Root Systems: Classification, Construction, and Consequences

Callum Laval

Contents

1 Introduction

Historically, root systems came from the study of objects called Lie algebras which is a big motivation towards the study of root systems. So to motivate us as to why we want to look at root systems we will look at an example of a Lie algebra based on an example from section 12.2.2 of the Math426 lecture notes [1, p.105]. A definition of a Lie algebra can be found in Definition 3.1

Consider the set of 3-by-3 matrices over the complex numbers whose diagonal entries sum to equal zero i.e. the set of 3-by-3 trace-free matrices over $\mathbb{C}$. This set is traditionally denoted by $\mathfrak{sl}(3, \mathbb{C})$ because of its relationship to the set of special linear matrices (matrices whose determinant equals 1) [2, p.2]. For more details about this relationship look at Example 4.5(ii).

We can write a general element of $\mathfrak{sl}(3, \mathbb{C})$ as follows

$$\begin{pmatrix} x & a & b \\ d & y & c \\ e & f & -x-y \end{pmatrix}$$

with all entries being complex numbers.
It is clear that $\mathfrak{sl}(3, \mathbb{C})$ forms a vector space since it contains the zero matrix, additive inverses and is closed under matrix addition and scalar multiplication, so let us find a basis for it.
Let us define M_{ij} to be the 3-by-3 matrix with zeros everywhere except in the i, j position for $1 \leq i \leq 3$ and $1 \leq j \leq 3$ for example

$$M_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and define

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then a basis for $\mathfrak{sl}(3, \mathbb{C})$ is

$$\{H_1,\ H_2,\ M_{1,2},\ M_{2,1},\ M_{1,3},\ M_{3,1},\ M_{2,3},\ M_{3,2}\}.$$

An important operation on $\mathfrak{sl}(3,\mathbb{C})$ is the commutator bracket denoted by $[,]$ where for $X, Y \in \mathfrak{sl}(3,\mathbb{C})$ we have $[X,Y] = XY - YX$. We can see that this operation is closed on $\mathfrak{sl}(3,\mathbb{C})$ as if we take $X, Y \in \mathfrak{sl}(3,\mathbb{C})$ we can use standard properties of the trace to get

$$tr([X,Y]) = tr(XY - YX) = tr(XY) - tr(YX) = tr(XY) - tr(XY) = 0,$$

hence $[X,Y] \in \mathfrak{sl}(3,\mathbb{C})$. In fact, $\mathfrak{sl}(3,\mathbb{C})$ together with the commutator bracket form a Lie algebra [2, p.2].

The subspace $\mathfrak{h}$ of all diagonal matrices in $\mathfrak{sl}(3,\mathbb{C})$ forms something called a Cartan subalgebra of $\mathfrak{sl}(3,\mathbb{C})$ and is very important in constructing root systems from Lie algebras [3, p.211]. A definition of a Cartan subalgebra can be found in Definition 3.12.

We have that a general element $H \in \mathfrak{h}$ will have the form

$$H = \begin{pmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{pmatrix},$$

where $\theta_1 + \theta_2 + \theta_3 = 0$.

Now let us look at what happens when we take the commutator bracket of H with the basis elements of $\mathfrak{sl}(3,\mathbb{C})$. Since H_1, H_2 and H are all diagonal matrices, they commute with each other so

$$\begin{aligned} [H, H_1] &= HH_1 - H_1H = HH_1 - HH_1 = 0, \\ [H, H_2] &= HH_2 - H_2H = HH_2 - HH_2 = 0. \end{aligned}$$

For $M_{1,2}$ we have

$$\begin{aligned} [H, M_{1,2}] &= \begin{pmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \theta_1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & \theta_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= (\theta_1 - \theta_2) M_{1,2}. \end{aligned}$$

In fact through similar calculations we get that

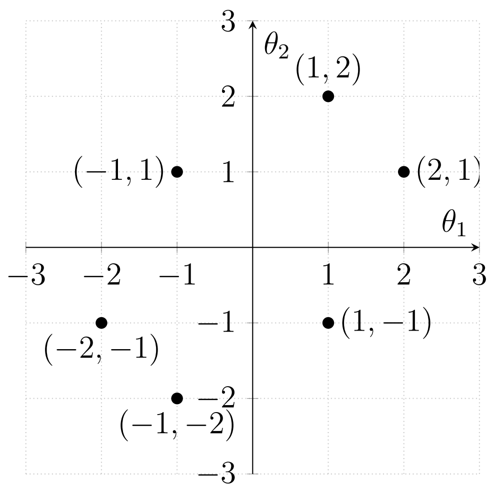
$$[H, M_{i,j}] = (\theta_i - \theta_j) M_{i,j},$$

and since $\theta_3 = -\theta_1 - \theta_2$ we have

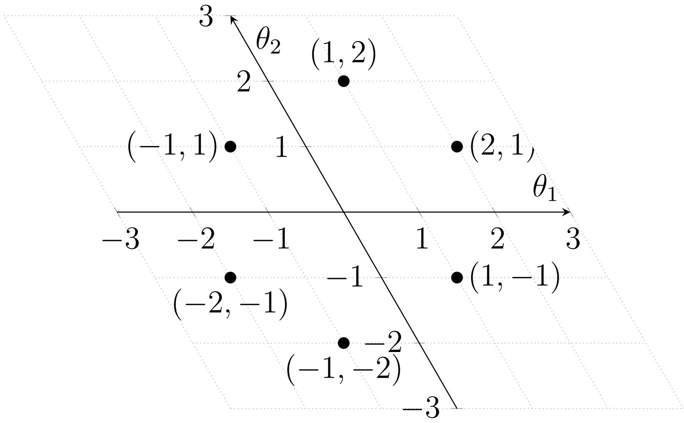
$$\begin{aligned}
[H, M_{1,2}] &= (\theta_1 - \theta_2)M_{1,2}, \\
[H, M_{2,1}] &= (-\theta_1 + \theta_2)M_{2,1}, \\
[H, M_{1,3}] &= (\theta_1 - \theta_3)M_{1,3} = (2\theta_1 + \theta_2)M_{1,3} \\
[H, M_{3,1}] &= (-\theta_1 + \theta_3)M_{3,1} = (-2\theta_1 - \theta_2)M_{3,1} \\
[H, M_{2,3}] &= (\theta_2 - \theta_3)M_{2,3} = (\theta_1 + 2\theta_2)M_{2,3} \\
[H, M_{3,2}] &= (-\theta_2 + \theta_3)M_{3,2} = (-\theta_1 - 2\theta_2)M_{3,2}.
\end{aligned}$$

Notice how each $M_{i,j}$ are actually eigenvectors of H with respect to the commutator bracket instead of matrix multiplication with eigenvalue $\theta_i - \theta_j$.

Now if we plot each of the non-zero eigenvalues on a coordinate axis where we fix a basis $\{\theta_1, \theta_2\}$ and write the eigenvalues with respect to this basis, we get the following picture.



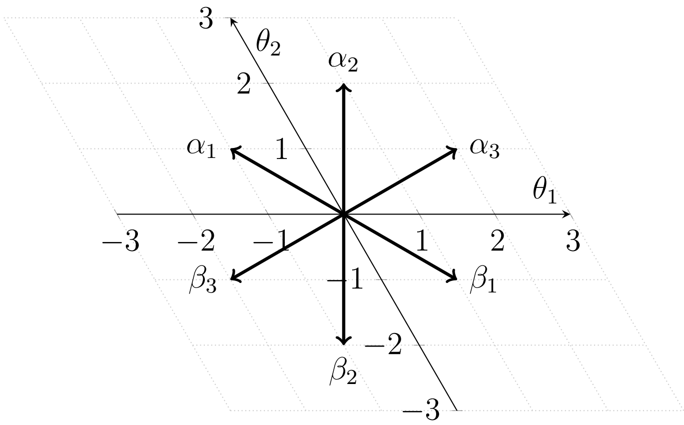
This is not a very enlightening picture as the points seem to not have much symmetry and form an outline of an irregular hexagon.
But now let us plot these points where the axes are $\frac{2\pi}{3}$ radians apart instead of $\frac{\pi}{2}$ radians.



Now we can see the points form a regular hexagon. Details on why we plot the axis in this way can be found in Example 3.20

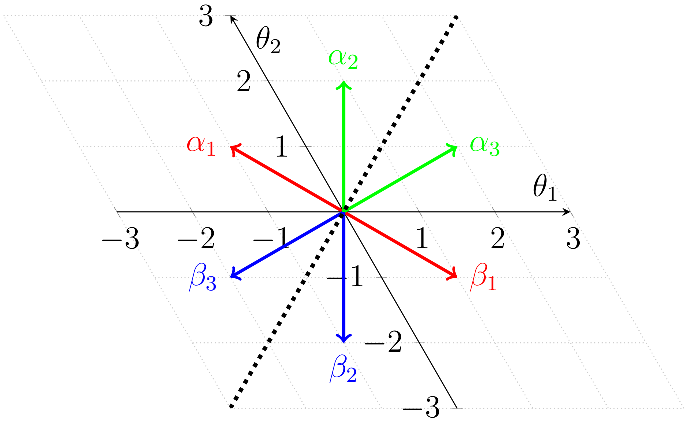
Let us consider these points as vectors in $\mathbb{R}^2$ and let Φ denote the collection of these vectors. The vectors don't form a subspace of $\mathbb{R}^2$ since it does not contain the zero vector although it is clear that they span $\mathbb{R}^2$. Let us discuss some nice properties of Φ.

First let us label the vectors like so.



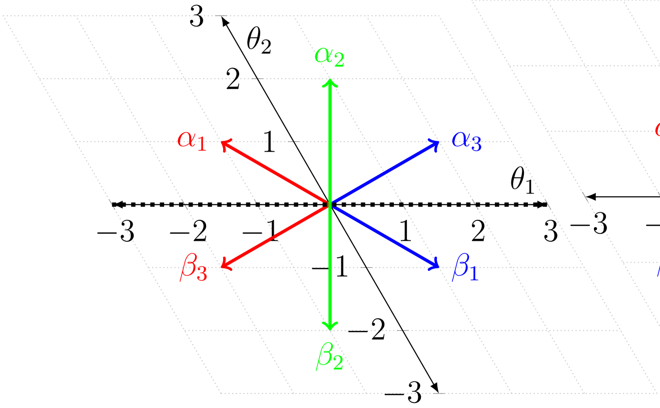
We have labelled them in this way as we have $\beta_i = -\alpha_i$. So our first property is that if $\alpha \in \Phi$ then $-\alpha \in \Phi$.

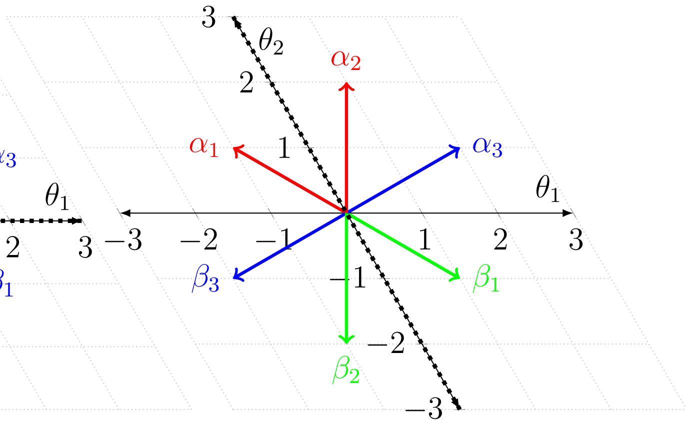
Now let us consider the reflections of Φ in the line orthogonal to α_1 and β_1 i.e the line $\theta_1 = \theta_2$.



So α_1 is sent to β_1, α_2 is sent to α_3 and β_3 is sent to β_2. Hence a reflection of a vector in Φ in this line will send it to another vector in Φ.

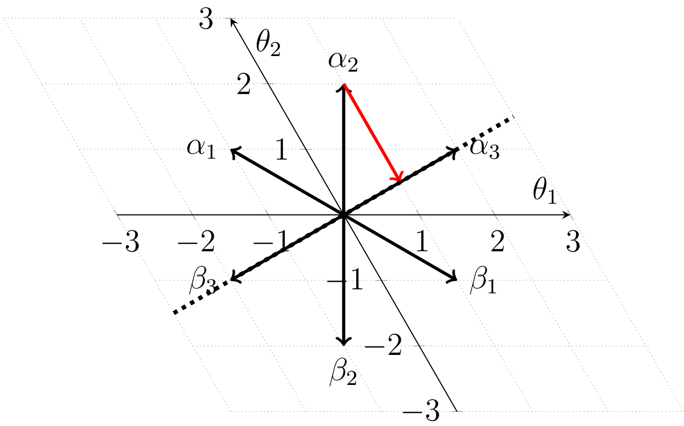
The same thing happens if we look at the reflections in other lines orthogonal to vectors in Φ.





Hence Φ is unchanged by a reflection in a line orthogonal to a vector in Φ.

Our final property is that if we consider the orthogonal projection of a vector $\alpha \in \Phi$ onto the line through a different root $\beta \in \Phi$ and the origin, then the projected α is equal to a half integer multiple of β.



So in the picture, the red arrow represents the orthogonal projection of α_2 onto the line through α_3 and the origin. We can see that α_2 is being projected to $\frac{1}{2}\alpha_3$.

In fact Φ is a root system called A_2. We can apply similar methods to other Lie algebras to construct other root systems. These root systems actually determines the Lie algebra up to isomorphism and since root systems tend to be simpler than Lie algebras it can make certain questions such as how to classify Lie algebras easier to answer. We can also construct Lie algebras from root systems which lead to the construction of Lie algebras that would have otherwise been very hard to define such as the Lie algebra G_2.

Since root systems are simpler we will focus on root systems by themselves in the next chapter and then in chapter 3 discuss Lie algebras and their relationship with root systems. In chapter 4 we will then go over some applications of root systems and Lie algebras in other areas of maths and physics.

2 Root Systems

This section is mainly based around the ideas from sources [2] and [5]. [2] follows a structure of briefly discussing reflections then going into the axioms of root systems and constructing the 2D examples. After that it discusses simple systems (or bases as the source calls them) and the Weyl group which leads to Coxeter graphs and Dynkin Diagrams which leads to the classification theorem.

The structure for this section will be the same but the approaches to of some of the topics will differ by bringing ideas from [5]. In particular we will go into more detail on reflections by talking about more results and our discussion of the Weyl group will mainly take ideas from [5]. We will use the approach of proving the existence of simple systems from [5] which differs from the approach in [2]. Finally we will approach the proof of the classification theorem differently from [2] by taking ideas from the proof and introducing them as their own lemmas or propositions.

2.1 Axioms

Before we define a root system we should first discuss what exactly we mean by a reflection since a key property of root systems is its reflectional symmetry.

Let V be a finite dimensional real inner product space. That is V is a vector space over $\mathbb{R}$ with a map $(,) : V \times V \to \mathbb{R}$ that is bilinear, symmetric and positive-definite which we call an inner product. If V is 2-dimensional then a reflection in V is a linear transformation that fixes a line in the vector space and sends any vector orthogonal to that line to its negative. These properties uniquely define the linear map [2, p.42].

Now we want to define reflections in higher dimensions which is where the notion of a hyperplane comes in. A hyperplane of V is a subspace of V whose dimension is one less then that of V [4]. So for 1 dimension the hyperplane is a point, for 2 dimensions it is a line and for 3 dimensions it is a plane.

Hence a reflection can be generally defined as a linear transformation on V that fixes some hyperplane in V and sends any vector orthogonal to that hyperplane to its negative.
Using the inner product $(,)$ on V we can define an explicit formula for reflections as follows.

Definition 2.1. [2, p.42][5, p.3] Let $\alpha \in V$, let $P_\alpha = \{\beta \in V : (\beta, \alpha) = 0\}$ be the hyperplane in V orthogonal to α, and let σ_α be the transformation on V representing the reflection in P_α.
Then σ_α acts on $\beta \in V$ by

$$\sigma_\alpha(\beta) = \beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha,$$

i.e. β is sent to $\sigma_\alpha(\beta)$ when reflecting in P_α.

We will use the constant $\frac{2(\beta,\alpha)}{(\alpha,\alpha)}$ quite frequently so to simplify things let $\langle\beta,\alpha\rangle = \frac{2(\beta,\alpha)}{(\alpha,\alpha)}$ so that the formula becomes

$$\sigma_\alpha(\beta) = \beta - \langle\beta,\alpha\rangle\alpha.$$

One property of reflections is that they preserve the lengths of vectors and angles of between vectors. Since the inner product determines both of these, we can state this property of reflections as the following.

Proposition 2.2. *Let $\alpha,\beta,\gamma \in V$. Then $(\sigma_\alpha\beta,\sigma_\alpha\gamma) = (\beta,\gamma)$ i.e. reflections are orthogonal transformations.*

Proof. Using the fact that inner products are bilinear and symmetric we get

$$\begin{aligned}
(\sigma_\alpha\beta,\sigma_\alpha\gamma) &= (\beta - \langle\beta,\alpha\rangle\alpha, \gamma - \langle\gamma,\alpha\rangle\alpha) \\
&= (\beta,\gamma) - \langle\gamma,\alpha\rangle(\beta,\alpha) - \langle\beta,\alpha\rangle(\alpha,\gamma) + \langle\beta,\alpha\rangle\langle\gamma,\alpha\rangle(\alpha,\alpha) \\
&= (\beta,\gamma) - 2\frac{(\gamma,\alpha)(\beta,\alpha)}{(\alpha,\alpha)} - 2\frac{(\beta,\alpha)(\alpha,\gamma)}{(\alpha,\alpha)} + 4\frac{(\beta,\alpha)(\gamma,\alpha)}{(\alpha,\alpha)} \\
&= (\beta,\gamma).
\end{aligned}$$

□

Another intuitive property of reflection is that a reflection applied to a vector twice does not change the vector i.e. a reflection composed with itself gives the identity map.

Proposition 2.3. *Let $\alpha \in V$. Then $\sigma_\alpha\sigma_\alpha$ is the identity map on V i.e. for all $\beta \in V$ we have $\sigma_\alpha\sigma_\alpha(\beta) = \beta$.*

Proof. Let $\beta \in V$. Since inner products are bilinear then $\langle,\rangle$ is linear in the first argument hence

$$\begin{aligned}
\sigma_\alpha\sigma_\alpha(\beta) &= \sigma_\alpha(\beta - \langle\beta,\alpha\rangle\alpha) \\
&= \beta - \langle\beta,\alpha\rangle\alpha - \langle\beta - \langle\beta,\alpha\rangle\alpha,\alpha\rangle\alpha \\
&= \beta - \langle\beta,\alpha\rangle\alpha - \langle\beta,\alpha\rangle\alpha + \langle\beta,\alpha\rangle\langle\alpha,\alpha\rangle\alpha \\
&= \beta - 2\langle\beta,\alpha\rangle\alpha + \langle\beta,\alpha\rangle\langle\alpha,\alpha\rangle\alpha \\
&= \beta - \frac{4(\beta,\alpha)}{(\alpha,\alpha)}\alpha + \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\frac{2(\alpha,\alpha)}{(\alpha,\alpha)}\alpha \\
&= \beta - \frac{4(\beta,\alpha)}{(\alpha,\alpha)}\alpha + \frac{4(\beta,\alpha)}{(\alpha,\alpha)}\alpha \\
&= \beta.
\end{aligned}$$

□

Another property of reflections are that they are injective

Proposition 2.4. *Let $\alpha \in V$. Then σ_α is an injective map.*

Proof. Let $\beta_1, \beta_2 \in V$. Then we have

$$\sigma_\alpha \beta_1 = \sigma_\alpha \beta_2 \implies \sigma_\alpha \sigma_\alpha \beta_1 = \sigma_\alpha \sigma_\alpha \beta_2 \implies \beta_1 = \beta_2.$$

Hence σ_α is injective. $\square$

In our initial example we saw that the root system A_2 has the following properties.

- All the roots are non-zero and span $\mathbb{R}^2$.
- If α is a root then $-\alpha$ is also a root.
- The set of roots is invariant under reflections in lines orthogonal to roots.
- The orthogonal projection of a root α onto a line though a different root β and the origin will be a half integer multiple of β.

Now let us axiomatise these properties.

Definition 2.5. [2, p.42] Let V be a finite dimensional vector space over $\mathbb{R}$ with inner product denoted by $(,)$. $\Phi \subset V$ is a root system if the following hold.

(R1) Φ is a finite set that spans V and does not contain 0.

(R2) If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and Φ contains no other multiples of α.

(R3) If $\alpha \in \Phi$ then the reflection in the hyperplane orthogonal to α leaves Φ invariant.

(R4) If $\alpha, \beta \in \Phi$ then $2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.

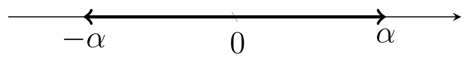
For such a set Φ we call its elements roots.

Remark 2.6. Sometimes when root systems are define, (R2) is omitted and what we defined above is called a reduced root system [2, p.42]. Similarly sometimes (R4) is omitted and a root system with axiom (R4) is called a crystallographic root system [5, p.39]. For our purposes we will simply call anything satisfying the above definition a root system.

Remark 2.7. (R4) is the formalised version of the final property we saw in the initial example. This was the property that the orthogonal projection of a root α onto a line though a different root β will be a half integer multiple of β.

From now on we will assume that V is a real inner product space and $\Phi \subset V$ is a root system. So using these axioms let us construct some root systems.

Example 2.8. If V is 1-dimensional then by (R2) if α is a root then so is $-\alpha$ and Φ contains no other multiples of α. Since V is 1-dimensional it is just the real number line so any other vectors in V will be some multiple of α and so cannot be roots hence the only possible 1-dimensional root system is the root system only containing the roots α and $-\alpha$ which is called A_1 [2, p.43].



A_1

Now let us find the 2-dimensional examples.
Recall that for $\alpha \in V$, $|\alpha| = (\alpha, \alpha)^{\frac{1}{2}}$ defines a norm on V and for $\alpha, \beta \in V$ we have the formula $(\beta, \alpha) = |\beta||\alpha| \cos\theta$, where θ is the angle between β and α. Hence

$$\langle \beta, \alpha \rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = \frac{2|\beta||\alpha| \cos\theta}{|\alpha||\alpha| \cos(0)} = 2\frac{|\beta|}{|\alpha|} \cos\theta,$$

and so

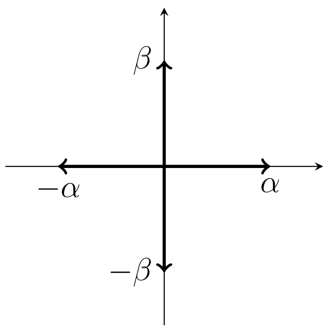
$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 2\frac{|\alpha|}{|\beta|} \cos\theta \times 2\frac{|\beta|}{|\alpha|} \cos\theta = 4\cos^2\theta.$$

By (R4), $\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle \in \mathbb{Z}$ and since $0 \leq 4\cos^2\theta \leq 4$, α and β must have the same sign and their product must be at most 4.
So let us assume that $\alpha \neq \pm\beta$ and $|\beta| \geq |\alpha|$. Then we can consider the pair of integers satisfying these conditions and use the formulas $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2\theta$ and $\langle \beta, \alpha \rangle = 2\frac{|\beta|}{|\alpha|} \cos\theta$ to find θ and $\frac{|\beta|}{|\alpha|}$.
This gives the following table [2, p.45].

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ	$\frac{\lvert\beta\rvert}{\lvert\alpha\rvert}$	Example
0	0	$\frac{\pi}{2}$	undetermined	Example 2.9
1	1	$\frac{\pi}{3}$	1	Example 2.10
-1	-1	$\frac{2\pi}{3}$	1	Example 2.10
1	2	$\frac{\pi}{4}$	$\sqrt{2}$	Example 2.11
-1	-2	$\frac{3\pi}{4}$	$\sqrt{2}$	Example 2.11
1	3	$\frac{\pi}{6}$	$\sqrt{3}$	Example 2.12
-1	-3	$\frac{5\pi}{6}$	$\sqrt{3}$	Example 2.12
2	2	0	1	None
-2	-2	0	1	None

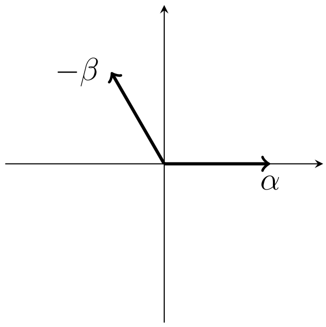
Note that the case where $\langle \alpha, \beta \rangle = \pm 2$ implies that $\beta = \pm\alpha$ meaning that this cannot be a 2-dimensional root system since if it was a root system, it would only contain α and $-\alpha$ which cannot span $\mathbb{R}^2$ and so does not satisfy (R1). Hence we may discard these two cases.

Example 2.9. [2, p.44] In the case where $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle = 0$ we can scale the lengths however we want. So if we let $|\beta| = |\alpha|$ then since $\theta = \frac{\pi}{2}$ we have $\beta \in P_\alpha$ and $\alpha \in P_\beta$ so the only other vectors we get when reflecting in P_α and P_β are $-\alpha$ and $-\beta$.
Hence we have 2 copies of A_1 so the root system we get is $A_1 \times A_1$.

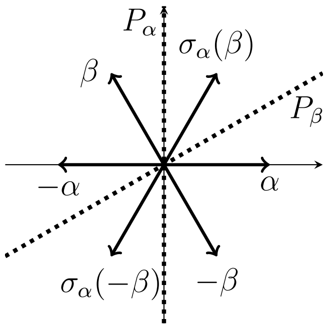


$A_1 \times A_1$

Example 2.10. [2, p.44] In the case where $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle = -1$ we have that β and α are the same length. So if we plot α on the x-axis then the angle between α and β needs to be $\frac{2\pi}{3}$ so we get the following.



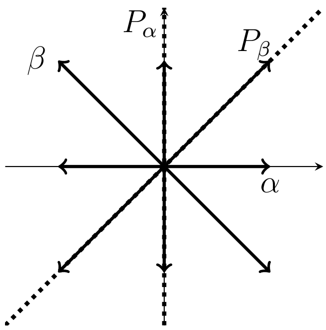
We can then find the other roots by looking at the reflections in P_α and P_β of each root. This then gives our initial example of A_2.



A_2

We will also get A_2 for the case where $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle = 1$

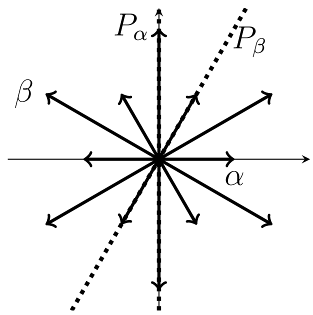
Example 2.11. [2, p.44] Using the same method as above for $\langle \alpha, \beta \rangle = -1$ and $\langle \beta, \alpha \rangle = -2$ we get the root system called B_2.



B_2

We also get B_2 when $\langle \alpha, \beta \rangle = 1$ and $\langle \beta, \alpha \rangle = 2$.

Example 2.12. [2, p.44] Then for $\langle \alpha, \beta \rangle = -1$ and $\langle \beta, \alpha \rangle = -3$ we get the root system called G2.



G_2

We also get G_2 when $\langle \alpha, \beta \rangle = 1$ and $\langle \beta, \alpha \rangle = 3$.

2.2 Simple Systems

When we constructed the 2-dimensional examples, we only needed the roots α and β in order to determine the rest of the root system as we can find the other roots by considering the reflections in P_α and P_β of $\pm\alpha$ and $\pm\beta$.
If we recall the formula for the reflection in P_α of β i.e.

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha,$$

then since $\langle \beta, \alpha \rangle \in \mathbb{Z}$, we can see that this means that in each of our examples every root can be written as a $\mathbb{Z}$-linear combination of α and β. This then motivates us to make the following definition.

Definition 2.13. [2, p.47][5, p.8] Let Φ be a root system in a real inner product space V. Then $\Delta \subset \Phi$ is a simple system with its elements called simple roots if

(S1) Δ is a basis for V.

(S2) Each $\alpha \in \Phi$ is a $\mathbb{Z}$-linear combination of Δ such that each coefficient has the same sign.

Remark 2.14. Sometimes a simple system is called a base. [2, p.47]

Example 2.15. In A_2 we have

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta + \alpha,$$
$$\sigma_\alpha(-\beta) = -\beta + \langle \beta, \alpha \rangle \alpha = -\beta - \alpha,$$

so the roots are $\alpha, \beta, -\alpha, -\beta, \alpha + \beta, -\alpha - \beta$. Hence a simple system for A_2 is α, β.

Now the natural question to ask is for any root system do we have a simple system? This would be very useful if we do as then we only need to consider a subset of the roots to determine properties for the whole root system. While currently we do not know if simple systems always exist, something we do know always exists are positive systems.

Definition 2.16. [5, p.8] Choose an ordered basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V and let $\beta \in \Phi$ be a root such that $\beta = \sum_{i=0}^{n} \lambda_i v_i$ for $\lambda_i \in \mathbb{R}$. Then we call β positive with respect to $\mathcal{B}$ if $0 < \lambda_k$, where k is the smallest index i such that $\lambda_i \neq 0$. If $\lambda_k < 0$ then we call β negative with respect to $\mathcal{B}$

Remark 2.17. Since V is finite we may assume that $\dim(V) = n$ for $n \in \mathbb{N}$ and since by (R1) $\beta \neq 0$ for all $\beta \in \Phi$, a smallest index k such that $\lambda_k \neq 0$ always exists.

We then denote the collection of positive roots with respect to some basis by Φ^+ and call Φ^+ a positive system and similarly we let Φ^- denote the collection of negative roots and call it a negative system. It is clear that for any root system we have a positive system since we always have a basis of V.

Proposition 2.18. *[5, p.8] Let Φ be a root system with positive and negative systems Φ^+ and Φ^- with respect to some basis. Then*

(i) $\Phi^+ \cap \Phi^- = \emptyset$ *and* $\Phi = \Phi^+ \cup \Phi^-$ *i.e.* Φ^+ *and* Φ^- *form a disjoint union of* Φ.

(ii) $|\Phi^+| = |\Phi^-|$. *Moreover* $\Phi^- = -\Phi^+ := \{-\alpha : \alpha \in \Phi^+\}$

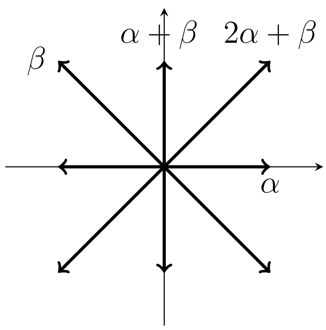
Proof. (i) Clearly the zero vector is the only vector that is both positive and negative but by (R1), Φ does not contain the zero vector so $\Phi^+ \cap \Phi^- = \emptyset$. Moreover, since Φ does not contain the zero vector, every element of Φ must be positive or negative and belong to at most one of the sets Φ^+ and Φ^-. This means that we have the disjoint union $\Phi = \Phi^+ \cup \Phi^-$.

(ii) By (R2) roots come in pairs $\{\alpha, -\alpha\}$ and if $\alpha \in \Phi^+$ then $-\alpha \in \Phi^-$ so the result follows. □

Proposition 2.19. *If Δ is a simple system in Φ then there exists a unique positive system containing Δ.*

Proof. We refer to the proof from [5, p.8]. □

Example 2.20. A simple system in B_2 is $\Delta = \{\alpha, \beta\}$ so then we get the following positive system.



B_2

Since we are on the topic of positive systems, let us take a short digression from showing the existence of simple systems to how reflections act on positive systems. If $\alpha \in \Phi^+$ then we have $\sigma_\alpha \alpha = -\alpha \in \Phi^-$ so positive systems are not invariant under reflections but it turns out that if we remove α from the positive system then this set is invariant under the reflection σ_α.

Proposition 2.21. *[5, p.10] Let Δ be a simple system of Φ contained in a positive system Φ^+. If $\alpha \in \Delta$, then $\sigma_\alpha(\Phi^+ \backslash \{\alpha\}) = \Phi^+ \backslash \{\alpha\}$.*

Proof. We give will give a sketch of this proof, a full proof can be found here [5, p.10].
Firstly, if we let $\alpha \in \Delta$, then the premise of the proof is to take an element $\beta \in \Phi^+ \backslash \{\alpha\}$ and show that σ_α maps it into $\Phi^+ \backslash \{\alpha\}$.

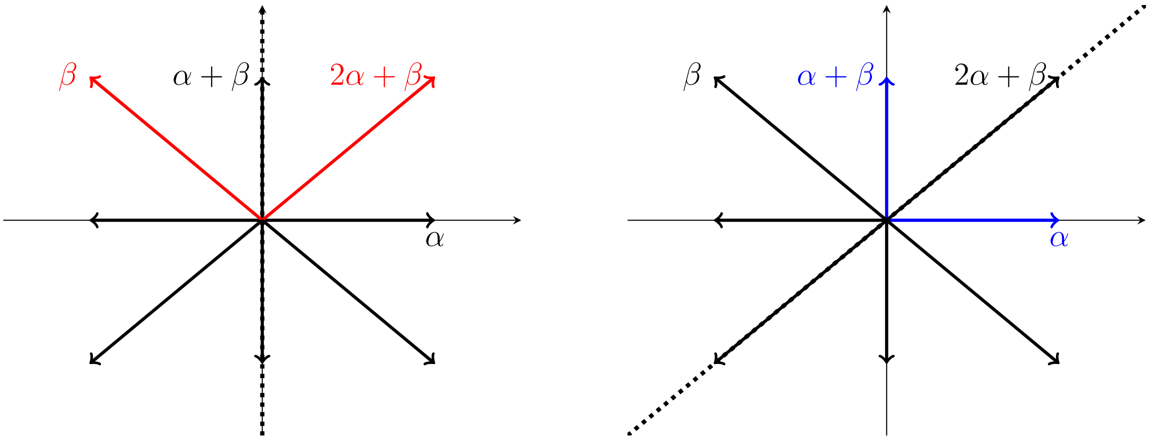
We can do this by first writing β as a linear combination of simple roots i.e. $\beta = \sum_{\gamma \in \Delta} \lambda_\gamma \gamma$. Since $\beta \neq \pm\alpha$, there exists $\delta \in \Delta \backslash \{\alpha\}$ such that $\lambda_\delta > 0$.
We can then do the following calculation giving an expression of $\sigma_\alpha \beta$ as a linear combination of Δ (note that this calculation is not in [5, p.10]).

$$\sigma_\alpha \beta = \beta - \langle \beta, \alpha \rangle \alpha = \sum_{\gamma \in \Delta} \lambda_\gamma \gamma - \langle \beta, \alpha \rangle \alpha = (\lambda_\alpha - \langle \beta, \alpha \rangle)\alpha + \lambda_\delta \delta + \sum_{\gamma \in \Delta \backslash \{\alpha, \delta\}} \lambda_\gamma \gamma.$$

$\lambda_\delta > 0$ implies that $\sigma_\alpha \beta$ is positive and it can be shown by contradiction that $\sigma_\alpha \beta \neq \alpha$ so this proves the claim. □

Another way to characterise this proposition is that if σ_α maps a positive root β to a negative root then we must have $\beta = \alpha$.

Example 2.22. Returning back to B_2, in the pictures below we have highlighted the positive roots and the lines of reflections for σ_α and σ_β where α and β are simple roots. We can see that these reflections follow Proposition 2.21.



B_2

We know that every simple system determines a unique positive system so to show existence of simple systems we want to show the other direction holds i.e. every positive system determines a unique simple system. So given a positive system Φ^+, how would we go about construct a simple system? A naive way to construct one would be just to take a subset Γ of Φ of minimal size such that each root in Φ^+ can be expressed as a non-negative linear combination of Γ. It is clear that such a subset exists and it turns out that Γ is in fact a simple system but before we can show that we need the following result.

Lemma 2.23. *[5, p.9] Let Γ be as stated above and let $\alpha, \beta \in \Gamma$ such that $\alpha \neq \beta$. Then $(\alpha, \beta) \leq 0$.*

Proof. We refer to the proof from [5, p.9].

□

Since simple systems clearly satisfies the condition of Γ we also get the following corollary.

Corollary 2.24. *Let Δ be a simple system in Φ and let $\alpha, \beta \in \Delta$ such that $\alpha \neq \beta$. Then $(\alpha, \beta) \leq 0$.*

We now have all the tools we need to show that simple systems exist.

Theorem 2.25. *Every positive system in Φ contains a unique simple system.*

Proof. We will give a sketch of a proof, a full proof can be found here [5, p.8].
The proof involves letting Δ be a subset of Φ of minimal size such that each root in Φ^+ can be expressed as a non-negative linear combination of Δ. The

only thing we then need to do to show that Δ is a simple system is to show it is linearly independent. If we assume that Δ is linearly dependent then we can write $\sum_{\beta\in\Delta_1}\lambda_\beta\beta-\sum_{\gamma\in\Delta_2}\mu_\gamma\gamma=0$ where Δ_1 and Δ_2 are disjoint subsets of Δ and λ_β and μ_γ are strictly positive. Positive definiteness and Lemma 2.23 implies that $0\leq(\sum_{\beta\in\Delta_1}\lambda_\beta\beta,\sum_{\gamma\in\Delta_2}\mu_\gamma\gamma)\leq 0$ so $\sum_{\beta\in\Delta_1}\lambda_\beta\beta=0$ which is a contradiction.

□

So from now on we will assume that Φ has a simple system Δ and positive system Φ^+ with respect to Δ.

2.3 The Weyl Group

One of the important properties of root systems is the reflectional symmetry which can be expressed as a group called the Weyl group.

Definition 2.26. [2, p.43] The Weyl group of Φ denoted by W is the subgroup of $GL_n(\mathbb{R})$ generated by the reflections σ_α for $\alpha\in\Phi$.

This group will help us prove certain properties of root systems and also connects root systems to other areas of mathematics such as group theory.

Proposition 2.27. *[5, p.6] If $w\in W$ and $\alpha\in\Phi$ then $w\sigma_\alpha w^{-1}=\sigma_{w\alpha}$.*

Proof. We have

$$w\sigma_\alpha w^{-1}(w\alpha)=w\sigma_\alpha\alpha=w(-\alpha)=-w\alpha,$$

so $w\sigma_\alpha w^{-1}$ sends wα to its inverse so we only need to show that $w\sigma_\alpha w^{-1}$ fixes the hyperplane $P_{w\alpha}$.
Now let $\lambda\in P_\alpha$ meaning that σ_α fixes λ. By Proposition 2.2, we have

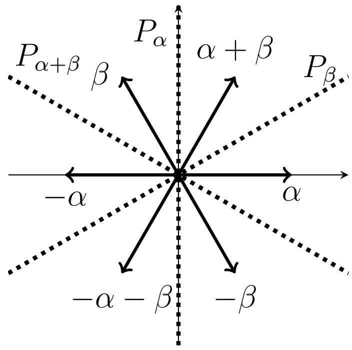
$$(w\lambda,w\alpha)=(\lambda,\alpha)=0,$$

so $w\lambda$ lies in $P_{w\alpha}$ if and only if λ lies in P_α.
Hence

$$w\sigma_\alpha w^{-1}(w\lambda)=w\sigma_\alpha\lambda=w\lambda$$

so $w\sigma_\alpha w^{-1}$ fixes $P_{w\alpha}$ hence $w\sigma_\alpha w^{-1}=\sigma_{w\alpha}$. □

Example 2.28. Let us look at A_2 with the roots and hyperplanes labelled.



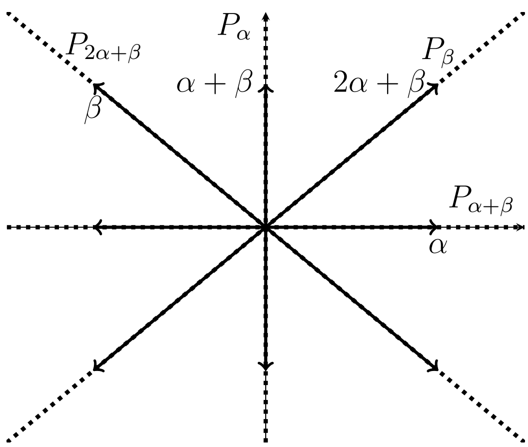
A_2

Note that the hyperplane orthogonal to $\gamma \in \Phi^+$ is the same as the hyperplane orthogonal to $-\gamma \in \Phi^-$ so $\sigma_\gamma = \sigma_{-\gamma}$ hence W is equal to the group generated by the reflections $\sigma_\alpha, \sigma_\beta$ and $\sigma_{\alpha+\beta}$.
Then according to the above Proposition we should have
$\sigma_\alpha \sigma_\beta \sigma_\alpha = \sigma_{\sigma_\alpha(\beta)} = \sigma_{\alpha+\beta}$.
So let us consider the following.

$$\sigma_\alpha \sigma_\beta \sigma_\alpha(\alpha+\beta) = \sigma_\alpha \sigma_\beta(\beta) = \sigma_\alpha(-\beta) = -\alpha - \beta,$$
$$\sigma_\alpha \sigma_\beta \sigma_\alpha(\beta) = \sigma_\alpha \sigma_\beta(\alpha+\beta) = \sigma_\alpha(\alpha+\beta) = \beta.$$

So $\sigma_\alpha \sigma_\beta \sigma_\alpha$ sends $\alpha+\beta$ to its negative and fixes the roots orthogonal to $\alpha+\beta$ which is exactly what we expect according to the Proposition.

Example 2.29. Let us look at B_2.



B_2

We have that W is generated by the reflections $\sigma_\alpha, \sigma_\beta, \sigma_{\alpha+\beta}$ and $\sigma_{2\alpha+\beta}$.

But notice that

$$
\begin{aligned}
\sigma_{\alpha+\beta}(\alpha) &= \alpha = \sigma_\alpha \sigma_\alpha(\alpha), \\
\sigma_{\alpha+\beta}(\beta) &= -2\alpha - \beta = \sigma_\alpha \sigma_\beta(\beta), \\
\sigma_{\alpha+\beta}(\alpha+\beta) &= -\alpha - \beta = \sigma_\beta \sigma_\alpha \sigma_\beta(\alpha+\beta), \\
\sigma_{\alpha+\beta}(2\alpha+\beta) &= -\beta = \sigma_\beta \sigma_\alpha(2\alpha+\beta).
\end{aligned}
$$

Similarly we can write the reflection $\sigma_{2\alpha+\beta}$ in terms of σ_α and σ_β. Hence σ_α and σ_β generates all the reflections of B_2 and so they generate W.

Since the roots in B_2 outline a square let us consider the dihedral group of order 8 i.e. the symmetry group of a square denoted by D_8. The group can be thought as being generated by a rotation $\frac{\pi}{2}$ anticlockwise about the origin denoted by a and a reflection in a line of symmetry of the square denoted b. Then in terms of reflections we can write these generators as $a = \sigma_\alpha \sigma_\beta$ and $b = \sigma_\alpha$. So $W = \langle \sigma_\alpha, \sigma_\beta \rangle = \langle \sigma_\alpha, \sigma_\alpha \sigma_\beta \rangle = \langle a, b \rangle = D_8$ hence $W = D_8$ as we would expect.

Motivated by the previous example, our next goal is to show that in general the Weyl group of Φ can be generated by σ_α for $\alpha \in \Delta$. Phrased another way, if we let W' be the subgroup of W generated by σ_α for $\alpha \in \Delta$ then we want to show that $W' = W$. The following definition and lemma will help us show this.

Definition 2.30. Let $\beta \in \Phi$ and write $\beta = \sum_{\alpha \in \Delta} \lambda_\alpha \alpha$. The height of β is $ht\beta = \sum_{\alpha \in \Delta} \lambda_\alpha$.

Lemma 2.31. *Let $\beta \in \Phi^+$ and choose $\gamma \in W'\beta \cap \Phi^+$ such that γ has the smallest height in $W'\beta \cap \Phi^+$. Then $\gamma \in \Delta$.*

Proof. We refer to the proof from [5, p.11]. □

Remark 2.32. Note that $W'\beta \cap \Phi^+$ is non-empty as W' contains the identity element since it is a subgroup of W so $\beta \in W'\beta \cap \Phi^+$. This means it makes sense for us to choose an element in that set.

Theorem 2.33. *[5, p.11] W is generated by the reflections σ_α for $\alpha \in \Delta$, i.e. $W' = W$.*

Proof. Let $\beta \in \Phi^+$. Lemma 2.31 implies that there exists $w \in W'$ such that $w\beta = \alpha$ for some $\alpha \in \Delta$. Hence $\beta = w^{-1}\alpha \in W'\Delta$ so $\Phi^+ \subset W'\Delta$.

Now let $\beta \in \Phi^-$. Then similarly there exists $w \in W'$ such that $-w\beta = \alpha$ for some $\alpha \in \Delta$. Then

$$\beta = -w^{-1}\alpha = w^{-1}(-\alpha) = w^{-1}\sigma_\alpha\alpha \in W'\Delta.$$

Hence $\Phi^- \subset W'\Delta$ meaning that $\Phi \subset W'\Delta$.

Now let $\sigma_\beta \in W$. Since $\beta \in \Phi \subset W'\Delta$, we know there exists $w \in W'$ and $\alpha \in \Delta$ such that $\beta = w\alpha$. Then by Proposition 2.27 we have

$$\sigma_\beta = \sigma_{w\alpha} = w\sigma_\alpha w \in W'$$

hence $W \subset W'$ so $W = W'$. $\square$

This proof gives us the following corollary.

Corollary 2.34. *For every $\beta \in \Phi$ there exists $w \in W$ such that $w\beta \in \Delta$.*

2.4 Coxeter Graphs and Dynkin Diagrams

Our next goal is to classify all possible root systems up to isomorphism. By isomorphism we mean that root systems Φ and Φ' in respective vector spaces V and V' are isomorphic if there exists a vector space isomorphism from V to V' that maps Φ onto Φ'.

If we recall our example of $A_1 \times A_1$ then this root system can be broken down into two copies of A_1 whilst it seems that there are no clear ways to decompose A_1, A_2, B_2 or G_2 into other root systems. So it seems there are certain root systems that cannot be broken down further and can be used to build other root systems. So we can just classify the root systems that cannot be broken down further which will then allow us to construct all root systems by combining them. This then motivates our next definition.

Definition 2.35. [2, p.52] A root system Φ is irreducible if Φ cannot be partitioned into the union of two proper subsets $\Phi = \Phi_1 \cup \Phi_2$ such that $(\alpha, \beta) = 0$ for all $\alpha \in \Phi_1$ and $\beta \in \Phi_2$.

Now a problem that we are going to run into if we want to classify all irreducible root systems is that we can't visualise root systems with dimension greater than three so in order to represent these higher dimensional root system we will simplify them.

For a root system Φ, if we know the number of simple roots in Φ along with their lengths and the angle between them, then we know what the Weyl group is since it is generated by σ_α for $\alpha \in \Delta$ and so we can find the other roots by applying the Weyl group to the simple roots. It turns out that if we encode this information into a graph then we will be able to determine the entire root system up to isomorphism.

Definition 2.36. [2, p.56] Let Δ be a simple system of a root system Φ with $|\Delta| = n$. Let us order the elements of Δ such that $\Delta = \{\alpha_1, \ldots, \alpha_n\}$. Then the Coxeter graph of Φ with respect to Δ is the graph containing n vertices each labelled from 1 to n with the ith vertex joined to the jth vertex by $\langle\alpha_i, \alpha_j\rangle\langle\alpha_j, \alpha_i\rangle$ edges for $i \neq j$.

From our earlier table we know that if α and β are distinct positive roots then $\langle\alpha, \beta\rangle\langle\beta, \alpha\rangle = 0, 1, 2$ or 3. So using the formula $\langle\alpha, \beta\rangle\langle\beta, \alpha\rangle = 4\cos^2\theta$ we can find the corresponding angles between α and β.

Theorem 2.37. *Let $\Phi \subset V$ and $\Phi' \subset V'$ be two root systems with the same Coxeter graph. Then Φ and Φ' are isomorphic.*

Proof. We will give a sketch of a proof, a full proof can be found here [2, p.55]. Note that the proof is actually a proof showing that if two root systems have the same Cartan matrix then they are isomorphic. The Cartan matrix of a root system is defined on the same page and it should be clear that this is equivalent to proving the above statement.

For a sketch of the proof, since Φ and Φ' have the same Coxeter graphs, they must have the same number of simple roots so let $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ and $\Delta' = \{\alpha'_1, \ldots, \alpha'_n\}$ be the ordered simple systems of Φ and Φ' respectively. Then we can define the vector space isomorphism $\phi : V \to V'$ sending α_i to α'_i. From this, we can define an isomorphism from the Weyl group of Φ to the Weyl group of Φ' which can then be used to show that ϕ maps Φ onto Φ'.

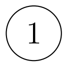
□

Hence instead of saying "the Coxeter graph of Φ with respect to Δ" we may simply say "the Coxeter graph of Φ" as the Coxeter graph is independent of the choice of Δ.

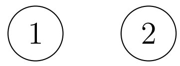
Let us draw the Coxeter graphs of the root systems we have seen so far.

Example 2.38. [2, p.56]

(i) A_1 has only one simple root so the Coxeter graph is just a single vertex.

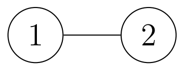


(ii) $A_1 \times A_1$ has two simple roots which are $\frac{\pi}{2}$ radians apart hence
$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2(\frac{\pi}{2}) = 0$.
So the Coxeter graph is two disconnected vertices.

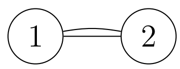


Which is two copies of A_1 as we would expect.

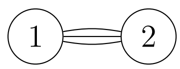
(iii) A_2 has two simple roots which are $\frac{2\pi}{3}$ radians apart hence
$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2(\frac{2\pi}{3}) = 1$.
So the Coxeter graph is two vertices connected by 1 edge.



(iv) B_2 has two simple roots which are $\frac{3\pi}{4}$ radians apart hence
$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2(\frac{3\pi}{4}) = 2$.
So the Coxeter graph is two vertices connected by 2 edges.



(v) G_2 has two simple roots which are $\frac{5\pi}{6}$ radians apart hence
$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2(\frac{5\pi}{6}) = 3$.
So the Coxeter graph is two vertices connected by 3 edges.



Proposition 2.39. *Φ is irreducible if and only if its Coxeter graph is connected.*

Proof. This follows straight from the definition since the Coxeter graph of Φ is connected if and only if Φ cannot be split into proper subsets $\Phi = \Phi_1 \cup \Phi_2$ such that for all $\alpha \in \Phi_1$ and for all $\beta \in \Phi_2$ we have $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 0$ implying

$$\langle \alpha, \beta \rangle = 2\frac{(\alpha, \beta)}{(\beta, \beta)} = 0, \text{ or } \langle \beta\, \alpha \rangle = 2\frac{(\beta, \alpha)}{(\alpha, \alpha)} = 0,$$

so $(\alpha, \beta) = 0$ meaning Φ is not irreducible. $\square$

The only issue is that Coxeter graphs do not tell us whether a simple root is longer or shorter than another simple root. But this can be easily fixed with Dynkin diagrams.

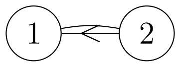
Definition 2.40. [2, p.57] The Dynkin diagram of Φ is the Coxeter graph of Φ with an arrow pointing to the shorter of the two roots whenever a double or triple edge occurs.

Remark 2.41. If two roots are connected by a single edge in the Coxeter graph of Φ then they have the same length. We can see this since a single edge between $\alpha, \beta \in \Phi$ implies $\langle \beta, \alpha \rangle = \pm 1$ and $\theta = \frac{\pi}{3}$ or $\frac{2\pi}{3}$ so then by using the previous formula $\langle \beta, \alpha \rangle = 2\frac{|\beta|}{|\alpha|} \cos\theta$ we get

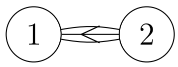
$$\left| \frac{\beta}{\alpha} \right| = \left| \frac{\langle \beta, \alpha \rangle}{2\cos\theta} \right| = \left| \pm \frac{1}{2 \times \frac{1}{2}} \right| = 1.$$

This is why we only need arrows on double or triple edges.

Example 2.42. The Dynkin diagrams for A_1, $A_1 \times A_1$ and A_2 are the same as their Coxeter graphs.
For the Dynkin diagram of B_2 since we only have two vertices it does not matter which way the arrow points so the Dynkin diagram is the following.



Similarly the Dynkin diagram for G_2 is the following



2.5 Classification

In this section we will finish our goal from the previous chapter in classifying all irreducible root systems up to isomorphism.
To allow more flexibility in proving this, instead of just working with roots we will work something called an admissable set.

Definition 2.43. [2, p.60] Let V be a finite real inner product space. A set of vectors $\mathfrak{u} = \{\epsilon_1, \dots \epsilon_n\}$ in V is called admissible if $\mathfrak{u}$ is a set of n linearly independent unit vector such that $(\epsilon_i, \epsilon_j) \leq 0$ and $4(\epsilon_i, \epsilon_j)^2 = 0, 1, 2$, or 3 for $i \neq j$.
For such a set $\mathfrak{u}$ we will attach a graph Γ with n vertices labelled from 1 to n with vertices i and j connected by $4(e_i, e_j)^2$ edges.

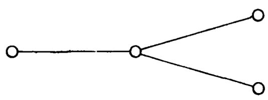
Example 2.44. [2, p.60] For any root system, if we take the set of simple roots divided by their lengths then this forms an admissable set. This shows how we can easily associate an admissable set to a root system.

Some important examples of graphs that come from admissable sets are the following

Example 2.45. [2, p.60] A simple chain is a graph of the following form.

Picture source [2, p.60].

Example 2.46. A branch point or a node is a graph of the following form.



Picture source [2, p.62].

To help us classify all Coxeter graphs let us give some results on admissable sets.

Lemma 2.47. *Let $\mathfrak{u} = \{\epsilon_1, \dots, \epsilon_n\}$ be an admissable set with an associated graph Γ.*

(i) If $\mathfrak{u}'$ is a non-empty subset of $\mathfrak{u}$ then $\mathfrak{u}'$ is also an admissable set whose graph can be obtained by omitting the corresponding vertices and incident edges.

(ii) $0 < n + 2\sum_{i<j}^{n}(\epsilon_i, \epsilon_j)$.

(iii) The number of pairs of vertices connected by at least one edge is strictly less than n.

Proof. (i) follows straight from the definition of an admissable set. The following proofs for (ii) and (iii) are based on [2, p.60] but with more details given to the calculations.

Let $\epsilon = \sum_{i=1}^{n} \epsilon_i$.
Since each ϵ_i are unit vectors and linearly independent with the fact that the inner product is positive-definite and symmetric we get that

$$\begin{aligned}
0 < (\epsilon, \epsilon) &= (\sum_{i=1}^{n} \epsilon_i, \sum_{i=1}^{n} \epsilon_i), \\
&= \sum_{i=1}^{n}(\epsilon_i, \epsilon_i) + \sum_{i<j}^{n}(\epsilon_i, \epsilon_j) + \sum_{i>j}^{n}(\epsilon_i, \epsilon_j) \\
&= n + \sum_{i<j}^{n}(\epsilon_i, \epsilon_j) + \sum_{i<j}^{n}(\epsilon_j, \epsilon_i) \\
&= n + \sum_{i<j}^{n}(\epsilon_i, \epsilon_j) + \sum_{i<j}^{n}(\epsilon_i, \epsilon_j) \\
&= n + 2\sum_{i<j}^{n}(\epsilon_i, \epsilon_j).
\end{aligned}$$

proving the claim.

For (iii) let the ith and jth vertices be connected by at least one edge where $i \neq j$.
This means that $(\epsilon_i, \epsilon_j) \neq 0$ so since $\mathfrak{u}$ is an admissable set and $i \neq j$ we have that $4(\epsilon_i, \epsilon_j)^2 = 1, 2,$ or 3.
Then since $(\epsilon_i, \epsilon_j) \leq 0$ for $i \neq j$ we have

$$\begin{aligned}
4(\epsilon_i, \epsilon_j)^2 \geq 1 &\implies (\epsilon_i, \epsilon_j)^2 \geq \frac{1}{4} \\
&\implies (\epsilon_i, \epsilon_j) \geq -\frac{1}{2} \\
&\implies 2(\epsilon_i, \epsilon_j) \leq -1.
\end{aligned}$$

So if there existed at least n pairs of vertices connected by at least one edge we would have $2\sum_{i<j}^{n}(\epsilon_i,\epsilon_j) \leq -n$, contradicting the inequality from (ii). This contradiction then proves the claim.

$\square$

Corollary 2.48. *[2, p.60] Let Γ be an associated graph of an admissable set $\mathfrak{u}$. Then Γ contains no cycles.*

Proof. A cycle of Γ would be a graph Γ' associated to a non-empty subset $\mathfrak{u}'$ of $\mathfrak{u}$ which by part (i) of Lemma 2.47 is also an admissable set.

Assume the Γ' is such a cycle with the corresponding admissable set $\mathfrak{u}'$ having m elements. Since Γ' is a cycle, we can relabel the vertices such that each pair of vertices $(i, i+1)$ along with $(m, 1)$ are connected by at least one edge. This means there are at least m pairs of vertices in Γ' that are connected by at least one edge contradicting part (iii) of Lemma 2.47. $\square$

Lemma 2.49. *Let Γ be an associated graph of an admissable set $\mathfrak{u}$. Then no more than three edges can occur at any given vertex of Γ.*

Proof. We will give a sketch of a proof, a full proof can be found here [2, p.60].

The proof starts by supposing there exists $\epsilon, \eta_1, \ldots, \eta_k \in \mathfrak{u}$, all distinct, such that $\eta_1, \ldots, \eta_k$ are connected to ϵ by $1, 2$ or 3 edges. Then using the linear independence of $\mathfrak{u}$ and the fact that these are unit vectors we can get the inequality $\sum_{i=1}^{k} 4(\epsilon,\eta_i)^2 < 4$. Since $4(\epsilon,\eta_i)^2$ is the number of edges joining ϵ and η_i in Γ, this proves the claim. $\square$

Proposition 2.50. *[2, p.60] Let Γ be graph associated with an admissable set $\mathfrak{u} = \{\epsilon_1, \ldots, \epsilon_n\}$ with $S = \{\epsilon_1, \ldots, \epsilon_k\} \subset \mathfrak{u}$ forming a simple chain in Γ. Then if we define $\epsilon = \sum_{i=1}^{k} \epsilon_i$ we have that the set $\mathfrak{u}' = (\mathfrak{u} \setminus S) \cup \{\epsilon\}$ is an admissable set.*

Remark 2.51. The graph of $\mathfrak{u}'$ can be obtained by shrinking the simple chain in Γ to a single point.

Proof. From Lemma 2.47(i), we know that $\mathfrak{u} \setminus S$ is an admissable set. Hence we just need to check the properties of an admissable set still hold true when considering the exceptional cases of $\mathfrak{u}'$ involving ϵ.

Since $\mathfrak{u}$ is linearly independent we have that

$$\begin{aligned} 0 &= \lambda_0\epsilon + \lambda_1\epsilon_{k+1} + \cdots + \lambda_n\epsilon_n \\ &= \lambda_0\epsilon_1 + \cdots + \lambda_0\epsilon_k + \lambda_1\epsilon_{k+1} + \cdots + \lambda_n\epsilon_n, \end{aligned}$$

implies that $\lambda_i = 0$ for all i hence $\mathfrak{u}'$ is linearly independent.

For $1 \leq i \leq k-1$ we have that vertices i and $i+1$ are connected by one edge. So we have $4(\epsilon_i, \epsilon_{i+1})^2 = 1$ which implies that $2(\epsilon_i, \epsilon_{i+1}) = -1$.
We also know that if $2 \leq h-j \leq k-1$ then vertices j and h are disconnected (as otherwise we would have a cycle) meaning $(\epsilon_j, \epsilon_h) = 0$.
Using the formula $(\epsilon, \epsilon) = k + 2\sum_{1\leq i<j\leq k}(\epsilon_i, \epsilon_j)$ from Lemma 2.47(iii) we have

$$\begin{aligned}(\epsilon, \epsilon) &= k + 2\sum_{1\leq i<j\leq k}(\epsilon_i, \epsilon_j)\\ &= k + \sum_{i=1}^{k-1} 2(\epsilon_i, \epsilon_{i+1}) + \sum_{2\leq h-j\leq k-1}(\epsilon_j, \epsilon_h)\\ &= k - (k-1) = 1,\end{aligned}$$

so ϵ is a unit vector.

Let $\epsilon_u \in \mathfrak{u} \setminus S$. Vertex u can be connected to at most one of the the vertices $1, \ldots, k$ as otherwise we would have a cycle in Γ contradicting Corollary 2.48. Hence either $(\epsilon_u, \epsilon) = 0$ or $(\epsilon_u, \epsilon) = (\epsilon_u, \epsilon_i)$ for $1 \leq i \leq k$.
This means that $4(\epsilon_u, \epsilon)^2 = 0$ or $4(\epsilon_u, \epsilon)^2 = 4(\epsilon_u, \epsilon_i)^2 = 0, 1, 2$ or 3 since $\epsilon_u, \epsilon_i \in \mathfrak{u}$.

So $\mathfrak{u}'$ is a linearly independent set of unit vectors such that for $\epsilon_i, \epsilon_j \in \mathfrak{u}'$ we have $4(\epsilon_i, \epsilon_j)^2 = 0, 1, 2$ or 3.
Hence $\mathfrak{u}'$ is an admissable set.

$\square$

This proposition is useful for identifying graphs which cannot come from admissable set.

Example 2.52. Consider the following graphs containing simple chains.

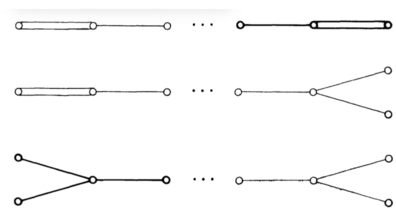
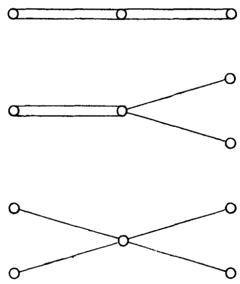


Figure 1. Picture source [2, p.61].

If we shrink the simple chains to a single point like in Proposition 2.50 we get the graphs pictured.



Picture source [2, p.61].

We can see that each of these graphs contain a point which has a vertex with four edges meaning that these graphs are not associated to admissable sets. So then from the contrapositive of the previous proposition we know that our initial graphs are also not associated to admissable sets.

Corollary 2.53. *Let Γ be a graphs associated with an admissable set. Then Γ contains none of the graphs from Figure 1.*

Now using this corollary we can give the following result which greatly limits what connected Coxeter graphs we can have.

Proposition 2.54. *[2, p.61] Let Γ be a* ***connected*** *graph associated to an admissable set.*

(i) If Γ has a triple edge then Γ is the graph G_2.

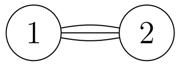
(ii) Γ can have at most one double edge.

(iii) Γ cannot contain both a double edge and a branch point.

(iv) Γ can have at most one branch point.

(v) If Γ contains only single edges and no branch points, then it is a simple chain.

Proof. (i) If we recall the Coxeter graph of G_2 then Lemma 2.49 shows that if we add any other connected vertices or edges to the graph then we get a graph no longer associated to an admissable set.



Hence if we have a triple edge then any other vertex cannot be connected to this triple edge and so must be connected. So any connected graph with a triple edge contains only that triple edge which is G_2.

(ii) Assume that Γ has two double edges. Since Γ is connected, these double edges must be connected by some subgraph of Γ.
By part (i), Γ contains no triple edges and these double edges cannot be directly connected to each other or connect by double edges since then we would have four edges coming from a single vertex contradicting Lemma 2.49. So they must be connected by a simple chain or a simple chain containing one or more branch points.
But this contradicts the Corollary 2.53 as then Γ contains at least one of the graphs in Figure 1 as a subgraph. This contradiction proves the claim.

(iii) Similarly from above, if Γ contains a double edge and a branch point they must be connected by a simple chain or a simple chain with one or more branch points which contradicts Corollary 2.53 so this cannot happen.

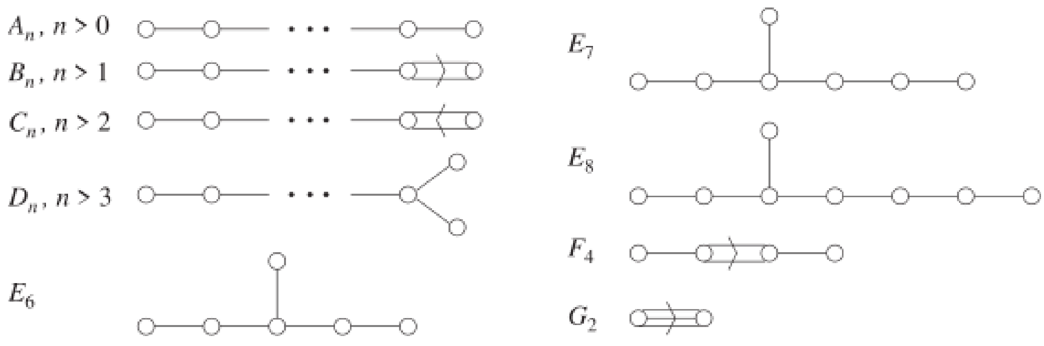
(iv) If Γ contains at least two branch points then by (iii) Γ does not have a double edge so the branch points must be connected by simple chains. But then Γ contains the third graph from Figure 1 so Γ must contain at most one branch point.

(v) Assume that Γ contains only single edges and no branch points and is not a simple chain.
Since Γ only has single edges, it is clear that it must contain a simple chain and since Γ is connected there must exist some vertex that is connected to this simple chain but is not part of it.
If this vertex is connected to a single vertex in the chain then Γ has a branch point and if this vertex is connected to multiple vertices in the chain then we have a cycle in Γ.
Hence this is a contradiction so we have proved the claim.

□

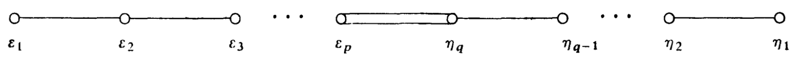
We can now prove the classification Theorem, one of the most important theorems in the study of root systems.

Theorem 2.55. *[2, p.57] If Φ is an irreducible root system with* $\dim V = |\Delta| = n$*, then its Dynkin diagram is one of the following:*



Picture source [6].

Proof. Let Γ be a graph associated to an admissable set.
If Γ has a triple edge then from the Proposition 2.54 we know that Γ is G_2.
If Γ has a double edge then we know that it does not contain a triple edge, it does not contain another double edge and it does not contain a branch point. Hence the rest of its edges must be single edges and since it has no cycles it must have the following form.



Picture source [2, p.61].

So we have $2(\varepsilon_i, \varepsilon_{i+1}) = -1 = 2(\eta_j, \eta_{j+1})$ for $1 \leq i \leq p-1$ and $1 \leq j \leq q-1$ with all other distinct pairs being orthogonal except (ε_p, η_q). Setting $\varepsilon = \sum_{i=1}^{p} i\varepsilon_i$ we get that

$$\begin{aligned}
(\varepsilon, \varepsilon) = (\sum_{i=1}^{p} i\varepsilon_i, \sum_{i=1}^{p} i\varepsilon_i) &= \sum_{i=1}^{p} i^2(\varepsilon_i, \varepsilon_i) + \sum_{i=1}^{p-1} i(i+1)(\varepsilon_i, \varepsilon_{i+1}) + \sum_{2 \leq j-i \leq p-1} ij(\varepsilon_i, \varepsilon_j) \\
&= \sum_{i=1}^{p} i^2 - \sum_{i=1}^{p-1} i(i+1) \\
&= p^2 - \sum_{i=1}^{p-1} i \\
&= p^2 - \frac{(p-1)p}{2} \\
&= \frac{p}{2}(p+1).
\end{aligned}$$

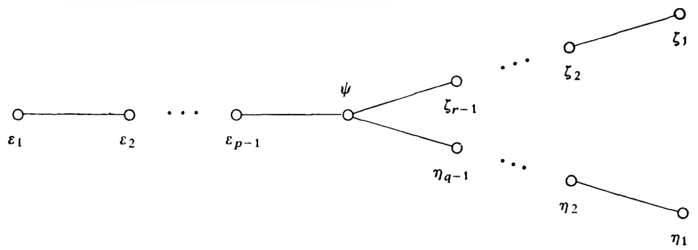
Similarly if we set $\eta = \sum_{i=1}^{q} i\eta_i$ we get $(\eta, \eta) = \frac{q}{2}(q+1)$.

Since $4(\varepsilon_p, \eta_q)^2 = 2$ and $(\varepsilon_i, \eta_j) = 0$ for $i \neq p$ and $j \neq q$ we have $(\varepsilon, \eta)^2 = p^2q^2(\varepsilon_p, \eta_q)^2 = \frac{p^2q^2}{2}$.
The Cauchy-Schwarz inequality states that

$$\begin{aligned}(\varepsilon, \eta)^2 < (\varepsilon, \varepsilon)(\eta, \eta) &\implies \frac{p^2q^2}{2} < \frac{pq}{4}(p+1)(q+1) \\ &\implies 2pq < pq + p + q + 1 \\ &\implies pq - p - q < 1 \\ &\implies (p-1)(q-1) < 2.\end{aligned}$$

Since p and q are positive integers, one solution of the inequality is $p = q = 2$ giving the graph for F_4. The only other solutions are when $p = 1$ or $q = 1$, then we can choose any value for the other variable which gives the graphs B_n or C_n.
Hence F_4, B_n and C_n are the only graphs with a double edge.

If Γ has a branch point then the rest of its edges must be single edges and it cannot contain any other branch points so it must be of the form below. Similarly from before set $\varepsilon = \sum_{i=1}^{p-1} i\varepsilon_i, \eta = \sum_{i=1}^{q-1} i\eta_i$ and $\zeta = \sum_{i=1}^{r-1} i\zeta_i$.



Picture source [2, p.61].

So $(\varepsilon, \varepsilon) = \frac{p}{2}(p-1), (\eta, \eta) = \frac{q}{2}(q-1)$ and $(\zeta, \zeta) = \frac{r}{2}(r-1)$.
Let $\theta_1, \theta_2, \theta_3$ be the respective angles between ψ and ε, η, ζ.

Then we have

$$\begin{aligned}\cos^2\theta_1 &= \frac{(\varepsilon,\psi)^2}{(\varepsilon,\varepsilon)(\psi,\psi)} \\ &= \frac{(p-1)^2(\varepsilon_{p-1},\psi)^2}{(\varepsilon,\varepsilon)\times 1} \\ &= \frac{(p-1)^2(\frac{1}{4})}{\frac{1}{2}p(p-1)} \\ &= \frac{1}{2}(1-\frac{1}{p}).\end{aligned}$$

Similarly we have $\cos^2\theta_2 = \frac{1}{2}(1-\frac{1}{q})$ and $\cos^2\theta_3 = \frac{1}{2}(1-\frac{1}{r})$.

From [2, p.60] we have the inequality $\cos^2\theta_1+\cos^2\theta_2+\cos^2\theta_3 < 1$ (this comes from the inequality $\sum_{i=1}^{k} 4(\epsilon,\eta_i)^2 < 4$ discussed in the proof of Lemma 2.49) which gives

$$\frac{1}{2}(1-\frac{1}{p}+1-\frac{1}{q}+1-\frac{1}{r}) < 1 \implies \frac{1}{p}+\frac{1}{q}+\frac{1}{r} > 1.$$

If any of p, q or r equals 1 then Γ does not have a branch point hence from changing labels we may assume that $\frac{1}{p} \leq \frac{1}{q} \leq \frac{1}{r} \leq \frac{1}{2}$.
Then we have

$$1 < \frac{1}{p}+\frac{1}{q}+\frac{1}{r} \leq \frac{3}{r} \leq \frac{3}{2},$$

hence $r = 2$.
We also have

$$\frac{1}{2} = 1-\frac{1}{r} < \frac{1}{p}+\frac{1}{q} \leq \frac{2}{q} \leq 1,$$

so $2 \leq q < 4$.
If $q = 3$ then

$$\frac{1}{6} = 1-\frac{1}{r}-\frac{1}{q} < \frac{1}{p} \leq \frac{1}{2},$$

so $p < 6$, and if $q = 2$ then

$$0 = 1-\frac{1}{r}-\frac{1}{q} < \frac{1}{p} \leq \frac{1}{2},$$

so $p \geq 2$.

Hence the possible triples (p, q, r) are $(p, 2, 2)$ corresponding to the graph D_n, $(3, 3, 2)$ corresponding to E_6, $(4, 3, 2)$ corresponding to E_7 and $(5, 3, 2)$ corresponding to E_8 meaning that the possible graphs with branch points are D_n, E_6, E_7 and E_8.

The only other graph we have not consider is a graph Γ with no triple edges, no double edges and no branch points which we know from the Proposition 2.54 implies that Γ is a simple chain which corresponds to the graphs A_n.

So we have shown that graphs $A-G$ are the only graphs that can come from an admissable set which implies that these are the only possible Coxeter graphs.
Each Coxeter graph determines the Dynkin diagram except for B_n and C_n which can be distinguished by the direction of the arrow on their double edge.
Hence the Theorem follows.

□

3 Lie Algebras

Now we have gone in detail on roots systems we will now return to Lie algebras and how root systems come from them. We start by defining and giving examples of Lie algebras and then work our way to defining semisimple Lie algebras. This takes ideas from multiple sources.

Then in sections 3.2 and 3.3, we define and give examples of Cartan subalgebras which leads us into finding root systems of specific Lie algebras. Both of these subsections are based on ideas from [3] in particular the examples in section 3.3 of finding the root systems of specific Lie algebras are based on the source's treatment of the general cases with more detail given to the calculations.

3.1 Definition

Definition 3.1. [3, p.108] Let $\mathfrak{g}$ be a vector space with a map $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the Lie bracket such that this map satisfies the following properties for $x, y, z \in \mathfrak{g}$

(L1) The map is bilinear.

(L2) The map is skew-symmetric meaning $[x, y] = -[y, x]$.

(L3) The map satisfies $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$ (this identity is called the Jacobi identity).

Then $\mathfrak{g}$ is called a Lie algebra.

For our purposes we will always be looking at complex Lie algebras consisting of matrices which means that we can just take the Lie bracket to be the commutator bracket.

Example 3.2. [7, p.61] Some examples of Lie algebras are the following.

(i) The Lie algebra $\mathfrak{gl}(n)$ denotes the set of n-by-n complex matrices.

(ii) From our initial example in chapter 1 we have the Lie algebra $\mathfrak{sl}(n)$ denoting the set of n-by-n trace-free matrices.

(iii) $\mathfrak{so}(n)$ denotes the set of n-by-n antisymmetric matrices i.e. the set of n-by-n matrices M such that $M^T=-M$ where M^T is the transpose of M.

(iv) $\mathfrak{sp}(2n)$ denotes the set of $2n$-by-$2n$ matrices M such that $\Omega M+M^T\Omega=0$ for the block matrix $\Omega=\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}$. [3, p.239]

(v) $\mathfrak{u}(n)$ denotes the set of n-by-n skew-Hermitian matrices i.e. the set of n-by-n matrices M such that $M^\dagger=-M$ where $M^\dagger$ is the conjugate transpose of M.

(vi) $\mathfrak{su}(n)$ denotes the set of n-by-n skew-Hermitian matrices whose trace equals 0. So we can define this Lie algebra as $\mathfrak{su}(n)=\mathfrak{u}(n)\cap\mathfrak{sl}(n)$.

In fact any Lie algebra is a complex inner product space. Note that this is different from a real inner product space which we defined at the start of chapter 2.

Definition 3.3. [2, p.4][2, p.21] Let $\mathfrak{g}$ be a Lie algebra and define the map $ad_x:\mathfrak{g}\times\mathfrak{g}\to\mathbb{C}$ for $x\in\mathfrak{g}$ where $ad_x(y)=[x,y]$ (This is called the adjoint form or adjoint representation of $\mathfrak{g}$).
Then we can define a map $K(,):\mathfrak{g}\times\mathfrak{g}\to\mathbb{C}$ where $K(x,y)=tr(ad_x\circ ad_y)$ such that this map is conjugate symmetric, linear in the first argument and is positive definite implying that $K(,)$ is a complex inner product.
This is called the Killing form on $\mathfrak{g}$.

Naturally we want to classify all Lie algebras up to isomorphism[1] just like we did for root systems. In chapter 1 we briefly outlined a way to get a root system from the Lie algebra $\mathfrak{sl}(3)$ and mentioned how it determines the Lie algebra up to isomorphism. In general we can always retrieve a root system from a certain type of Lie algebra called a semisimple Lie algebra, which we will define later (Definition 3.9), and this root system determines the semisimple Lie algebra up to isomorphism. Hence, since we have classified all root systems up to isomorphism we have already classified all semisimple Lie algebras up to isomorphism. So while we do not have a classification for Lie algebras in general, we do have a classification for semisimple Lie algebras.

In order to define what a semisimple Lie algebra is we need the following definitions.

Definition 3.4. [2, p.1][2, p.4] $\mathfrak{s}$ is a subalgebra of a Lie algebra $\mathfrak{g}$ if it is a subspace of $\mathfrak{g}$ such that $[x, y] \in \mathfrak{s}$ for all $x, y \in \mathfrak{s}$.
Such a subalgebra s is called abelian if $[x, y] = 0$ for all $x, y \in \mathfrak{s}$.

Remark 3.5. If we consider the vector space $\mathbb{C}$, which is also a field and hence abelian, we have for all $x, y \in \mathbb{C}$ that the commutator bracket gives

$$[x, y] = xy - yx = xy - xy = 0.$$

This is the motivation behind this definition of abelian.

Definition 3.6. [2, p.6] Let I be a subspace of a Lie algebra $\mathfrak{g}$.
I is called an ideal of $\mathfrak{g}$ if $[x, y] \in I$ for all $x \in I$ and $y \in \mathfrak{g}$.

This is analogous to ideals in rings or normal subgroups in groups and just like how we have trivial normal subgroups or trivial ring ideals we have trivial Lie algebra ideals.

Example 3.7. For any Lie algebra $\mathfrak{g}$ we clearly have that $\mathfrak{g}$ itself is an ideal, any ideal of $\mathfrak{g}$ that is not equal to $\mathfrak{g}$ itself we call a proper ideal. We also have that the subspace $\{0\}$ is an ideal as by bilinearity we have

$$[0, g] = 0[0, g] = 0$$

for all $g \in \mathfrak{g}$.
$\{0\}$ is called the trivial ideal of $\mathfrak{g}$.

[1] By isomorphism between Lie algebras we mean a vector space isomorphism that preserves the Lie bracket.

Then just like how simple groups are groups with no proper non-trivial normal subgroups we have the following definition.

Definition 3.8. [3, p.122] A Lie algebra $\mathfrak{g}$ is called simple if $\dim \mathfrak{g} > 1$ and $\mathfrak{g}$ contains no proper non-trivial ideals.

Now we can define a semisimple Lie algebra.

Definition 3.9. [8, p.300] A Lie algebra $\mathfrak{g}$ is called semisimple if it is the direct product of simple Lie algebras.

Remark 3.10. Sources may differ on the definition of a semisimple Lie algebra due to there being a few equivalent definitions. For more details look at chapter 20 of [8].

Example 3.11. $\mathfrak{sl}(n)$ is semisimple for $n \geq 2$, $\mathfrak{so}(n)$ is semisimple for $n \geq 3$ and $\mathfrak{sp}(2n)$ is semisimple for $n \geq 1$. [9, p.170]

3.2 Cartan Subalgebras

We now generalise a method for getting a root system from a Lie algebra. Recall in chapter 1 that we referred to something called a Cartan subalgebra which in the case of $\mathfrak{sl}(3)$ was the subspace of diagonal matrices. This is very important for finding root systems so let us properly define it.

Definition 3.12. [2, p.80][2, p.17] Let $\mathfrak{h}$ be a subalgebra of a matrix Lie algebra $\mathfrak{g}$. $\mathfrak{h}$ is a toral subalgebra if all of its elements are diagonalizable. Then if $\mathfrak{h}$ is a maximal toral subalgebra, that is $\mathfrak{h}$ is a toral subalgebra that is not contained in any other toral subalgebra of $\mathfrak{g}$, then $\mathfrak{h}$ is a Cartan subalgebra.

Remark 3.13. Again sources may differ for the definition of Cartan subalgebras and toral subalgebras. A reason why they may differ is if they want to work with Lie algebras over a general field $\mathbb{F}$. Since we are only interested in Lie algebras over $\mathbb{C}$ this definition is suitable.

Proposition 3.14. *Every finite dimensional complex Lie algebra contains a Cartan subalgebra.*

Proof. For any Lie algebra there always exists a toral subalgebra since the zero subalgebra is a toral subalgebra. If a complex Lie algebra is finite dimensional then since there exists toral subalgebras, there must exist a toral subalgebra of maximum dimension. □

Example 3.15. (i) [3, p.211] As seen in chapter 1, for $\mathfrak{sl}(n)$, a Cartan subalgebra is the subspace of diagonal matrices,

$$\begin{aligned}\mathfrak{h} &= \left\{ \begin{pmatrix} \theta_1 & 0 & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_n \end{pmatrix} : \theta_1 + \theta_2 + \cdots + \theta_n = 0 \right\} \\ &= \{\mathrm{diag}(\theta_1, \theta_2, \dots, \theta_n) : \theta_1 + \theta_2 + \cdots + \theta_n = 0\}.\end{aligned}$$

(ii) [3, p.239] For $\mathfrak{sp}(2n)$, a Cartan subalgebra is the following set of $2n$-by-$2n$ diagonal matrices,

$$\begin{aligned}\mathfrak{h} &= \left\{ \begin{pmatrix} \theta_1 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & \theta_n & 0 & \dots & 0 \\ 0 & \dots & 0 & -\theta_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -\theta_n \end{pmatrix} : \theta_i \in \mathbb{C} \right\} \\ &= \{\mathrm{diag}(\theta_1, \dots, \theta_n, -\theta_1, \dots, -\theta_n) : \theta_i \in \mathbb{C}\}.\end{aligned}$$

(iii) For $\mathfrak{so}(2n)$ a Cartan subalgebra is the following [10, p.87]

$$\begin{aligned}\mathfrak{h} &= \left\{ \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_n \end{pmatrix} : M_i = \begin{pmatrix} 0 & \theta_i \\ \theta_i & 0 \end{pmatrix}, \theta_i \in \mathbb{C} \right\} \\ &= \{\mathrm{diag}(M_1, M_2, \dots, M_n)\}.\end{aligned}$$

But in order to make computing the roots easier we want the Cartan subalgebra to contain diagonal matrices like the previous Cartan subalgebras.

In order to do this consider the map which sends $X \in \mathfrak{sl}(2n)$ to MX where $M = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$.

This map is bijective since it is self-inverse and we have that $(MX)^T = -XM$. So let us redefine $\mathfrak{so}(2n)$ to be the following which is isomorphic to our original definition, [3, p.268]

$$\mathfrak{so}(2n) = \left\{ X \in \mathfrak{gl}(2n) : MX + X^T M = 0, \text{where } M = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} \right\}.$$

Then the Cartan subalgebra of $\mathfrak{so}(2n)$ is the same as the Cartan subalgebra above for $\mathfrak{sp}(2n)$.[3, p.269]

(iv) [3, p.269] Similarly, we can redefine $\mathfrak{so}(2n+1)$ to be

$$\mathfrak{so}(2n+1) = \left\{ X \in \mathfrak{gl}(2n+1) : MX + X^T M = 0, \text{where } M = \begin{pmatrix} 0 & I_n & 0 \\ I_n & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}.$$

Then the Cartan subalgebra is the following

$$\begin{aligned} \mathfrak{h} &= \left\{ \begin{pmatrix} \theta_1 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \theta_n & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & -\theta_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -\theta_n & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} : \theta_i \in \mathbb{C} \right\} \\ &= \{\mathrm{diag}(\theta_1, \dots, \theta_n, -\theta_1, \dots, -\theta_n, 0) : \theta_i \in \mathbb{C}\}. \end{aligned}$$

The Cartan subalgebra allows us to deconstruct the Lie algebra and find the 'roots' of a Lie algebra which corresponds to the roots of a root system.

Proposition 3.16. *[2, p.35][11, p.126] Let $\mathfrak{g}$ be a semisimple finite dimensional complex Lie algebra with Cartan subalgebra $\mathfrak{h}$.*
$\mathfrak{g}$ can be decomposed as $\mathfrak{g} = \oplus\, \mathfrak{g}_\alpha$ for $\mathfrak{g}_\alpha = \{x \in \mathfrak{g} : [h, x] = \alpha(h)x,\ \ h \in \mathfrak{h}\}$ where $\alpha \in \mathfrak{h}^* = \{\phi : \mathfrak{h} \to \mathbb{C} : \phi \text{ is a linear map}\}$.
This is called the root space decomposition of $\mathfrak{g}$ and the nonzero α's are called the roots of $\mathfrak{g}$.

From our original definition of a root system, our roots live in a real inner product space V. The roots of a Lie algebra live in $\mathfrak{h}^*$ which is also an inner product space.

Definition 3.17. [2, p.39] Let $\mathfrak{g}$ be a finite dimensional complex Lie algebra with Cartan subalgebra $\mathfrak{h}$ and root space decomposition $\mathfrak{g} = \oplus \mathfrak{g}_\alpha$.
$\mathfrak{h}^* = \{\phi : \mathfrak{h} \to \mathbb{C} : \phi \text{ is a linear map}\}$ is called the dual space of $\mathfrak{h}$ and is an inner product space with inner product $K^*(\alpha, \beta) = K(t_\alpha, t_\beta)$ where for $\phi \in \mathfrak{h}^*$, $t_\phi \in \mathfrak{h}$ is the unique element such that $\phi(h) = K(t_\phi, h)$ for all $h \in H$. We will call the inner product $K^*(,)$ the dual killing form of $\mathfrak{h}^*$.

Remark 3.18. For a proof of the existence and uniqueness of the element $t_\phi \in \mathfrak{h}$, look at proposition 8.2 in and its immediate corollary in [2, p.36].

Now we give the following theorem which states the link between Lie algebras and root systems.

Theorem 3.19. *[9, p.237] Let $\mathfrak{g}$ be a finite dimensional complex Lie algebra with Cartan subalgebra $\mathfrak{h}$ and root space decomposition $\mathfrak{g} = \oplus\, \mathfrak{g}_\alpha$.*

(i) The set $\phi = \{\alpha \in \mathfrak{h}^* : \mathfrak{g}_\alpha \neq 0\}$ is a root system with respect to the dual Killing form.
We call ϕ the root system of $\mathfrak{g}$ with respect to $\mathfrak{h}$.

(ii) Let ϕ be the root system of $\mathfrak{g}$ with respect to $\mathfrak{h}$.
If $\mathfrak{g}'$ is a finite dimensional complex Lie algebra with root system ϕ' with respect to some Cartan subalgebra $\mathfrak{h}'$ such that $\phi \cong \phi'$, then $\mathfrak{g} \cong \mathfrak{g}'$.
i.e The root system of a Lie algebra determines the Lie algebra up to isomorphism.

(iii) If ϕ is a root system then there exists a Lie algebra $\mathfrak{g}'$ and a Cartan subalgebra $\mathfrak{h}'$ of $\mathfrak{g}'$ such that the root system of $\mathfrak{g}'$ with respect to $\mathfrak{h}'$ is isomorphic to ϕ.

Proof. Proofs for each part can be found in [2].
For (i) look at chapter 8, for (ii) look at section 14.2 and for (iii) see section 18. □

Hence this shows that classifying root systems up to isomorphism is the same as classifying semisimple Lie algebras.

3.3 Root Systems From Semisimple Lie Algebras

Now let us give some examples of finding root systems from semisimple Lie algebras.

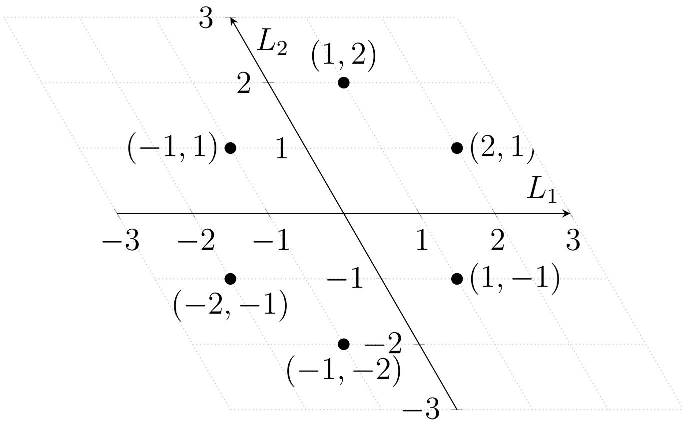
Example 3.20. We have already found a root system from a Lie algebra in our initial example of $\mathfrak{sl}(3)$. From example 3.15(i), the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{sl}(3)$ is the set of 3-by-3 diagonal matrices whose trace equal zero. Recall that when we take the commutator bracket of a general element $H = \begin{pmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{pmatrix} \in \mathfrak{h}$ where $\theta_3 = -\theta_1 - \theta_2$, with a basis element $M_{i,j}$ of $\mathfrak{sl}(3)$ then we have the formula

$$[H, M_{i,j}] = (\theta_i - \theta_j)M_{i,j}.$$

In the initial example we then plotted $(\theta_i - \theta_j)$ but from Proposition 3.16 we know that the roots of Lie algebras are linear maps so let us define the constant functions $L_i : \mathfrak{h} \to \mathbb{C}$; $H \mapsto \theta_i$. Then the roots of $\mathfrak{sl}(3)$ are the linear functions $L_i - L_j$ for $i, j \in \{1, 2, 3\}$ where $i \neq j$. But since $L_3 = -L_1 - L_2$, we get the full list of roots to be

$$L_1 + L_2,\ -L_1 + L_2,\ 2L_1 + L_2,\ -2L_1 - L_2,\ L_1 + 2L_2,\ -L_1 - 2L_2.$$

Then if we plot these roots on a coordinate axis with respect to the basis $\{L_1, L_2\}$ and where the axes are $\frac{2\pi}{3}$ radians apart then we get the familiar image of the root system A_2.



Now let us briefly discuss the reasoning behind having the axes $\frac{2\pi}{3}$ radians apart. In vector spaces, an inner product can be used to see the geometry of the space. For example, if we consider the vector space $\mathbb{R}^2$ with the inner product of the dot product, the standard basis of $\mathbb{R}^2$ is $\{(1,0),(0,1)\}$ and their dot product equals zero corresponding to the angle between them being $\pi/2$ and so we would draw an axis for $\mathbb{R}^2$ with the axes perpendicular.

Since roots live in the dual space of the Cartan subalgebra which has the inner product of the dual Killing form, we can take a basis of the dual space of the Cartan subalgebra for $\mathfrak{sl}(3)$ then we can find the angle between them with respect to the dual Killing form to figure how we should plot our axis and it turns out that this angle is $\frac{2\pi}{3}$ justifying our diagrams. More details on finding a basis and calculating the angle between them can be found in chapter 17 of the Math426 lecture notes [1, p.159].

In general we have that the root system of $\mathfrak{sl}(n+1)$ is A_n.

Example 3.21. [3, p.270]

Let us find the root system of
$$\mathfrak{so}(5) = \left\{ X \in \mathfrak{gl}(5) : MX + X^T M = 0, \text{where } M = \begin{pmatrix} 0 & I_2 & 0 \\ I_2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}.$$
First we need to find a basis of $\mathfrak{so}(5)$.
From solving the equation $MX + X^T M = 0$ for a general 5-by-5 matrix, we get that a general element of $\mathfrak{so}(5)$ is of the form

$$\begin{pmatrix} a_1 & a_2 & 0 & a_3 & 0 \\ b_1 & b_2 & -a_3 & 0 & 0 \\ 0 & c_1 & -a_1 & -b_1 & 0 \\ -c_1 & 0 & -a_2 & -b_2 & 0 \\ d_1 & d_2 & d_3 & d_4 & d_5 \end{pmatrix}.$$

Let $M_{i,j}$ be a matrix with a 1 in position i,j and zeroes in all other entries like in our initial example from chapter 1. Then we can define a basis of $\mathfrak{so}(5)$ as the following set

$$\{M_{1,1} - M_{3,3},\ M_{2,2} - M_{4,4},\ M_{1,2} - M_{4,3},\ M_{2,1} - M_{3,4},\ M_{1,4} - M_{2,3},\ M_{3,2} - M_{4,1},$$
$$M_{5,1},\ M_{5,2},\ M_{5,3},\ M_{5,4},\ M_{5,5}\}.$$

To simplify things let us define the matrix $X_{i,j} = M_{i,j} - M_{2+j,2+i}$, for $1 \leq i,j \leq 2$ so we can rewrite the basis as

$$\{X_{1,1},\ X_{2,2},\ X_{1,2},\ X_{2,1},\ M_{1,4} - M_{2,3},\ M_{3,2} - M_{4,1},$$
$$M_{5,1},\ M_{5,2},\ M_{5,3},\ M_{5,4},\ M_{5,5}\}.$$

A general element in the Cartan subalgebra of $\mathfrak{so}(5)$ is

$$H = \begin{pmatrix} \theta_1 & 0 & 0 & 0 & 0 \\ 0 & \theta_2 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & 0 & 0 \\ 0 & 0 & 0 & \theta_4 & 0 \\ 0 & 0 & 0 & 0 & \theta_5 \end{pmatrix} \text{ where } \theta_3 = -\theta_1,\ \theta_4 = -\theta_2 \text{ and } \theta_5 = 0.$$

So if we calculate the commutator bracket of H with the basis elements we get

$$
\begin{aligned}
[H, M_{i,j}] &= HM_{i,j} - M_{i,j}H \\
&= \theta_j M_{i,j} - \theta_i M_{i,j} \\
&= (\theta_j - \theta_i) M_{i,j}. \\
[H, X_{i,j}] &= HX_{i,j} - X_{i,j}H \\
&= HM_{i,j} - HM_{2+j,2+i} - M_{i,j}H + M_{2+j,2+i}H \\
&= \theta_j Mi, j - \theta_{2+i} M_{2+j,2+i} - \theta_i M_{i,j} + \theta_{2+j} M_{2+j,2+i} \\
&= (\theta_j - \theta_i) M_{i,j} + (\theta_{2+j} - \theta_i 2 + i) M_{2+j,2+i} \\
&= (\theta_j - \theta_i) M_{i,j} - (\theta_j - \theta_i) M_{2+j,2+i} \\
&= (\theta_j - \theta_i) X_{i,j}.
\end{aligned}
$$

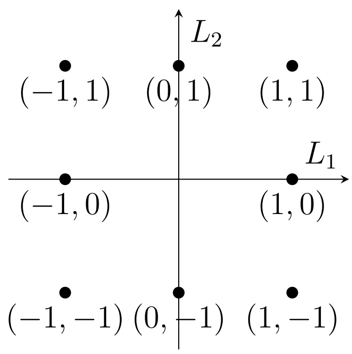
Similarly we get,

$$
\begin{aligned}
[H, M_{1,4} - M_{2,3}] &= (-\theta_1 - \theta_2)(M_{1,4} - M_{2,3}). \\
[H, M_{3,2} - M_{4,1}] &= (\theta_1 + \theta_2)(M_{1,4} - M_{2,3}).
\end{aligned}
$$

Hence defining the function $L_i(H) = \theta_i$, we get the non-zero roots to be

$$-L_1 + L_2,\ L_1 - L_2,\ -L_1 - L_2,\ L_1 + L_2,\ L_1, L_2,\ -L_1, -L_2.$$

Then plotting the roots on an axis we get the root system B_2.



In general, the root system of $\mathfrak{so}(2n+1)$ is B_n for $n \geq 2$ with the root system of $\mathfrak{so}(3)$ being A_1. In fact, since $\mathfrak{so}(3)$ and $\mathfrak{sl}(2)$ have the same root system and root systems determine semisimple Lie algebras up to isomorphism we have the isomorphism $\mathfrak{so}(3) \cong \mathfrak{sl}(2)$.

Example 3.22. [3, p.240] Now let us find the root system of

$$\mathfrak{sp}(6) = \left\{ M \in \mathfrak{gl}(6) : \Omega M + M^T \Omega = 0 \text{ for } \Omega = \begin{pmatrix} 0 & I_3 \\ -I_3 & 0 \end{pmatrix} \right\}$$

From the definition we get a general element of $\mathfrak{sp}(6)$ to be

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & a_5 & b_4 & b_5 \\ c_1 & c_2 & c_3 & a_6 & b_5 & c_4 \\ d_1 & d_2 & d_3 & -a_1 & -b_1 & -c_1 \\ d_2 & e_1 & e_2 & -a_2 & -b_2 & -c_2 \\ d_3 & e_2 & f_1 & -a_3 & -b_3 & -c_3 \end{pmatrix}$$

Like in the previous example let us define the following matrices to simplify the basis

$$X_{i,j} = M_{i,j} - M_{3+j,3+i},\ Y_{i,j} = M_{i,3+j} + M_{j,3+i},\ Z_{i,j} = M_{3+i,j} + M_{3+j,i}.$$

Then we can define a basis as the following

$$\{X_{1,1},\ X_{1,2},\ X_{1,3},\ M_{1,4},\ Y_{1,2},\ Y_{1,3},\ X_{2,1},\ X_{2,2},\ X_{2,3},\ M_{2,5},\ Y_{2,3},\\ X_{3,1},\ X_{3,2},\ X_{3,3},\ M_{3,6},\ M_{4,1},\ Z_{1,2},\ Z_{1,3},\ M_{5,2},\ Z_{2,3},\ M_{6,3}\}.$$

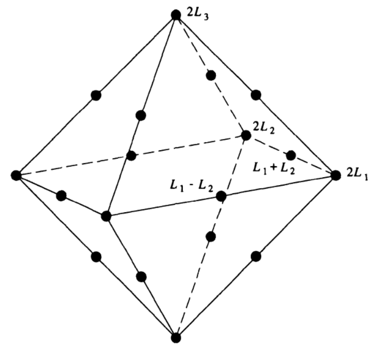
A general element of the Cartan subalgebra is

$$H = \begin{pmatrix} \theta_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\theta_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\theta_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\theta_3 \end{pmatrix}$$

We have

$$\begin{aligned} [H, X_{i,j}] &= HM_{i,j} + HM_{3+j,3+i} - M_{i,j}H - M_{3+j,3+i}H \\ &= \theta_j M_{i,j} - \theta_i M_{3+j,3+i} - \theta_i M_{i,j} + \theta_j M_{3+j,3+i} \\ &= (\theta_j - \theta_i)X_{i,j} \\ [H, Y_{i,j}] &= HM_{i,3+j} - HM_{j,3+i} - M_{i,3+j}H + M_{j,3+i}H \\ &= -\theta_j M_{i,3+j} + \theta_i M_{j,3+i} - \theta_i M_{i,3+j} + \theta_j M_{j,3+i} \\ &= (-\theta_i - \theta_j)Y_{i,j} \\ [H, Y_{i,j}] &= HM_{3+i,j} - HM_{3+j,i} - M_{3+i,j}H + M_{3+j,i}H \\ &= \theta_j M_{3+i,j} - \theta_i M_{3+j,i} + \theta_i M_{3+i,j} - \theta_j M_{3+j,i} \\ &= (\theta_i + \theta_j)Y_{i,j} \\ [H, M_{1,4}] &= -2\theta_1 M_{1,4},\ [H, M_{2,5}] = -2\theta_2 M_{2,5},\ [H, M_{3,6}] = -2\theta_3 M_{3,6} \\ [H, M_{4,1}] &= 2\theta_1 M_{4,1},\ [H, M_{5,2}] = 2\theta_2 M_{5,2},\ [H, M_{6,3}] = 2\theta_3 M_{6,3} \end{aligned}$$

Hence the nonzero roots of $\mathfrak{sp}(6)$ are $\pm L_i \pm L_j$ for $i \neq j$ and $\pm 2L_i$. Then plotting these roots on a coordinate axis with three axes, we get that the roots correspond to the vertices and midpoints of the edges of an octahedron like the following picture.



Picture source [3, p.254].

This root system corresponds to C_3 and in general we have that the root system of $\mathfrak{sp}(2n)$ is C_n for $n \geq 3$ with the root system of $\mathfrak{sp}(2)$ being A_1 and the root system of $\mathfrak{sp}(4)$ being B_2. So we have the isomorphisms $\mathfrak{sp}(2) \cong \mathfrak{so}(3) \cong \mathfrak{sl}(2)$ and $\mathfrak{sp}(4) \cong \mathfrak{so}(5)$.

Example 3.23. [3, p.270] Now let us find the root system of

$$\mathfrak{so}(4) = \left\{ X \in \mathfrak{gl}(4) : MX + X^T M = 0, \text{where } M = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \right\}.$$

A general element of $\mathfrak{so}(4)$ is

$$\begin{pmatrix} a_1 & a_2 & 0 & a_3 \\ b_1 & b_2 & -a_3 & 0 \\ 0 & c_1 & -a_1 & -b_1 \\ -c_1 & 0 & -a_2 & -b_2 \end{pmatrix}$$

So then the following matrices form a basis of $\mathfrak{so}(4)$

$$e_1 = M_{1,1} - M_{3,3},\ e_2 = M_{1,2} - M_{4,3},\ e_3 = M_{1,4} - M_{2,3}$$
$$e_4 = M_{2,1} - M_{3,4},\ e_5 = M_{2,2} - M_{4,4},\ e_6 = M_{3,2} - M_{4,1}.$$

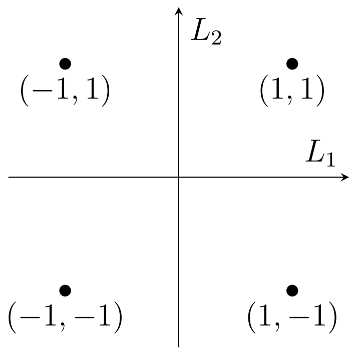
A general element of the Cartan subalgebra is

$$H = \begin{pmatrix} \theta_1 & 0 & 0 & 0 \\ 0 & \theta_2 & 0 & 0 \\ 0 & 0 & -\theta_1 & 0 \\ 0 & 0 & 0 & -\theta_2 \end{pmatrix}.$$

So then taking the commutator brackets of H with the basis elements we get

$$[H, e_1] = 0,\ [H, e_2] = (\theta_1 - \theta_2)e_2,\ [H, e_3] = (\theta_1 + \theta_2)e_3,$$
$$[H, e_4] = (-\theta_1 + \theta_2)e_4,\ [H, e_5] = 0,\ [H, e_6] = (-\theta_1 - \theta_2)e_6.$$

Hence the non-zero roots are $\pm L_1 \pm L_2$ which when we plot on a coordinate axis we get the root system $A_1 \times A_1$. Note that this means that while $\mathfrak{so}(4)$ is semisimple, it is not simple.



More generally we have that the root system of $\mathfrak{so}(2n)$ is D_n for $n \geq 4$ and the root system for $\mathfrak{so}(6)$ is A_3. So we have the isomorphisms $\mathfrak{so}(4) \cong \mathfrak{sl}(2) \oplus \mathfrak{sl}(2)$ and $\mathfrak{so}(6) \cong \mathfrak{sl}(4)$.

Let us recap the results from these examples

Proposition 3.24. *[3, p.326]*
For $n \geq 1$, the root system of $\mathfrak{sl}(n+1)$ is A_n.
For $n \geq 2$, the root system of $\mathfrak{so}(2n+1)$ is B_n.
For $n \geq 3$, the root system of $\mathfrak{sp}(2n)$ is C_n.
For $n \geq 4$, the root system of $\mathfrak{so}(2n)$ is D_n.

Proposition 3.25. *[3, p.326-p.327] We have the following isomorphisms between semisimple Lie algebras*

$$\mathfrak{sl}(2) \cong \mathfrak{so}(3) \cong \mathfrak{sp}(2),$$
$$\mathfrak{so}(5) \cong \mathfrak{sp}(4),$$
$$\mathfrak{sl}(2) \oplus \mathfrak{sl}(2) \cong \mathfrak{so}(4),$$
$$\mathfrak{sl}(4) \cong \mathfrak{so}(6).$$

These are sometimes called exceptional isomorphisms.

So now let us classify all simple complex Lie algebras.

Theorem 3.26. *Let $\mathfrak{g}$ be a simple complex Lie algebra. Then $\mathfrak{g}$ is isomorphic to $\mathfrak{sl}(n)$, $\mathfrak{so}(n)$ or $\mathfrak{sp}(2n)$ for some $n \in \mathbb{N}$ or it is isomorphic to a Lie algebra whose root system is E_6, E_7, E_8, F_4 or G_2.*

4 Applications

We finish by briefly discussing where root systems and Lie algebras are used in other areas of maths and physics.

4.1 Lie Groups

The study of Lie algebras comes from the study of Lie groups so let us define a Lie group.

Definition 4.1. [1, p.30] Let $GL(n)$ denote the group of invertible n-by-n matrices. Then a matrix Lie group is a topologically closed subgroup of $GL(n)$.
This means that a subgroup G of $GL(n)$ is a Lie group if any convergent sequence whose elements are in G has limit belonging to G.

Remark 4.2. There exists Lie groups which are not matrix groups but the only Lie groups we will discuss are matrix Lie groups so for simplicity this definition will suffice.

So how do Lie groups relate to Lie algebras? It turns out that they are linked by the exponential map.

Definition 4.3. [3, p.116] The exponential of a n-by-n matrix M is

$$\exp(M) = I_n + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \cdots = \sum_{n=1}^{\infty} \frac{1}{n!}M^n$$

Proposition 4.4. *[3, p.116] Let G be a matrix Lie group. Then the set*

$$\mathfrak{g} = \{g \in \mathfrak{gl}(n) : exp(tg) \in G \text{ for all } t \in \mathbb{R}\}$$

forms a Lie algebra and is called the Lie algebra of G.

So Lie algebras can be seen as a linearised form of Lie groups which can make studying the Lie algebra of a Lie group simpler then studying the Lie group itself and then studying the root system of a Lie algebra can be even simpler than studying the Lie algebra itself.

Example 4.5. [9, p.5-p.9] Some examples of Lie groups and their Lie algebras are the following.

(i) (General linear matrices). The exponential of any n-by-n matrix is invertible meaning that the Lie algebra of $GL(n)$ is $\mathfrak{gl}(n)$.

(ii) (Special linear matrices). $SL(n)$ denotes the set of n-by-n matrices with determinant 1. This is a Lie group whose Lie algebra is $\mathfrak{sl}(n)$. This is the relationship we mentioned in our initial example between $\mathfrak{sl}(3)$ and the special linear matrices hence why we use this notation.

(iii) (Special orthogonal matrices). $SO(n)$ denotes the set of n-by-n matrices whose transpose equals its inverse and have determinant 1. This is a Lie group whose Lie algebra is $\mathfrak{so}(n)$.

(iv) (Symplectic matrices). $Sp(2n)$ denotes the set of $2n$-by-$2n$ matrices M that satisfy

$$M^T \Omega M = \Omega, \text{ for } \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

This is a Lie group with Lie algebra $\mathfrak{sp}(2n)$.

(v) (Unitary matrices). $U(n)$ denotes the set of matrices whose conjugate transpose equals its inverse. This is a Lie group whose Lie algebra is $\mathfrak{u}(n)$.

(vi) (Special unitary matrices). $SU(n)$ denotes the subgroup of $U(n)$ where every matrix element has determinant 1. This is a Lie group whose Lie algebra is $\mathfrak{su}(n)$.

4.2 Differential Equations

The study of Lie groups came from the study of differential equations. In particular Lie groups are used by 'acting' on a solution of a differential equation. So let us first define what we mean by a group acting on a set.

Definition 4.6. [12][p.144] Let S be a set and let G be a group. A group action of G on the set S is a map from $G \times S$ to S which sends (g, s) to an element in S, denoted by $g.s$, such that

(i) If e is the identity element of G then $e.s = s$ for all $s \in S$, and

(ii) $(g_1 g_2).s = g_1.(g_2.s)$ for all $g_1, g_2 \in G$ and for all $s \in S$.

If we have a group action between a group G and a set S we often say that G acts on or is acting on S and say that $g \in G$ acts on or is acting on $s \in S$ to refer to the element $g.s$.

So a group action for G allows us to map the elements of a set to itself in such a way that preserves some of the structure or properties of G.
Now let us return to differential equations

Definition 4.7. [13, p.93] Suppose that we have a system of ordinary differential equations with p real independent variables and q real dependent variables has solution $u = f(x)$.
A symmetry group of the system is a Lie group G acting on an open subset of $\mathbb{R}^p \times \mathbb{R}^q$ such that for $g \in G$, $u = g.f(x)$ is also a solution of the system whenever $g.f$ is defined.

Example 4.8. [13, p.91] Let us consider the simple differential equation $\frac{d^2u}{dx^2} = 0$ where $p = q = 1$. Clearly if $u = f(x)$ is a solution then f is just a linear function i.e. $u = f(x) = ax + b$ for some $a, b \in \mathbb{R}$ hence the solutions are of the form $(x, ax + b)$.
Now let us consider the group $SO(2)$ which consists of rotations meaning that the matrices are of the form $M = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ for some angle θ.
We can define a group action of $SO(2)$ on the set $\mathbb{R}^2$ by

$$M.(x, u) = (x\cos\theta - u\sin\theta, x\sin\theta + u\cos\theta).$$

So if we have $M \in SO(2)$ act on a solution then

$$M.(x, ax + b) = (x\cos\theta - (ax + b)\sin\theta, x\sin\theta + (ax + b)\cos\theta).$$

So let us define $(x', u') = M.(x, ax + b)$ and try and write u' in terms of just x'.
So we have $x' = x\cos\theta - (ax + b)\sin\theta$, which provided that $\cot\theta \neq a$ can be rearranged to get

$$x = \frac{x' + b\sin\theta}{\cos\theta - a\sin\theta},$$

and we can substitute this into $u' = x\sin\theta + (ax + b)\cos\theta$ and rearrange to get

$$u' = \frac{sin\theta + a\cos\theta}{\cos\theta - a\sin\theta}x' + \frac{b}{\cos\theta - a\sin\theta},$$

which is of the form $a'x' + b'$ and so is a solution of $\frac{d^2u}{dx'^2} = 0$.
Hence $SO(2)$ is a symmetry group for the differential equation $\frac{d^2u}{dx^2} = 0$.

So if we know that symmetry group for a system of differential equations, it allows us to construct new solutions from known solutions.
One useful result coming from the study of Lie groups acting on solutions of differential equations allows us to reduce the order of a differential equation by one if its symmetry group is a certain kind of Lie group called an r-parameter Lie group.

Definition 4.9. [13, p.15] An r-parameter Lie group is an r-dimensional Lie group G such that both the group operation

$$m : G \times G \to G, m(g,h) = gh$$

and the inversion map

$$i : G \to G, i(g) = g^{-1}$$

are both smooth maps meaning they are both infinitely differentiable.

Proposition 4.10. *[13, p.142] Any n-th order ordinary differential equation that has a 1-parameter symmetry group is equivalent to an $(n-1)$th order equation.*

Proof. This follows from results proven in [13]. Namely Proposition 2.18 on page 86 and Corollary 2.54 on page 141 □

4.3 Hydrogen Atoms

Finally, we look at an application in Physics. In quantum mechanics, the Hamiltonian $\hat{H}$ is an operator that acts on a quantum state of a system. This is shown in the time independent Schrödinger equation $\hat{H}|\Psi\rangle = E|\Psi\rangle$, where $|\Psi\rangle$ is a quantum state of a system and E is the energy of the system in that state [14, p.28].

Let $|\Psi\rangle$ be a quantum state of a hydrogen atom with energy E meaning that $\hat{H}|\Psi\rangle = E|\Psi\rangle$, and let $g \in SO(3)$. Then $\hat{H}(g|\Psi\rangle) = E(g|\Psi\rangle)$ i.e. the energy level of the hydrogen atom in a certain quantum state is the same as the energy of a hydrogen atom in the quantum state being acted on by an element in $SO(3)$ [15, p.228]. This means that multiple quantum states of a hydrogen atom share the same energy levels which in physics is described as the hydrogen atom having degenerate energy levels [14, p.29].

There are more associations between particle physics and semisimple Lie groups (Lie groups whose Lie algebras are semisimple) such as $SO(n)$ and $Sp(n)$ as they can help describe the symmetries that various particles have. Then root systems are useful in clarifying how two symmetry groups are related to each. More information on connections between Lie groups and Lie algebras in quantum physics can be found here [16], in particular chapter 8 discusses roots and Dynkin diagrams.

5 References

[1] Jonny Evans. Math426: Lie groups and lie algebras. `https://modules.lancaster.ac.uk/course/view.php?id=35794`. Accessed: 28.11.2022.

[2] James E. Humphreys. *Introduction to Lie Algebras and Representation Theory*. Graduate texts in mathematics. Springer, 1972.

[3] William Fulton and Joe Harris. *Representation theory : a first course*. Graduate texts in mathematics ; 129. Springer, New York, corr. 3rd print. edition, 1996.

[4] Eric W. Weisstein. Hyperplane wolfram mathworld article. `https://mathworld.wolfram.com/Hyperplane.html`. Accessed: 16.12.2022.

[5] James E. Humphreys. *Reflection groups and coxeter groups*. Cambridge studies in advanced mathematics ; 29. Cambridge University Press, Cambridge, 1990.

[6] Todd Rowland. Dynkin diagram wolfram mathworld article. `https://mathworld.wolfram.com/DynkinDiagram.html`. Accessed: 16.12.2022.

[7] Arthur A. Sagle and Ralph E. Walde. *Introduction to Lie groups and Lie algebras*. Pure and applied mathematics (Academic Press) ; 51. Academic Press, New York, 1973.

[8] Patrice Tauvel and Robert W. T Yu. *Lie algebras and algebraic groups*. Springer monographs in mathematics. Springer, New York, 2005.

[9] Brian C Hall. *Lie groups, Lie algebras, and representations an elementary introduction*. Graduate texts in mathematics ; 222. Springer, Cham ; New York, second edition. edition, 2015.

[10] J. Frank (John Frank) Adams. *Lectures on Lie groups*. Mathematics lecture note series. W. A. Benjamin, New York, 1969.

[11] L. (Luc) Frappat, Antonino Sciarrino, and Paul Sorba. *Dictionary on Lie algebras and superalgebras*. Academic Press, San Diego, 2000.

[12] M. Eie and S. Chang. *A Course On Abstract Algebra*. World Scientific Publishing Company, 2010.

[13] Peter J Olver. *Applications of Lie groups to differential equations*. Graduate texts in mathematics ; 107. Springer-Verlag, New York, 1986.

[14] L. D. Landau and L. M. Lifshitz. *Quantum Mechanics Non-Relativistic Theory, Second Edition: Volume 3.* Pergamon Press, 1965.

[15] Robert Gilmore. *Lie groups, physics, and geometry : an introduction for physicists, engineers and chemists.* Cambridge University Press, Cambridge, 2008.

[16] Howard M Georgi. *Lie algebras in particle physics; 2nd ed.* Frontiers in physics. Perseus Books, 1999.