in V$, and then

$$0 = b(v + w, v + w) = b(v, v) + b(v, w) + b(w, v) + b(w, w) = b(v, w) + b(w, v) = 2b(v, w)$$

using the fact that b is symmetric. Then since we are not in characteristic 2, the equation above implies that $b(v, w) = 0$. ☐

Definition 110. Let b be a nondegenerate symmetric bilinear form, nondegenerate hermitian form, or nondegenerate skew-hermitian form on a vector space V. The **Witt index** or **isotropy index** of b is the maximum dimension of a totally isotropic subspace. Traditionally we use q for the Witt index.

Remark 111. The Witt index of an anisotropic form is zero.

Example 112. Return to the quadratic form $q(x, y) = x^2 - y^2$ on $\mathbb{R}^2$, with associated bilinear form $b((x_1, y_1), (x_2, y_2)) = x_1x_2 - y_1y_2$. As previously noted, q is isotropic, because $(1, 1)$ is an isotropic vector, but it is not totally isotropic. So the maximum dimension of a totally isotropic subspace is 1.

Lemma 113. *The Witt index of a form on a vector space V is at most half the dimension of V.*

Definition 114. A vector space with a form is **split** or **metabolic** if the Witt index is exactly half the dimension.

Example 115. Let k be an arbitrary field, of characteristic $\neq 2$. The quadratic space $(k^2, x^2 - y^2)$ is split, because it has dimension 2 and Witt index 1. This particular quadratic space is called the **hyperbolic plane** (over k).

Exercise 116. Let h be the skew-hermitian form represented by the matrix

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

Show that h has Witt index 2, i.e. h is split.

Theorem 117. *Assuming* $\operatorname{char} k \neq 2$, *every split quadratic space is a direct sum of copies of the hyperbolic plane.*

Exercise 118. Let h be the hermitian form represented by the matrix

$$H = \begin{pmatrix} & I_q & \\ \varepsilon I_q & & \\ & & C \end{pmatrix}$$

where I_q is the $q \times q$ identity matrix and $\varepsilon = \pm 1$ and C is an invertible diagonal matrix. Show that the Witt index of h is q. Similarly, if b is the symmetric bilinear form represented by

$$B = \begin{pmatrix} & I_q & \\ I_q & & \\ & & C \end{pmatrix}$$

then show the Witt index of b is q.

Theorem 119 (Witt decomposition)**.** *Let (V, q) be a quadratic space over a field k. Then V can be written (essentially uniquely) in the form*

$$V \cong V_t \oplus V_a \oplus V_h$$

where V_t is totally isotropic ($q|_{V_t} = 0$), V_a is anisotropic, and V_h is split. We can also emphasize the forms more by writing this as

$$(V, q) \cong (V_t, 0) \oplus (V_a, q_a) \oplus (V_h, q_h)$$

Also, given this decomposition, the Witt index is $\frac{1}{2} \dim V_h$, the number of hyperbolic planes comprising V_h.

4 Pinnings

One goal of this document is to describe similar pinning structures on special orthogonal groups and special unitary groups, as we have done for special linear groups above. This section includes background material required to build up to that.

4.1 Algebraic groups

Definition 120. An **algebraic variety** is the set of solutions to a polynomial equation (or list of polynomial equations). More precisely, we need to fix a field k and talk about an **algebraic variety over** k, which is the set of solutions in k to a polynomial equation.

Example 121. A circle is an algebraic variety because it is the set of points (x, y) which are solutions to $x^2 + y^2 = 1$. More precisely, the usual circle is a variety over the field $\mathbb{R}$. If we take a bigger field, such as $\mathbb{C}$, we might get more solutions like $(i, \sqrt{2})$, or if we take a smaller field like $\mathbb{Q}$ we will get fewer solutions (just the points on the circle with rational coordinates).

Example 122. For any polynomial function $f(x)$, the curve $y = f(x)$ is an algebraic variety.

Definition 123. Let V_1, V_2 be algebraic varieties. A **regular map** or **morphism of algebraic varieties** is, very roughly, a map $V_1 \to V_2$ which can be described (locally) in terms of polynomials. (You can approximate regular by thinking of "smooth" or "infinitely differentiable" or "essentially a polynomial function.")

Definition 124. An **algebraic group** is a set which has both a group structure and the structure of an algebraic variety (over a base field k), and the group operations are regular maps. That is, the multiplication map

$$m : G \times G \to G \qquad m(x, y) = xy$$

and the inversion map

$$i : G \to G \qquad i(x) = x^{-1}$$

are regular maps.

Example 125. The special linear group is an algebraic group.

$$\mathrm{SL}_n(k) = \{X \in \mathrm{M}_n(k) : \det(X) = 1\}$$

Example 126. Let b be a nondegenerate symmetric bilinear form with Witt index q and matrix B (in a fixed basis). The **special orthogonal group** of b is

$$\mathrm{SO}_{n,q}(k, b) = \left\{X \in \mathrm{SL}_n(k) : X^T B X = B\right\}$$

This is a subgroup of $\mathrm{SL}_n(k)$, and is itself an algebraic group.

Example 127. Let h be a nondegenerate hermitian or skew-hermitian form with Witt index q, and with matrix H (in a fixed basis). The **special unitary group** of h is

$$\mathrm{SU}_{n,q}(L, h) = \{X \in \mathrm{SL}_n(L) : X^* H X = H\}$$

where $X^* = \overline{X}^T$. The special unitary group is a subgroup of $\mathrm{SL}_n(L)$, and is itself an algebraic group. Note that for technical reasons, to view $\mathrm{SU}_{n,q}(L, h)$ as an algebraic variety, the base field needs to be k, not L.

Definition 128. Let $G(k), G'(k)$ be algebraic groups. A **morphism of algebraic groups** is a group homomorphism $G(k) \to G'(k)$ which is also a regular map. An **abstract homomorphism** is a group homomorphism $G(k) \to G'(k)$ (but not necessarily regular).

4.2 Lie algebras

For this whole section, let k be a field and assume $\operatorname{char} k \neq 2$.

Definition 129. An **algebra** over k or k**-algebra** is a k-vector space A equipped with an additional multiplication map $A \times A \to A$, where scalar multiplication is compatible with multiplication. First, since A is a vector space, it has an addition operation which makes it an abelian group.

$$A \times A \to A \qquad (a, b) \mapsto a + b$$

A has a zero element and additive inverses for every element, and addition is commutative.

$$a + 0 = 0 + a = a \qquad a + (-a) = (-a) + a = 0 \qquad a + b = b + a$$

for every $a, b \in A$. Also, A has a scalar multiplication map.

$$k \times A \to A$$

This scalar multiplication map distributes over addition, meaning

$$\lambda(a + b) = \lambda a + \lambda b \qquad \text{and} \qquad (\lambda + \mu)a = \lambda a + \lambda b$$

for all $\lambda, \mu \in k$ and $a, b \in A$. A also has a multiplication map, which distributes over addition.

$$A \times A \to A \qquad (a, b) \mapsto ab$$

Distributing over addition means that

$$a(b + c) = ab + ac \qquad (a + b)c = ac + bc$$

for every $a, b, c \in A$. Note that we do NOT assume that multiplication is associative, that is, we do not assume that $(ab)c = a(bc)$. If you have this property, then A is called an **associative algebra**, and in this case A is a ring under its addition and multiplication operations. But in general, A may not be associative.

Note that we do NOT assume that A has a unity, or that multiplication is commutative. A also has the property that multiplication is compatible with scalars, in the sense that

$$(\lambda a)(\mu b) = (\lambda \mu)ab$$

for every $\lambda, \mu \in k$ and $a, b \in A$.

Definition 130. A **Lie algebra** over k is an algebra L over k with two additional properties. First, traditionally in a Lie algebra the operation is denoted with brackets, like so.

$$[-, -] : L \times L \to L$$

so the multiplication of two elements x, y is written $[x, y]$ instead of xy. The two additional equations that make L into a Lie algebra are

$$[x, x] = 0 \qquad \forall x \in L$$

and

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$

This second one is called the **Jacobi identity**.

Exercise 131. Show that in any Lie algebra, $[x, y] = -[y, x]$ by applying the fact that $[x+y, x+y] = 0$ and using the distributive property.

Exercise 132. Show that if L is an algebra (not necessarily a Lie algebra) in which $[x, y] = -[y, x]$ and the characteristic of k is not 2, then $[x, x] = 0$.

Example 133. $\mathbb{R}^3$ with the cross product operation is a Lie algebra. Verifying the Jacobi identity is tedious, I do not recommend.

Example 134. Fix $n \geq 1$ and let $\mathrm{M}_n(k)$ be the set of $n \times n$ matrices with entries in k. Then $\mathrm{M}_n(k)$ is a Lie algebra under the operation

$$[A, B] = AB - BA$$

This is the most important example of a Lie algebra. All of the Lie algebras studied in this document are some form of a subalgebra of this Lie algebra.

4.3 Lie algebras associated to classical groups

Every algebraic group has an associated Lie algebra. This is somewhat complicated to describe, so rather than talk about the general definition, I will just describe the specific Lie algebra associated with each of the groups we are dealing with. All of these are matrix Lie algebras that are subalgebras of $\mathrm{M}_n(k)$, meaning they have the same bracket operation but are just a particular subset.

Definition 135. The Lie algebra associated to the special linear group $\mathrm{SL}_n(k)$ is

$$\mathfrak{sl}_n(k) = \{X \in \mathrm{M}_n(k) : \mathrm{tr}(X) = 0\}$$

The bracket operation is the same as on $\mathrm{M}_n(k)$. That is,

$$[X, Y] = XY - YX$$

where XY refers to matrix multiplication.

Exercise 136. Show that if $\mathrm{tr}(X) = \mathrm{tr}(Y) = 0$, then $\mathrm{tr}[X, Y] = 0$. In other words, $\mathfrak{sl}_n(k)$ is actually closed under the bracket operation.

Definition 137. Let b be a nondegenerate symmetric bilinear form with Witt index q and matrix B, and let $\mathrm{SO}_{n,q}(k, B)$ be the associated special orthogonal group. The **special orthogonal Lie algebra** associated to B is

$$\mathfrak{so}_{n,q}(k, B) = \left\{X \in \mathrm{M}_n(k) : X^T B + BX = 0 \text{ and } \mathrm{tr}(X) = 0\right\}$$

again with the usual matrix bracket on $\mathrm{M}_n(k)$.

Definition 138. Let h be a nondegenerate hermitian or skew-hermitian form with Witt index q and matrix H, and let $\mathrm{SU}_{n,q}(L,H)$ be the associated special unitary group. The **special unitary Lie algebra** associated to H is

$$\mathfrak{su}_{n,q}(L,H) = \{X \in \mathrm{M}_n(L) : X^*H + HX = 0 \text{ and } \mathrm{tr}(X) = 0\}$$

with the usual matrix bracket on $\mathrm{M}_n(L)$. Note that $\mathfrak{su}_{n,q}(L,H)$ is a k-vector space, but not an L-vector space.

4.4 Roots and root spaces

For some computations in this section, we need to assume that k is a sufficiently large field. There are occasionally issues when k is a field with only 2 or 3 elements.

Definition 139. Let $G(k)$ be an algebraic group that is one of $\mathrm{SL}_n(k), \mathrm{SO}_{n,q}(k,B)$, or $\mathrm{SU}_{n,q}(L,H)$, and let $\mathfrak{g}(k)$ be the associated Lie algebra (either $\mathfrak{sl}_n(k), \mathfrak{so}_{n,q}(k,B)$, or $\mathfrak{su}_{n,q}(L,H)$). Let $S(k)$ be the torus subgroup of $G(k)$ (the full subgroup of $\mathrm{SL}_n(k)$, or the subgroup we previously specified for $\mathrm{SO}_{n,q}(k,b)$ or $\mathrm{SU}_{n,q}(L,h)$). Then $S(k)$ acts on $\mathfrak{g}(k)$ by conjugation. That is, the map

$$S(k) \times \mathfrak{g}(k) \to \mathfrak{g}(k) \qquad (s,x) \mapsto sxs^{-1}$$

is a well-defined group action. (In the expression sxs^{-1}, we are just doing matrix multiplication.)

Exercise 140. Verify this is well-defined in the case of $\mathrm{SL}_n(k)$. That is, show that if $s \in \mathrm{SL}_n(k)$ is a diagonal matrix, and $X \in \mathfrak{sl}_n(k)$ is a matrix with trace zero, then sXs^{-1} has trace zero.

Definition 141. Let $\alpha \in X^*(S)$ be a character (not necessarily a root). That is, α is a group homomorphism $S(k) \to k^\times$. Define

$$\mathfrak{g}_\alpha = \left\{X \in \mathfrak{g}(k) : sXs^{-1} = \alpha(s)X, \ \forall s \in S(k)\right\}$$

The expression $\alpha(s)x$ can be confusing. What it means is we are plugging in s into α, since α is a function with domain $S(k)$, so $\alpha(s)$ is just a scalar, an element of $k^\times$. Then thinking of $x \in \mathfrak{g}(k)$ as a matrix, we multiply each individual entry of the matrix x by the scalar $\alpha(s)$.

Conceptually, $\mathfrak{g}_\alpha$ is kind of like the set of "simultaneous eigenvectors" for all of $S(k)$, if we think of each $s \in S(k)$ as a linear map from $\mathfrak{g}(k)$ to itself (which is possible using the group action). If this is confusing, just focus on the definition above.

Exercise 142. Show that $\mathfrak{g}_\alpha$ is a vector subspace of $\mathfrak{g}(k)$.

Example 143. For most characters, the set $\mathfrak{g}_\alpha$ is just the zero subspace. To illustrate, we consider the character α_1 for the group $\mathrm{SL}_2(k)$ with its full diagonal subgroup $S(k)$. Recall that

$$S(k) = \left\{ \begin{pmatrix} s_1 & \\ & s_1^{-1} \end{pmatrix} : s_1 \in k^\times \right\}$$

and α_1 is the map

$$\alpha_1 : S(k) \to k^\times \qquad \alpha_1 \begin{pmatrix} s_1 & \\ & s_1^{-1} \end{pmatrix} = s_1$$

In this case, the Lie algebra $\mathfrak{g}(k)$ is the set of 2×2 matrices with entries in k and trace zero, so concretely

$$\mathfrak{g}(k) = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} : a, b, c \in k \right\}$$

We will try to determine $\mathfrak{g}_{\alpha_1}$. Let

$$x = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{pmatrix}$$

If $x \in \mathfrak{g}(k)$ belongs to $\mathfrak{g}_{\alpha_1}$, then $sxs^{-1} = s_1 x$ for every $s \in S(k)$. This means that

$$\begin{aligned} \begin{pmatrix} s_1 & \\ & s_1^{-1} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{pmatrix} \begin{pmatrix} s_1 & \\ & s_1^{-1} \end{pmatrix}^{-1} &= s_1 \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{pmatrix} \\ \begin{pmatrix} s_1 x_{11} & s_1 x_{12} \\ s_1^{-1} x_{21} & -s_1^{-1} x_{11} \end{pmatrix} \begin{pmatrix} s_1^{-1} & \\ & s_1 \end{pmatrix} &= \begin{pmatrix} s_1 x_{11} & s_1 x_{12} \\ s_1 x_{21} & -s_1 x_{11} \end{pmatrix} \\ \begin{pmatrix} x_{11} & s_1^2 x_{12} \\ s_1^{-2} x_{21} & -x_{11} \end{pmatrix} &= \begin{pmatrix} s_1 x_{11} & s_1 x_{12} \\ s_1 x_{21} & -s_1 x_{11} \end{pmatrix} \end{aligned}$$

Remember that for x to be in $\mathfrak{g}_\alpha$, this has to be true for every s. So it has to be true for every $s_1 \in k^\times$. Comparing the $(1,1)$ entries, this implies that

$$x_{11} = s_1 x_{11}$$

If this is true for every s_1, then it must be that $x_{11} = 0$. (This is precisely where we need to assume that k is not too small of a set. If k has only 2 or 3 elements, this deduction is not valid.) Similarly, comparing the $(1,2)$ entries we get

$$s_1^2 x_{12} = s_1 x_{12}$$

which again is only possibly true for every s_1 if $x_{12} = 0$. And a similar argument shows that $x_{21} = 0$ as well. So the only possible matrix x belonging to $\mathfrak{g}_{\alpha_1}$ is the zero matrix.

Example 144. In contrast with the previous example, $\mathfrak{g}_\alpha$ is not always zero. Take the character $2\alpha_1$ in the same scenario as above.

$$2\alpha_1 : S(k) \to k^\times \qquad 2\alpha_1 \begin{pmatrix} s_1 & \\ & s_1^{-1} \end{pmatrix} = s_1^2$$

Then we repeat the calculation. Suppose $x \in \mathfrak{sl}_2(k)$ belongs to $\mathfrak{g}_{2\alpha_1}$. Then reusing the left-hand side of the calculation above, we conclude that

$$sxs^{-1} = \begin{pmatrix} x_{11} & s_1^2 x_{12} \\ s_1^{-2} x_{21} & -x_{11} \end{pmatrix} = s_1^2 x = \begin{pmatrix} s_1^2 x_{11} & s_1^2 x_{12} \\ s_1^2 x_{21} & -s_1^2 x_{11} \end{pmatrix}$$

Again by comparing the $(1,1)$ entries, we can conclude that x_{11} must be zero, since $x_{11} = s_1^2 x_{11}$ can only be true for every $s_1 \in k^\times$ if $x_{11} = 0$. And we can do a similar argument to conclude that $x_{21} = 0$. However, the $(1,2)$ entry does not work out quite the same way, because it gives

$$s_1^2 x_{12} = s_1^2 x_{12}$$

which is just always true, regardless of $s_1 \in k^\times$ and $x_{12} \in k$. As a result, we conclude that

$$\mathfrak{g}_{2\alpha_1} = \left\{ \begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix} : u \in k \right\}$$

which is a subspace of $\mathfrak{sl}_2(k)$ of dimension 1.

Definition 145. A **root of $\mathfrak{g}$ with respect to** S is a character α so that the subspace $\mathfrak{g}_\alpha$ is nonzero.

Exercise 146. Show that $-2\alpha_1$ is also a root for $\mathfrak{sl}_2(k)$ with root space

$$\mathfrak{g}_{-2\alpha_1} = \left\{ \begin{pmatrix} 0 & 0 \\ u & 0 \end{pmatrix} : u \in k \right\}$$

Definition 147. Let X be an $n \times n$ nilpotent matrix, i.e. some power of X is the zero matrix. The **matrix exponential** of X is

$$e^X = \exp(X) = \sum_{i=0}^{\infty} \frac{X^i}{i!} = I_n + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots$$

If X were not nilpotent, we would need to worry about convergence, but since X is nilpotent only finitely many terms are nonzero. For all of the matrices we'll deal with in these notes, the series terminates after just two or three nonzero terms.

Example 148. Let X be an element of the root space $\mathfrak{g}_{2\alpha_1}$ for $\mathfrak{sl}_2(k)$, so

$$X = \begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix}$$

for some $u \in k$. Then $X^2 = 0$, so

$$\exp(X) = I_2 + X = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$$

Definition 149. Let $G(k)$ be a special linear, special orthogonal, or special unitary group, with Lie algebra $\mathfrak{g}(k)$ and torus $S(k)$. Let α be a root, with nonzero root space $\mathfrak{g}_\alpha$. The **root subgroup** associated to α is

$$U_\alpha = \exp(\mathfrak{g}_\alpha) = \{\exp(x) : x \in \mathfrak{g}_\alpha\}$$

Example 150. The root subgroup associated to $2\alpha_1$ for $\mathrm{SL}_2(k)$ is

$$U_{2\alpha_1} = \left\{ \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} : u \in k \right\}$$

Determining pinning maps.
Let $G(k)$ be a special linear, special orthogonal, or special unitary group with Lie algebra $\mathfrak{g}(k)$ and torus $S(k)$. Let α be a root, with nonzero root space $\mathfrak{g}_\alpha$. The associated pinning map X_α is the exponential map from $\mathfrak{g}_\alpha$ to $G(k)$.

Example 151. Let $G = \mathrm{SL}_2(k)$ and $\alpha = 2\alpha_1$. The pinning map $X_{2\alpha_1}$ is just the exponential map

$$\exp : \mathfrak{g}_{2\alpha_1} \to U_{2\alpha_1}$$

(followed by inclusion into $\mathrm{SL}_2(k)$).

Remark 152. In general, we don't really want to do the whole calculation we did in Example 143 for a bunch of potential roots, so it is usually faster to go about things as follows. Let X be an arbitrary element of the Lie algebra $\mathfrak{g}(k)$, and let s be an arbitrary element of the torus subgroup $S(k)$. Then compute sXs^{-1}, and look at the various coefficients that appear. Are any of these equal to $\alpha(s)$ for some character α? Those are usually the roots.

4.5 Root systems

As a first approximation, a root system is a "very symmetrical" finite set of vectors in $\mathbb{R}^n$.

Definition 153. Fix a positive integer n, and let $E = \mathbb{R}^n$. Let $(-,-)$ denote the usual dot product on E. Let $\alpha, \beta \in E$. We say α and β are **orthogonal** if $(\alpha, \beta) = 0$. Also define

$$\langle \alpha, \beta \rangle = \frac{2(\alpha, \beta)}{(\beta, \beta)}$$

Note that $\langle -,- \rangle$ is linear in the first component, but not in the second. Also note that $\langle \alpha, \beta \rangle = 0 \iff (\alpha, \beta) = 0$ (assuming $\beta \neq 0$).

Definition 154. Let E be as above, and let $\alpha \in E$. The **hyperplane orthogonal to** α is the set of all vectors orthogonal to α, and is called P_α.

$$P_\alpha = \{\beta \in E : (\alpha, \beta) = 0\}$$

Definition 155. For $\alpha \in E$, let $\sigma_\alpha : E \to E$ be the reflection across P_α. In terms of a formula,

$$\sigma_\alpha : E \to E \qquad \sigma_\alpha(\beta) = \beta - \langle \alpha, \beta \rangle \alpha$$

This is a reflection in a geometric sense, and also in the sense that $\sigma_\alpha^2 = \mathrm{Id}_E$.

Definition 156. Let $E = \mathbb{R}^n$. A **root system** in E is a finite set $\Phi \subset E$ such that

(R1) $0 \notin \Phi$ and the span of Φ is E

(R2) For $\alpha \in \Phi$, the only multiples of α that are in Φ are $\pm\alpha$ and $\pm 2\alpha$.

(R3) For $\alpha \in \Phi$, $\sigma_\alpha(\Phi) = \Phi$.

(R4) For $\alpha, \beta \in \Phi$, $\langle \alpha, \beta \rangle \in \mathbb{Z}$.

Technically speaking, (R2) is redundant because one can prove that (R2) is a consequence of (R1), (R3), and (R4). However, we just include it in the definition. The root system Φ is called **reduced** if (R2) can be replaced by the stronger statement

(R2') For $\alpha \in \Phi$, the only multiples of α that are in Φ are $\pm\alpha$.

An element of a root system is called a **root**.

Remark 157. There is a complete classification of root systems via something called "Dynkin diagrams." The details of the proof of this classification are not important for us, and really we only need to care about type $\mathsf{A}, \mathsf{B}, \mathsf{C}$, and BC root systems. There are also type $\mathsf{D}, \mathsf{E}, \mathsf{F}, \mathsf{G}$ root systems.

Example 158. We describe the A_{n-1} root system. Let $E = \mathbb{R}^n$, and let $\alpha_i = (0, \dots, 1, \dots, 0)$ be the vector with 1 in the ith position and 0 elsewhere. Let $\alpha_{ij} = \alpha_i - \alpha_j$. Then set

$$\Phi = \{\alpha_{ij} : 1 \leq i, j \leq n, i \neq j\}$$

Then Φ is a reduced root system, called the type A_{n-1} root system. Note that Φ does not actually span all of E – the span of Φ is actually the subspace orthogonal to the vector $(1, 1, 1, \dots, 1)$. So technically speaking, Φ is a root system not in E, but rather in this $(n-1)$-dimensional subspace of E.

Definition 159. The **rank** of Φ is the dimension of the space spanned by Φ.

Example 160. The rank of A_{n-1} is $n-1$. In general, for any root system the subscript denotes the rank, so for example D_4 is the type D root system of rank 4.

Example 161. We describe the $\mathsf{B}_n, \mathsf{C}_n$, and BC_n root systems. Let $E = \mathbb{R}^n$, and $\alpha_i = (0, \dots, 1, \dots, 0)$ as before. Then

$$\mathsf{B}_n = \{\pm\alpha_i : 1 \leq i \leq n\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \leq i < j \leq n\}$$
$$\mathsf{C}_n = \{\pm 2\alpha_i : 1 \leq i \leq n\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \leq i < j \leq n\}$$

Note that the second part is identical, the only difference is the 2 in the first set of roots. Both B_n and C_n are reduced root systems (in E). The BC_n root system is simply the union of the two.

$$\mathsf{BC}_n = \{\pm\alpha_i, \pm 2\alpha_i : 1 \leq i \leq n\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \leq i, j \leq n, i \neq j\}$$

BC_n is not reduced. In fact, the root systems BC_n are the only non-reduced root systems (we omit the proof). The shortest roots $\pm\alpha_i$ in BC_n are called **multipliable** because their double is also a root.

Definition 162. A root system is **reducible** if it can be written as a disjoint union of two subsets which are orthogonal to each other. A root system is **irreducible** if it is not reducible.

Remark 163. All of the example root systems above ($\mathsf{A}_n, \mathsf{B}_n, \mathsf{C}_n, \mathsf{BC}_n$) are irreducible.

Remark 164. Despite the unfortunate similarity of terminology, there is no relationship between a root system being "reduced/non-reduced" and "reducible/irreducible." For example, it is possible to have a root system which is non-reduced and irreducible (BC_n is the key example).

Lemma 165. *Any root system Φ can be written uniquely as a disjoint union of irreducible root systems. These are called the* ***irreducible components*** *of Φ.*

Because of the lemma, it is sufficient to study irreducible root systems.

Definition 166. Let Φ be a root system of rank n in $E = \mathbb{R}^n$. For any $\alpha \in \Phi$, we can consider the reflection σ_α as an invertible linear map $E \to E$. By (R3), if we restrict σ_α to Φ, the output is Φ. So we can think of σ_α as a map $\Phi \to \Phi$. The **Weyl group** of Φ, denoted $W(\Phi)$, is the subgroup of invertible linear maps $E \to E$ generated by all of the σ_α. Since all the generators take Φ to itself, every element of $W(\Phi)$ sends Φ to itself.

Remark 167. In general, there are elements of $W(\Phi)$ which are not σ_α for any α.

Despite the fact that $W(\Phi)$ is generated by elements of order 2, not every element of $W(\Phi)$ has to have order 2. However, since Φ is a finite set and $W(\Phi)$ is a subgroup of the permutation group of Φ, $W(\Phi)$ is a finite group. Usually, $W(\Phi)$ is not the full permutation group of Φ, since $\sigma_\alpha(\beta)$ must have the same length as β. So $W(\Phi)$ must permute the set of roots of the same length. For example, in B_n, this means that $W(\Phi)$ can only send roots of the form $\pm\alpha_i$ to other roots of this form, never to a root such as $\alpha_1 + \alpha_2$.

Also, if Φ is reducible, then $W(\Phi)$ must only send roots in one irreducible component within that irreducible component. However, within that restriction, $W(\Phi)$ "acts transitively," meaning that for any two roots α, β of the same length in the same irreducible component, there exists a $w \in W(\Phi)$ such that $w(\alpha) = \beta$. Since most of the time we will deal with irreducible root systems, the main restriction here is that α, β have the same length.

4.6 Pinnings

Definition 168. Let k be a field, and let $G = G(k)$ be an algebraic group, with a maximal k-split torus subgroup $S = S(k)$[2]. Let $\Phi = \Phi(G, S)$ be the root system of G with respect to S. A **pinning** of G (with respect to S) is a set of maps

$$X_\alpha : V_\alpha(k) \to G(k)$$

for each $\alpha \in \Phi$, such that:

1. Each $V_\alpha(k)$ is a vector space over k.
2. X_α is a pseudo-homomorphism, meaning that for each $u, v \in V_\alpha(k)$,

   $$X_\alpha(u) \cdot X_\alpha(v) = X_\alpha(u+v) \cdot \Big(\text{some correction factor}\Big)$$

   More precisely, the correction factor is 1 unless $2\alpha \in \Phi$. The full formula is

   $$X_\alpha(u) \cdot X_\alpha(v) = X_\alpha(u+v) \cdot \prod_{\substack{i>1 \\ i\alpha \in \Phi}} X_{i\alpha}\Big(\varphi^i_\alpha(u, v)\Big)$$

3. (Torus conjugation formula) For every $s \in S(k)$,

   $$s \cdot X_\alpha(u) \cdot s^{-1} = X_\alpha\Big(\alpha(s)u\Big)$$

[2] You don't need to know exactly what a maximal k-split torus is, just imagine S as some generalization of the diagonal subgroup of $\mathrm{SL}_n(k)$.

4. (Commutator formula) For every $\alpha, \beta \in \Phi$ with $\alpha \neq c\beta$ for any scalar c, there are polynomial maps
$$N_{ij}^{\alpha\beta} : V_\alpha(k) \times V_\beta(k) \to V_{i\alpha+j\beta}(k)$$
such that for any $u \in V_\alpha(k)$ and $v \in V_\beta(k)$, we have
$$\Big[X_\alpha(u), X_\beta(v)\Big] = \prod_{\substack{i,j\geq 1\\ i\alpha+j\beta\in\Phi}} X_{i\alpha+j\beta}\Big(N_{ij}^{\alpha\beta}(u,v)\Big)$$
When we need to worry about the order of the product on the right, we order it first by $i+j$ in increasing order, then secondarily by i in increasing order.

Definition 169. A **Weyl group representative** is an element $w \in G$ which normalizes the torus S. That is,
$$w \cdot S \cdot w^{-1} = S$$
or equivalently
$$w \cdot s \cdot w^{-1} \in S \qquad \forall s \in S$$
The true Weyl group is the quotient of the normalizer subgroup of S by the centralizer subgroup of S, but it is often easier to work with elements of the group G, so rather than work in that quotient we work with representative lifts from that quotient, i.e. elements of the normalizer subgroup of S.

A **pinning compatible with the Weyl group** is a pinning with an additional property:

5. (Weyl group conjugation formula) For every $\alpha \in \Phi$, there is a Weyl group element $w_\alpha \in G(k)$ which belongs to the subgroup generated by $X_\alpha(V_\alpha(k))$ and $X_{-\alpha}(V_{-\alpha}(k))$ and such that for every $\beta \in \Phi$ and $v \in V_\beta(k)$,
$$w_\alpha \cdot X_\beta(v) \cdot w_\alpha^{-1} = X_{\sigma_\alpha(\beta)}\Big(\varphi_{\alpha\beta}(v)\Big)$$
where $\varphi_{\alpha\beta} : V_\beta \to V_{\sigma_\alpha(\beta)}$ is some invertible linear function.

Remark 170. By general theory that we are not going into, under appropriate assumptions about the group G, G always contains a maximal k-split torus S, and there is always a root system $\Phi(G, S)$, and a pinning of G that is compatible with the Weyl group always exists.

5 Special orthogonal groups

In this section we discuss the special orthogonal group associated to a bilinear form.

Definition 171. Let b be a nondegenerate symmetric bilinear form on a k-vector space V of dimension n, with matrix B (with respect to a fixed basis of V). The **orthogonal group** of B (or of b) is the group of $n \times n$ invertible matrices X with entries in k satisfying $X^T B X = B$.

$$\mathrm{O}_n(k, b) = \left\{ X \in \mathrm{GL}_n(k) : X^T B X = B \right\}$$

The **special orthogonal group** of B is the subgroup of $\mathrm{O}_n(k, b)$ of matrices with determinant equal to one.

$$\mathrm{SO}_n(k, b) = \left\{ X \in \mathrm{SL}_n(k) : X^T B X = B \right\}$$

In the future, we will also want to specify the Witt index of b in the notation, so let q be the Witt index of b (or B equivalently). Then we denote these groups by

$$\mathrm{O}_{n,q}(k, b) \qquad \text{and} \qquad \mathrm{SO}_{n,q}(k, b)$$

When it is clear from context, we can also just write this as $\mathrm{SO}(k, b)$ or even just $\mathrm{SO}(k)$.

Exercise 172. Verify that $\mathrm{O}_n(k, b)$ is a group under matrix multiplication and that $\mathrm{SO}_n(k, b)$ is a subgroup.

Exercise 173. Show that if b and b' are equivalent bilinear forms, then the associated orthogonal groups $\mathrm{O}_{n,q}(k, b)$ and $\mathrm{O}_{n,q}(k, b')$ are isomorphic. Hint: if $B' = P^T B P$, then

$$X^T B' X = B' \iff X^T \left(P^T B P \right) X = P^T B P \iff \left((P^T)^{-1} X^T P^T \right) B \left(P X P^{-1} \right) = B$$

and it is relevant that $(P^T)^{-1} = (P^{-1})^T$.

Exercise 174. Verify that the isomorphism you found in the previous exercise also gives an isomorphism between $\mathrm{SO}_{n,q}(k, b)$ and $\mathrm{SO}_{n,q}(k, b')$.

Remark 175. The main takeaway from the exercise is: changing a bilinear form to an equivalent bilinear form doesn't change the group (up to isomorphism), so if we just want to study the group, then we are free to use whichever equivalent bilinear form we like. Because of Theorem 68, we can always choose a basis so that our nondegenerate symmetric bilinear form is diagonal, which makes it have an associated diagonal matrix. Then because of Exercise 173, changing our basis in this way doesn't affect the group. So any orthogonal or special orthogonal group can be described using an invertible diagonal matrix (every diagonal matrix is symmetric, so now symmetric becomes redundant). However, it is not always best to assume that B is completely diagonalized.

5.1 Split, quasi-split, and otherwise

Let k be an arbitrary field of characteristic $\neq 2$. If it helps, you can usually pretend that $k = \mathbb{R}$ or $k = \mathbb{Q}$ without running into any problems.

Let V be a vector space over k of dimension n, and let b be a nondegenerate symmetric isotropic bilinear form on V. Fix a basis of V, and let B be the matrix of b in that basis. Then B is a symmetric, invertible matrix. We could use Theorem 68 to assume that B is a diagonal matrix, but we don't actually want to do this.

Let Q be the quadratic form associated to b, and let q be the Witt index of Q. Since b is isotropic, $q \geq 1$. Note that $q \leq \frac{n}{2}$. By the Witt decomposition theorem, we can write V as

$$V = V_t \perp V_o \perp V_h$$

where $V_t = \operatorname{rad}(V)$ is the totally isotropic radical, V_o is anisotropic, and V_h is a direct sum of q hyperbolic planes (so $\dim V_h = 2q$). Since b is nondegenerate, $V_t = \{0\}$, so

$$V = V_o \perp V_h$$

We can write out V_h as an orthogonal direct sum of 2-dimensional hyperbolic planes, so that

$$V = V_o \perp V_h = V_o \perp H_1 \perp H_2 \perp \cdots \perp H_q$$

where each H_i has dimension 2. The value of q determines the following: [3]

- If q is as big as possible, which is to say, $q = \frac{n}{2}$ (equivalently $n = 2q$) or $q = \frac{n-1}{2}$ (equivalently $n = 2q+1$) then the special orthogonal group $\mathrm{SO}(k,b)$ is called "k-split" which for our purposes means it's already understood.

- If q is almost as big as possible, which is to say, $q = \frac{n}{2} - 1$ (equivalently $2q + 2 = n$ so notably n must be even) then $\mathrm{SO}(k,b)$ is quasi-split.

- In the remaining case, $n > 2q + 2$ and $\mathrm{SO}(k,b)$ is neither split nor quasi-split.

We will always work in the case $q \geq 1$, since otherwise the group is what is called **anisotropic**. Using the Witt decomposition, we can choose a basis of V so that in that basis the bilinear form has the matrix

$$B = \begin{pmatrix} & I_q & \\ I_q & & \\ & & C \end{pmatrix}$$

where I_q is the $q \times q$ identity matrix and C is an $(n-2q) \times (n-2q)$ invertible diagonal matrix. When working on examples, the following list of minimal cases can be helpful:

- Split cases: $(n,q) = (4,2), (5,2), (6,3), (7,3)$

- Quasi-split cases: $(n,q) = (4,1), (6,2), (8,3)$

- Non-quasi-split cases: $(n,q) = (5,1), (6,1), (7,2)$

In the minimal quasi-split case $(n,q) = (4,1)$, the root system is B_1. The B_1 and A_1 root systems are isomorphic. B_1 has just 2 elements.

$$\Phi = \{\pm\alpha_1\}$$

[3] See section 34.6 (page 222) of Humphreys' book on linear algebraic groups, or page 256 of Borel.

In the next smallest quasi-split example $(n, q) = (6, 2)$, the root system is B_2. Incidentally, B_2 and C_2 are isomorphic. B_2 has 8 elements.

$$\Phi = \{\pm\alpha_1, \pm\alpha_2, \pm\alpha_1 \pm \alpha_2\}$$

The next smallest quasi-split example is $(n, q) = (8, 3)$, where the root system is B_3, which has 18 elements. (Note that B_3 is not isomorphic to C_3.)

$$\Phi = \{\pm\alpha_1, \pm\alpha_2, \pm\alpha_3, \pm\alpha_1 \pm \alpha_2, \pm\alpha_1 \pm \alpha_3, \pm\alpha_2 \pm \alpha_3\}$$

5.2 Lie algebra

Definition 176. Let B be the $(n \times n)$ matrix of a nondegenerate symmetric bilinear form with Witt index $q \geq 1$, and let $\mathrm{SO}_{n,q}(k, b)$ be the associated special orthogonal group. The **special orthogonal Lie algebra** associated to B is

$$\mathfrak{so}_{n,q}(k, b) = \left\{X \in \mathrm{M}_n(k) : X^T B + BX = 0\right\}$$

Exercise 177. Recall that we can write B in the form

$$B = \begin{pmatrix} & I_q & \\ I_q & & \\ & & C \end{pmatrix}$$

where C is a $(n - 2q) \times (n - 2q)$ invertible matrix. Using this form, show that if $X \in \mathfrak{so}_{n,q}(k, b)$, then X has the block form

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$$

where the blocks have the same size as the corresponding blocks of B, and these blocks satisfy the matrix equations

$$\begin{aligned} X_{21}^T + X_{21} &= 0 \\ X_{11}^T + X_{22} &= 0 \\ X_{31}^T C + X_{23} &= 0 \\ X_{22}^T + X_{11} &= 0 \\ X_{12}^T + X_{12} &= 0 \\ X_{32}^T C + X_{13} &= 0 \\ X_{23}^T + C X_{31} &= 0 \\ X_{13}^T + C X_{32} &= 0 \\ X_{33}^T C + C X_{33} &= 0 \end{aligned}$$

Alternately, we can remove the redundant equations and do some manipulation to rewrite this as

$$\begin{aligned}
X_{21} &= -X_{21}^T \\
X_{12} &= -X_{12}^T \\
X_{22} &= -X_{11}^T \\
X_{23} &= -X_{31}^T C \\
X_{13} &= -X_{32}^T C \\
X_{33} &= -C^{-1} X_{33}^T C
\end{aligned}$$

so X has the form

$$X = \begin{pmatrix} X_{11} & X_{12} & -X_{32}^T C \\ X_{21} & -X_{11}^T & -X_{31}^T C \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$$

where X_{11}, X_{31}, X_{32} are arbitrary and X_{21}, X_{12}, X_{33} are subject to the previous constraint equations.

5.3 Torus and characters

Let $q \geq 1$ and $n > 2q$, and let B be the matrix of a nondegenerate symmetric bilinear form on k^n with Witt index q. The assumption $n > 2q$ ensures that C is not an empty block.

$$B = \begin{pmatrix} & I_q & \\ I_q & & \\ & & C \end{pmatrix}$$

C is an $(n-2q) \times (n-2q)$ diagonal matrix with entries $c_1, \ldots, c_{n-2q} \in k^\times$. Then consider the associated special orthogonal group $\mathrm{SO}(k, b)$.

$$\mathrm{SO}_{n,q}(k, b) = \mathrm{SO}_n(k, b) = \left\{X \in \mathrm{SL}_n(k) : X^T B X = B\right\}$$

Exercise 178. Suppose $s \in \mathrm{SO}_{n,q}(k, b)$ is a diagonal matrix. Show that s has the form

$$s = \begin{pmatrix} s_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & s_q & & & & & & \\ & & & s_1^{-1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & s_q^{-1} & & & \\ & & & & & & \pm 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & \pm 1 \end{pmatrix}$$

for some $s_1, \ldots, s_q \in k^\times$. (Each ± 1 in the middle block has an independent sign.)

Definition 179. The **diagonal torus** or **torus subgroup** of $\mathrm{SO}_{n,q}(k,b)$ is the following subgroup of diagonal matrices.

$$S(k) = \left\{ \begin{pmatrix} s_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & s_q & & & & & & \\ & & & s_1^{-1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & s_q^{-1} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} : s_1, \ldots, s_q \in k^\times \right\}$$

In other words, $S(k)$ is a subgroup of $\mathrm{SO}_{n,q}(k,b)$, but it does not consist of all diagonal matrices in $\mathrm{SO}(k,b)$. Rather, it only consists of those for which the bottom right diagonal block is the identity matrix.

Definition 180. For $i = 1, \ldots, n$, let $\alpha_i : S(k) \to k^\times$ be the map which picks off the ith diagonal entry of a diagonal matrix $s \in S(k)$. (The notation in Borel [Bor91] section 23.9 uses y_i instead of α_i, but means exactly the same thing.) You can verify that α_i is a group homomorphism.

Definition 181. The **character group** of $S(k)$ is the set of group homomorphisms from $S(k)$ to $k^\times$, and is denoted $X^*(S)$ or $\mathrm{Hom}(S(k), k^\times)$. $X^*(S)$ is an abelian group, and we call the operation "character addition." For example, α_1 and α_2 are characters and $\alpha_1 + \alpha_2$ is another character described by the formula

$$(\alpha_1 + \alpha_2)(s) = \alpha_1(s)\alpha_2(s)$$

or in more concrete matrix terms,

$$(\alpha_1 + \alpha_2) \begin{pmatrix} s_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & s_q & & & & & & \\ & & & s_1^{-1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & s_q^{-1} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} = s_1 s_2$$

Exercise 182. Assume $n > 2q$. Show that the root system of $\mathrm{SO}_{n,q}(k,b)$ is

$$\Phi = \Phi\Big(\mathrm{SO}_{n,q}(k,b), S(k)\Big) = \{\pm\alpha_i : 1 \le i \le q\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \le i, j \le q, i \ne j\}$$

This is the type B_q root system. (Note that if $n = 2q$, the root system is type D_q.) Elements of Φ are also elements of the character group $X^*(S)$, but we call them roots.

Exercise 183. Show that if $q \ge 2$, then Φ as above contains $2q^2$ roots.

5.4 Quasi-split groups

5.4.1 The minimal example ($n = 4, q = 1$)

Exercise 184. Let B be the matrix of the bilinear form in the minimal quasi-split case $n = 4, q = 1$.

$$B = \begin{pmatrix} & 1 & & \\ 1 & & & \\ & & c_1 & \\ & & & c_2 \end{pmatrix}$$

Assume that the fixed basis of $V = k^n$ is $\{e_1, \ldots, e_4\}$ where $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and so on. Show that

- The associated bilinear form b_C for the matrix C is

$$b_C\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = c_1 x_1 y_1 + c_2 x_2 y_2$$

- The associated bilinear form b_B for the matrix B is

$$b_B\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}\right) = x_2 y_1 + x_1 y_2 + c_1 x_3 y_3 + c_2 x_4 y_4$$

- The associated quadratic form for the bilinear form b_C is

$$q_C\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = c_1 x_1^2 + c_2 x_2^2$$

- The associated quadratic form for the bilinear form b_B is

$$q_B\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = 2x_1 x_2 + c_1 x_3^2 + c_2 x_4^2$$

- An arbitrary element of $\mathfrak{so}_{4,1}(k, b)$ has the form

$$X = \begin{pmatrix} x_{11} & 0 & -c_1 x_{32} & -c_2 x_{42} \\ 0 & -x_{11} & -c_1 x_{31} & -c_2 x_{41} \\ x_{31} & x_{32} & 0 & x_{34} \\ x_{41} & x_{42} & -\frac{c_1}{c_2} x_{34} & 0 \end{pmatrix}$$

  where $x_{11}, x_{31}, x_{41}, x_{32}, x_{42}, x_{34}$ are arbitrary elements of k.

Exercise 185. Generalize the previous descriptions of the bilinear and quadratic forms of B and C to the general case where B has the form

$$B = \begin{pmatrix} & I_q & \\ I_q & & \\ & & C \end{pmatrix}$$

where $C = \operatorname{diag}(c_1, \ldots, c_{n-2q})$ and $q \geq 1$ and $n > 2q$.

Example 186. In this example we work out the roots, root spaces, and root subgroups for $\mathrm{SO}_{4,1}(k, b)$. Let X be an arbitrary element of $\mathfrak{so}_{4,1}(k, b)$ as above. Let $s \in S(k)$ be an arbitrary element of the torus subgroup.

$$s = \begin{pmatrix} s_1 & & & \\ & s_1^{-1} & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

We simplify the conjugation sXs^{-1}.

$$\begin{aligned} sXs^{-1} &= \begin{pmatrix} s_1 & & & \\ & s_1^{-1} & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} x_{11} & 0 & -c_1x_{32} & -c_2x_{42} \\ 0 & -x_{11} & -c_1x_{31} & -c_2x_{41} \\ x_{31} & x_{32} & 0 & x_{34} \\ x_{41} & x_{42} & -\frac{c_1}{c_2}x_{34} & 0 \end{pmatrix} \begin{pmatrix} s_1 & & & \\ & s_1^{-1} & & \\ & & 1 & \\ & & & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} s_1x_{11} & 0 & -s_1c_1x_{32} & -s_1c_2x_{42} \\ 0 & -s_1^{-1}x_{11} & -s_1^{-1}c_1x_{31} & -s_1^{-1}c_2x_{41} \\ x_{31} & x_{32} & 0 & x_{34} \\ x_{41} & x_{42} & -\frac{c_1}{c_2}x_{34} & 0 \end{pmatrix} \begin{pmatrix} s_1^{-1} & & & \\ & s_1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \\ &= \begin{pmatrix} x_{11} & 0 & -s_1c_1x_{32} & -s_1c_2x_{42} \\ 0 & -x_{11} & -s_1^{-1}c_1x_{31} & -s_1^{-1}c_2x_{41} \\ s_1^{-1}x_{31} & s_1x_{32} & 0 & x_{34} \\ s_1^{-1}x_{41} & s_1x_{42} & -\frac{c_1}{c_2}x_{34} & 0 \end{pmatrix} \end{aligned}$$

If this is going to be equal to $\alpha(s)X$ for some α, with X being nonzero, this is only going to be possible if $\alpha(s) = s_1$ or $\alpha(s) = s_1^{-1}$. That is, the roots are α_1 and $-\alpha_1$, and

$$\mathfrak{g}_{\alpha_1} = \left\{ \begin{pmatrix} 0 & & -c_1x_{32} & -c_2x_{42} \\ & 0 & & \\ & x_{32} & 0 & \\ & x_{42} & & 0 \end{pmatrix} : x_{32}, x_{42} \in k \right\}$$

$$\mathfrak{g}_{-\alpha_1} = \left\{ \begin{pmatrix} 0 & & & \\ & 0 & -c_1x_{31} & -c_2x_{41} \\ x_{31} & & 0 & \\ x_{41} & & & 0 \end{pmatrix} : x_{31}, x_{41} \in k \right\}$$

Thus it makes sense to define maps as below:

$$X_{\alpha_1} : k^2 \to \mathfrak{so}_{4,1}(k,b) \qquad X_{\alpha_1}(x_{32},x_{42}) = \begin{pmatrix} 0 & & -c_1x_{32} & -c_2x_{42} \\ & 0 & & \\ & x_{32} & 0 & \\ & x_{42} & & 0 \end{pmatrix}$$

$$X_{-\alpha_1} : k^2 \to \mathfrak{so}_{4,1}(k,b) \qquad X_{-\alpha_1}(x_{31},x_{41}) = \begin{pmatrix} 0 & & & \\ & 0 & -c_1x_{31} & -c_2x_{41} \\ x_{31} & & 0 & \\ x_{41} & & & 0 \end{pmatrix}$$

Definition 187. We develop some notation to make referring to these rows in particular blocks easier. Given a row vector $v = (v_1, \ldots, v_m) \in k^m$, let $E_{i,\ell \leq j \leq \ell+m}(v)$ denote the matrix with the vector v in the entries $E_{i\ell}, \ldots, E_{i,\ell+m}$. Similarly, $E_{\ell \leq i \leq \ell+m,j}(v^T)$ is the matrix with the column vector v^T in the entries $E_{\ell j}, \ldots, E_{\ell+m,j}$. Using this notation,

$$X_{\alpha_1}(v_1,v_2) = X_{\alpha_1}(v) = E_{3\leq i\leq 4,2}(v^T) - E_{1,3\leq 4\leq j}(vC)$$
$$X_{-\alpha_1}(v_1,v_2) = X_{-\alpha_1}(v) = E_{3\leq i\leq 4,1}(v^T) - E_{2,3\leq j\leq 4}(vC)$$

We return to the previous example of $n = 4, q = 1$. There are two things going on with $\mathrm{SO}_{4,1}$ that did not happen in $\mathrm{SL}_2(k)$. First, $\mathfrak{g}_{\pm\alpha_1}$ is two-dimensional, not one-dimensional. Second, elements of $\mathfrak{g}_{\alpha_1}$ and $\mathfrak{g}_{\alpha_2}$ do not square to zero.

$$\begin{pmatrix} 0 & & -c_1x_{32} & -c_2x_{42} \\ & 0 & & \\ & x_{32} & 0 & \\ & x_{42} & & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & -\left(c_1x_{32}^2 + x_2x_{42}^2\right) & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} = \begin{pmatrix} 0 & -Q_C\left(x_{32},x_{42}\right) & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

Interestingly, the thing that shows up is actually the output of the quadratic form Q_C associated to C applied to the vector (x_{32}, x_{42}). Setting this aside for now, we see that

$$\begin{aligned} U_{\alpha_1} = \exp(\mathfrak{g}_{\alpha_1}) &= \left\{ \begin{pmatrix} 1 & -\frac{1}{2}\left(c_1x_{32}^2 + c_2x_{42}^2\right) & -c_1x_{32} & -c_2x_{42} \\ & 1 & & \\ & x_{32} & 1 & \\ & x_{42} & & 1 \end{pmatrix} : x_{32}, x_{42} \in k \right\} \\ &= \left\{ \begin{pmatrix} 1 & -\frac{1}{2}Q_C\left(x_{32},x_{42}\right) & -c_1x_{32} & -c_2x_{42} \\ & 1 & & \\ & x_{32} & 1 & \\ & x_{42} & & 1 \end{pmatrix} \right\} \\ U_{-\alpha_1} = \exp(\mathfrak{g}_{-\alpha_1}) &= \left\{ \begin{pmatrix} 1 & & & \\ -\frac{1}{2}\left(c_1x_{31}^2 + x_2x_{41}^2\right) & 1 & -c_1x_{31} & -c_2x_{41} \\ x_{31} & & 1 & \\ x_{41} & & & 1 \end{pmatrix} : x_{31}, x_{41} \in k \right\} \\ &= \left\{ \begin{pmatrix} 1 & & & \\ -\frac{1}{2}Q_C\left(x_{31},x_{41}\right) & 1 & -c_1x_{31} & -c_2x_{41} \\ x_{31} & & 1 & \\ x_{41} & & & 1 \end{pmatrix} \right\} \end{aligned}$$

Thus the pinning maps for $\mathrm{SO}_{4,1}(k,b)$ should be

$$x_{\alpha_1} : k^2 \to \mathrm{SO}_{4,1}(k,b) \qquad x_{\alpha_1}(x_{32}, x_{42}) = \begin{pmatrix} 1 & -\frac{1}{2}\left(c_1 x_{32}^2 + c_2 x_{42}^2\right) & -c_1 x_{32} & -c_2 x_{42} \\ & 1 & & \\ & x_{32} & 1 & \\ & x_{42} & & 1 \end{pmatrix}$$

$$x_{-\alpha_1} : k^2 \to \mathrm{SO}_{4,1}(k,b) \qquad x_{-\alpha_1}(x_{31}, x_{41}) = \begin{pmatrix} 1 & & & \\ -\frac{1}{2}\left(c_1 x_{31}^2 + c_2 x_{41}^2\right) & 1 & -c_1 x_{31} & -c_2 x_{41} \\ x_{31} & & 1 & \\ x_{41} & & & 1 \end{pmatrix}$$

5.4.2 Another example ($n = 6, q = 2$)

In this case, the matrix of our bilinear form is

$$B = \begin{pmatrix} & & 1 & & & \\ & & & 1 & & \\ 1 & & & & & \\ & 1 & & & & \\ & & & & c_1 & \\ & & & & & c_2 \end{pmatrix}$$

and the root system is type B_2.

$$\Phi = \{\pm\alpha_1, \pm\alpha_2, \pm\alpha_1 \pm \alpha_2\}$$

Exercise 188. Using the calculations from Exercise 177, write down the general form of an element of the special orthogonal Lie algebra $\mathfrak{so}_{6,2}(k,b)$.

Exercise 189. Mimic the calculation of Example 186 to determine the root spaces for the Lie algebra $\mathfrak{so}_{6,2}(k,b)$. That is, take $s \in S(k)$ an arbitrary element of the diagonal torus subgroup, and $X \in \mathfrak{so}_{6,2}(k,b)$ an arbitrary element of the Lie algebra, and calculate sXs^{-1}. Then compare the result of that calculation with $\alpha(s)X$ to determine which characters α have nonzero root space $\mathfrak{g}_\alpha$. Confirm that the roots are

$$\Phi = \{\pm\alpha_1, \pm\alpha_2, \pm\alpha_1 \pm \alpha_2\}$$

(the B_2 root system). Write out general forms for elements of the root spaces, and use these to determine what the maps X_α and x_α look like.

Remark 190. In the previous exercise, you should find that the dimensions of the root spaces are

Root α	Dimension of $\mathfrak{g}_\alpha$
$\pm\alpha_i \pm \alpha_j$	1
$\pm\alpha_i$	$n - 2q = 2$

5.4.3 The general quasi-split case ($n = 2q + 2$)

Note that this includes the cases $(n, q) = (4, 1)$ and $(n, q) = (6, 2)$ that we studied above. Let $q \geq 1$ and let $n = 2q + 2$, and let b be a symmetric nondegenerate bilinear form on k^n with Witt index q, and choose a basis of k^n so that the matrix of b has the form

$$B = \begin{pmatrix} & I_q & & \\ I_q & & & \\ & & c_1 & \\ & & & c_2 \end{pmatrix}$$

where I_q is the $(q \times q)$ identity matrix, and $c_1, c_2 \in k^\times$.

Exercise 191. Generalize exercise 188 to the case $n = 2q + 2$. That is, reuse the calculation from Exercise 177 to write down the general form of an element of the special orthogonal Lie algebra $\mathfrak{so}_{2q+2,q}(k, b)$.

Exercise 192. Generalize exercise 189 to the case $n = 2q + 2$. That is, take $s \in S(k)$ an arbitrary element of the diagonal torus subgroup, and $X \in \mathfrak{so}_{2q+2,q}(k, b)$ an arbitrary element of the Lie algebra, calculate sXs^{-1}, and compare the result of that calculation with $\alpha(s)X$ to determine which characters α have nonzero root space $\mathfrak{g}_\alpha$. Confirm that the set of roots is the B_q root system as described below.

$$\Phi = \Phi\Big(\mathrm{SO}_{2q+2,q}(k, b), S(k)\Big) = \{\pm\alpha_i : 1 \leq i \leq q\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \leq i, j \leq q, i \neq j\}$$

Write out general forms for elements of the root spaces, and use these to determine what the maps X_α and x_α look like.

Remark 193. In the previous exercise, you should find that the dimensions of the root spaces are

Rootα	Dimension of $\mathfrak{g}_\alpha$
$\pm\alpha_i \pm \alpha_j$	1
$\pm\alpha_i$	$n - 2q = 2$

5.5 General non-split case ($n \geq 2q + 2$)

Let b be a symmetric nondegenerate bilinear form on k^n with Witt index q, and choose a basis of k^n so that the matrix of b is

$$B = \begin{pmatrix} & I_q & \\ I_q & & \\ & & C \end{pmatrix}$$

where I_q is the $(q \times q)$ identity matrix and $C = \mathrm{diag}(c_1, c_2, \ldots, c_{n-2q})$ is an $(n-2q) \times (n-2q)$ invertible diagonal matrix with entries in $k^\times$. There is not much to say about generalizing Exercise 191 to the general non-split case, since there isn't much more to do than just copy down the general result of the calculations from Exercise 177.

We can generalize Exercise 192 to the general non-split case, by starting with $s \in S(k)$ and $X \in \mathfrak{so}_{n,q}(k, b)$ and computing sXs^{-1} to determine the roots and root spaces for $\mathrm{SO}_{n,q}(k, b)$. The dimensions of the root spaces are

Root α	Dimension of $\mathfrak{g}_\alpha$
$\pm\alpha_i \pm \alpha_j$	1
$\pm\alpha_i$	$n - 2q$

6 Special unitary groups

In this section, we examine special unitary groups in more detail. For the duration of this section, fix a quadratic field extension $L/k = k(\sqrt{d})/k$ and let $\sigma(x) = \overline{x}$ be the standard involution

$$\overline{a + b\sqrt{d}} = a - b\sqrt{d}$$

(for $a, b \in k$). Also fix n, and fix a hermitian or skew-hermitian form h on $V = L^n$. Fix a basis of V, and let H be the matrix of h in a fixed basis. Let q be the Witt index of h. Via a hermitian analog of the Witt decomposition theorem, it is possible to choose a basis of V so that the matrix H has the form

$$H = \begin{pmatrix} & I_q & \\ \varepsilon I_q & & \\ & & C \end{pmatrix}$$

where $\varepsilon = \pm 1$ is 1 if h is hermitian and -1 if h is skew-hermitian, and $C = \operatorname{diag}(c_1, \ldots, c_{n-2q})$ is an invertible diagonal matrix corresponding to the anisotropic part of h. Note that when h is hermitian, the entries of C must lie in k, while if h is skew-hermitian the entries of C must be "purely imaginary," i.e. of the form $a\sqrt{d}$ for some $a \in k$. This can be described in both cases by saying that

$$\overline{C} = \varepsilon C$$

Remark 194. It is possible to assume $\varepsilon = 1$ or $\varepsilon = -1$ without any loss of generality. However, we will write things out for both cases as much as possible.

Definition 195. The **special unitary group** associated to h is

$$\mathrm{SU}_{n,q}(L, h) = \{X \in \mathrm{SL}_n(L) : X^* H X = H\}$$

Conceptually, we can think of $\mathrm{SU}_{n,q}(L, h)$ as linear transformations $V \to V$ which preserve h. But we will never make serious use of this point of view; rather we will focus on the concrete matrix equation $X^* H X = H$.

Remark 196. If $n = 2q$, then $\mathrm{SU}_{n,q}$ is quasi-split, but if $n > 2q$ then $\mathrm{SU}_{n,q}$ is not quasi-split (and hence also not split). The case $n = 2q$ was previously studied in [RR22], although in that paper we had H in a different form than our general form in these notes. In particular, in that paper we put H in the block diagonal form below.

$$\begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}$$

6.1 Quasi-split case ($n = 2q$)

Let L/k be a quadratic extension and $\sigma(x) = \overline{x}$ be the conjugation operation on L. Let $q \geq 1$ and let $n = 2q$ and let V be an n-dimensional L-vector space, fix a basis of V, and let h be the nondegenerate hermitian or skew-hermitian form on V with matrix

$$H = \begin{pmatrix} & I_q \\ \varepsilon I_q & \end{pmatrix}$$

Note that h has Witt index q, half the dimension of V. Since the Witt index can never be more than half the dimension, this equivalent to saying that the Witt index is maximal. The associated special unitary group to h (and H) is

$$\mathrm{SU}_{n,q}(L,h) = \mathrm{SU}_{2q,q}(L,h) = \{X \in \mathrm{SL}_{2q}(L) : X^* H X = H\}$$

Exercise 197. Show that if $t \in \mathrm{SU}_{2q,q}(L,h)$ is a diagonal matrix, then t has the form

$$t = \begin{pmatrix} t_1 & & & & & \\ & \ddots & & & & \\ & & t_q & & & \\ & & & \bar{t}_1^{-1} & & \\ & & & & \ddots & \\ & & & & & \bar{t}_q^{-1} \end{pmatrix} \qquad t_1, \ldots, t_q \in L^\times$$

Definition 198. The set $T(k)$ of all diagonal matrices in $\mathrm{SU}_{2q,q}(L,h)$ is a maximal torus, but we are more interested in the maximal split torus $S(k)$ described below.

$$S(k) = \left\{ \begin{pmatrix} s_1 & & & & & \\ & \ddots & & & & \\ & & s_q & & & \\ & & & s_1^{-1} & & \\ & & & & \ddots & \\ & & & & & s_q^{-1} \end{pmatrix} : s_i \in k^\times \right\}$$

For $i = 1, 2, \ldots, q$ let $\alpha_i : S(k) \to k^\times$ be the map which picks off the ith diagonal entry.

Lemma 199. *$X^*(S)$ is isomorphic to $\mathbb{Z} \oplus \cdots \oplus \mathbb{Z} = \mathbb{Z}^q$ with basis $\{\alpha_1, \alpha_2, \ldots, \alpha_q\}$.*

Exercise 200. The root system of $\mathrm{SU}_{2q,q}(L,h)$ (with respect to S) has type C_q.

$$\Phi_k = \{\pm 2\alpha_i\} \cup \{\pm\alpha_i \pm \alpha_j : i \neq j\} \qquad 1 \leq i, j \leq q$$

How big is the set Φ_k?

Example 201. Consider the minimal example $(n,q) = (4,2)$. The root system is type C_2.

$$\Phi_k = \{\pm 2\alpha_1, \pm 2\alpha_2, \pm\alpha_1 \pm \alpha_2\}$$

Define

$$p_\varepsilon = \begin{cases} 1 & \varepsilon = -1 \\ \sqrt{d} & \varepsilon = 1 \end{cases} \qquad\qquad k_\varepsilon = kp_\varepsilon = \begin{cases} k & \varepsilon = -1 \\ k\sqrt{d} & \varepsilon = 1 \end{cases}$$

The pinning maps for long roots are

$$\begin{aligned} x_{2\alpha_i} &: k \to \mathrm{SU}_{n,q}(L,h) & x_{2\alpha_i}(v) &= 1 + E_{i,q+i}(p_\varepsilon v) \\ x_{-2\alpha_i} &: k \to \mathrm{SU}_{n,q}(L,h) & x_{-2\alpha_i}(v) &= 1 + E_{q+i,i}(p_\varepsilon v) \end{aligned}$$

$$x_{2\alpha_1}(v) = \begin{pmatrix} 1 & & p_\varepsilon v & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \qquad x_{-2\alpha_1}(v) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ p_\varepsilon v & & 1 & \\ & & & 1 \end{pmatrix}$$

$$x_{2\alpha_2}(v) = \begin{pmatrix} 1 & & & \\ & 1 & & p_\varepsilon v \\ & & 1 & \\ & & & 1 \end{pmatrix} \qquad x_{-2\alpha_2}(v) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & p_\varepsilon v & & 1 \end{pmatrix}$$

The pinning maps for medium roots are

$$\begin{aligned} x_{\alpha_i-\alpha_j} &: L \to \mathrm{SU}_{n,q}(L,h) & x_{\alpha_i-\alpha_j}(v) &= 1 + E_{ij}(v) - E_{q+j,q+i}(\overline{v}) \\ x_{\alpha_i+\alpha_j} &: L \to \mathrm{SU}_{n,q}(L,h) & x_{\alpha_i+\alpha_j}(v) &= 1 + E_{i',q+j'}(v) - \varepsilon E_{j',q+i'}(\overline{v}) \\ x_{-(\alpha_i+\alpha_j)} &: L \to \mathrm{SU}_{n,q}(L,h) & x_{-(\alpha_i+\alpha_j)}(v) &= 1 + E_{q+j',i'}(v) - \varepsilon E_{q+i',j'}(\overline{v}) \end{aligned}$$

where $i' = \min(i,j)$ and $j' = \max(i,j)$. Explicitly, these look like

$$x_{\alpha_1-\alpha_2}(v) = \begin{pmatrix} 1 & v & & \\ & 1 & & \\ & & 1 & \\ & & -\overline{v} & 1 \end{pmatrix} \qquad x_{\alpha_2-\alpha_1}(v) = \begin{pmatrix} 1 & & & \\ v & 1 & & \\ & & 1 & -\overline{v} \\ & & & 1 \end{pmatrix}$$

$$x_{\alpha_1+\alpha_2}(v) = \begin{pmatrix} 1 & & & v \\ & 1 & -\varepsilon\overline{v} & \\ & & 1 & \\ & & & 1 \end{pmatrix} \qquad x_{-\alpha_1-\alpha_2}(v) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & -\varepsilon\overline{v} & 1 & \\ v & & & 1 \end{pmatrix}$$

Exercise 202. Let $\alpha = 2\alpha_1$ and $\beta = -\alpha_1 + \alpha_2$.

(a) Verify that $\alpha + \beta$ and $\alpha + 2\beta$ are in Φ_k, but that these are the only positive integer linear combinations of the form $i\alpha + j\beta$ which are in Φ_k. In other words, on the right hand hand side of the commutator formula, we get something like

$$\prod_{\substack{i,j \geq 1 \\ i\alpha + j\beta \in \Phi_k}} x_{i\alpha+j\beta}\left(N_{ij}^{\alpha\beta}(u,v)\right) = x_{\alpha+\beta}\left(?\right) \cdot x_{\alpha+2\beta}\left(?\right)$$

(b) Calculate the left hand side of the commutator formula in this case:

$$\left[x_{2\alpha_1}(u), x_{-\alpha_1+\alpha_2}(v)\right] = ?$$

(c) Using your answer for part (b), work out what the inputs to $x_{\alpha+\beta}$ and $x_{\alpha+2\beta}$ should be in the formula in part (a). In other words, determine the two expressions below.

$$N_{11}^{2\alpha_1,-\alpha_1+\alpha_2}(u,v) = ?$$
$$N_{12}^{2\alpha_1,-\alpha_1+\alpha_2}(u,v) = ?$$

6.2 Non-quasi-split case ($n > 2q$)

Let $L/k = k(\sqrt{d})/k$ be a quadratic extension with conjugation operation $\sigma(x) = \overline{x}$, and let V be an n-dimensional L-vector space. Fix a basis of V, and let h a nondegenerate isotropic (meaning $q \geq 1$) hermitian or skew-hermitian form on V with Witt index q, with associated (skew-)hermitian matrix H. Then the associated special unitary group is

$$\mathrm{SU}_{n,q}(L,h) = \{X \in \mathrm{SL}_n(L) : X^* H X = H\}$$

Note that we are no longer assuming that h has maximal Witt index. Instead, we just know that $1 \leq q \leq \frac{n}{2}$.

Example 203. The smallest non-quasi-split case is $(n,q) = (5,2)$. In this case, V is a 5-dimensional L-vector space, and h is the hermitian or skew-hermitian form on V with matrix

$$H = \begin{pmatrix} & I_2 & \\ \varepsilon I_2 & & \\ & & c_1 \end{pmatrix} = \begin{pmatrix} & & 1 & & \\ & & & 1 & \\ \varepsilon 1 & & & & \\ & \varepsilon 1 & & & \\ & & & & c_1 \end{pmatrix}$$

Note that $c_1 \in k$ in the hermitian case and $c_1 \in k(\sqrt{d})$ in the skew-hermitian case, and $c_1 \neq 0$ in either case. A general diagonal matrix in $\mathrm{SU}_{5,2}(L,h)$ has the form

$$t = \begin{pmatrix} t_1 & & & & \\ & t_2 & & & \\ & & \overline{t}_1^{-1} & & \\ & & & \overline{t}_2^{-1} & \\ & & & & 1 \end{pmatrix} \qquad t_1, t_2 \in L^\times$$

Let $T(k)$ be the subgroup of such matrices. The maximal k-split torus we need is the subgroup $S(k) \subset T(k)$ where the diagonal entries lie in k.

$$S(k) = \left\{ \begin{pmatrix} s_1 & & & & \\ & s_2 & & & \\ & & s_1^{-1} & & \\ & & & s_2^{-1} & \\ & & & & 1 \end{pmatrix} : s_1, s_2 \in k^\times \right\}$$

For $i = 1,2$ let $\alpha_i : S(k) \to k^\times$ be the map which picks off the ith diagonal entry. The character group of S is the group of group homomorphisms $S(k) \to k^\times$, and it is isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$ with basis $\{\alpha_1, \alpha_2\}$. The root system of $\mathrm{SU}_{5,2}(L,h)$ (with respect to S) is the following set of 12 characters.

$$\Phi_k = \{\pm\alpha_1, \pm\alpha_2, \pm 2\alpha_1, \pm 2\alpha_2, \pm(\alpha_1 - \alpha_2), \pm(\alpha_1 + \alpha_2)\}$$

This is the root system of type BC_2. This root system is called "non-reduced" because it contains roots which are proportional but not negatives of each other, e.g. α_1 and $2\alpha_1$. Depending on where you look, this does not meet the definition of a root system. However, we do consider it a root system here, and the traditional definition is a "reduced" root system. Note that BC_q is actually the only non-reduced root system, so expanding the definition in this way adds only one case.

We return to the general non-quasi-split case ($n > 2q$). Choose a basis of V in which the matrix H is

$$H = \begin{pmatrix} & I_q & \\ \varepsilon I_q & & \\ & & C \end{pmatrix}$$

with C diagonal. When H has this form, the full diagonal subgroup of $\mathrm{SU}_{n,q}(L,h)$ is

$$T(k) = \left\{ \begin{pmatrix} T_1 & & \\ & \overline{T}_1^{-1} & \\ & & D \end{pmatrix} \;\middle|\; \begin{array}{l} T_1 = \mathrm{diag}(s_1, \dots, s_q) \in \mathrm{GL}_q(L), \\ D = \mathrm{diag}(d_1, \dots, d_{n-2q}),\ \mathrm{N}_{L/k}(d_j) = 1, \\ \displaystyle\prod_{i=1}^{q} s_i \overline{s_i}^{-1} \cdot \prod_{j=1}^{n-2q} d_j = 1 \end{array} \right\}$$

where T_1 is an arbitrary $q \times q$ invertible diagonal matrix with entries in L, each $d_j \in L^\times$ satisfies $\mathrm{N}_{L/k}(d_j) = 1$, and determinant of the overall matrix is 1. But the maximal k-split torus subgroup that we need is

$$S(k) = \left\{ \begin{pmatrix} S_1 & & \\ & S_1^{-1} & \\ & & I_{n-2q} \end{pmatrix} = \begin{pmatrix} s_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & s_q & & & & & & \\ & & & s_1^{-1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & s_q^{-1} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} : s_i \in k^\times \right\}$$

A major additional complication in the case $n > 2q$ is that the root system $\Phi_k = \Phi(\mathrm{SU}_{n,q}(L,h), S)$ is no longer "reduced," because it contains some scalar multiples of its elements other than ± 1. Thankfully, it only includes scalar multiples of $\pm 1, \pm 2, \pm \frac{1}{2}$. The root system is called BC_q, which is

$$\Phi_k = \{\pm\alpha_i, \pm 2\alpha_i : 1 \le i \le q\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \le i, j \le q, i \ne j\}$$

The Lie algebra of $\mathrm{SU}_{n,q}(L,h)$ is

$$\mathfrak{su}_{n,q}(L,h) = \{X \in \mathfrak{sl}_n(L) : X^* H + H X = 0\}$$

We will work out what a general element of $\mathfrak{su}_{n,q}(L,h)$ looks like. Let $X \in \mathfrak{su}_{n,q}(L,h)$, and write it in block form

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$$

with the block sizes corresponding to those in H. That is, $X_{11}, X_{12}, X_{21}, X_{22}$ are all $q \times q$, X_{13} and X_{23} have q rows and $(n-2q)$ columns, while X_{31} and X_{32} have q columns and $(n-2q)$ rows, and X_{33}

is $(n-2q)\times(n-2q)$. Then X^*H+HX becomes

$$\begin{aligned}
X^*H+HX &= \begin{pmatrix} X_{11}^* & X_{21}^* & X_{31}^* \\ X_{12}^* & X_{22}^* & X_{32}^* \\ X_{13}^* & X_{23}^* & X_{33}^* \end{pmatrix} \begin{pmatrix} & I_q & \\ \varepsilon I_q & & \\ & & C \end{pmatrix} + \begin{pmatrix} & I_q & \\ \varepsilon I_q & & \\ & & C \end{pmatrix} \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} \\
&= \begin{pmatrix} \varepsilon X_{21}^* & X_{11}^* & X_{31}^*C \\ \varepsilon X_{22}^* & X_{12}^* & X_{32}^*C \\ \varepsilon X_{23}^* & X_{13}^* & X_{33}^*C \end{pmatrix} + \begin{pmatrix} X_{21} & X_{22} & X_{23} \\ \varepsilon X_{11} & \varepsilon X_{12} & \varepsilon X_{13} \\ CX_{31} & CX_{32} & CX_{33} \end{pmatrix}
\end{aligned}$$

so for X to belong to $\mathfrak{su}_{n,q}$ we get the following nine equations, with some redundancies.

$$\begin{aligned}
\varepsilon X_{21}^* + X_{21} &= 0 \\
X_{11}^* + X_{22} &= 0 \\
X_{31}^*C + X_{23} &= 0 \\
\varepsilon X_{22}^* + \varepsilon X_{11} &= 0 \\
X_{12}^* + \varepsilon X_{12} &= 0 \\
X_{32}^*C + \varepsilon X_{13} &= 0 \\
\varepsilon X_{23}^* + CX_{31} &= 0 \\
X_{13}^* + CX_{32} &= 0 \\
X_{33}^*C + CX_{33} &= 0
\end{aligned}$$

Removing redundancies and rearranging, combined with the facts that C is invertible, diagonal, $\overline{C} = \varepsilon C$, we get the following conditions for $X \in \mathfrak{su}_{n,q}$.

$$\begin{aligned}
X_{22} &= -X_{11}^* \\
X_{12} &= -\varepsilon X_{12}^* \\
X_{21} &= -\varepsilon X_{21}^* \\
X_{13} &= -X_{32}^*\overline{C} = -\varepsilon X_{32}^*C \\
X_{23} &= -X_{31}^*C = -\varepsilon X_{31}^*\overline{C} \\
X_{33} &= C^{-1}X_{33}^*C
\end{aligned}$$

So an arbitrary element of $\mathfrak{su}_{n,q}$ looks like

$$X = \begin{pmatrix} X_{11} & X_{12} & -\varepsilon X_{32}^*C \\ X_{21} & -X_{11}^* & -X_{31}^*C \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$$

with X_{12} and X_{21} hermitian or skew-hermitian depending on ε, $X_{33} = C^{-1}X_{33}^*C$, and X_{11}, X_{31}, X_{32} arbitrary. In particular, the diagonal entries of X_{12} and X_{21} are either in k or "purely imaginary" depending on ε.

Let $X \in \mathfrak{su}_{n,q}(L,h)$ be a generic element of the Lie algebra and let $s \in S(k)$ be a generic torus element. We want to work out the root spaces in $\mathfrak{su}_{n,q}$ by comparing sXs^{-1} and $\alpha(s)X$.

$$sXs^{-1} = \begin{pmatrix} s_1 &&&&&&&& \\ & \ddots &&&&&&& \\ && s_q &&&&&& \\ &&& s_1^{-1} &&&&& \\ &&&& \ddots &&&& \\ &&&&& s_q^{-1} &&& \\ &&&&&& 1 && \\ &&&&&&& \ddots & \\ &&&&&&&& 1 \end{pmatrix} \left(x_{ij}\right) \begin{pmatrix} s_1^{-1} &&&&&&&& \\ & \ddots &&&&&&& \\ && s_q^{-1} &&&&&& \\ &&& s_1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& s_q &&& \\ &&&&&& 1 && \\ &&&&&&& \ddots & \\ &&&&&&&& 1 \end{pmatrix}$$

$$= \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$$

where each X_{ij} is a block matrix as described below.

$$X_{11} = \left(s_i s_j^{-1} x_{ij}\right) = \begin{pmatrix} x_{11} & \cdots & s_1 s_q^{-1} x_{1q} \\ \vdots & \ddots & \vdots \\ s_q s_1^{-1} x_{q1} & \cdots & x_{qq} \end{pmatrix}$$

$$X_{12} = \left(s_i s_j x_{i,q+j}\right) = \begin{pmatrix} s_1^2 x_{1,q+1} & \cdots & s_1 s_q x_{1,2q} \\ \vdots & \ddots & \vdots \\ s_q s_1 x_{q,q+1} & \cdots & s_q^2 x_{q,2q} \end{pmatrix}$$

$$X_{13} = \left(s_i x_{i,2q+j}\right) = \begin{pmatrix} s_1 x_{1,2q+1} & \cdots & s_1 x_{1n} \\ \vdots & \ddots & \vdots \\ s_q x_{q,2q+1} & \cdots & s_q x_{qn} \end{pmatrix}$$

$$X_{21} = \left(s_i^{-1} s_j^{-1} x_{q+i,j}\right) = \begin{pmatrix} s_1^{-2} x_{q+1,1} & \cdots & (s_1 s_q)^{-1} x_{q+1,q} \\ \vdots & \ddots & \vdots \\ (s_q s_1)^{-1} x_{2q,1} & \cdots & s_q^{-2} x_{2q,q} \end{pmatrix}$$

$$X_{22} = \left(s_i^{-1} s_j x_{q+i,q+j}\right) = \begin{pmatrix} x_{q+1,q+1} & \cdots & s_1^{-1} s_q x_{q+1,2q} \\ \vdots & \ddots & \vdots \\ s_q^{-1} s_1 x_{2q,q+1} & \cdots & x_{2q,2q} \end{pmatrix}$$

$$X_{23} = \left(s_i^{-1} x_{q+i,2q+j}\right) = \begin{pmatrix} s_1^{-1} x_{q+1,2q+1} & \cdots & s_1^{-1} x_{q+1,n} \\ \vdots & \ddots & \vdots \\ s_q^{-1} x_{2q,2q+1} & \cdots & s_q^{-1} x_{2q,n} \end{pmatrix}$$

$$X_{31} = \left(s_j^{-1} x_{2q+i,j}\right) = \begin{pmatrix} s_1^{-1} x_{2q+1,1} & \cdots & s_q^{-1} x_{2q+1,q} \\ \vdots & \ddots & \vdots \\ s_1^{-1} x_{n,1} & \cdots & s_q^{-1} x_{nq} \end{pmatrix}$$

$$X_{32} = \left(s_j x_{2q+i,q+j}\right) = \begin{pmatrix} s_1 x_{2q+1,q+1} & \cdots & s_q x_{2q+1,2q} \\ \vdots & \ddots & \vdots \\ s_1 x_{n,q+1} & \cdots & s_q x_{n,2q} \end{pmatrix}$$

$$X_{33} = \left(x_{2q+i,2q+j}\right) = \begin{pmatrix} x_{2q+1,2q+1} & \cdots & x_{2q+1,n} \\ \vdots & \ddots & \vdots \\ x_{n,2q+1} & \cdots & x_{nn} \end{pmatrix}$$

By inspecting the above matrices, we can see that the roots are

$$\Phi = \{\pm\alpha_i, \pm 2\alpha_i : 1 \leq i \leq q\} \cup \{\pm\alpha_i \pm \alpha_j : 1 \leq i, j \leq q, i \neq j\}$$

which is the non-reduced root system of type BC_q. The root spaces are described in the tables below. In the table, we assume $i \neq j$, and that all dimensions are over k.

Root	Blocks	Dimension of root space
$+\alpha_i$	X_{13}, X_{32}	$2(n-2q)$
$-\alpha_i$	X_{31}, X_{23}	$2(n-2q)$
$+2\alpha_i$	X_{12}	1
$-2\alpha_i$	X_{21}	1
$\alpha_i + \alpha_j$	X_{12}	2
$-(\alpha_i + \alpha_j)$	X_{21}	2
$\pm(\alpha_i - \alpha_j)$	X_{11}, X_{22}	2

There are three root lengths: short, medium, and long. The root spaces for long roots are easiest to describe: they lie along the diagonal of the X_{12} and X_{21} blocks, each root space occupying a single matrix entry. Since the X_{12} and X_{21} blocks each satisfy $X = -\varepsilon X^*$, these diagonal entries satisfy $x = -\varepsilon\overline{x}$. If $\varepsilon = -1$, this means the diagonal entries in question lie in k, while if $\varepsilon = 1$ they lie in $k\sqrt{d}$, the "purely imaginary" elements of the quadratic extension $k(\sqrt{d})/k$. To denote this, let

$$p_\varepsilon = \begin{cases} 1 & \varepsilon = -1 \\ \sqrt{d} & \varepsilon = 1 \end{cases} \qquad\qquad k_\varepsilon = kp_\varepsilon = \begin{cases} k & \varepsilon = -1 \\ k\sqrt{d} & \varepsilon = 1 \end{cases}$$

Then we can describe the root spaces as

$$\mathfrak{g}_{2\alpha_i} = \{E_{i,i+q}(x) : x \in k_\varepsilon\} = \{E_{i,i+q}(p_\varepsilon v) : v \in k\}$$
$$\mathfrak{g}_{-2\alpha_i} = \{E_{i+q,i}(x) : x \in k_\varepsilon\} = \{E_{i+q,i}(p_\varepsilon v) : v \in k\}$$

Consequently, the root space maps and root subgroup maps are simple to describe.

$$\begin{aligned} X_{2\alpha_i} &: k \to \mathfrak{su}_{n,q} & X_{2\alpha_i}(v) &= E_{i,q+i}(p_\varepsilon v) \\ X_{-2\alpha_i} &: k \to \mathfrak{su}_{n,q} & X_{-2\alpha_i}(v) &= E_{q+i,i}(p_\varepsilon v) \\ x_{2\alpha_i} &: k \to \mathrm{SU}_{n,q} & x_{2\alpha_i}(v) &= 1 + E_{i,q+i}(p_\varepsilon v) \\ x_{-2\alpha_i} &: k \to \mathrm{SU}_{n,q} & x_{-2\alpha_i}(v) &= 1 + E_{q+i,i}(p_\varepsilon v) \end{aligned}$$

Next, we describe the root spaces and root subgroup maps for medium length roots, those of the form $\pm\alpha_i \pm \alpha_j$ with $i \neq j$. Their root spaces live in the X_{11}, X_{12}, X_{21}, and X_{22} blocks. Since X_{12} and X_{21} are hermitian or skew-hermitian, the root spaces for $\pm(\alpha_i + \alpha_j)$ occur in transpose-pairs of entries. Since $X_{22} = -X_{11}^*$, the root spaces for $\alpha_i - \alpha_j$ occur in pairs of entries, one from X_{11} and the accompanying entry of X_{22} determined by that entry.

Root	Block	Specific entries
$\alpha_i - \alpha_j$	X_{11}, X_{22}	$x_{ij}, x_{q+j,q+i}$
$\alpha_i + \alpha_j$	X_{12}	$x_{i,q+j}, x_{j,q+i}$
$-\alpha_i - \alpha_j$	X_{21}	$x_{q+i,j}, x_{q+j,i}$

For each of these roots, the specific listed entries of an arbitrary element $X = (x_{ij}) \in \mathfrak{su}_{n,q}$ can be anything, besides their pairwise dependency. So each root space has dimension 1 over L, hence dimension 2 over k.

$$\begin{aligned} X_{\alpha_i-\alpha_j} &: L \to \mathfrak{su}_{n,q} & X_{\alpha_i-\alpha_j}(v) &= E_{ij}(v) - E_{q+j,q+i}(\overline{v}) \\ X_{\alpha_i+\alpha_j} &: L \to \mathfrak{su}_{n,q} & X_{\alpha_i+\alpha_j}(v) &= E_{i',q+j'}(v) - E_{j',q+i'}(\varepsilon\overline{v}) \\ X_{-\alpha_i-\alpha_j} &: L \to \mathfrak{su}_{n,q} & X_{-\alpha_i-\alpha_j}(v) &= E_{q+j',i'}(v) - E_{q+i',j'}(\varepsilon\overline{v}) \end{aligned}$$

where $i' = \min(i,j)$ and $j' = \max(i,j)$. The conventions on $X_{\pm(\alpha_i+\alpha_j)}$ are chosen so that

$$X_{-\alpha_i-\alpha_j}(v) = X_{\alpha_i+\alpha_j}(v)^T$$

Finally, we describe the root spaces and root subgroup maps for short roots, those of the form $\pm\alpha_i$. For $+\alpha_i$, the root space lives in the X_{13} and X_{32} blocks, while for $-\alpha_i$ it lives in the X_{23} and X_{31} blocks. The root space takes up a whole row or column of each block. Since $X_{13} = -\varepsilon X_{32}^* C$, the entries of a row in X_{13} are determined by the entries of a corresponding column in X_{32}. There are $n - 2q$ of these entries, which can be any element of L, so the root space associated to $\pm\alpha_i$ has dimension $2(n-2q)$ over k.

$$\mathfrak{g}_{\alpha_i} = \left\{ \begin{pmatrix} 0 & 0 & X_{13} \\ 0 & 0 & 0 \\ 0 & X_{32} & 0 \end{pmatrix} \right\}$$

where

$$X_{32} = \begin{pmatrix} 0 & v_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & v_{n-2q} & 0 \end{pmatrix} \qquad X_{13} = -\varepsilon X_{32}^* C = \begin{pmatrix} 0 & \cdots & 0 \\ -\varepsilon c_1 \overline{v}_1 & \cdots & -\varepsilon c_{n-2q} \overline{v}_{n-2q} \\ 0 & \cdots & 0 \end{pmatrix}$$

with the column vector $(v_1, \ldots, v_{n-2q})$ occuring as the ith column of X_{32} and the ith row of X_{13}, and the elements $v_1, \ldots, v_{n-2q}$ can be anything from L. Similarly,

$$\mathfrak{g}_{-\alpha_i} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & X_{23} \\ X_{31} & 0 & 0 \end{pmatrix} \right\}$$

where

$$X_{31} = \begin{pmatrix} 0 & v_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & v_{n-2q} & 0 \end{pmatrix} \qquad X_{23} = -X_{31}^* C = \begin{pmatrix} 0 & \cdots & 0 \\ -c_1\overline{v}_1 & \cdots & -c_{n-2q}\overline{v}_{n-2q} \\ 0 & \cdots & 0 \end{pmatrix}$$

with $(v_1, \ldots, v_{n-2q})$ being the ith column of X_{31} and corresponding to the ith row of X_{23}.

Root	Block	Specific entries	Range
α_i	X_{32}	$x_{2q+j,q+i}$	$1 \leq j \leq n-2q$
	X_{13}	$x_{i,2q+j}$	$1 \leq j \leq n-2q$
$-\alpha_i$	X_{31}	$x_{2q+j,i}$	$1 \leq j \leq n-2q$
	X_{23}	$x_{q+i,2q+j}$	$1 \leq j \leq n-2q$

References

[Bor91] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[BT73] Armand Borel and Jacques Tits. Homomorphismes "abstraits" de groupes algébriques simples. *Ann. of Math. (2)*, 97:499–571, 1973.

[Con] Keith Conrad. Bilinear forms. `https://kconrad.math.uconn.edu/blurbs/linmultialg/bilinearform.pdf`.

[Gre64] Marvin J. Greenberg. Algebraic rings. *Trans. Amer. Math. Soc.*, 111:472–481, 1964.

[KS09] Martin Kassabov and Mark V. Sapir. Nonlinearity of matrix groups. *J. Topol. Anal.*, 1(3):251–260, 2009.

[Rap11] Igor A. Rapinchuk. On linear representations of Chevalley groups over commutative rings. *Proc. Lond. Math. Soc. (3)*, 102(5):951–983, 2011.

[Rap13] Igor A. Rapinchuk. On abstract representations of the groups of rational points of algebraic groups and their deformations. *Algebra Number Theory*, 7(7):1685–1723, 2013.

[Rap19] Igor A. Rapinchuk. On abstract homomorphisms of Chevalley groups over the coordinate rings of affine curves. *Transform. Groups*, 24(4):1241–1259, 2019.

[RR22] Igor Rapinchuk and Joshua Ruiter. On abstract homomorphisms of some special unitary groups. *Proc. Amer. Math. Soc.*, 150(10):4241–4258, 2022.