UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2018:30

Classification of Root Systems

Filip Koerfer

Examensarbete i matematik, 15 hp
Handledare: Jakob Zimmermann
Examinator: Martin Herschend
Juni 2018

Department of Mathematics
Uppsala University

Contents

Introduction

As the title suggests, this paper explores root systems and their classification by Dynkin diagrams. It is based to a vast majority on chapters 11 and 13 in [EW].

However, before we delve into root systems and Dynkin diagrams, let us begin by explaining why they are important. Root systems are important mainly because of Lie algebras, we will not define what a Lie algebra is in this text, nor will we study them in any way. They are, however, very important in modern mathematics and they have many applications in physics, in particular particle physics. Because of the importance of Lie algebras we would like to classify them, and it turns out that a certain type of Lie algebras, namely semi-simple Lie algebras over algebraically closed fields, are completely classified by root systems. This essentially means that if we know what every root system looks like, then we also know what every Lie algebra of this type looks like.

Now this is where Dynkin diagrams come in. We do not know what every root system looks like, but we can classify all the root systems using Dynkin diagrams. Because finding every Dynkin diagram associated with a root system is fairly simple to do. So when we know what every Dynkin diagram associated with a root system looks like, then we also know what every root system looks like, up to isomorphism. Therefore we can derive what every semi-simple Lie algebra over an algebraically closed field looks like, again up to isomorphism. This is one reason why root systems and Dynkin diagrams are important in modern mathematics.

In the first section we will explore some linear algebra which will be used extensively throughout this entire paper. We will see what an inner-product space over a field is and look at several properties of these spaces. In the second section we define what a root system is and continue by looking at the restrictions on root systems which are necessary in order to classify them. We also define what a base of a root system is and show that every root system has such a basis. We move on to the third section where the Weyl group is introduced and we explore how the Weyl group is used in order to recover a root system by only looking at its basis. In section four we start by looking at how different bases of the same root system compare, and later define what we mean when we say that two root systems are isomorphic. With this we are able to introduce Dynkin diagrams associated with root systems, which we in section five will classify which in turn classifies all root systems.

1 Vector Spaces

In order to study root systems we first need to define what a vector space is, and look at some properties of vector spaces.

1.1 Fields

Before we can define a vector space we need to know what a field is.

Definition 1.1.1. A *field* is a set K together with

- a binary operation $+ : K \times K \to K$ (addition),
- a binary operation $\cdot : K \times K \to K$ (multiplication),
- an element $0_K \in K$,
- and element $1_K \in K$ which is not equal to 0_K

with the following properties for all $x, y, z \in K$

$(x + y) + z = x + (y + z)$	(associativity of $+$)
$(x \cdot y) \cdot z = x \cdot (y \cdot z)$	(associativity of $\cdot$)
$x + y = y + x$	(commutativity of $+$)
$x \cdot y = y \cdot x$	(commutativity of $\cdot$)
$x \cdot (y + z) = x \cdot y + x \cdot z$	(distributivity)
$0_K + x = x$	(neutral element for $+$)
$1_K \cdot x = x$	(neutral element for $\cdot$)
$\forall a \exists b : a + b = 0_K$	(additive inverse)
$\forall a \exists b : a \cdot b = 1_K$	(multiplicative inverse)

We denote this field as $(K, +, \cdot)$ and use the shorthand notation $K = (K, +, \cdot)$ where it creates no ambiguity.

Example 1.1.2. Two of the most common examples of infinite fields are the real numbers denoted by $\mathbb{R}$ and the complex numbers denoted as $\mathbb{C}$. An example of a finite field is $\mathbb{Z}_3 = \mathbb{Z}/3\mathbb{Z}$ with normal addition and multiplication for integers modulo 3.

Remark. As a handy convention we shall write multiplication of field elements as $a \cdot b = ab$ where it is clear what is meant.

Remark. A field $(K, +, \cdot)$ can be specified as two abelian groups, namely $(K, +)$ and $(K \setminus \{0\}, \cdot)$ that fulfill the distributivity property of one over the other.

1.2 Vector Spaces and Linear Maps

Definition 1.2.1. Let K be a field. A *vector space* over K is a set V together with

- a binary operation $+ : V \times V \to V$ (addition),
- an action $. : K \times V \to V$ (scalar multiplication),
- an element $\overline{0} \in V$ (the zero vector)

with the following properties for all $u, v, w \in V$ and $x, y \in K$:

$(u+v)+w = u+(v+w)$	(associativity of $+$)
$(x \cdot y).v = x.(y.v)$	(associativity of .)
$u+v = v+u$	(commutativity of $+$)
$\overline{0}+v = v$	(neutral element for $+$)
$1_K.v = v$	(neutral element for .)
$x.(u+v) = x.v + x.u$	(distributivity 1)
$(x+y).v = x.v + y.v$	(distributivity 2)
$\forall \alpha \exists \beta : \alpha + \beta = \overline{0}$	(additive inverse).

Elements of V are called *vectors* and elements of K are called *scalars*.

Example 1.2.2. A vector space which is commonly considered is $\mathbb{R}^n$ over $\mathbb{R}$ with the usual addition and scalar multiplication. We do also, however, have the set of all $m \times n$ matrices over a field K is a vector space where the addition is normal matrix addition and scalar multiplication is just multiplying each entry of the matrix by the same scalar. A final example of a vector space is K over K for any field K, this is indeed a vector space.

Remark. If there is no ambiguity we shall write $x.v = xv$ for any scalar x and any vector v in the associated vector space.

While we will not be studying linear maps in great detail in this text. We will, however, need a way of comparing vector spaces to see when they are "similar" or not, this is referred to as vector spaces being isomorphic. In order to do this, let us begin by defining what a linear map is.

Definition 1.2.3. Let V and W be vector spaces over K. A function $T : V \to W$ is a *linear map* if for every $u, v \in V$ and $x, y \in K$ we have;

$$T(xu + yv) = xT(u) + yT(v).$$

Remark. Note that on the left hand side of the equality the scalar multiplication and vector addition is in V, but on the right hand side it is in W.

We say that a linear map $T : V \to W$ is *injective* if for $u, v \in V$ we have that $T(u) = T(v)$ implies that $u = v$. We say that T is *surjective* if $T(V) = W$. If T is both injective and surjective we say that T is *bijective* or that T is a *bijection*.

Definition 1.2.4. Let V and W be two vector spaces. A linear map $T : V \to W$ is called an *isomorphism* if T is a bijection. If there exists an isomorphism between two vector spaces, we say that they are *isomorphic*.

We can also talk about bilinear maps. Bilinear maps are functions from the Cartesian product of two vector spaces into a third vector space satisfying the linearity property in both arguments. An example of such a map which we will use extensively throughout this entire paper is defined below.

Definition 1.2.5. Let V be a vector space over $\mathbb{R}$, a map $(-,-) : V \times V \to \mathbb{R}$ is called an *inner product* if for all $u, v, w \in V$ and $x, y \in \mathbb{R}$ it satisfies

- $(u, v) = (v, u)$ (symmetry),
- $(x.u + y.v, w) = x \cdot (u, w) + y \cdot (v, w)$ (linearity in the first argument),
- $(u, u) \geq 0$ and $(u, u) = 0$ if and only if $u = \overline{0}$ (positive-definiteness).

A vector space V endowed with an inner product is called an *inner-product space*.

Remark. Note that because of the symmetry property we have linearity in both arguments.

1.3 Basis of Vector Spaces

In order to study root systems properly we need to know what a basis of a vector space is, and what dimension a vector space has. This is a natural point of interest as we shall see later in this text.

Definition 1.3.1. Let V be a vector space over K and A be a subset of V, the *linear span* (or simply *span*) of A is the set

$$\operatorname{span}(A) = \left\{ \sum_{i=1}^{k} x_i v_i \ : \ k \in \mathbb{N}, v_i \in A, x_i \in K \right\}$$

of all finite linear combinations of elements in A.

Definition 1.3.2. Let V be a vector space and $A \subset V$, a vector $v \in V$ is *linearly independent* of A if $v \notin \operatorname{span}(A)$. If v is not linearly independent it is said to be *linearly dependent* of A. A set $B \subset V$ is *linearly dependent* if there is a vector $w \in B$ which is linearly dependent of $B \setminus \{w\}$, otherwise B is called *linearly independent.*

Remark. Note that this is equivalent to saying that $A = \{a_1, \ldots, a_n\} \subset V$ is linearly independent if $\sum_{i=1}^{n} x_i a_i = 0$ if and only if $x_i = 0$, for all x_i where x_i are elements of the underlying field of V. This can be realized by assuming that $a_k \in \operatorname{span}(A \setminus \{a_k\})$, then there exists a sum $s := x_1 a_1 + \cdots + x_{k-1} a_{k-1} + x_{k+1} a_{k+1} + \cdots + x_n a_n = a_k$. Now take $a_k - s$, this is equal to zero, is of the form described earlier, and not all $x_i = 0$. This is not a proof of the equivalence, but it might give an idea of how the proof would look.

Definition 1.3.3. A linearly independent subset B of a vector space V is called a *basis* if $\operatorname{span}(B) = V$.

Definition 1.3.4. The *dimension* $\dim(V)$ of a vector space V is the cardinality $n \in \mathbb{N}$ of a basis of V.

Remark. Note that the dimension of a vector space is defined invariant of what basis we have chosen, in other words, given two bases B_1, B_2 of a vector space V the cardinalities of B_1 and B_2 are indeed equal.

1.4 Hyperplanes and Reflections

Definition 1.4.1. Let V be a vector space over a field K, a *subspace* of V is a subset $U \subset V$ such that for all $v, u \in V$ and $x \in K$ the following holds true:

- $\bar{0} \in U$,
- if $u, v \in U$ then $u + v \in U$,
- if $u \in U$ then $xu \in U$.

Remark. Note that a subspace of any vector space is also a vector space.

Definition 1.4.2. Let V be an n-dimensional inner-product space over a field K, a *hyperplane* H in V is a subspace of dimension $n - 1$. A vector $v \in V$ is said to be *normal* to H if $v \neq \bar{0}$ and $(v, h) = 0_K \ \ \forall h \in H$.

Example 1.4.3. For example in the vector space $\mathbb{R}^3$ over $\mathbb{R}$ the hyperplanes would be two dimensional subspaces, which corresponds to the intuition we have for what a plane is. And in $\mathbb{R}^2$ over $\mathbb{R}$ the hyperplanes are just lines.

Let V be a finite-dimensional inner-product space over the real numbers. Given any vector $u \in V,\ u \neq \overline{0}$ we denote the reflection in the hyperplane normal to u as s_u, we note that $s_u(u) = -u$ and for any vector v belonging to the hyperplane $s_u(v) = v$ by construction. It can easily be realized that when $V = \mathbb{R}^n$, the reflection of any given vector $x \in \mathbb{R}^n$ is

$$s_u(x) = x - \frac{2(x,u)}{(u,u)}u.$$

We see that our two desired properties are fulfilled; $s_u(u) = u - 2\frac{(u,u)}{(u,u)}u = u - 2u = -u$ and $s_u(v) = v - 2\frac{(v,u)}{(u,u)}u = v$ because $(v,u) = 0$ as they per construction are orthogonal.

Lemma 1.4.4. Let V be a finite-dimentional inner-product space over $\mathbb{R}$. For any $x, y, v \in \mathbb{R}^n$ with $v \neq \overline{0}$ the reflection s_v preserves the inner product: $(s_v(x), s_v(y)) = (x, y)$.

Proof. We use that the inner product in a real vector space is bilinear and positive definite to expand our expression and get the desired equality.

$$\begin{aligned}
(s_v(x), s_v(y)) &= \left(x - 2\frac{(x,v)}{(v,v)}v, y - 2\frac{(y,v)}{(v,v)}v\right) = \\
&= (x,y) + \left(x, -2\frac{(y,v)}{(v,v)}v\right) + \left(-2\frac{(x,v)}{(v,v)}v, y\right) + \left(-2\frac{(x,v)}{(v,v)}v, -2\frac{(y,v)}{(v,v)}v\right) = \\
&= (x,y) - 2\frac{(y,v)}{(v,v)}(x,v) - 2\frac{(x,v)}{(v,v)}(v,y) + 2\frac{(y,v)}{(v,v)} \cdot 2\frac{(x,v)}{(v,v)}(v,v) = \\
&= (x,y) - 4\frac{(x,v)(y,v)}{(v,v)} + 4\frac{(x,v)(y,v)}{(v,v)} = (x,y)
\end{aligned}$$

□

With this lemma we know all the necessary linear algebra which is required in order to define and study root systems to the extent which is done in this text.

2 Root Systems

In this section all vector spaces are assumed to be over $\mathbb{R}$ unless specifically stated otherwise. We will also use the convention

$$\langle x, v\rangle := \frac{2(x, v)}{(v, v)}$$

as it will clear up the future text. Note that $\langle x, v\rangle$ is linear with respect to the first argument, and not the second.

Definition 2.0.1. Let V be a real inner-product space, a subset R of V is called a *root system* if it satisfies the following axioms:

(R1) R is finite, it spans V and it does not contain $\overline{0}$.

(R2) If $\alpha \in R$, then the only scalar multiples of α in R are $\pm\alpha$.

(R3) If $\alpha \in R$, then the reflection s_α permutes the elements of R.

(R4) If $\alpha, \beta \in R$, then $\langle \alpha, \beta\rangle \in \mathbb{Z}$.

The elements of R are called *roots*.

We can see that the fourth axiom (R4) is quite restricting (together with the other three axioms), and to show exactly how restrictive it is we will prove the so called Finiteness Lemma. This will give us a great way to start with our classification.

2.1 Classifying Properties Using the Finiteness Lemma

Lemma 2.1.1. (Finiteness Lemma) Let R be a root system in the real inner-product space V. Let $\alpha, \beta \in R$ with $\alpha \neq \pm\beta$. Then $\langle \alpha, \beta\rangle\langle \beta, \alpha\rangle \in \{0, 1, 2, 3\}$.

Proof. We know by (R4) that the product is an integer as both $\langle \alpha, \beta\rangle$ and $\langle \beta, \alpha\rangle$ are integers, so we must establish the bounds for which integers their product could be. We also know by (R1) that α, β are non-zero, and for any two non-zero vectors $u, v \in V$ the angle θ between u and v is given by $(u, v)^2 = (u, u)(v, v)\cos^2\theta$. Using this equality and the properties of the inner product we get

$$\begin{aligned}\langle \alpha, \beta\rangle\langle \beta, \alpha\rangle &= 2\frac{(\alpha, \beta)}{(\beta, \beta)} \cdot 2\frac{(\beta, \alpha)}{(\alpha, \alpha)} = \\ &= 4\frac{(\alpha, \beta)^2}{(\alpha, \alpha)(\beta, \beta)} = 4\frac{(\alpha, \alpha)(\beta, \beta)\cos^2\theta}{(\alpha, \alpha)(\beta, \beta)} = \\ &= 4\cos^2\theta.\end{aligned}$$

This is clearly less than or equal to 4 because $0 \leq \cos^2\theta \leq 1$. Now assume that $\cos^2\theta = 1$, then θ must be an integer multiple of π which gives us $\alpha = a\beta$ for some $a \in \mathbb{R}$ contradicting our assumption that $\alpha \neq \pm\beta$ together with the root system axiom (R2). Thus $\langle \alpha, \beta\rangle\langle \beta, \alpha\rangle \in \{0, 1, 2, 3\}$. □

This lemma will prove to be very handy when starting to classify root systems as it shows us that our options are very limited. Using the Finiteness Lemma we will now explore exactly what possibilities we have for the integers $\langle \alpha, \beta \rangle$. Take any two roots α, β in a root system R with $\alpha \neq \pm\beta$, without loss of generality we may chose the labelling so that $(\beta, \beta) \geq (\alpha, \alpha)$ and thus

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

With some fairly simple calculations using $4\cos^2\theta = \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$ we see that the possibilities are:

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ	$\frac{(\beta,\beta)}{(\alpha,\alpha)}$
0	0	$\pi/2$	undefined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

This table holds a lot of information, for instance it shows us when sums and differences of two roots also lie in our root system. Let us explore this more carefully.

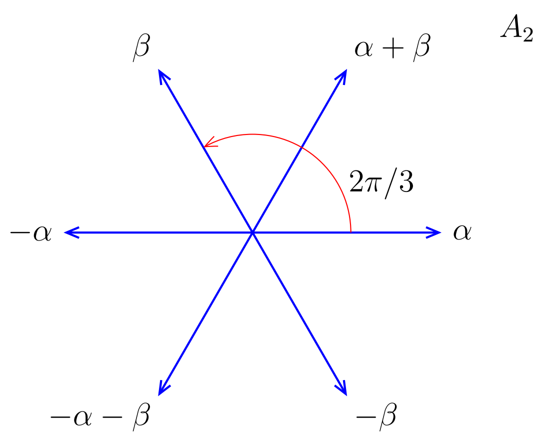
Proposition 2.1.2. Let R be a root system and α, β be roots in R.

(a) If the angle between α and β is strictly obtuse (greater than $\pi/2$), then $\alpha + \beta \in R$.

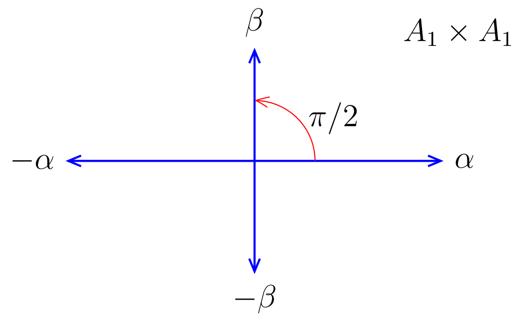
(b) If the angle between α and β is strictly acute (smaller than $\pi/2$) and $(\beta, \beta) \geq (\alpha, \alpha)$, then $\alpha - \beta \in R$.

Proof. Without loss of generality we may chose the labeling so that $(\beta, \beta) \geq (\alpha, \alpha)$. By axiom (R3) we know that $s_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta$ lies in R. The table above shows that when θ is strictly acute $\langle \alpha, \beta \rangle = 1$ which implies that $\alpha - \beta \in R$, and if θ is strictly obtuse then $\langle \alpha, \beta \rangle = -1$ which implies that $\alpha + \beta \in R$. □

Example 2.1.3. A standard example of a root system in the inner-product space $\mathbb{R}^2$ over $\mathbb{R}$ is as shown in the picture below. This root system is called a *type* A_2 system where the 2 refers to the dimension of the vector space that it spans. Here each arrow is a root.



Example 2.1.4. Suppose now that β is perpendicular to α. These two vectors gives us a root system, namely the root system known as $A_1 \times A_1$. Because $(\alpha, \beta) = 0$ the reflection s_α sends β to $-\beta$ and s_β sends α to $-\alpha$, so there is no interaction between $\pm\alpha$ and $\pm\beta$. This means that we have no information about how long β is if we know the length of α.



Naturally looking at these two examples one might ask; can we construct a basis for these systems? The short answer is yes, and it is exactly what we will do.

2.2 Basis of Root Systems

Before defining what a basis of a root system is we will firstly see if we can reduce a root system into several smaller root systems. This is of natural interest if we want to classify our root systems, because then we only need to look at the root systems that we can not reduce. This suggests the following definition.

Definition 2.2.1. The root system R is *irreducible* if R can not be expressed as a disjoint union of the two non-empty subsets $R_1 \cup R_2$ such that $(\alpha, \beta) = 0$ for all $\alpha \in R_1$ and $\beta \in R_2$.

Remark. Note that if a decomposition of R into R_1 and R_2 exists, then R_1 and R_2 are root systems in their respective spans. Example 2.1.4 is a perfect example of this, because as the name entails we can write $A_1 \times A_1 = \{\pm\alpha, \pm\beta\} = \{\pm\alpha\} \cup \{\pm\beta\}$. It is easy to check the axioms of root systems to verify that $\{\pm\alpha\}$ is a root system, and that it is perpendicular to $\{\pm\beta\}$ which is also a root system in its span.

Lemma 2.2.2. Let R be a root system in the real inner-product space V. We may write R as a disjoint union $R = R_1 \cup R_2 \cup \cdots \cup R_n$, where each R_i is an irreducible root system in the space V_i spanned by R_i, and V is a direct sum of the orthogonal subspaces $V_1, \ldots, V_n$.

Proof. Firstly we notice that the family of root system $\{R_i\}_{i=1}^n$ is a partition of R and can thus be described by an equivalence relation. Let $\sim$ be the relation on R defined as $\alpha \sim \beta$ if there exists roots $\gamma_1, \ldots, \gamma_s \in R$, where $\alpha = \gamma_1$ and $\beta = \gamma_s$ such that $(\gamma_i, \gamma_{i+1}) \neq 0$, for $1 \leq i < s$. For $\sim$ to be an equivalence relation it must be reflexive, symmetric and transitive.
To show the reflexive property of $\sim$ we let $\gamma_1 = \gamma_s = \alpha$. Then α is non-zero since it is a root. Moreover, because our inner product is positive definite $(\alpha, \alpha) \neq 0$ so $\alpha \sim \alpha$.
In order to show symmetry of $\sim$ we let $\alpha \sim \beta$. Then by definition of $\sim$ there exists $\alpha = \gamma_1, \gamma_2, \ldots, \gamma_s = \beta$ such that $(\gamma_i, \gamma_{i+1}) \neq 0$, when $1 \leq i < s$. Because the roots are in a real inner-product space $(\gamma_i, \gamma_{i+1}) = (\gamma_{i+1}, \gamma_i)$, for $1 \leq i < s$, holds, and thus $\beta \sim \alpha$.
Lastly to show transitivity let $\alpha_1 \sim \alpha_2$ and $\alpha_2 \sim \alpha_3$, per definition we then have $\alpha_1 = \gamma_1, \gamma_2, \ldots, \gamma_s = \alpha_2$ as seen before, we also have $\alpha_2 = \gamma_s, \gamma_{s+1}, \ldots, \gamma_t = \alpha_3$ for some $t \geq s$. Here is it clear that $(\gamma_i, \gamma_{i+1}) \neq 0$ hold true for $1 \leq i < s$, and, for $s \leq i < t$,

and thus also for $1 \leq i < t$.

So the relation $\sim$ is an equivalence relation and its equivalence classes constitute a partition of R, let these equivalence classes be the R_i. For each R_i to be a root system they must satisfy the four root system axioms. (R1), (R2) and (R4) all follow trivially because the vectors in R_i are themselves roots of R. For (R3) to be satisfied the reflection s_α must permute the elements of R_i when $\alpha \in R_i$, so it is enough to show that given any $\alpha \in R_i$ it is true that $s_\alpha(\beta) \in R_i$ for all $\beta \in R_i$. Given α and β in R_i we know that $\alpha \sim \beta$ which implies the existence of $\alpha = \gamma_1, \gamma_2, \ldots, \gamma_s = \beta$ all in R_i such that $(\gamma_i, \gamma_{i+1}) \neq 0$, when $1 \leq i < s$, we now need to show that $\alpha \sim \beta$ implies $\alpha \sim s_\alpha(\beta)$. Given any α and β in R_i, let us study $(\alpha, s_\alpha(\beta))$.

$$\begin{aligned}(\alpha, s_\alpha(\beta)) &= \left(\alpha, \beta - 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha\right) = \\ &= (\alpha,\beta) - 2\frac{(\alpha,\beta)}{(\alpha,\alpha)}(\alpha,\alpha) = -(\alpha,\beta)\end{aligned}$$

Assume that $(\alpha, \beta) \neq 0$, then clearly $(\alpha, s_\alpha(\beta)) \neq 0$, so given our γ_i we can see that if $(\gamma_{s-1}, \gamma_s) \neq 0$, then $(\gamma_{s-1}, s_\alpha(\gamma_s)) \neq 0$ for all $\alpha \in R_i$, which shows that when $\beta \in R_i$, then $s_\alpha(\beta) \in R_i$.

Since R_i was chosen arbitrarily we can deduce that all R_i are root systems, and that each R_i is an irreducible root system follows directly from the construction of our equivalence relation and thus our construction of R_i. Because $\bigcup_{i=1}^n R_i = R$, we can be sure that each root in R appears in some R_i and thus also in every V_i, so the sum of all V_i spans V. Thus we are left to show that the set $\{v_1, \ldots, v_k\}$, where each v_i is taken from V_i, is linearly independent. So, suppose that $v_1 + \cdots + v_k = 0$, now taking the inner product with v_j we get $0 = (v_1, v_j) + \cdots + (v_j, v_j) + \cdots + (v_k, v_j) = (v_j, v_j)$ as $(\alpha, \beta) = 0$ per construction when $\alpha \in V_a,\ \beta \in V_b,\ a \neq b$, so every $v_j = 0$ and therefore $V = V_1 \oplus \cdots \oplus V_k$. □

This lemma tells us that every root system can be decomposed into irreducible root systems. So it is indeed enough for us to classify all irreducible root systems in order to classify all root systems, and in order to do this we begin by defining what we mean by a base of a root system.

Definition 2.2.3. A subset B of R is a *base* for the root system R if

(B1) B is a vector space basis for the inner-product space V spanned by R;

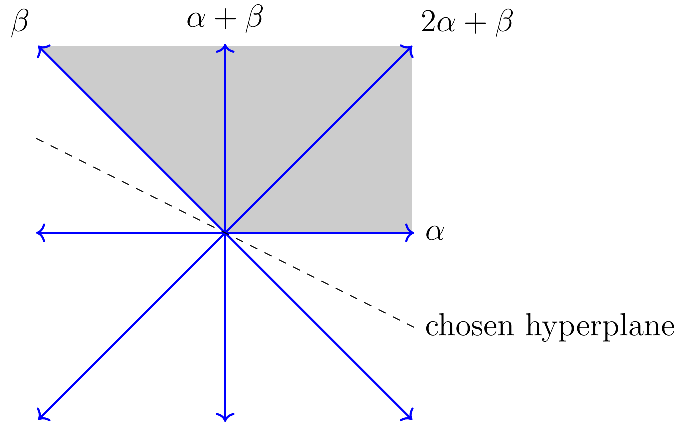
(B2) every $\beta \in R$ can be written as $\beta = \sum_{\alpha \in B} k_\alpha \alpha$ where $k_\alpha \in \mathbb{Z}$ where all the non-zero k_α have the same sign.

Lemma 2.2.4. Let R be a root system in the real inner-product space V with base B, then the angle between two distinct elements of B is obtuse.

Proof. Assume that there are two distinct roots $\alpha, \beta \in B$ with the angle $\theta < \frac{\pi}{2}$ between them Also assume without loss of generality that $(\alpha, \alpha) \leq (\beta, \beta)$. Then by Proposition 2.1.2 there exists $\gamma \in R$ such that $\alpha - \beta = \gamma$, this root γ can not belong to the basis B because B is a vector space basis and can not be linear dependent. So $\gamma \in R \setminus B$, but according to (B2) γ must then be equal to the sum $\sum_{i=1}^{\dim V} k_i \alpha_i$ where $k_i \in \mathbb{Z}$ and all non-zero k_i have the same sign. By assumption $\gamma = \alpha - \beta$ for $\alpha, \beta \in B$, but here we have one positive k_i, and one negative k_i which contradicts the definition of a base. □

Definition 2.2.5. A root $\beta \in R$ is said to be *positive with respect to B* if the coefficients given in (B2) are positive, and otherwise β is *negative with respect to B*.

Example 2.2.6. Let us look at an example to give us a clearer understanding of what really is going on here. Take the root system B_2 (shown in the picture below) and a hyperplane in the vector space spanned by our root system which does not contain any of the roots in B_2. Now we can label all the roots on one side of the hyperplane as positive and those on the other side as negative. Now if we label the roots appropriately, we can see that there is a base on the positive side of the hyperplane, namely $B = \{\alpha, \beta\}$.



Note also that the base consists of the roots which are the closest to the hyperplane.

Now we would like to see if every root system has a basis, and in order to do this we first need to prove the following lemma.

Lemma 2.2.7. Let V be a real inner-product space of dimension $n \geq 2$. Then V is not the union of finitely many hyperplanes of dimension $n - 1$.

Proof. We know by previous results that V is isomorphic with $\mathbb{R}^n$ as vector spaces over $\mathbb{R}$ because V is n-dimensional. So we shall identify V with $\mathbb{R}^n$ and consider the set

$$S = \{(1, t, \dots, t^{n-1}) \in \mathbb{R}^n \, : \, t \in \mathbb{R}\}.$$

Now any hyperplane H_i consists of points $(x_1, \dots, x_n)$ which are zeros to some non-trivial equation $a_1 x_1 + \cdots + a_n x_n = 0$ where $a_i \in \mathbb{R}$, therefore the number of elements of $S \cap H_i$ is equal to the number of solutions $t \in \mathbb{R}$ of $a_1 + a_2 t + \cdots + a_n t^{n-1} = 0$. Thus $S \cap H_i$ has less than n elements because a non-zero polynomial of degree $n - 1$ over a field has at most $n - 1$ zeros. Therefore we would need an infinite amount of hyperplanes to cover S, and since $S \subset \mathbb{R}^n$ we clearly need to cover all of S to be able to cover all of $\mathbb{R}^n$. Now because this holds for $\mathbb{R}^n$ it also holds for any V isomorphic with $\mathbb{R}^n$. □

Theorem 2.2.8. Every root system has a base.

Proof. Let R be a root system in the inner-product space V.

If V has dimension 1 we can chose the basis of the root system to be any of the two vectors in R. This is because R does not contain the 0 vector and any vector in R will therefore span V, so (B1) is satisfied.
Now pick any vector $\beta \in R$, then by (R2) together with (R3) we have $R = \{\beta, -\beta\}$. Assume that $B = \{\beta\}$ is a basis of R, then $\beta = \sum_{\alpha \in B} k_\alpha \alpha = 1 \cdot \beta$ where $k_\alpha \in \mathbb{Z}$ (in this case $k_\beta = 1$) and $-\beta = \sum_{\alpha \in B} k_\alpha \alpha = (-1) \cdot \beta$, thus (B2) is satisfied.

If V has dimension $n \geq 2$ we may choose a vector $z \in V$ which does not lie in the orthogonal space of any root in R. Such a vector must exist due to lemma 2.2.7, because we only have a finite amount of roots and thus a finite amount of orthogonal spaces of codimension 1 (hyperplanes).
Let R_z^+ denote the set of all roots $\alpha \in R$ such that $(z, \alpha) > 0$. We claim that

$$B := \{\alpha \in R_z^+ \ : \ \alpha \text{ is not the sum of two elements in } R_z^+\}$$

is a base for R.
We shall begin by showing (B2). If $\beta \in R$, then either β or $-\beta$ is in R_z^+, so we can without loss of generality say that $\beta \in R_z^+$. It is also sufficient to show (B2) for β as we only need to multiply every term of the sum by -1 to receive $-\beta$. Assume for a contradiction that there exists $\beta \in R_z^+$ which can not be expressed as $\beta = \sum_{\alpha \in B} k_\alpha \alpha$ for some $k_\alpha \in \mathbb{Z}$ with $k_\alpha \geq 0$. We may choose β such that (z, β) is as small as possible. Because β can not be expressed as $\sum_{\alpha \in B} k_\alpha \alpha$ we can see that $\beta \notin B$ holds, thus there must exist $\beta_1, \beta_2 \in R_z^+$ such that $\beta = \beta_1 + \beta_2$. By linearity of inner products we get

$$(z, \beta) = (z, \beta_1) + (z, \beta_2)$$

where $0 < (z, \beta_i) < (z, \beta)$ for $i = 1, 2$. Note that at least one of β_1 and β_2 can not be expressed as $\sum_{\alpha \in B} k_\alpha \alpha$, because otherwise $\beta_1 + \beta_2 = \beta$ could be expressed as such a sum as well. Assume without loss of generality that it is β_1 which can not be expressed as $\sum_{\alpha \in B} k_\alpha \alpha$, this contradicts the choice of β as $(z, \beta_1) < (z, \beta)$. Thus there can exist no $\beta \in R_z^+$ which can not be expressed as $\beta = \sum_{\alpha \in B} k_\alpha \alpha$, so (B2) holds.
It remains to show (B1). Because R spans all of V we know by (B2) that our proposed base B spans V because B spans R, so we must show that B is linearly independent. Firstly we note that any two distinct roots α and β in B form an obtuse angle between them by Lemma 2.2.4. Suppose that $\sum_{\alpha \in B} r_\alpha \alpha = 0$, where $r_\alpha \in \mathbb{R}$. Now we can split this sum into a sum with positive r_α and a sum for the negative. Call the sum of positive terms

$$x := \sum_{\alpha \in B \,:\, r_\alpha > 0} r_\alpha \alpha = \sum_{\beta \in B \,:\, r_\beta < 0} (-r_\beta)\beta$$

and note that the positive sum must be equal to -1 times the negative sum because of $\sum_{\alpha \in B} r_\alpha \alpha = 0$. Hence

$$(x, x) = \left(\sum_{\alpha \in B \,:\, r_\alpha > 0} r_\alpha \alpha, \sum_{\beta \in B \,:\, r_\beta < 0} (-r_\beta)\beta \right) =$$

by linearity in both arguments we get

$$= \sum_{\substack{\alpha \in B \,:\, r_\alpha > 0 \\ \beta \in B \,:\, r_\beta < 0}} r_\alpha (-r_\beta)(\alpha, \beta) \leq 0$$

and thus $(x, x) = 0$ which implies $x = 0$, because $(x, x) \geq 0$ by positive definiteness and $(x, x) \leq 0$ by the inequalities above. Therefore

$$0 = (x, z) = \sum_{\alpha \in B \,:\, r_\alpha > 0} r_\alpha (\alpha, z),$$

where each $(\alpha, z) > 0$ as $\alpha \in R_z^+$, so we must have $r_\alpha = 0$ for all α, and similarly $r_\beta = 0$ for all β. Thus our proposed B is linearly independent and spans V so it is a vector space basis, and thus it satisfies (B1). ☐

In the future we will let R^+ denote the set of all positive roots in a root system R with respect to a base B, and let R^- denote the set of all negative roots. Then $R = R^+ \cup R^-$ is a disjoint union, and the set $B \subseteq R$. We call the elements of B *simple roots* and the reflections s_α for $\alpha \in B$ *simple reflections*.

3 Group Theory and the Weyl Group

Before we study the Weyl group we need to know what a group is. We will introduce groups in a similar fashion as we did with the necessary linear algebra in the first section. That is, we will list all definitions and properties needed so that we can refer back to them later.

3.1 Group Theory

Definition 3.1.1. A *group* is an ordered pair $(G, \cdot)$ of a set G and a binary operation on G such that

(G1) the operation is associative;

(G2) there is an identity element, $1_G = 1$;

(G3) every element x of G has an inverse in G, denoted x^{-1}.

Remark. While $(G, \cdot)$ is the correct notation of a group with the operation $\cdot$ it is customary to only write G if there is no risk of misunderstanding. In the definition, however, G is the *underlying set* of the group. Moreover, we will often write $x \cdot y$ as xy.

Definition 3.1.2. A *subgroup* of a group $(G, \cdot)$ is a subset H of the set G that inherits the binary operation from $(G, \cdot)$ such that

(1) $1 \in H$;

(2) $x \in H$ implies $x^{-1} \in H$;

(3) $x, y \in H$ implies $x \cdot y \in H$.

Definition 3.1.3. A *group homomorphism* is a mapping φ from a group G to a group H such that (written multiplicatively) $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in G$.

Note that injectivity of group homomorphisms is defined precisely as injectivity of linear maps.

Definition 3.1.4. The subgroup $\langle X \rangle$ of a group G *generated by* a subset X of G is the set of all products in G (including the empty product and one-term products) of elements of X and inverses of elements of X. A group is *generated* by a set X when $\langle X \rangle = G$.

Definition 3.1.5. Let $\varphi : G \to H$ be a group homomorphism. The *image* of φ is

$$\operatorname{Im} \varphi = \{\varphi(x) \,:\, x \in G\} \subseteq H.$$

The *kernel* of φ is

$$\operatorname{Ker} \varphi = \{x \in G \,:\, \varphi(x) = 1_H\} \subseteq G.$$

Lemma 3.1.6. A group homomorphism $\varphi : G \to H$ is injective if and only if its kernel only contains the identity element.

Proof. Assume that φ is injective. First note that $\varphi(1_G) = 1_H$. Let $\varphi(x) = \varphi(1_G) = 1_H$, then by injectivity of φ we have $x = 1_G$, so $\operatorname{Ker}(\varphi) = \{1_G\}$.
Assume that $\operatorname{Ker}(\varphi) = \{1_G\}$. For $x, y \in G$ such that $\varphi(x) = \varphi(y)$ we have $\varphi(xy^{-1}) = \varphi(x)\varphi(y)^{-1} = \varphi(x)\varphi(x)^{-1} = 1_H$, we use that φ is a group homomorphism twice in the first equality and $\varphi(x) = \varphi(y)$ in the second. Thus $xy^{-1} \in \operatorname{Ker}(\varphi)$, so by assumption $xy^{-1} = 1_G$, which implies $x = y$ and thus φ is injective. □

Definition 3.1.7. A left *group action* of a group G on a set X is a mapping $G \times X \to X$ by $(g, x) \mapsto g \star x$, such that $1 \star x = x$ and $g \star (h \star x) = (gh) \star x$, for all $g, h \in G$ and $x \in X$. We say that G *acts* on the left on X.

Definition 3.1.8. For a left group action of a group G on a set X, the *orbit* of $x \in X$ is $\mathrm{Orb}_G(x) = \{y \in X \, : \, y = g \star x \text{ for some } g \in G\}$.

3.2 The Weyl Group

For each α in a root system R we have defined s_α to be the reflection about the hyperplane normal to α, so s_α acts as an invertible linear map on $V = \mathrm{span}(R)$. We may therefore construct the group generated by the reflections s_α on V for $\alpha \in R$. This is known as the *Weyl group* of R and is denoted by W or $W(R)$.

Lemma 3.2.1. The Weyl group W associated to R is finite.

Proof. By axiom (R3) the elements of W permute R, so there exists a group homomorphism from W into the group of all permutations of R associating each reflection with a permutation. We claim that this homomorphism is injective and thus W is finite.
We know by Lemma 3.1.6 that our homomorphism is injective if the kernel only contains the identity map. So, suppose that $g \in W$ is in the kernel of this homomorphism, then by definition g fixes the roots in R. But because R spans a vector space V, it follows that g must also fix all elements of V, and thus g must be the identity map.
Now because R is finite the group of all permutations of R is also finite, which together with the injectivity of our homomorphism implies that W is finite. □

Now that we have defined what the Weyl group is we want to use it. Let us begin by considering a subgroup of the Weyl group $W(R)$, namely $W_0(B) := \langle s_\gamma \, : \, \gamma \in B \rangle$ of W, where B is the base of R. We would now like to show that if the base of a root system is given, then we can recover the entire root system from it by using $W_0(B)$. Again, as it is customary to do with the Weyl group, we will write $W_0(B)$ as W_0 when there is no ambiguity.

Lemma 3.2.2. If $\alpha \in B$, then the reflection s_α permutes the set of positive roots other than α.

Proof. Suppose that $\beta \in R^+$ and $\beta \neq \alpha$. We know that $\beta = \sum_{\gamma \in B} k_\gamma \gamma$ for some $k_\gamma \geq 0$. Since $\beta \neq \alpha$ and $\beta \in R$, there is some $\gamma \in B$ with $k_\gamma \neq 0$ and $\gamma \neq \alpha$. We know by (R3) that $s_\alpha(\beta) \in R$; and from $s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$ we get the following:

$$\begin{aligned} s_\alpha(\beta) &= \sum_{\gamma \in B} k_\gamma \gamma - \left\langle \sum_{\gamma \in B} k_\gamma \gamma, \alpha \right\rangle \alpha = \\ &= \sum_{\gamma \in B} k_\gamma \gamma - \sum_{\gamma \in B} k_\gamma \langle \gamma, \alpha \rangle \alpha = \\ &= \sum_{\gamma \in B} k_\gamma (\gamma - \langle \gamma, \alpha \rangle \alpha) \end{aligned}$$

We know that $k_\gamma \geq 0$ for all γ, so it remains to show that $\gamma - \langle \gamma, \alpha \rangle \alpha$ is positive (or zero). Because both γ and α are simple roots we know by Lemma 2.2.4 that the angle between them is obtuse for all γ, and thus by the Finiteness Lemma we see that $\langle \gamma, \alpha \rangle \geq 0$ for all γ. This tells us that all the non-zero coefficients of $s_\alpha(\beta)$ as a linear combination of base elements are positive, and thus $s_\alpha(\beta) \in R^+$. □

Proposition 3.2.3. Suppose that $\beta \in R$. Then there exists $\alpha \in B$ and g in W_0 such that $\beta = g(\alpha)$.

Proof. Suppose that $\beta \in R^+$ and that $\beta = \sum_{\gamma \in B} k_\gamma \gamma$ with $k_\gamma \in \mathbb{Z}$ and $k_\gamma \geq 0$. We now define the height of β as

$$\mathrm{ht}(\beta) := \sum_{\gamma \in B} k_\gamma,$$

and proceed by using induction on the height of β. If $\mathrm{ht}(\beta) = 1$ then $\beta \in B$ so we may take $\alpha = \beta$ and g to be the identity map. For the inductive step, suppose that $\mathrm{ht}(\beta) = n \geq 2$. Because of (R2) at least two of the k_γ are strictly positive. I claim that there is a $\gamma \in B$ such that $(\beta, \gamma) > 0$, because if no γ like that exists, then all $\gamma \in B$ has the property $(\beta, \gamma) \leq 0$. This would imply that

$$(\beta, \beta) = \left(\beta, \sum_{\gamma \in B} k_\gamma \gamma\right) = \sum_{\gamma \in B} k_\gamma (\beta, \gamma) \leq 0,$$

but because inner products are positive definite we see that $(\beta, \beta) \leq 0$ gives $\beta = 0$ which is a contradiction as $0 \notin R$. We may therefore choose some $\gamma \in B$ such that $(\beta, \gamma) > 0$. Then

$$\mathrm{ht}(s_\gamma(\beta)) = \mathrm{ht}(\beta) - \mathrm{ht}(\langle \beta, \gamma \rangle \gamma) = \mathrm{ht}(\beta) - \langle \beta, \gamma \rangle.$$

We also know by the previous lemma that $s_\gamma(\beta) \in R^+$. Since $\mathrm{ht}(s_\gamma(\beta)) \leq \mathrm{ht}(\beta)$ the induction hypothesis tells us that there exist $\alpha \in B$ and $h \in W_0$ such that $s_\gamma(\beta) = h(\alpha)$. Thus we get $\beta = s_\gamma(s_\gamma(\beta)) = s_\gamma(h(\alpha))$, so we may choose $g = s_\gamma h$ which lies in W_0.
Now suppose that $\beta \in R^-$, so $-\beta \in R^+$. By the first part, $-\beta = g(\alpha)$ for some $g \in W_0$ and $\alpha \in B$. Because g is linear we get $\beta = g(-\alpha) = g(s_\alpha(\alpha))$ where $gs_\alpha \in W_0$. □

This proposition gives us exactly what we need. Now we know that any root from a root system can be recovered by only considering a basis of the associated root system along with W_0. So what does this have to do with the Weyl group, why are we only studying W_0, and not W? As we will prove using the next two lemmas, it turns out that W_0 and W are the same group.

Lemma 3.2.4. If α is a root and g is in the Weyl group W, then $gs_\alpha g^{-1} = s_{g\alpha}$.

Proof. Take any root β and we see that

$$\begin{aligned} gs_\alpha g^{-1}(\beta) &= g\left(g^{-1}(\beta) - \langle g^{-1}(\beta), \alpha \rangle \alpha\right) = \\ &= gg^{-1}(\beta) - g\left(\langle g^{-1}(\beta), \alpha \rangle \alpha\right) = \\ &= \beta - \langle \beta, g(\alpha) \rangle\, g(\alpha) = \\ &= s_{g\alpha}(\beta). \end{aligned}$$

The equality $\langle g^{-1}(\beta), \alpha \rangle = \langle \beta, g(\alpha) \rangle$ is due to Lemma 1.4.4 since g is a reflection. □

Lemma 3.2.5. Let R be a root system and B be a base of R. The Weyl group is generated by s_α for $\alpha \in B$. In other words $W_0 = W$.

Proof. By definition W is generated by the reflections s_β for $\beta \in R$, so it is sufficient to prove that $s_\beta \in W_0$ for any $\beta \in R$. By Proposition 3.2.3 we know that given β we can find $g \in W_0$ and $\alpha \in B$ such that $\beta = g(\alpha)$. Now by Lemma 3.2.4 we know that $s_\beta = s_{g(\alpha)} = gs_\alpha g^{-1}$. Now $gs_\alpha g^{-1}$ is in W_0 since $g \in W_0$ and g^{-1} is also in W_0 because W_0 is a group. And since $\alpha \in B$ we know per definition that s_α is also in W_0, and because W_0 is a group and thus closed under its operation we get $gs_\alpha g^{-1} \in W_0$. Therefore for any $\beta \in R$ we get $s_\beta \in W_0$, so $W = W_0$. □

4 First Step in Classifying Root Systems

We are equipped with the Weyl group, knowledge about root systems and their associated bases, but how do different bases of the same root system compare? And can we find any way to describe a root system without making a choice of basis? These are all questions which we shall answer in this section, and even expand on.

4.1 Cartan Matrices and Dynkin Diagrams

Firstly, let us begin by showing that different bases of a root system are essentially the same in a geometrical sense. To do this, however, we need to prove the following lemma.

Lemma 4.1.1. Let P and Q be matrices, all of whose entries are non-negative integers. If $PQ = I$, where I is the identity matrix, then P and Q are permutation matrices. That is, each row and column of P and Q has exactly one non-zero entry, namely 1.

Proof. Let $P = (p_{ij}) \in \mathbb{Z}_{\geq 0}^{m\times n}$ and $Q = (q_{ij}) \in \mathbb{Z}_{\geq 0}^{m\times n}$ where all p_{ij} and q_{ij} are non-negative integers. Then if $PQ = I$ we have that $\sum_{k=1}^{n} p_{ik}q_{kj}$ is the ij-th element of I. Now assume that we have one row in P with two non-zero entries, p_{ab} and p_{ac}, and that $PQ = I$. We know because of $PQ = I$ that at every row and column in both P and Q needs to have at least one non-zero element. However when we calculate the a-th row in I we only want q_{ba} or q_{ca} to be non-zero, not both, let's say without loss of generality that $q_{ba} \neq 0$. Then all other columns in Q must be zero in the c-th position, because otherwise we would have two non-zero entries in the a-th row in the identity matrix, which is not allowed. However if every column in Q is zero in the c-th position, then all elements of the c-th row in Q are equal to zero, also contradiction that $PQ = I$. We can also see that q_{ba} and p_{ab} must be equal to 1, because if one is bigger than 1 their product will also be bigger than one due to them being integers.
This argument can be made with assuming two elements in a column being non-zero in P, and respectively for Q as well. Therefore both P and Q must be permutation matrices. □

Theorem 4.1.2. Let R be a root system and suppose that B and B' are two bases of R. Then there exists an element g in the Weyl group $W(R)$ such that $B' = g(B)$.

Proof. Let $B = \{\alpha_1, \dots, \alpha_n\}$ and α be a root. Then for any $w \in W(R)$, because the Weyl group permutes roots, $w^{-1}(\alpha)$ is also a root. Suppose that $w^{-1}(\alpha) = \sum_{i=1}^{n} k_i\alpha_i$ where all coefficients k_i have the same sign. Then $\alpha = w\left(\sum_{i=1}^{n} k_i\alpha_i\right) = \sum_{i=1}^{n} k_i w(\alpha_i)$ with the same coefficients and because α and w were chosen arbitrarily $w(B)$ must also be a base. This shows that the Weyl group permutes the collection of all bases of R.
Let R^+ denote the set of positive roots of R with respect to B. Let $B' = \{\alpha'_1, \dots, \alpha'_n\}$ be another base of R. Let R'^+ denote the positive roots of R with respect to B', and let R'^- denote the negative roots of R with respect to B'. Note that all of these sets have the size $|R|/2$.
We shall now do the proof by using induction on the size of $|R^+ \cap R'^-|$.

The base case is when $R^+ \cap R'^- = \emptyset$, in this case $R^+ = R'^+$ so the bases B and B' give the same positive roots. Each element of B' is a positive root with respect to B, so we may define a matrix P with entries p_{ij} by

$$\alpha'_j = \sum_{i=1}^{n} p_{ij}\alpha_i,$$

whose coefficients are all non-negative integers. Similarly we may define a matrix Q with entries p_{ij} by

$$\alpha_k = \sum_{i=1}^{n} q_{ik}\alpha_i'.$$

Note that because $R^+ = R'^+$ we know that for each α_j' there exists an α_i such that $\alpha_j' = \alpha_i$, and therefore $q_{ji} = 1$ as well as $p_{ij} = 1$. The rest of the terms in the respective row and column will be 0, and therefore only the jj-th term in the matrix product PQ will survive and be equal to $q_{ji} \cdot p_{ij} = 1$. Because this is true for all α_j' we see that $PQ = I = QP$, and by Lemma 4.1.1 Q and P must be permutation matrices. Hence the sets B and B' coincide, and for the element w we can choose the identity and get $w(B) = B'$.

Now suppose that $|R^+ \cap R'^-| = n > 0$. Then $B \cap R'^- \neq \emptyset$ since otherwise $B \subseteq R'^+$, which implies $R^+ \subseteq R'^+$ and therefore $R^+ = R'^-$ because they have the same size. Take some $\alpha \in B \cap R'^-$, then by Lemma 3.2.2 we know that $s_\alpha(R^+)$ is the set of roots obtained from R^+ by replacing α with $-\alpha$, namely $s_\alpha(R^+) = (R^+ \setminus \{\alpha\}) \cup \{-\alpha\}$. So the intersection $s_\alpha(R^+) \cap R'^-$ has $n-1$ elements. The set $s_\alpha(R^+)$ is the set of positive roots with respect to the base $s_\alpha(B)$. By the induction hypothesis, there is some w_1 in the Weyl group such that $w_1(s_\alpha(B)) = B'$. Now take $w = w_1 s_\alpha$; this sends B to B'. □

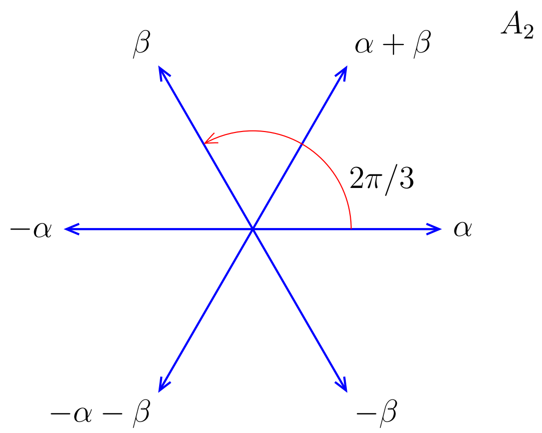
Note that because all elements $g \in W(R)$ are themselves reflections they preserve, by Lemma 1.4.4, the inner product of roots in R. Therefore by $(\alpha, \beta) = |\alpha||\beta| \cos\theta$, the angle between roots is also preserved, and in particular the angle between simple roots are preserved. So with this we note that all different bases of the same root system are the same, up to the labelling of the simple roots. We will now explore this concept a little bit deeper.

Definition 4.1.3. Let B be a base in a root system R and fix an order on the elements of B, say $(\alpha_1, \dots, \alpha_n)$. The *Cartan matrix* of R is the $n \times n$ matrix with ij-th entry $\langle \alpha_i, \alpha_j \rangle$.

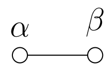
For any roots $\beta, \alpha_i, \alpha_j$ we have, by Lemma 1.4.4, that $\langle s_\beta(\alpha_i), s_\beta(\alpha_j) \rangle = \langle \alpha_i, \alpha_j \rangle$. Now we see that together with Theorem 4.1.2 the Cartan matrix is independent of the base chosen for R. However the Cartan matrix still depends on the ordering chosen on the elements of B. Thus a Cartan matrix of a root system is not free of choice, in other words there is not a unique Cartan matrix for each root system.

To find a unique way to represent the base elements of a root system R we look at graphs $\Delta = \Delta(R)$ defined as follows. The vertices of Δ are labelled by the simple roots of R, and between the vertices we draw $d_{\alpha\beta}$ lines where $d_{\alpha\beta} := \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}$. When $d_{\alpha\beta} > 1$ the two roots have different lengths and are not orthogonal, so we draw an arrow pointing from the longer root to the shorter one. This diagram is called the *Dynkin diagram* of R and we will later show in Proposition 4.2.2 that it is unique for R. The same graph without any arrows is called the *Coxeter graph* of R.

Example 4.1.4. Let us consider the root system A_2 again.



Here we can clearly see, as noted previously, that $\{\alpha, \beta\}$ is a base of this root system. So let us see what the Dynkin diagram of A_2 looks like, we know by the consequences of the Finiteness Lemma that we have $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 1 = d_{\alpha\beta}$ edges between α and β. So we get the Dynkin diagram;



.

4.2 Isomorphisms of Root Systems

Now that we know how to represent the base of a root system as a Dynkin diagram, let us see how we can compare root systems. In particular we will see how root systems compare when they have the same Dynkin diagram.

Definition 4.2.1. Let R and R' be root systems in the real inner-product spaces V and V', respectively. We say that R and R' are isomorphic if there is a vector space isomorphism $\varphi : V \to V'$ such that

(a) $\varphi(R) = R'$, and

(b) for any two roots $\alpha, \beta \in R$, $\langle \alpha, \beta \rangle = \langle \varphi(\alpha), \varphi(\beta) \rangle$.

Remark. Note that if θ is the angle between two roots α and β, then $4\cos^2\theta = \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$, so condition (b) ensures that angles between roots are preserved.

It follows from the definition of root system isomorphisms that if R and R' are isomorphic, then they have the same size due to vector space isomorphisms being bijections together with criterion (a). Now because the Dynkin diagram of a root system is constructed using the integers $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$ we see that by criterion (b), they will also have the same Dynkin diagram. We will now prove the converse.

Proposition 4.2.2. Let R and R' be root systems in the real vector spaces V and V', respectively. If the Dynkin diagram of R and R' are the same, then the root systems are isomorphic.

Proof. We may choose bases $B = \{\alpha_1, \ldots, \alpha_n\}$ in R and $B' = \{\alpha'_1, \ldots, \alpha'_n\}$ in R' so that for all i, j we have

$$\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle.$$

Let $\varphi : V \to V'$ be the linear map which maps α_i to α_i'. By definition, this is a vector space isomorphism satisfying condition (b) in Definition 4.2.1. We must therefore show that $\varphi(R) = R'$.

Let $v \in V$ and $\alpha_i \in B$, then

$$\begin{aligned}\varphi(s_{\alpha_i}(v)) &= \varphi\left(v - \langle v, \alpha_i\rangle \alpha_i\right) = \\ &= \varphi(v) - \langle v, \alpha_i\rangle \alpha_i'.\end{aligned}$$

I claim now that $\langle v, \alpha_i\rangle = \langle \varphi(v), \alpha_i'\rangle$. We use that $\langle -, -\rangle$ is linear in its first component, and that B is a vector space basis of V, and that we therefore can express v as $\sum_{\alpha\in B} k_\alpha \alpha$ where all $k_\alpha \in \mathbb{R}$.

$$\begin{aligned}\langle v, \alpha_i\rangle = \left\langle \sum_{\alpha\in B} k_\alpha \alpha, \alpha_i \right\rangle &= \sum_{\alpha\in B} k_\alpha \langle \alpha, \alpha_i\rangle = \\ &[\text{Now we use that } \langle \alpha_i, \alpha_j\rangle = \langle \alpha_i', \alpha_j'\rangle] \\ &= \sum_{\alpha\in B} k_\alpha \langle \alpha', \alpha_i'\rangle = \left\langle \sum_{\alpha\in B} k_\alpha \alpha', \alpha_i' \right\rangle = \\ &= \left\langle \sum_{\alpha\in B} k_\alpha \varphi(\alpha), \alpha_i' \right\rangle = \left\langle \varphi\left(\sum_{\alpha\in B} k_\alpha \alpha\right), \alpha_i' \right\rangle = \\ &= \langle \varphi(v), \alpha_i'\rangle\end{aligned}$$

Therefore we may write $\varphi(s_{\alpha_i}(v)) = s_{\alpha_i'}(\varphi(v))$.
By Lemma 3.2.5 the simple reflections s_{α_i} generate the Weyl group of R. So we have that for any $g \in W(R)$, g can be written as $g(v) = s_{\alpha_{\sigma(1)}} \cdots s_{\alpha_{\sigma(k)}}(v)$ for some permutation σ of $\{1, \ldots, k\}$, where $k \leq n$. Now, by $\varphi(s_{\alpha_i}(v)) = s_{\alpha_i'}(\varphi(v))$ we have

$$\begin{aligned}\varphi(\mathrm{Orb}_{W(R)}(v)) \ni \varphi(g(v)) &= \varphi(s_{\alpha_{\sigma(1)}} \cdots s_{\alpha_{\sigma(k)}}(v)) = \\ &= s_{\alpha'_{\sigma(1)}} \varphi(s_{\alpha_{\sigma(2)}} \cdots s_{\alpha_{\sigma(k)}}(v)) = \\ &\vdots \\ &= s_{\alpha'_{\sigma(1)}} \cdots s_{\alpha'_{\sigma(k)}} \varphi(v) \in \mathrm{Orb}_{W(R')}(\varphi(v)).\end{aligned}$$

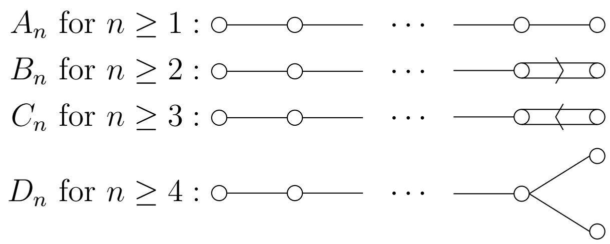
Hence $\varphi(\mathrm{Orb}_{W(R)}(v)) \subseteq \mathrm{Orb}_{W(R')}(\varphi(v))$. We also know that $\varphi(B) = B'$ because B is the chosen vector space basis of V and B' is the chosen vector space basis of V' and φ is a vector space isomorphism from V to V'. All of this, together with Proposition 3.2.3, tells us that we must have $\varphi(R) \subseteq R'$.
Applying the same argument on φ^{-1} gives us $\varphi^{-1}(R') \subseteq R$, therefore $\varphi(R) = R'$, as required. $\square$

With this proposition we only need to classify every Dynkin diagram associated to root systems in order to classify all root systems, up to isomorphism.

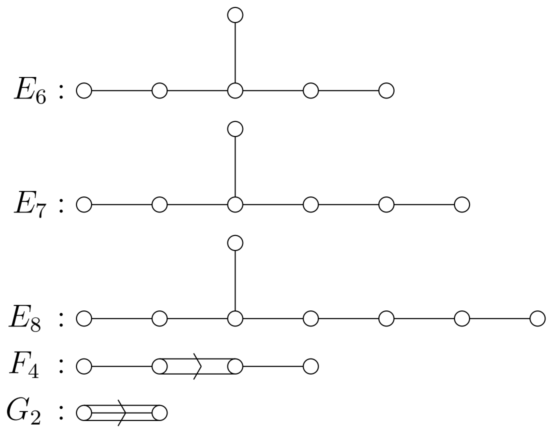
5 Classification of Dynkin diagrams

The goal of this section is to prove the following theorem.

Theorem 5.0.1. Given an irreducible root system R, the unlabelled Dynkin diagram associated to R is either of one of the four families;



where each of the diagrams above has n vertices, or one of the five exceptional diagrams



Note that there is no repetition in the list. For example, because C_2 would be the same diagram as B_2 it is not included.

Definition 5.0.2. Let V be a real inner-product space with inner product $(-,-)$. A subset A of V consisting of linearly independent vectors $v_1, v_2, \ldots, v_n$ is said to be *admissible* if it satisfies the following conditions:

(a) $(v_i, v_i) = 1$ for all i and $(v_i, v_j) \leq 0$ if $i \neq j$.

(b) If $i \neq j$, then $4(v_i, v_j)^2 \in \{0, 1, 2, 3\}$.

To the admissible set A we associate the graph Γ_A with vertices labelled by the vectors $v_1, \ldots, v_n$, and with $d_{ij} := 4(v_i, v_j)^2$ edges between v_i and v_j for $i \neq j$.

We now want to show that for each base of a root system, we can find an associated admissible set. Because then it will be enough to study the graphs associated with admissible sets, and later show that every such graph is associated with a root system.

Lemma 5.0.3. Suppose that B is a base of a root system and consider the set defined as $A := \{\alpha/\sqrt{(\alpha, \alpha)} : \alpha \in B\}$. Then A is admissible and the graph Γ_A is the Coxeter graph of B.

Proof. Let $\alpha, \beta \in B$, then $\alpha/\sqrt{(\alpha,\alpha)} \in A$ and $\beta/\sqrt{(\beta,\beta)} \in A$. Then

$$\left(\frac{\alpha}{\sqrt{(\alpha,\alpha)}}, \frac{\alpha}{\sqrt{(\alpha,\alpha)}}\right) = \frac{(\alpha,\alpha)}{\sqrt{(\alpha,\alpha)}\sqrt{(\alpha,\alpha)}} =$$

$$= \frac{(\alpha,\alpha)}{(\alpha,\alpha)} = 1,$$

so property (a) of admissible sets is fulfilled. Let us check property (b), here we use the symmetry of $(-,-)$ in the first equality:

$$4\left(\frac{(\alpha,\beta)}{\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}}\right)^2 = \frac{4(\alpha,\beta)(\beta,\alpha)}{\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}} =$$

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

Which by the Finiteness Lemma is in $\{0,1,2,3\}$, so property (b) is also fulfilled.

Now we must show that the graph Γ_A is indeed the Coxeter graph of B. This follows directly from our previous equality showing that

$$4\left(\frac{(\alpha,\beta)}{\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}}\right)^2 = \langle\alpha,\beta\rangle\langle\beta,\alpha\rangle$$

because $4\left((\alpha,\beta)/\left(\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}\right)\right)^2$ is exactly the amount of edges between the vertices α and β in Γ_A, and $\langle\alpha,\beta\rangle\langle\beta,\alpha\rangle$ is exactly the amount of edges between α and β in the Coxeter graph of B. □

Before the next lemma, note that any subset of an admissible set is also an admissible set. This is very easy to realize by looking at the definition of admissible sets, so we will not prove this.

Lemma 5.0.4. The number of pairs of vertices joined by at least one edge is at most $|A| - 1$.

Proof. Suppose that $A = \{v_1 \ldots, v_n\}$. Set $v = \sum_{i=1}^n v_i$. As A is linearly independent we know that $v \neq 0$. Hence

$$\begin{aligned}(v,v) &= \left(\sum_{i=1}^n v_i, \sum_{i=1}^n v_i\right) \\ &= \sum_{i=1}^n (v_i,v_i) + 2\sum_{i<j}(v_i,v_j) \\ &= n + 2\sum_{i<j}(v_i,v_j),\end{aligned}$$

because $(v_i, v_i) = 1$ for all i. Thus

$$n > \sum_{i<j} -2(v_i,v_j) = \sum_{i<j}\sqrt{d_{ij}} \geq N,$$

where $d_{ij} = 4(v_i, v_j)^2$ for $i \neq j$ is the number of vertices between v_i and v_j. Thus N is the number of pairs $\{v_i, v_j\}$ such that $d_{ij} \geq 1$, this is the number we are looking for and it satisfies $n = |A| > N$. □

Corollary 5.0.5. The graph Γ_A does not contain any cycles.

Proof. Suppose that Γ_A has a cycle. Let A' be the subset of A consisting of all the vectors involved in the cycle. Then A' is also an admissible set with the corresponding graph $\Gamma_{A'}$, this graph has at least as many edges as vertices, which is a contradiction to the previous lemma. □

Lemma 5.0.6. No vertex of Γ_A is incident to four or more edges.

Proof. Take a vertex v of Γ_A, and let $v_1, v_2, \ldots, v_n$ be all the vertices in Γ_A joined to v. Since Γ_A can not contain any cycles, $(v_i, v_j) = 0$ for all $i \neq j$. Consider the subspace U with the basis $v_1, \ldots, v_n, v$, the Gram-Schmidt process allows us to extend $v_1, \ldots, v_n$ to an orthonormal base of U by adding a vector v_0 satisfying $(v, v_0) \neq 0$. Now we may express v as a linear combination of our new basis elements:

$$v = \sum_{i=0}^{n} (v, v_i) v_i$$

By assumption v comes from an admissible set, and therefore $1 = (v, v) =$
$= (\sum_{i=0}^{n} (v, v_i) v_i, \sum_{i=0}^{n} (v, v_i) v_i)$ where only the (v_i, v_i) terms survive in the sums for all i. We can therefore rewrite it as

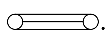
$$\begin{aligned}(v, v) &= \left(\sum_{i=0}^{n} (v, v_i) v_i, \sum_{j=0}^{n} (v, v_j) v_j \right) = \\ &= \sum_{\substack{i=1 \\ j=1}}^{n} ((v, v_i) v_i, (v, v_j) v_j) = \\ &= \sum_{i=0}^{n} (v, v_i)^2 (v_i, v_i) = \\ &[(v_i, v_i) = 1 \text{ per definition of admissible sets.}] \\ &= \sum_{i=0}^{n} (v, v_i)^2 = 1\end{aligned}$$

Because $(v, v_0)^2 > 0$ we see that

$$\sum_{i=1}^{n} (v, v_i)^2 < 1.$$

Now as A is admissible and $(v, v_i) \neq 0$, we know by property (b) of Definition 5.0.2 that $(v, v_i)^2 \geq \frac{1}{4}$ for $1 \leq i \leq n$. Hence $n \leq 3$. □

Then the following corollary follows trivially.

Corollary 5.0.7. If Γ is connected and has a triple edge, then $\Gamma =$ . □

Lemma 5.0.8. (Shrinking Lemma) Suppose Γ_A has a subgraph which is a *line*, that is, of the form

where there are no multiple edges between the vertices. Define $A' = (A \backslash \{v_1, v_2, \ldots, v_k\}) \cup \{v\}$ where $v = \sum_{i=1}^{k} v_i$. Then A' is admissible and the graph $\Gamma_{A'}$ is obtained from Γ_A by shrinking the line to a single vertex.

Proof. It is clear that A' is linearly independent, so it remains to check the properties of the inner product on vectors in A'. By the properties of admissible sets we see that

$$(v_i, v_{i+1}) = \begin{cases} -1 & \text{if } 1 \leq i < k, \\ 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

This allows us to calculate (v, v) as follows:

$$(v, v) = \left(\sum_{i=1}^{k} v_i, \sum_{j=1}^{k} v_j \right) = \sum_{\substack{i=1 \\ j=1}}^{k} (v_i, v_j) =$$
$$= \sum_{i=1}^{k} (v_i, v_i) + \sum_{i=1}^{k-1} (v_i, v_{i+1}) = k - (k-1) = 1$$

So property (a) of admissible sets holds for A'.

Suppose that $w \in A$ and $w \neq v_i$, for $1 \leq i \leq k$. Then w is joined to at most one of $v_1, \ldots, v_k$, because otherwise there would be a cycle. Therefore either $(w, v) = 0$ if w is not joined to any of the vectors $v_1, \ldots, v_k$, or

$$(w, v) = \left(w, \sum_{1=1}^{k} v_i \right) = \sum_{i=1}^{k} (w, v_i) = (w, v_j)$$

where v_j for some $j \in \{1, \ldots, k\}$ is the vertex that w is joined with. Then $4(w, v)^2 = 4(w, v_j)^2 \in \{0, 1, 2, 3\}$, so property (b) of admissible sets holds and we conclude that A' is an admissible set. □

Definition 5.0.9. Let Γ be a graph associated with an admissible set, then we say that a vertex in Γ is a *branch vertex* if it is incident to three or more edges.

Remark. By Lemma 5.0.6 we know that a branch edge is incident to exactly three edges.

Lemma 5.0.10. The graph Γ has

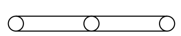
(i) no more than one double edge;

(ii) no more than on branch vertex; and

(iii) not both a double edge and a branch vertex.

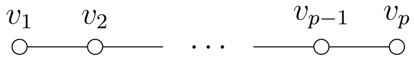
Proof. Assume that Γ has more than one double edge, then because Γ is connected, there exists a subgraph consisting of two double edges connected by a line as shown below.

By the Shrinking Lemma, we obtain an admissible set with graph



which contradicts Lemma 5.0.6. The proofs of (ii) and (iii) follow the exact same idea. □

Lemma 5.0.11. Suppose that Γ has a line as a subgraph:



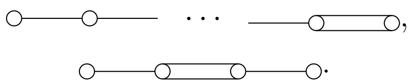
Let $v = \sum_{i=1}^{p} i v_i$. Then $(v, v) = \frac{p(p+1)}{2}$.

Proof. As in the Shrinking Lemma we note that $2(v_i, v_{i+1}) = -1$, for $1 \le i < p$, and that $(v_i, v_j) = 0$, for $i \ne j$ otherwise, due to the shape of Γ. So,

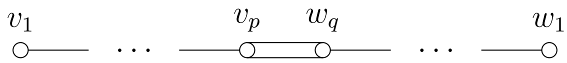
$$\begin{aligned}
(v, v) &= \left(\sum_{i=1}^{p} i v_i, \sum_{j=1}^{p} j v_j \right) = \sum_{\substack{i=1 \\ j=1}}^{p} ij(v_i, v_j) = \\
&= \sum_{i=1}^{p} i^2 (v_i, v_i) + \sum_{i=1}^{p-1} 2i(i+1)(v_i, v_{i+1}) = \\
&= \sum_{i=1}^{p} i^2 - \sum_{i=1}^{p-1} i^2 + i = \sum_{i=1}^{p} i^2 - \sum_{i=1}^{p-1} i^2 - \sum_{i=1}^{p-1} i = \\
&= p^2 - \sum_{i=1}^{p-1} i = p^2 - \frac{p(p-1)}{2} = \\
&= \frac{2p^2 - p^2 + p}{2} = \frac{p(p+1)}{2}.
\end{aligned}$$

□

Proposition 5.0.12. If Γ has a double edge, then Γ is one of



Proof. By Lemma 5.0.10, any such Γ has the form



where without loss of generality $p \ge q$. Let $v = \sum_{i=1}^{p} i v_i$ and $w = \sum_{i=1}^{q} i w_i$, then by Lemma 5.0.11 we have

$$(v, v) = \frac{p(p+1)}{2}, \quad (w, w) = \frac{q(q+1)}{2}.$$

We can see by the form of the graph that $4(v_p, w_q)^2 = 2$ and $(v_i, w_j) = 0$ in all other cases. Therefore by linearity of the inner product we have

$$(v, w)^2 = (p v_p, q w_q)^2 = \frac{p^2 q^2}{2}.$$

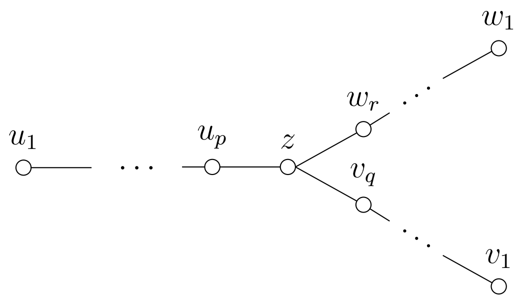
As v and w are linearly independent, the Cauchy-Schwartz inequality implies $(v, w)^2 < (v, v)(w, w)$. Through some substitution and simplification of the previous inequality we get $2pq < (p+1)(q+1)$, and hence

$$2pq < pq + p + q + 1 \quad \text{which gives us} \quad pq - p - q < 1.$$

So either $q = 1$ and $p \ge 1$ or $p = q = 2$. □

Proposition 5.0.13. If Γ has a branch vertex, then either Γ is D_n for some $n \geq 4$ or Γ is E_6, E_7 or E_8.

Proof. By Lemma 5.0.10, any such Γ has the form



where, without loss of generality, $p \geq q \geq r$. We must show that either $q = r = 1$ or $q = 2$, $r = 1$, and $p \leq 4$.

Let $u = \sum_{i=1}^{p} iu_i$, $v = \sum_{i=1}^{q} iv_i$ and $w = \sum_{i=1}^{r} iw_i$. Then we have $(u, v) = (v, w) = (u, w) = 0$, so u, v and w are pairwise orthogonal. Let $\hat{u} = u/\sqrt{(u, u)}$, $\hat{v} = v/\sqrt{(v, v)}$ and $\hat{w} = w/\sqrt{(w, w)}$. Because u, v, w and z are linearly independent, the space spanned by u, v, w, z has an orthonormal basis

$$\{\hat{u}, \hat{v}, \hat{w}, z_0\}$$

for some choice of z_0 which satisfies $(z, z_0) \neq 0$. We may write $z = (z, \hat{u})\hat{u} + (z, \hat{v})\hat{v} + (z, \hat{w})\hat{w} + (z, z_0)z_0$. By taking the inner product with z on both sides of the equality above we get

$$(z, z) = (z, \hat{u})^2 + (z, \hat{v})^2 + (z, \hat{w})^2 + (z, z_0)^2 = 1,$$

where due to the properties of admissible sets $(z, z) = 1$. Now because $(z, z_0) \neq 0$, we get

$$(z, \hat{u})^2 + (z, \hat{v})^2 + (z, \hat{w})^2 < 1.$$

By Lemma 5.0.11 we know the lengths of u, v, w. So we have

$$(z, \hat{u})^2 = \left(z, \frac{u}{\sqrt{(u, u)}}\right)^2 = \frac{(z, pu_p)^2}{(u, u)} = $$
$$= \frac{p^2(z, u_p)^2}{\frac{p(p+1)}{2}} = \frac{2p^2}{4p(p+1)},$$

and similarly for v and w. Substituting this into our previous inequality gives

$$\frac{2p^2}{4p(p+1)} + \frac{2q^2}{4q(q+1)} + \frac{2r^2}{4r(r+1)} < 1.$$

By expanding and re-writing this inequality in several steps we get that it is equivalent to

$$\frac{1}{p+1} + \frac{1}{q+1} + \frac{1}{r+1} > 1.$$

Since $p \geq q \geq r \geq 1$, we have $\frac{1}{p+1} \leq \frac{1}{q+1} \leq \frac{1}{r+1} \leq \frac{1}{2}$. Therefore, by using the inequality above we see that $\frac{3}{r+1} > 1$, and hence $r < 2$, so we must have $r = 1$. Now by using that $r = 1$ we get that $\frac{2}{q+1} + \frac{1}{2} > 1$, and therefore $\frac{1}{q+1} > \frac{1}{4}$, so $q < 3$. Assume that $q = 2$, then we have $\frac{1}{p+1} + \frac{1}{3} + \frac{1}{2} > 1$, which is equivalent to $\frac{1}{p+1} > \frac{1}{6}$, so $p < 5$. On the other hand, if $q = 1$ we get $\frac{1}{p+1} > 0$, so we have no restriction on p when $q = r = 1$. $\square$

We have now found all connected graphs which come from admissible sets. Let us return to our Dynkin diagram in Theorem 5.0.1. Call this Dynkin diagram Δ, we saw in Lemma 5.0.3 that the Coxeter graph of Δ, say $\bar{\Delta}$, must appear in our collection. If Δ has no multiple edges, then, by Proposition 5.0.13, $\Delta = \bar{\Delta}$ is one of the graphs listed in Theorem 5.0.1.

If Δ has a double edge, then Proposition 5.0.12 tells us that there are two possibilities for $\bar{\Delta}$. In the case of B_2 and F_4 it does not matter in which way we put the arrow, otherwise there are two different choices, giving B_n and C_n for $n \geq 2$.

Finally if Δ has a triple edge, then Corollary 5.0.7 tells us that $\Delta = G_2$. This completes the proof of Theorem 5.0.1.

We are now done with the goal of this paper, we have studied root systems and shown their relation to the Weyl group and Dynkin diagrams. In turn we classified all graphs associated with admissible sets and showed that admissible sets are correlated with root systems, and that the graphs associated with admissible sets are indeed the Dynkin diagrams associated with root systems. Therefore this concludes our classification of root systems.

References

[EW] Erdmann, Karin; Wildon, Mark J. *Introduction to Lie Algebras.* Springer Undergraduate Mathematics Series. Springer-Verlag London, Ltd., London, 2006.