Representations of the Lie algebras $\mathfrak{sl}_2\mathbb{C}$ and $\mathfrak{sl}_3\mathbb{C}$ and physical applications

Luuk Hendrikx

July 21, 2013

Contents

Introduction

Symmetry in nature is one of the many connections between mathematics and physics. One may develop a algebraic way to work with symmetry in mathematics. One example of this are Lie groups and algebras. Ever since the 20th century symmetry became more important in physics. From Noether's theorem, to the theory of spin of particles, symmetry is important in many parts of physics.

As soon as technology allowed physicist to build sufficiently large detectors to observe a wide range of exotic particles, the search began for a way to classify the particles. One way to do this is by using the symmetry properties of the specific system. This is done by looking at the symmetry group of the Hamiltonian, thereby constructing so called multiplets. In this thesis we focus specifically on the way in which elementary particles fit into symmetry multiplets of several symmetry groups, as first described by physicist Gell-Mann.

We start by exploring the general theory of Lie groups and Lie algebras and their representations. We focus on a particular subgroup of Lie algebras, the semisimple ones. These semisimple Lie algebras have particularly nice properties whih make them easier to study. Then we focus on representations of two important Lie algebras, $\mathfrak{sl}_2\mathbb{C}$ and $\mathfrak{sl}_3\mathbb{C}$. Especially their irreducible representations will be of great interest for their applications in the classification of particles. We aim to therefore fully classify and understand these irreducible representations.

We then turn our attention to the symmetry in physics by considering elementary particles. Using empirical data, one may classify these particles into groups with similar properties. We then apply our knowledge of the representation theory of Lie algebras to provide a explanation for this structure. We will focus on symmetry groups $SU(2)$ and $SU(3)$. Therefore the knowledge obtained from the latter part will prove very useful. Using the notion of symmetry and representations one might classify the elementary particles in so-called multiplets. Though this profound connection may seem suprising and somewhat arbitrary, one may set up a theory (quark model) in which an explanation is given for the multiplet structure of the elementary particles.

We then proceed to apply this model to several kinds of particles, the baryons and mesons. We classify them into multiplets using the structure of the Lie algebra $\mathfrak{sl}_3\mathbb{C}$.

Finally, we quickly go over the disadvantages of the quark model. Furthermore, we briefly explain how one might extend the machinery developed in this thesis to study larger Lie algebras and thereby larger symmetry groups resulting in a more general setup applicable to many physical theories. The quark model is an elementary example of the use of representation theory and symmetry in physics; many theories are extensions or modifications of some of the theory presented here.

1 Lie Groups and Lie Algebras

1.1 Lie Groups

Lie groups are important objects in many parts of mathematis and physics. Intuitively, Lie groups combine the notion of a group and a manifold into one object. We will define them, and give some examples. Then we look at what a representation is and what these representations look like for certain Lie groups. We start by defining what a Lie group is.

Definition 1. *A Lie group G is a smooth manifold endowed with the structure of a group such that this structure is compatible with the manifold structure. That is to say multiplication $\times : G \times G \to G$ and inversion $\iota : G \to G$ are smooth maps.*

Many terms from both groups and manifolds can be applied to Lie groups by regarding the associated properties. We can therefore talk about connected Lie groups in terms of its manifold structure, or normal subgroups of a Lie group by studying its group structure. In similar fashion one may define maps between Lie groups to preserve both of these structures. We then take

Definition 2. *A map $\rho : G \to H$ between Lie groups is a map that is both differentiable and a group homomorphism: $\rho(g)\rho(h) = \rho(gh)$ for all $g, h \in G$*

Examples of Lie groups are plenty. Most Lie groups we discuss in this thesis will be realized as matrix groups, hence we will focus on these. A first example would be $GL_n\mathbb{K}$. This is the Lie group of invertible $n \times n$ matrices over a field $\mathbb{K}$. We will often use $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. It inherits a manifold structure from the $n \times n$ matrices by taking the entries as coordinates and is an open subset of the $n \times n$ matrices by continuity of the determinant function. The smoothness of multiplication and inversion is easily checked. Multiplication arises as polynomials in the coordinates, which is smooth, while the smoothness of the inverse operation follows from Cramer's rule for computing inverses. All of the groups that we will be using can be viewed as subgroups of $GL_n\mathbb{R}$. They arise in two different ways.

First, one can describe a subgroup using an equation defined on the entries of the matrix. Let us take $\mathbb{K} = \mathbb{R}$ for the moment. An example of a subgroup constructed in this way is the following one. Consider the subgroup of $GL_n\mathbb{R}$ of invertible $n \times n$ matrices by taking all matrices $X \in GL_n\mathbb{R}$ satisfying $det(X) = 1$. One may check that this is indeed a Lie group. It is often denoted $SL_n\mathbb{R}$.

Alternatively, we may describe them as subgroups of this $GL_n\mathbb{R}$, preserving some structure on $\mathbb{K}^n$. Denote $\mathbb{R}^n = V$. The group $SO_n\mathbb{R}$ is defined as the subgroup of $SL_n\mathbb{R}$ preserving a bilinear symmetric non-degenerate form $Q : V \times V \to \mathbb{C}$. By this we mean that $Q(v, w) = Q(Av, Aw)$ for all $v, w \in V$. We can make this more explicit by writing Q as $Q(v, w) = v^t \cdot M \cdot w$ for some matrix M. Using this definition, preservation of Q is equivalent to $v^t \cdot M \cdot w = v^t \cdot A^t \cdot M \cdot A \cdot w$ for all $v, w \in V$ and all $A \in SO_n\mathbb{R}$. This means $M = A^t \cdot M \cdot A$. Now taking $M = Id$ the form Q has all the desired properties, and we may describe
$SO_n\mathbb{R} = \{A \in Mat_{n\times n} : det(A) = 1, A^t \cdot A = Id\}$ which is often a more convenient description.

1.2 Representations of Lie groups

We will primarily study representations of Lie groups and algebras. We start by defining what a representation on a group is.

Definition 3. *Given a group G and a finite dimensional (complex) vector space V, a representation of G on V is a homomorphism $\rho : G \to GL(V)$*

Often we will call V the representation. This proves convenient for concise notation, while the meaning is in most cases obvious. We can then define maps between different representations. These maps will have to be G-linear maps of vector spaces to be maps of representations. That is, given two representations V and W of a group G, a map $\psi : V \to W$ is a map of representations the following diagram commutes for all $g \in G$

$$\begin{tikzcd} V \arrow[r, "\psi"] \arrow[d, "g"] & W \arrow[d, "g"] \\ V \arrow[r, "\psi"] & W \end{tikzcd} \tag{1}$$

We will need to develop some more terminology. First of all there is the notion of a subrepresentation.
Given a representation V a *subrepresentation* W is a vector subspace of V that is invariant under the action of G: $\rho(W) \subset W$. Whether a representation contains such subrepresentations is a interesting subject. Of course, any representation V has two subrepresentations: V itself and the trivial one $\{0\}$.
If the representation contains no other subrepresentations we call it *irreducible.* If it does, we call V *reducible.* We will later see that any reducible representation may be decomposed as a direct sum of irreducible subrepresentations in which case we call the representation V *completely reducible.*

Another basic notion is how one may manipulate representations to create new representations. These are operations on the underlying vector spaces that create new vector spaces. These are representations by defining the action of G accordingly. We will discuss a few examples of such operations and their associated representations. We can later use this information.

- Given two representations V and W their tensor product $V \otimes W$ is a representation where the homomorphism is defined as $\rho_V \otimes \rho_W : G \to GL(V \otimes W)$ by $\rho_V \otimes \rho_W(g)(v \otimes w) = \rho_V(g)(v) \otimes \rho_W(g)(w) \in V \otimes W$
- Given two representations V and W their direct sum $V \oplus W$ is a representation where the homomorphism is defined as $\rho_V \oplus \rho_W : G \to GL(V \oplus W)$ by $\rho_V \oplus \rho_W(g)(v \oplus w) = \rho_V(g)(v) \oplus \rho_W(g)(w) \in V \oplus W$
- Given a representation V, its dual[1] V^* is a representation where the homomorphism is $\rho^* : G \to GL(V^*)$ by setting $\rho^*(g)(\alpha)(v) = \alpha(\rho(g^{-1})(v)$, where $\alpha \in V^*$. Another way to state this is to regard $\rho^*(g) \in GL(V^*)$ as a matrix in some basis of V^*. We then have in terms of the matrix $\rho(g) \in GL(V)$ that $\rho^*(g) = \rho(g^{-1})^t$.

It is easily checked that the representations as defined above indeed define genuine representations. Using these operations on representations, we can construct new representations from given ones. We will later construct such new representations and study their reducibility.

[1]The space dual space of V, denoted V^*, is defined as the vector space of all linear maps $\lambda : V \to \mathbb{K}$ where we will most often use $\mathbb{K} = \mathbb{C}$

1.3 Lie Algebras

Another subject of interest will be the so called Lie algebras. These are objects that correspond to Lie groups, in the sense we can associate to any matrix Lie group a Lie algebra in a convenient way. However, we will first define them, and then concentrate on the connection.

Definition 4. *A Lie algebra $\mathfrak{g}$ is a vector space over a field $\mathbb{K}$ together with a product map, sometimes called the (Lie) bracket, $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ with the following properties:*

- Bilinearity: $[\lambda X + \mu Y, Z] = \lambda[X, Z] + \mu[Y, Z]$ *for all* $\lambda, \mu \in \mathbb{K}$ *and all* $X, Y, Z \in \mathfrak{g}$
- Skew-symmetry: $[X, Y] = -[Y, X]$ *for all* $X, Y \in \mathfrak{g}$
- Jacobi identity: $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$ *for all* $X, Y, Z \in \mathfrak{g}$

Similar to the case of Lie groups, we want maps to retain the structure on the Lie algebra.
A map $\psi : \mathfrak{g} \to \mathfrak{h}$ is a *map of Lie algebras* if it preserves the bracket:

$$\psi([X, Y]) = [\psi(X), \psi(Y)] \text{ for all } X, Y \in \mathfrak{g}$$

Subalgebras are also defined similarly calling $\mathfrak{h} \subset \mathfrak{g}$ a *Lie subalgebra* if $[X, Y] \in \mathfrak{h}$ for all $X, Y \in \mathfrak{h}$. We can also define representations of these objects on vector spaces, as we did for Lie groups.

Definition 5. *Given a Lie algebra $\mathfrak{g}$ and a finite dimensional (complex) vector space V a representation of $\mathfrak{g}$ on V is a map of Lie algebras $\rho : \mathfrak{g} \to \mathfrak{g}l(V)$*

This means it is a linear map that preserves brackets: $\rho([X, Y]) = [\rho(X), \rho(Y)]$. Here, the bracket on the Lie algebra $\mathfrak{g}l(V)$, the endomorphisms on V, is just the commutator of matrices ($[X, Y] = XY - YX$), as we shall see in the next section. There are many Lie algebras with a broad variety of properties.
In this thesis we will focus on the so called "semisimple" Lie algebras. For completeness we give the definition of a semisimple Lie algebra, as we will prove and use two important properties of such Lie algebras. To define these semisimple Lie algebras, we need the notion of an ideal.

Definition 6. *Given a Lie algebra $\mathfrak{g}$, an* ideal *is a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ with the property that $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{h}$, $Y \in \mathfrak{g}$*

The idea of an ideal is similar to that of a normal subgroup in group theory, in the sense that a quotient by an ideal will again be an algebra.
We then define a special chain of ideals $\{\mathscr{D}^k\mathfrak{g}\}$ constructed inductively by setting $\mathscr{D}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and defining $\mathscr{D}^k\mathfrak{g} = [\mathscr{D}^{k-1}\mathfrak{g}, \mathscr{D}^{k-1}\mathfrak{g}]$. Then we call an algebra or ideal solvable if $\mathscr{D}^k\mathfrak{g} = 0$ for some k. Now one calls a Lie algebra $\mathfrak{g}$ semisimple is it has no solvable ideals. The Lie algebras we focus on will be semisimple ones because of two very convenient properties that we will see in due time.

1.4 Lie groups and their Lie algebras

In this section we will establish a relation between Lie groups and algebras. More specifically, we cam endow the tangent space at the identity of a Lie group with the structure of a Lie algebra. Let us look at the map given by conjugation by an element $g \in G$: :$\Psi_g : G \to G$, where

$h \mapsto ghg^{-1}$. If $\rho : G \to H$ is a homomorphism, we know the following diagram commutes.

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, "\Psi_g"'] & H \arrow[d, "\Psi_{\rho(g)}"] \\ G \arrow[r, "\rho"'] & H \end{tikzcd} \tag{2}$$

We now look at the differential of the conjugation map at the identity $e \in G$, and denote it

$$(d\Psi_g)_e = Ad(g) : T_eG \to T_eG, \text{ so } Ad : G \to Aut(T_eG)$$

As T_eG is a vector space, this defines a representation, the *adjoint representation.* Using this definition and the fact that ρ is a homomorphism, the following diagram also commutes.

$$\begin{tikzcd} T_eG \arrow[r, "(d\rho)_e"] \arrow[d, "Ad(g)"'] & T_eH \arrow[d, "Ad(\rho(g))"] \\ T_eG \arrow[r, "(d\rho)_e"'] & T_eH \end{tikzcd} \tag{3}$$

If we like dependence only on the differential $(d\rho)_e$, as Ad still depends on ρ itself, then we need to define the differential of the map Ad. We denote $ad : T_eG \to End(T_eG)$. Then the following diagram commutes

$$\begin{tikzcd} T_eG \arrow[r, "(d\rho)_e"] \arrow[d, "ad(d\rho(v))"'] & T_eH \arrow[d, "ad(d\rho(v)))"] \\ T_eG \arrow[r, "(d\rho)_e"'] & T_eH \end{tikzcd} \tag{4}$$

We thereby conclude that for any homomorphism ρ the differential $(d\rho)_e$ respects the adjoint action on the tangent space of G. Note that a representation ρ is in particular a map of Lie groups, thus a homomorphism. In fact, there is a 1-1-correspondence between representations of simply connected Lie groups and their associated Lie algebras. That is, the converse of the above construction also works given that the Lie group G is simply connected. A proof of this theorem might be found in [6] in section 3.6 for matrix Lie groups. We will implicitely use this as representations of Lie algebras are more convenient to work with.
Now we wish to make T_eG compatible with the structure of a Lie algebra. This is then the Lie algebra $\mathfrak{g}$ associated with a Lie group G. This can be accomplished by using the map ad above. Note ad is a multilinear map, as $ad(X) \in End(T_eG)$, hence $ad(X)(Y) \in T_eG$. We can therefore see ad as a product map $T_eG \times T_eG \to T_eG$. By setting $[X, Y] = ad(X)(Y)$ we have our bracket. We shall focus on the special case of matrix groups, for which we now show that the bracket given by the map ad is indeed the same as the commutator, as stated before.

All our matrix groups are subgroups of $GL_n\mathbb{K}$. Here, we can extend the conjugation map Ψ to the entire ambient space $End(\mathbb{K}^n)$. Its tangent space at e will again be $End(\mathbb{K}^n)$. This means the differential becomes a conjugation $Ad(g)(M) = gMg^{-1}$. We can now look at the differential of Ad by looking at paths.
A way to define the tangent space of a manifold is to define the tangent vectors as equivalence classes of speeds of the paths at time zero. For the tangent space T_xG, take paths $\gamma : [0, 1] \to G$ based at $x \in G$ and pick a chart $\phi : U \to \mathbb{R}^n$ where $x \in U$ is an open neighborhood. Furtheremore $\phi \circ \gamma$ is differentiable at 0. Then two such paths γ and δ are equivalent if the differentials at 0 of

$\phi \circ \gamma$ and $\phi \circ \delta$ coincide. The equivalence classes of these paths are then the tangent vectors and the collection of all tangent vectors forms the tangent space.
This means, given $X, Y \in T_e GL_n\mathbb{K}$ and a path $\gamma : [0,1] \to GL_n\mathbb{K}$, such that $\gamma(0) = e$ and $\dot{\gamma}(0) = X$. We then have

$$\begin{aligned}
[X, Y] &= ad(X)(Y) \\
&= \frac{d}{dt}|_{t=0} Ad(\gamma(t))(Y) \\
&= \frac{d}{dt}|_{t=0} \gamma(t) Y \gamma(t)^{-1} \\
&= \dot{\gamma}(0) Y \gamma(0) + \gamma(0) Y (-\gamma(0)^{-1} \dot{\gamma}(0) \gamma(0)^{-}1 \\
&= XY - YX
\end{aligned}$$

We see in the case that we are dealing with a subgroup of $GL_n\mathbb{K}$, the Lie bracket is the commutator of matrices. This is a very concrete way to describe the Lie bracket.
Using the definition of Lie algebras associated with a Lie group, we can describe several examples. First of all, we saw the example of the Lie group $SL_n\mathbb{C}$.

Secondly, using this notion, we may also define the tensor products and direct sums as we did for representations of Lie groups, only now for their corresponding algebras. Let $\gamma(t)$ be a path in G based at e, and $\dot{\gamma}(0) = X$ with $X \in \mathfrak{g}$, its associated Lie algebra. Let V and W be two representations of G. The induced representation of $\mathfrak{g}$ is defined by the action of X on V: $X(v) = \frac{d}{dt}|_{t=0}\gamma(t)(v)$. Then we have the following definitions for the tensor product, direct sum and dual space representations, where we use the derivative properties of these constructions.

- Given represenations V and W, the action of X on their tensor product $V \bigotimes W$ is given by $X(v \otimes w) = X(v) \otimes w + v \otimes X(w)$
- Given representations V and W the action of X on their direct sum $V \bigoplus W$ is given by $X(v \oplus w) = X(v) \oplus X(w)$
- Given a representation $\rho : \mathfrak{g} \to \mathfrak{g}l(V)$ on V, the action of X on its dual is given by $\rho' : \mathfrak{g} \to \mathfrak{g}l(V^*)$ where $\rho'(X) = -\rho(X)^t$.

When we discuss representations of several Lie algebras, the knowledge of how these operations construct new representations will prove to be useful.

1.5 Some properties of semisimple Lie algebras

We discuss two important properties of semisimple Lie algebras. The first property has to do with reducibility. One may prove that any representation of a semisimple Lie algebra is in fact completely reducible. However, in order to prove this important and fundamental theorem, we need to introduce some more mechanisms to work with Lie algebras.

We shall first study a bilinear form on the algebra $\mathfrak{g}$ called the *Killing form* K. Note that for any $X \in \mathfrak{g}$ we have the map $ad(X) : \mathfrak{g} \to \mathfrak{g}$. Then we have $ad(X) \in \mathfrak{g}l(\mathfrak{g})$. We can define the Killing form K on $X, Y \in \mathfrak{g}$ as

$$K(X, Y) = Tr(ad(X) \circ ad(Y))$$

One can immediatly see linearity from this definition, as it follows from linearity of ad and the trace. We will need a property of this Killing form which we will prove.

Lemma 1. *Given a Lie algebra* $\mathfrak{g}$ *and the letting* K *be the Killing form on the algebra defined as above. Then we have* $K([X,Y],Z] = K(X,[Y,Z])$ *for all* $X, Y, Z \in \mathfrak{g}$.

Proof. This is a consequence of properties of the trace of endomorphisms. We have for any endomorphisms R, S, T of a vector space that $Tr(SRT) = Tr(RTS)$ by the cyclic property of the trace. Then using linearity of the trace we have

$$\begin{aligned} Tr([R,S]T] &= Tr((RS - SR)T) \\ &= Tr(RST - SRT) \\ &= Tr(RST - SRT + RTS - RTS) \\ &= Tr(RST - RTS) = Tr(R[S,T]) \end{aligned}$$

Then this also holds for the Killing form as $ad(X)$ is a endomorphism of $\mathfrak{gl}(\mathfrak{g})$ for any $X \in \mathfrak{g}$ and the map ad preserves brackets. □

We will use this lemma in the proof of complete reducibility. We will also need a property which is harder to prove, but it will be useful. A full proof may be found in many books on representation theory (among which proposition C.10 in [1], page 20 in [2] and Theorem 9.2 in [3]), but we will only state it here. It is a consequence of a more general theorem known as Cartan's criterion.

Theorem 1. *A Lie algebra* $\mathfrak{g}$ *is semisimple if and only if its Killing form* K *is nondegenerate.*

Furthermore, we will need a lemma called Schur's lemma which is not so difficult to prove, but will prove to be very useful to extract properties of representations.

Lemma 2. Schur's lemma
Let V *and* W *be irreducible representations of a group* G*, and* $\phi : V \to W$ *a* G*-module homomorphism then either* ϕ *is an isomorphism, or* ϕ*=0. If* $V = W$*, then* $\phi = \lambda I$*, for* $\lambda \in \mathbb{C}$ *and* I *the identity.*

Now we are ready to prove one of the most fundamental theorems about semisimple Lie algebras; they are completely reducible. This will allow us to study these semisimple algebras in great detail.

Theorem 2. Complete Reducibility
Let V *be a finite dimensional representation of a semisimple Lie algebra* $\mathfrak{g}$*. Suppose there is a subspace* $W \subset V$ *that is invariant under the action of* $\mathfrak{g}$*:* $\rho(X)(w) \in W$ *for all* $X \in \mathfrak{g}$ *and all* $w \in W$*. Then there is a subspace* $U \subset V$ *such that* $V = W \oplus U$ *and where* U *is also invariant under the action of* $\mathfrak{g}$

To prove this theorem we need some more instruments.
Assume $\mathfrak{g} \subset \mathfrak{gl}(V)$. As $\mathfrak{g}$ is a vector space, let $U_1, \ldots, U_r$ be a basis. We also take a dual basis $U^1, \ldots, U^r$ in the sense that the Killing form has the property that $K(U_i, U^j) = \delta_i^j$ (K is nondegenerate by the preceding theorem).
We then define an operator on the vector space V called a Casimir operator defined as $C_V(v) = \Sigma U_i \cdot U^i \cdot v$. This operator has several properties. However, for concise notation in this proof we introduce the Einstein summation convention, in which summation over repeated indices is

implied. We can then write $C_V = U_i U^i$.
We first show that C_V commutes with the action of $\mathfrak{g}$. By linearity of the commutator, we are done if we can show $[U_n, C_V] = 0$ for all U_n.

Lemma 3. *Let C_V be the Casimir operator as defined above and $\{U_n\}$ a basis with $\{U^n\}$ its dual basis with respect to the Killing form. Then we have $[U_n, C_V] = 0$ for all U_n.*

Proof. For any U_i, U_j we have $[U_i, U_j] = c_{ij}^k U_k$. These coefficients c_{ij}^k are called structure constants. Note that by skew symmetry of the bracket we have $c_{ij}^k = -c_{ji}^k$. We wish to know how the basis $\{U_i\}$ and its dual $\{U^j\}$ commute. We do this using the above theorem and lemma.
We have $K(U_i, [U_j, U^k]) = -K([U_j, U_i], U^k) = -c_{ji}^k = -K(U_i, c_{jl}^k U^l)$. Now using nondegeneracy of K as $\mathfrak{g}$ is semisimple, we have $[U_j, U^k] = c_{lj}^k U^l$.
We compute

$$\begin{aligned}
[U_n, C_V] &= [U_n, U_i U^i] \\
&= U_n U_i U^i - U_i U^i U_n - U_i U_n U^i + U_i U_n U^i \\
&= [U_n, U_i] U^i + U_i [U_n, U^i] \\
&= c_{ni}^l U^i + c_{ln}^i U^l
\end{aligned}$$

The summation indices i and l are dummy variables, so switching their roles in the rightmost term we get $[U_n, C_V] = c_{ni}^l U^i + c_{in}^l U^i = 0$, as required. □

We use the Casimir for the actual proof of the complete reducibility theorem. We now give this proof.

Proof. Complete Reducibility
First of all, let us assume that the representation we are dealing with is faithful (that is, the group homomorphism is injective). If this is not the case, we can quotient out this kernel without affecting reducibility.

We start with the case where $W \subset V$ is an irreducible submodule of codimension 1. The Casimir operator acts as multiplication on W, by Schur's lemma. We have $Tr(C_V) \neq 0$; this is a nonzero scalar. Then V/W is 1-dimensional and a $\mathfrak{g}$-submodule. Now all 1-dimensional representation of semisimple Lie algebras are trivial as $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, so C_V acts on V/W as a scalar 0. This means for $C_V : V \to V$ has a 1-dimensional kernel; $Ker(C_V)$ is 1-dimensional, and we have

$$V = W \oplus Ker(C_V)$$

As C_V commutes with the U_n, $Ker(C_V)$ is a 1-dimensional $\mathfrak{g}$-module.

Now suppose W is not irreducible, but still of codimension 1. Then we have a (maximal) submodule $Z \subset W$. We then have that $W/Z \subset V/Z$ is an irreducible submodule, with codimension 1. Then using the previous result, we find a complementary submodule

$$V/Z = W/Z \oplus Y/Z$$

where Y/Z is a 1-dimensional submodule of V/Z. Now as $dim(Y) < dim(V)$, so by an induction argument on the dimension of V, which is trivial to prove when $dim(V) = 1$, we have that $Y = Z \oplus U$ by induction. Note immediatly that $U \cap W = \{0\}$. Then we have

$$V = W \oplus U$$

Let W now be any submodule of lower dimension. We saw reducibility does not matter by the previous argument finding a maximal subalgebra Z, so assume W irreducible. Consider the restriction map: $\rho : Hom(V, W) \to Hom(W, W)$. We can realize the action of $\mathfrak{g}$ on these spaces by taking for $X \in \mathfrak{g}$ and $f \in Hom(V, W)$ the action defined by $(Xf)(v) = X(f(v)) - f(X(v))$ for all $v \in V$.
Then consider the space of $\mathfrak{g}$ invariant homomorphisms, $Hom_{\mathfrak{g}}(W, W)$. Then $\rho^{-1}(Hom_{\mathfrak{g}}(W, W)) \subset Hom(V, W)$ are the homomorphisms that restrict to scalar multiplication on W. Let g be equal to the identity on W and zero elsewhere. Then given some element $h \in \rho^{-1}(Hom_{\mathfrak{g}}(W, W))$, it can be written as $h = f_0 + \lambda g$, where $f_0 \in Ker(\rho)$.
This space $Ker(\rho)$ then has codimension one in $\rho^{-1}(Hom_{\mathfrak{g}}(W, W))$. By the previous case, $\rho^{-1}(Hom_{\mathfrak{g}}(W, W)) = Ker(\rho) \oplus U$. By this composition, U is mapped surjectively into $Hom_{\mathfrak{g}}(W, W)$ (as ρ is surjective). We must therefore be able to pick ψ in U such that $\rho(\psi) = Id$. This implies $Ker(\psi) \cap W = 0$. Then we have

$$V = W \oplus Ker(\psi)$$

□

There is a different way to prove this using compact forms in [6].
In this way, any reducible representation V may be decomposed in a finite number of irreducible representations by repeating this process until we are left with only irreducible representations. Hence we can write $V = \bigoplus W_\alpha$ where each W_α is an irreducible subrepresentation of $\mathfrak{g}$.

The second important property is the conservation of Jordan-Chevalley decomposition (hereafter called Jordan decomposition). We first recall the term Jordan decomposition. An operator X_s is called semisimple when it is diagonizable. Equivalent (over algebraically closed fields) is to say V is spanned by eigenvectors of the operator X_s or that any X_s invariant subspace $W \subset V$ had an X_s invariant complement $W' \subset V$ such that $V = W \oplus W'$.
An operator X_n is called nilpotent if $(X_n)^k = 0$ for some k. Let X be any linear operator acting on a finite dimensional vector space V over $\mathbb{C}$ (it will work for any perfect field, of which $\mathbb{C}$ is an example).
Then we may decompose $X = X_s + X_n$ which is known as the Jordan-Chevalley decomposition of an operator. Here X_s is a semisimple operator and X_n a nilpotent one. Furthermore, this decomposition is unique, X_s and X_n commute and they may be expressed as polynomials in X.

If a semisimple Lie algebra $\mathfrak{g}$ is a subset of $\mathfrak{gl}(V)$ it is clear what we mean by its Jordan decomposition as it is an algebra of linear endomorphisms.
For general semisimple Lie algebras we can also define an abstract Jordan decomposition. Given $X \in \mathfrak{g}$, then we can decompose $ad(X) \in \mathfrak{gl}(\mathfrak{g})$ as $ad(X) = ad(S) + ad(N)$ where $ad(S)$ and $ad(N)$ are the semisimple and nilpotent parts of $ad(X)$ in the usual sense. This decomposition is also unique. In fact we have $S, N \in \mathfrak{g}$. This is explained in [2] at the beginning of section 5.6. We will now state the theorem.

Theorem 3. Preservation of Jordan decompostion
Let $\mathfrak{g}$ be a semisimple Lie algebra. Then for every $X \in \mathfrak{g}$, there exists a unique decomposition $X = S + N$ of X (abstract Jordan decompostion), such that for any representation on a finite

dimensional vector space V we have for $\rho(X) \in \mathfrak{gl}(V)$ that $\rho(X)_s = \rho(S)$ and $\rho(X)_n = \rho(N)$, where $\rho(X)_n$ and $\rho(X)_s$ are respectively the semisimple and nilpotent parts of $\rho(X) \in \mathfrak{gl}(V)$ (usual Jordan decomposition).

We will need two facts that are stated in [2]: lemma A in 4.2 and proposition 4.2.C. They say the following.

1. For $X \in \mathfrak{gl}(V)$ with usual Jordan decomposition $X = X_s + X_n$, then $ad(X) = ad(X_s) + ad(X_n)$ is the Jordan decomposition of $ad(X)$.
2. Given a vector space V, with $A \subset B \subset V$ subspaces and X maps B in to A, then both X_s and X_n map B into A

We can now prove the theorem. This will be done in two parts.

Proof. We will proceed in the following way. First we show that if $\mathfrak{g} \subset \mathfrak{gl}(V)$, then for $X \in \mathfrak{g}$ we also have $X_s, X_n \in \mathfrak{g}$. Note that by uniqueness of the decompositions and by virtue of fact (1), the abstract and usual Jordan decomposition for elements in $\mathfrak{g}$ coincide as it implies that both $ad(X_s)$ and $ad(X_n)$ are semisimple, hence X_s and X_n are the unique candidates for S and N given that they lie in $\mathfrak{g}$. So if we prove that X_s and X_n are indeed in $\mathfrak{g}$ we have that the two Jordan decompositions coincide. We are then able to prove the full theorem.

So, let $\mathfrak{g} \subset \mathfrak{gl}(V)$ and $X \in \mathfrak{g}$ an arbitrary element. For semisimple Lie algebras, we have $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$ and thereby $[X, \mathfrak{g}] \subset \mathfrak{g}$. Then we have immediatly $[X_s, \mathfrak{g}] \subset \mathfrak{g}$ and $[X_n, \mathfrak{g}] \subset \mathfrak{g}$ by fact (1). Define the normalizer of $\mathfrak{g}$ as $\mathfrak{n} = \{A \in \mathfrak{gl}(V) : [A, X] \in \mathfrak{g} \quad \forall X \in \mathfrak{g}\}$. Then we have $X_s, X_n \in \mathfrak{n}$. Note that $\mathfrak{n}$ includes $\mathfrak{g}$ as an ideal.

Now for any $\mathfrak{g}$-submodule $W \subset V$ (subspace invariant under the action of $\mathfrak{g}$), define the subalgebra of $\mathfrak{gl}(V)$

$$\mathfrak{s}_W = \{Y \in \mathfrak{gl}(V) : Y(W) \subset W, Tr(Y|_W) = 0\}$$

For example, if V were irreducible, the only subalgebra of this type would be $\mathfrak{s}_V = \mathfrak{sl}(V)$. For an arbitrary representation V there will generally be more such $\mathfrak{g}$-submodules W. The algebra $\mathfrak{g}$ is a subalgebra of the $\mathfrak{s}_W$ for any W by definition of $\mathfrak{g}$-submodule and the fact that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ forcing the elements of $\mathfrak{g}$ to be traceless. Furthermore, by fact (2), both X_s and X_n preserve the $\mathfrak{g}$-submodule W. Also X_n is traceless (as is any nilpotent endomorphism) which forces X_s to be traceless as well. Therefore X_n and X_s are both in $\mathfrak{s}_W$.

Let then $\mathfrak{g}' = \mathfrak{n} \cap (\bigcap_W \mathfrak{s}_W)$. Here the intersection runs over all $\mathfrak{g}$-submodules W. We will show $\mathfrak{g} = \mathfrak{g}'$. For any $X \in \mathfrak{g}'$ both X_s and X_n are in $\mathfrak{g}'$ as they are in $\mathfrak{n}$ and in all $\mathfrak{s}_W$. If we can prove that $\mathfrak{g}' = \mathfrak{g}$ we are done.

By complete reducibility, we have $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{h}$. This is the case as $\mathfrak{g}'$ is a finite dimensional $\mathfrak{g}$-module. However, we have $\mathfrak{g}' \subset \mathfrak{n}$. By definition of $\mathfrak{n}$ we then have $[\mathfrak{g}, \mathfrak{g}'] = \mathfrak{g}$, we must have $[\mathfrak{g}, \mathfrak{h}] = 0$. For any irreducible $\mathfrak{g}$-module, we have for any $Y \in \mathfrak{h}$ that $[\mathfrak{g}, Y] = 0$. By Schur's lemma Y restricted to W is a scalar, but $Y \in \mathfrak{s}_W$. This means $Tr(Y|_W) = 0$ and hence $Y|_W = 0$. This holds for any irreducible submodule, hence $Y = 0$. This means $\mathfrak{h} = 0$ and thus $\mathfrak{g} = \mathfrak{g}'$ and we are done.

The theorem now follows from this. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be the map of Lie algebras associated to a finite dimensional representation of a semisimple Lie algebra $\mathfrak{g}$. Then let us look at the algebra $\rho(\mathfrak{g}) \subset \mathfrak{gl}(V)$. By definition of semisimplicity, noting that $ad(S) \in \mathfrak{gl}(\mathfrak{g})$, we have that $\mathfrak{g}$ is spanned by eigenvectors of $ad(S)$, and therefore $\rho(\mathfrak{g})$ is spanned by eigenvectors of $ad(\rho(S))$

by preservation of the bracket. Then $ad(\rho(S))$ is semisimple.

By a similar argument (preservation of bracket) the nilpotency of $ad(N)$ assures that $ad(\rho(N))$ is also nilpotent. Furthermore $ad(\rho(S))$ and $ad(\rho(N))$ commute, by using the Jacobi identity and preservation of brackets under ρ. This means that $\rho(X) = \rho(S) + \rho(N)$ is the abstract Jordan decomposition of $\rho(X)$ by definition. Then knowing $\rho(X) \in \mathfrak{gl}(V)$ we showed that the decompositions coincide, hence this is also the usual Jordan decompostion. That is, $\rho(X)_s = \rho(S)$ and $\rho(X)_n = \rho(N)$. $\square$

This is in fact a very useful theorem. If we know some element can be diagonalized in some faithful representation, we know the particular choice of representation is not important. In any representation this element will be diagonizable. This means we can fully study the structure of semisimple Lie algebras by looking at specific faithful representations. The properties we distill from these representations will then often carry over to any arbitrary other representation, providing a very general analysis. We will use this fact in the representations of the Lie algebras we study.

2 Classification of representations of the Lie algebras $\mathfrak{s}l_2\mathbb{C}$ and $\mathfrak{s}l_3\mathbb{C}$

By using the above theorems for semisimple Lie algebras, we can study in great detail some of such Lie algebras. Two important Lie algebras are $\mathfrak{s}l_2\mathbb{C}$ and $\mathfrak{s}l_3\mathbb{C}$. Both are semisimple and both have many uses in the field of physics. One such application we shall discuss in this thesis in detail. In this section, we shall build the needed machinery to fully understand the representations of these Lie algebras. Special attention will be given to their irreducible representations, and how one may extract the decomposition of irreducible components from a given representation.

2.1 The Lie algebra $\mathfrak{s}l_2\mathbb{C}$

The Lie algebra $\mathfrak{s}l_2\mathbb{C}$ will prove of great importance within the scope of this thesis. It is one of the simplest cases of a semisimple Lie algebra, it is indeed even a simple algebra. Knowledge about its stucture can be used in many other semisimple Lie algebras, as will be demonstrated in due time. It is an important algebra as its representations are in 1-1 correspondence to representations of the group $SU(2)$ which is often used in physics. Representations of $SU(2)$ are in 1-1 correspondence with its Lie algebra $\mathfrak{s}u(2)$ as it is simply connected. Then these representations are in 1-1 correspondence to representations of its complexification, which is the algebra $\mathfrak{s}l_2\mathbb{C}$. We will not elaborate on these relations and just focus on the algebra $\mathfrak{s}l_2\mathbb{C}$.
The Lie algebra is $\mathfrak{s}l_2\mathbb{C}$ can be realized as the Lie algebra of traceless 2×2 matrices where the product is the commutator of matrices. We pick a basis for this algebra. The algebra itself will be three dimensional because of the restriction on the trace. We pick the following basis:

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

Every traceless 2×2 matrix may indeed be written as a linear combination of these three. Because the Lie bracket is bilinear and skew-symmetric, writing down the different brackets (also called commutation relations as in this context the bracket is the commutator) of the basis elements. We have:

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad [E, F] = H$$

Now we are ready to use one of the properties of semisimple Lie algebras. The Jordan decomposition is preserved, meaning that for any irreducible finite-dimensional representation V of $\mathfrak{s}l_2\mathbb{C}$, H will again be diagonizable.

Using this, one can immediatly decompose V accordingly into linear eigenspaces of the (diagonal) action of H. We denote these different eigenspaces by which eigenvalue H has on them. That is, we write $V = \bigoplus V_\alpha$, such that for $v \in V_\alpha$ we have $H(v) = \alpha v$. Here, the α are just a finite collection of complex numbers. Of course, we wish to know how the other basis elements act on the eigenspaces V_α. We can calculate this using the commutation relations. For example for F

we have:

$$\begin{aligned} H(F(v)) &= F(H(v)) + [H,F](v) \\ &= F(\alpha v) - 2F(v) \\ &= (\alpha - 2)F(v) \end{aligned}$$

Hence, we have that $F : V_\alpha \to V_{\alpha-2}$. Similarly, for E we find that $E : V_\alpha \to V_{\alpha+2}$. Because of finite dimensionality of V this must break at some point. That is, at some point we must have $E^k(v) = 0$ and $F^l(v) = 0$ for some k and l. We get several eigenspaces that are mapped to each by the action of the spaces E and F. We will call these eigenspaces a string.
Furthermore, what this tells us is that the space $V = \bigoplus_{n\in\mathbb{Z}} V_{\alpha+2n}$ is invariant under the action of $\mathfrak{sl}_2\mathbb{C}$, so it must be the whole space by irreducibility of V. So all V_α that appear can be reached by application of the operators E and F. The spaces form an unbroken string of eigenspaces: only at the ends of the string the action of E and F kills the spaces. We may summerize the action of the basis elements in a diagram:

$$\begin{tikzcd} {} \arrow[r, bend left, "E"] & V_{\alpha-2} \arrow[l, "F"'] \arrow[r, bend left, "E"] \arrow[loop below, "H"] & V_\alpha \arrow[l, "F"'] \arrow[r, bend left, "E"] \arrow[loop below, "H"] & V_{\alpha+2} \arrow[l, "F"'] \arrow[loop below, "H"] \end{tikzcd}$$

By finite dimensionality we know there must be some maximal α, such that $V_\alpha \neq 0$. Let us denote this α by n. Then for $v \in V_n$ we have $X(v) = 0$. Such v are sometimes called highest weights vectors. We will come back to this in greater detail for $\mathfrak{sl}_3\mathbb{C}$. We have several facts:

- The vectors $\{v, F(v), F^2(v), \ldots\}$ span V
- $E(F^m(v)) = m(n - m + 1)F^{m-1}(v)$
- All eigenspaces V_α are one dimensional

We can quickly illustrate the second fact. Let v be a highest weight vector with weight n. Then we have

$$\begin{aligned} E(F(v)) &= [E,F](v) + F(E(v)) \\ &= H(v) + F(0) \qquad\qquad\qquad = n \end{aligned}$$

Repeating this process the pattern becomes clear. However, we can just prove our statement by induction. The above provides the induction basis. We then compute

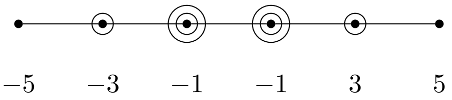
$$\begin{aligned} E(F^m(v)) &= E(F(F^{m-1}(v))) \\ &= [E,F](F^{m-1}(v)) + F(E(F^{m-1}(v))) \\ &= H(F^{m-1}(v)) + F((m-1)(n-m+2)F^{m-2}(v)) \\ &= ((n - 2(m-1)) + (m-1)(n-m+2))F^{m-1}(v) \\ &= m(n-m+1)F^{m-1}(v) \end{aligned}$$

Note that from the second fact we can draw another conclusion. By finite dimensionality there must be m such that $Y^m(v) = 0$. Then $E(F^m(v)) = m(n-m+1)F^{m-1}(v) = 0$, so $n-m+1 = 0$, so $n \in \mathbb{N}$. This brings us to the theorem:

Theorem 4. *For every $n \in \mathbb{N}$ there is a $(n+1)$-dimensional representation $D^{(n)}$, where H has integral eigenvalues $\{-n, -n+2, \dots, n-2, n\}$.*

We will just construct them. A first step would be to look at the standard representation of $\mathfrak{sl}_2\mathbb{C}$, where $V = \mathbb{C}^2$ and we let the matrices work on this vector space. Using the standard basis $\{e_1, e_2\}$, we note $H(e_1) = e_1$ and $H(e_2) = -e_2$. Then this representation is exactly $V^{(1)}$ as described above. The rest of the representations $D^{(n)}$ may now be realized as symmetric powers of V. The statement is that any irreducible representation of the Lie algebra $\mathfrak{sl}_2\mathbb{C}$ is a symmetric power[2]of the standard representation $V = \mathbb{C}^2$.

Let us look at an example, where we take a simple tensor product of some representations, and try to find how it decomposes. For $\mathfrak{sl}_2\mathbb{C}$ this might seems somewhat trivial, but it shows some machinery that we will need to decompose representations of $\mathfrak{sl}_3\mathbb{C}$, in particular a few tensor products which have important physical implications.
Let $V = \mathbb{C}^2$ the standard representation. How can we decompose $Sym^2V \otimes Sym^3V$? Using our knowledge of irreducible representations Sym^2V has eigenvalues $-2, 0, 2$ while Sym^3V has eigenvalues $-3, -1, 1, 3$. From the discussing of Lie algebras we know how to find eigenvalues of the tensor product. They will be pairwise sums of the possible eigenvalues of both the spaces. We get the following picture:



From the picture we deduce there are three different irreducible components. We can immediatly discern Sym^5V. Ignoring the eigenvalues of this representation, there is again a single highest representation Sym^3V. Continueing in this way, one concludes we have the decomposition: $Sym^2 \otimes Sym^3V = Sym^5V \oplus Sym^5V \oplus V$ In fact, one may generalize this statement to immediatly find how a tensor of two such spaces decomposes. In general we have $Sym^aV \otimes Sym^b = \bigoplus_{i=0}^{b} Sym^{a+b-2i}V$.

Let us summarize what we have found. Every irreducible representation of $\mathfrak{sl}_2\mathbb{C}$ on a vectorspace V is isomorphic to the representation $Sym^n\mathbb{C}^2$ for some integer n. We called these $D^{(n)}$. The eigenvalues of such a representation will be integral between n and $-n$, and differ by two: $-n, -n+2, \dots, n-2, n$. We will later use this short description of irreducible representations of $\mathfrak{sl}_2\mathbb{C}$ when we find subalgebras isomorphic to $\mathfrak{sl}_2\mathbb{C}$ in $\mathfrak{sl}_3\mathbb{C}$, at which point this knowledge will prove to be very useful.

[2]The symmetric power of a vector space is denoted $Sym^n(V)$. The symmetric power can be constructed as a quotient space of the tensor $V^{\otimes n}$ by the subspace $v_1 \otimes \dots \otimes v_n - v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(n)}$ where $\sigma \in S_n$, the symmetric group on n elements. It comes with a product, denoted " $\cdot$ ".
Given $\{e_i\}$ a basis for V, then

$$\{e_{i_1} \cdot e_{i_2} \cdot \dots \cdot e_{i_n} : i_1 \leq i_2 \leq \dots \leq i_n\}$$

is a basis for $Sym^n(V)$. Hence $Sym^n(V)$ can be viewed as the space of homogeneous polynomials in the variables e_i. For more information, see appendix B of [1].

2.2 The Lie algebra $\mathfrak{sl}_3\mathbb{C}$ and its adjoint action

This is the algebra that will be of great importance for our example of representation theory pertaining to particle physics. One may realize this Lie algebra as the algebra of traceless 3×3 matrices, and it is sometimes called A_2 in light of the full classification of Lie algebras. Indeed, we may again pick a basis, which will this time be 8-dimensional.
For the time being we wil soon pick a basis of this Lie algebra that will prove convenient to work with. However, in the section on particle physics, we may normalize somewhat differently as dictated by the literature. This however, changes nothing about the general theory developed here.
Note that in the case of $\mathfrak{sl}_2\mathbb{C}$ there was a specific element H that acted diagonally. By using this element we could decompose any representation in terms of the action of H. In the case of $\mathfrak{sl}_3\mathbb{C}$ the role of H will not be played by a single element, or a one dimensional subalgebra if you will, but by a bigger subalgebra. We will take this algebra to be the two dimensional subalgebra of diagonal matrices (three elements on the diagonal, together with tracelessness imposes two degrees of freedom). In more general cases, it is always important to find such an algebra; an abelian subalgebra that acts diagonally (appendix D of [1]). Such a subalgebra is called a Cartan subalgebra, and one may prove that it always exists. In the present case, we merely give this subalgebra and use it immediatly without worrying about the general case.
We will denote the Cartan subalgebra $\mathfrak{h}$, so in the present case $\mathfrak{h} \subset \mathfrak{sl}_3\mathbb{C}$ are the traceless diagonal 3×3 matrices. Now how do we compose with respect to the action of $\mathfrak{h}$? We define

Definition 7. *An eigenvector for $\mathfrak{h}$ is a vector $v \in V$ such that we can write $H(v) = \alpha(H)v$, where $\alpha \in \mathfrak{h}^*$ a linear functional on $\mathfrak{h}$ and $H \in \mathfrak{h}$.*

We then call such functionals α the eigenvalues. By complete reducibility we can now decompose V into spaces associated with the eigevalues α, so that we have for any finite-dimensional representation $V = \bigoplus V_\alpha$, where V_α is an eigenspace for $\mathfrak{h}$ and $\alpha \in \mathfrak{h}$. Some more terminology is in order. We call the functionals α occuring in a representation the *weights* of this representation. The associated eigenspaces V_α will be called *weight spaces*, and vectors within these spaces are the *weight vectors.*
A special role is played by the adjoint action, as we will see. We can do the above operations on this special representation, where the eigenspaces will then be subspaces of $\mathfrak{sl}_3\mathbb{C}$ itself. Using this adjoint representation, we can decompose $\mathfrak{sl}_3\mathbb{C} = \mathfrak{h} \oplus (\bigoplus \mathfrak{g}_\alpha)$. The eigenspaces and eigenvalues of the adjoint representation have special names. We call the weight spaces the *root spaces* and the weights *roots*. To make this explicit, we pick the following basis, where H_{12} and H_{23} will be the basis of the subalgebra $\mathfrak{h}$. Denote E_{ij} the matrix with a 1 in the i^{th} row and j^{th} collumn. Then take the basis

$$H_{12} = E_{11} - E_{22}, \qquad H_{23} = E_{22} - E_{33} \tag{5}$$
$$E_\alpha = E_{12}, \qquad E_{-\alpha} = E_{21} \tag{6}$$
$$E_\beta = E_{23}, \qquad E_{-\beta} = E_{32} \tag{7}$$
$$E_\gamma = E_{13}, \qquad E_{-\gamma} = E_{31} \tag{8}$$

Calculating commutation relations we find for a general $H \in \mathfrak{h}$, that is $H = diag(\lambda_1, \lambda_2, \lambda_3)$, that

$$[H, E_{ij}] = (\lambda_i - \lambda_j)E_{ij}$$

Define the fuctional $\omega_i \in \mathfrak{h}^*$ in the following way. Let $H \in \mathfrak{h}$, which means $H = diag(\lambda_1, \lambda_2, \lambda_3)$. We take $\omega_i(H) = \lambda_i$. This is just projection on the first matrix element.

Note that the span over $\mathbb{C}$ of $\{\omega_1, \omega_2, \omega_3\}$ is $\mathfrak{h}^*$. This is the case as $\mathfrak{h}^*$ is 2-dimensional. The three ω_i are not linearly independent, as $\omega_1 + \omega_2 + \omega_3 = 0$ as the matrices are traceless. However, any two ω_i and ω_j are linearly independent for $i \neq j$. So, this span is also two dimensional and a subspace of $\mathfrak{h}^*$, hence all of $\mathfrak{h}$.
Then by the above calculation, we have that the roots are $\omega_{ij} = \omega_i - \omega_j$, as these are the eigenvalues of the adjoint action of $\mathfrak{h}$, calculated from the commutator. Then to such a root ω_{ij} is associated its root space E_{ij}. We have thus decomposed $\mathfrak{sl}_3\mathbb{C}$ as follows: $\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha = \mathfrak{h} \oplus (\bigoplus_{i \neq j} \mathfrak{g}_{\omega_{ij}})$.
Let us draw a picture of the roots in $\mathfrak{h}^*$. In the picture, the dependent vectors ω_i are of equal length with a angle $\frac{\pi}{3}$ between them, such that $\omega_1 + \omega_2 + \omega_3 = 0$. The roots will then be given by just adding up vectors, in which way we produce a regular hexagon. This choice of angle may seem somewhat arbitrary. However, note that first of all, it works. Adding these vectors as described in the upcoming picture satisfies all the requirements we have for the ω_i. The reason why this should be an angle $\frac{\pi}{3}$ is actually much more involved. It turns out that for semisimple Lie algebras of rank two, there are only several possible configurations of roots of which this is one. This has to do with the Killing form but is quite involved so we will briefly come back to it in section 2.3.

Note that a weight will be a certain eigenvalue α which assigns values to the elements of $\mathfrak{h}$. We know the vectorspace $\mathfrak{h}$ is two dimensional, so let us take H_{12} and H_{23} as a basis. We can denote then $\alpha = (m_1, m_2)$ where $\alpha(H_1) = m_1$ and $\alpha(H_2) = m_2$. We can therefore also denote weights as ordered pairs of two numbers. This will turn out to be convenient when classifying different irreducible representations later on.

The adjoint action of $\mathfrak{h}$ is clear as it sends each root space to itself by multiplication by some scalar which is the eigenvalue. We can use this knowledge to find out how the rest of the Lie algebra acts by calculation. In general, we can come to the following conclusion. Given $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$, where the subscripts of the eigenspaces denote the corresponding eigenvalues in $\mathfrak{h}^*$. Then we have for any $H \in h$:

$$\begin{aligned}[H, [X, Y]] &= [X, [H, Y]] + [[H, X], Y] \\ &= [X, \beta(H)Y] + [\alpha(H)X, Y] \\ &= (\alpha(H) + \beta(H))[X, Y]\end{aligned}$$

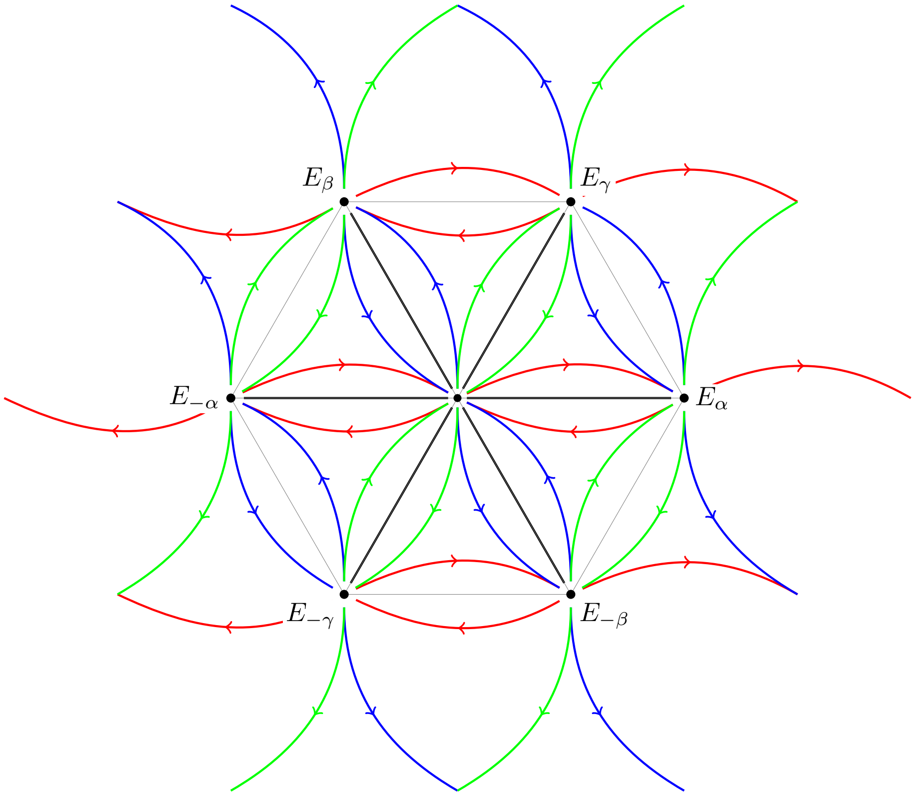
Therefore, $[X, Y]$ is an eigenvector for the adjoint action of $\mathfrak{h}$ with eigenvalue $\alpha + \beta$. Therefore, we can indeed conclude that the adjoint actions of the eigenspaces permute the eigenspaces themselves. That is

$$ad(\mathfrak{g}_\alpha) : \mathfrak{g}_\beta \to \mathfrak{g}_{\alpha+\beta}$$

In our case, we may calculate how every root space $E_\alpha, E_\beta, E_\gamma, E_{-\alpha}, E_{-\beta}, E_{-\gamma}$ permutes the spaces using the adjoint action. For example, we compute the adjoint action of E_α. We get

$$\begin{aligned}&[E_\alpha, E_\alpha] = 0, && [E_\alpha, E_{-\alpha}] = H_{12} \\ &[E_\alpha, E_\beta] = E_\gamma, && [E_\alpha, E_{-\beta}] = 0 \\ &[E_\alpha, E_\gamma] = 0, && [E_\alpha, E_{-\gamma}] = E_{-\beta}\end{aligned}$$

We can repeat this process for all eigenspaces and draw them in a diagram. Let us put the spaces on the vertices of a hexagon, with the eigenspace $\mathfrak{h}$ with eigenvalue 0 in the middle. We denote the adjoint action by an arrow. When the arrow leaves the diagram, we mean it maps to 0, killing the corresponding eigenspace. We use red arrows for the adjoint action of E_α and $E_{-\alpha}$, where the arrows going to the right denote the action of the first, and the arrows going to the left denote the action of the latter. We do this in blue for E_β and its counterpart, and in green for E_γ and its counterpart. We obtain the following diagram.



We see that the adjoint action of a particular eigenspace shifts all the other eigenspaces into a particular direction which is parallell to where the rootvector itself is pointing.

We can now use this knowlegde to compute how elements of a arbitrary representation shift the eigenspaces. The same picture applies here. Let V be a representation of $\mathfrak{sl}_3\mathbb{C}$. We know our Cartan subalgebra $\mathfrak{h}$ acts diagonally by virtue of the preservation of Jordan decomposition, so again we may decompose V into eigenspaces, as we saw before for the adjoint representation: $V = \bigoplus V_\alpha$, where $\alpha \in \mathfrak{h}^*$. We call the α the *weights* of the representation, and the spaces V_α the *weight spaces*. We now wonder how the rest of $\mathfrak{sl}_3\mathbb{C}$ acts on V. If we let $X \in \mathfrak{g}_\alpha$ and $v \in V_\beta$, how will X act on this vector v? This can be calculated this using the commutator, which is the adjoint action as our elements are matrices. Similarly, we compute the action of H on $X(v)$ to see in which eigenspace it is.

$$\begin{aligned} H(X(v)) &= X(H(v)) + [H, X](v) \\ &= X(\beta(H)v) + (\alpha(H)X)v \\ &= (\alpha(H) + \beta(H))X(v) \end{aligned}$$

This means that $X(v)$ is again eigenvector, only now with eigenvalue $\alpha + \beta$. In terms of our two numbers, we have for a weight $\beta = (b_1, b_2)$ and a root $\alpha = (a_1, a_2)$ that $X(v)$ is in the weightspace with weight $\beta + \alpha = (b_1 + a_1, b_2 + a_2)$. We can use these roots to generate a lattice in $\mathfrak{h}^*$, we shall denote Λ_R, the root lattice.

$$\Lambda_R = \{a\omega_{12} + b\omega_{23} + c\omega_{13} : a, b, c \in \mathbb{Z}\}$$

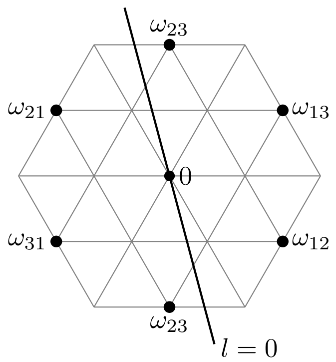
All weights α occuring in the representation will lie in some translate of this lattice as they are congruent modulo this lattice. If this were not the case, $W = \bigoplus_{\beta \in \Lambda_R} V_{\alpha+\beta}$ would be a subrepresentation, and this contradicts with irreducibility. We conclude all eigenvalues occuring in an irreducible representation of $\mathfrak{sl}_3\mathbb{C}$ differ by integral linear combinations of the roots $\omega_{ij} \in \mathfrak{h}^*$. We will use this property to describe general irreducible representations of $\mathfrak{sl}_3\mathbb{C}$.

2.3 The irreducible representations of $\mathfrak{sl}_3\mathbb{C}$

We now wish to find how the irreducible representations of $\mathfrak{sl}_3\mathbb{C}$ are structured. In the case of $\mathfrak{sl}_2\mathbb{C}$, we found a vector that generated the entire representation. This is in fact a very general principle that works for semisimple Lie algebras. However, we will state it specifically for $\mathfrak{sl}_3\mathbb{C}$ here. To do this, we need some kind of notion of a vector that lies in an eigenspace that is in some sense on the edge of our weight diagram. To specify what we mean, let us order the roots. This is done by picking a linear functional

$$l : \Lambda_R \to \mathbb{R}$$

We will define then positive roots as roots α such that $l(\alpha) > 0$ and negative roots when $l(\alpha) < 0$. Note that to avoid ambiquity, we need to pick l irrational with respect to the lattice so it can never intersect a root. Let us just pick some linear functional. We will later see it doesn't matter which one we picked. As it is defined on the root lattice, we take $l(\lambda_1\omega_1 + \lambda_2\omega_2 + \lambda_3\omega_3) = a\lambda_1 + b\lambda_2 + c\lambda_3$ with $a > b > c$ such that l is indeed irrational with respect to Λ_R. Drawing this l in a picture of $\mathfrak{h}^*$ with our roots, we get the following picture



The positive roots will now be ω_{12}, ω_{13} and ω_{23} for the specific l we have chosen. Note that they are not linearly independent. We will call a root *simple* if it cannot be expressed as a linear combination with positive coefficients of other positive roots. The simple roots will therefore be ω_{12} corresponding to the root space E_α and ω_{23} corresponding to E_β. We can express all roots in terms of these positive simple roots. This can be done by looking at what values they assign to H_1 and H_2, as these are a basis of $\mathfrak{h}$. We have $\omega_{12} = (2, -1)$ and $\omega_{23} = (-1, 2)$. Then we have:

$$\omega_{13} = (1, 1) \qquad \omega_{31} = (-1, -1)$$
$$\omega_{21} = (-2, 1) \qquad \omega_{32} = (1, -2)$$

2.3.1 The highest weights

Definition 8. *Let V be an irreducible finite-dimensional representation of $\mathfrak{sl}_3\mathbb{C}$. Then there is a vector $v \in V$ called the* highest weight vector *which has the property that*

- *The vector v is in some eigenspace V_λ, meaning it is an eigenvector for $\mathfrak{h}$*
- *The vector v is killed by the action of the positive root spaces, E_α, E_β and E_γ.*

Note that if E_α and E_β kill v then E_γ also kills v because $E_\gamma = [E_\alpha, E_\beta]$. The existence of such a vector v that is killed by the three positive root spaces might be somewhat intuitive by looking at the picture. This vector will in some sense be at the edge of the weight diagram, contained in a weightspace, such that our positive rootspaces act as shifts beyond the edge, our vector v is killed. In fact, the highest weight will be that for which the functional l we picked is maximal because any weight space beyond the highest weight space will be killed by action of the positive root spaces.
In a similar way to $\mathfrak{sl}_2\mathbb{C}$, this highest weight vector and its highest weight will prove very important to our discussion. As we saw in the case of $\mathfrak{sl}_2\mathbb{C}$, an entire irreducible representation was generated by applying the operator F to a weightspace with the highest weight. This highest weight was some integer, and fully determined the irreducible representation. We wish to do something similar for $\mathfrak{sl}_3\mathbb{C}$, but now we have two numbers to work with to find highest weights. We will later see any two integers (m_1, m_2) will be the highest weights of some irreducible representation. Thus for any two integers we get an irreducible representation of $\mathfrak{sl}_3\mathbb{C}$.
Let us first begin by stating in detail why the highest weight is so useful in the following theorem.

Theorem 5. *Let V be a finite-dimensional representation of $\mathfrak{sl}_3\mathbb{C}$, and $v \in V$ a highest weight vector. Then W, the subspace generated by successively applying the negative root spaces $E_{-\alpha}$, $E_{-\beta}$ and $E_{-\gamma}$ to v, is irreducible.*

Proof. We will first assume V to be irreducible. We generate W by applying the negative root spaces, and prove that W is invariant under $\mathfrak{sl}_3\mathbb{C}$. By irreducibility, W must be the entire V. It is enough to check that the root spaces E_α, E_β and E_γ map W to itself. Furthermore, as E_γ is the commutator of the other two, it is enough just to check for E_α and E_β. Of course, for v itself the statement is trivial, as it is in the kernel of the action of the root spaces.
We now proceed by induction.
Let w_n be a word on the letters $E_{-\alpha}$ and $E_{-\beta}$ with length smaller or equal to n. We take W_n the vector space spanned by such applying such words to v. Then $W = \bigcup_n W_n$. Now let us calculate how the spaces E_α and E_β act on $w_n(v)$. The above step gives the induction basis for the statement that E_α and E_β carry W_n into W_{n-1}. Let us proceed by induction for w_n. Note we can write any such $w_n(v)$ as $E_{-\alpha}(w_{n-1}((v))$ or $E_{-\beta}(w_{n-1}(v))$. Then for E_α:

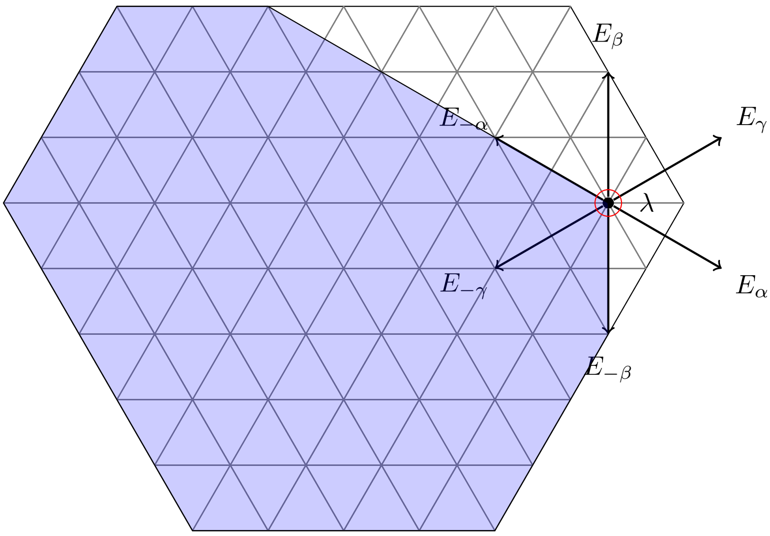
$$\begin{aligned} E_{\alpha(E_{-\beta}}(w_{n-1}(v))) &= E_{-\beta}(E_\alpha(w_{n-1}(v))) + [E_\alpha, E_{-\beta}](w_{n-1}(v)) \\ &\in E_{-\beta}(W_{n-2}) \\ &\subset W_{n-1} \end{aligned}$$

Similarly,

$$\begin{aligned} E_{\alpha(E_{-\beta}}(w_{n-1}(v))) &= E_{-\beta}(E_\alpha(w_{n-1}(v))) + [E_\alpha, E_{-\beta}](w_{n-1}(v)) \\ &\in E_{-\beta}(W_{n-2}) \\ &\subset W_{n-1} \end{aligned}$$

A similar calculation for E_β proves the theorem for reducible V. Let us now look at any representation V. Let λ be the weight associated to v ($v \in V_\lambda$). The above calculation establishes that W is a subrepresentation of V. We have that W_α is one dimensional. Suppose now, on the contrary, that W is not irreducible. Then $W = W' \oplus W''$ for representations W' and W''. Because they are representations, projections on either component commutes with the action of $\mathfrak{h}$. Therefore, we can decompose $W_\alpha = W'_\alpha \oplus W''_\alpha$. Because W_α is one-dimensional, so v is in either W' or W'', which W is either W' or W'', meaning it is irreducible. $\square$

Thus we know that the whole irreducible representation is, so to speak, contained below the eigenspace V_λ with highest weight containing the highest weight vector. As the action of the root vectors shift in particular directions, the above proof shows that only by shifting in the $E_{-\alpha}$, $E_{-\beta}$ and $E_{-\gamma}$ directions do we get to nontrivial spaces.
If λ is a highest weight, then $dim(V_\alpha) = 1$ as can be verified by taking another vector in $v' \in V_\lambda$. Then write v' in terms of the generating operators and v. By using the linear indepence of two of the three roots one may conclude that $v' = cv$. This shows that in fact, the highest weight is unique up to scalars as the highest weight space is one-dimensional, and therefore an irreducible representation is fully determined by its highest weight vector.
Furthermore, given an arbitrary representation, every highest weight vector we find corresponds to an irreducible subrepresentation, and the dimensionality of the weight space containing this highest weight vector gives the multiplicity of the irreducible subrepresentation in the decomposition of the representation. If we were to plot such an highest weight λ corresponding to a highest weight vector v, we get the following picture.



Here, λ is shown as a point in $\mathfrak{h}^*$ and we also show the directions in which the root spaces shift. As the positive root spaces kill the highest weight vector and the negative root spaces generate the entire irreducible representation, we know the weights are contained in the blue area of the diagram. We can now concentrate on what happens on the edges of the blue portion of the diagram to see if we can further confine the weights that will appear in our representation.However, let us first look in more detail at what these highest weights look like. For such highest weights λ we have the following characterization in terms of what the values of the highest weight is on H_{12} and H_{23}. This will turn out to be handy to actually give names to irreducible representations.

Lemma 4. *Given any two positive integers m_1, m_2 there is an irreducible representation of $\mathfrak{sl}_3\mathbb{C}$ having $\lambda = (m_1, m_2)$ as highest weight. In terms of $\{\omega_1, \omega_2, \omega_3\}$ the highest weight will be $m_1\omega_1 - m_2\omega_3$.*

We first prove a lemma about highest weights of tensor products.

Lemma 5. *Let $v \in V$ be highest weight vector and $w \in W$ a highest weight vector of representations V and W. This means both are killed by E_α, E_β and E_γ. Consider the representation $V \otimes W$. Then $v \otimes w$ is a highest weight vector for this representation.*

Proof. We compute the action of E_α on $v \otimes w \in V \otimes W$:

$$\begin{aligned} E_\alpha(v \otimes w) &= E_\alpha(v) \otimes w + v \otimes E_\alpha(w) \\ &= 0 \otimes w + v \otimes 0 \qquad\qquad = 0 \end{aligned}$$

Similar calculations hold for E_β and E_γ. So we see that indeed $v \otimes w$ will be a highest weight vector of $V \otimes W$. $\square$

Furthermore, we know a similar calculation for the weights corresponding to v and w that the weight corresponding to $v \otimes w$ will be the sum of the two weights for v and w. We can now prove the theorem.

Proof. Lemma 5
We then have highest weight $(1, 0)$ which belongs to the standard representation V of $\mathfrak{sl}_3\mathbb{C}$ on $\mathbb{C}^3$ by just letting the matrices act as operators. The eigenvectors will be the standard basis. Then note that E_α and E_β and E_γ all kill e_1 and that indeed e_1 had weight $(1, 0)$. In terms of the weights $\{\omega_1, \omega_2, \omega_3\}$ it corresponds to the weight ω_1.
Now, we already determined how the adjoint standard representation V^* works. An matrix X is represented by $-X^t$. We therefore our positive roots are represented by their negatie counterparts, and we have now that e_3 will be the highest weight vector, with weight $(0, 1)$ or $-\omega_3$. These representations with weights $(1, 0)$ and $(0, 1)$ are the so called fundamental representations. We will use them to build other representations with the required highest weights. We do this by taking tensor products of the fundamental representations. Let us take the representation $W = \bigotimes^{m_1} V \otimes \bigotimes^{m_2} V^*$.
We have that a highest weight of this representation will be $e_1 \otimes \ldots \otimes e_1 \otimes e_3{}^* \otimes \ldots \otimes e_3{}^*$ with weight (m_1, m_2) or $m_1\omega_1 - m_2\omega_3$. It is entirely unclear whether this representation is irreducible, but this doensn't matter. By the previous theorem we can generate an irreducible subrepresentation with highest weight (m_1, m_2). $\square$

Indeed we now know that any to integers m_1, m_2 give rise to an irreducible representation of $\mathfrak{sl}_3\mathbb{C}$.
Let use denote such representations with highest weight (m_1, m_2) as $D^{(m_1, m_2)}$. The question is, is this representation unique? It turns out we can indeed speak of an unique $D^{(m_1, m_2)}$. Suppose there are two irreducible representations V and W with highest weight α and highest weight

vectors $v \in V$ and $w \in W$.
By a simple calculation using the properties of the direct sum space $V \oplus W$ that we calculated before, we know (v, w) is a highest weight in $V \oplus W$ with weight α. Using this highest weight (v, w) generate a irreducible subrepresentation $U \subset V \oplus W$, by succesively applying the negative root spaces. Then we have projection maps $\pi_V : U \to V$ and $\pi_W : U \to W$. Now note that $\pi_V|_U$ is nontrivial as $\pi_V|_U(v \oplus w) = v$. Likewise $\pi_W|_U$ is nontrivial. By Schurs lemma, any nontrivial homomorphisms between two irreducible representations is an isomorphism. As V, W as well as U are all irreducible, we have $U \cong V$ and $U \cong W$, hence $V \cong W$ and we have uniqueness.

So we conclude that for every two positive integers m_1, m_2 we have a unique irreducible representation $D^{(m_1,m_2)}$ with highest weight (m_1, m_2) or more geometrically highest weight $m_1\omega_1 - m_2\omega_3$.

What we have not calculated is what exactly such a representation will precicely look like. First of all, we state a fact here that may be deduced from the so called Weyl character formula. This is a way to calculate multiplicities of weights within a representation. For more on the Weyl character formula see [2] chapter 24 section 3 or [1] chapter 24. For these representations of $\mathfrak{sl}_3\mathbb{C}$ (also called A_2 in [2]) we have

$$dim(D^{(m_1,m_2)}) = \frac{1}{2}(m_1 + 1)(m_2 + 1)(m_1 + m_2 + 2)$$

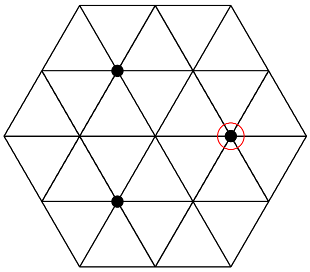
One may in fact prove that these representations can be realized as the kernel of the contraction map on the representation $Sym^{m_1}V \otimes Sym^{m_2}V^*$. However, the exact construction is not really insightful. We can just construct some representations as examples. First example will be the representation $D^{(1,0)}$. We know it is the standard representation on $V = \mathbb{C}^3$. Therefore, we immediatly know the weights and thereby the weight diagram. The weights will be our ω_1, ω_2 and ω_3 by the action of the matrices on the standard basis e_1, e_2 and e_3.

We will draw the roots as vectors of equal lengths with angles $\frac{\pi}{3}$ between them. The explanation for this is quite involved and pertains to a full classification of the semisimple Lie algebras, but we will briefly discuss it here. The Killing form may be used to define an inner product $< \ , \ >$ on the vector space of roots $\mathfrak{h}^*$.
Now define numbers $n_{\beta\alpha}$ with α and β in the positive simple roots as

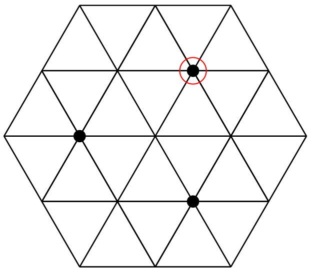
$$n_{\beta\alpha} = \frac{2 < \beta, \alpha >}{< \alpha, \alpha >}$$

It turns out that these $n_{\beta\alpha}$ are always integers. This then fixes a combination of length and angle between the positive simple roots. There are then only several possibilities. For Lie algebras of rank two, one possibility is the angle $\frac{2\pi}{3}$ between two positive roots of equal length. This is the case for $\mathfrak{sl}_3\mathbb{C}$, and is what we have drawn so far, considering ω_{12} and $omega_{23}$ are the positive simple roots.

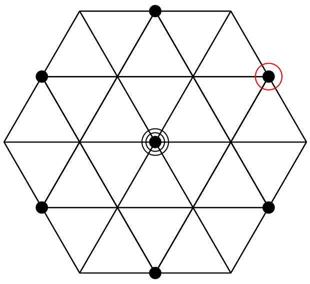
Let us now come back to the first irreducible representation of $\mathfrak{sl}_3\mathbb{C}$, which has the following weight diagram



We see the highest weight is ω_1. From now on we will circle the highest weight space in red. We can do the same for $D^{(0,1)}$, which will be the standard action on V^*. We get weight diagram



Here the highest weight is $-\omega_3$.
Now let us treat an important, more insightful representation which we will use later on. Let us take a look at the representation $D^{(0,1)} \otimes D^{(1,0)}$. We already calculated how these tensor products work, that is by adding the weights of the respective representations. However, we now get several ways in which to reach the 0-functional (that is, the spaces associated with $\mathfrak{h}$). We get the following diagram



Here the highest weight is $\omega_1 - \omega_3$. We can immediatly see the subrepresentation $D^{(1,1)}$ generated by the sum of the highest weights of the two representations. This will be the hexagon plus two middle points as one may readily calculate by generating the irreducible subrepresentation using the negative root spaces. We are left with a single dot with weight 0. We then see that $D^{(1,0)} \otimes D^{(0,1)} = D^{(1,1)} \oplus D^{(0.0)}$. We can continue building bigger representations like this, and deduce from them their irreducible components by finding highest weight vectors and acting on them with the negative root spaces. However, this requires lenghty calulations.
In fact we can generalize how the weight diagrams look, together with the multiplicity of each weight space in general. However, it will be easier to explain this once we have a more geometric picture of how general weight diagrams look for $\mathfrak{sl}_3\mathbb{C}$.

We now continue to look at how the weight diagram develops geometrically. This will prove to be insightful in how the weight diagrams of irreducible representations look. Let us return to

our highest weight λ somewhere in the weight diagram, and look at what happens when we study the eigenspaces occuring by acting with a single root space. We have three directions to choose from, being ω_{21}, ω_{32} and ω_{31} by acting with $E_{-\alpha}$, $E_{-\beta}$ and $E_{-\gamma}$ respectively. Let us move along the border by taking $E_{-\alpha}$ and apply it successively to the eigenspace V_λ. We get a string in the ω_{21}-direction. That is, we generate the spaces $V_{\lambda+k(\omega_{21})}$. By finite dimensionality of the representation this must end at some point, hence k is some finite integer.
Now let us use our knowledge of $\mathfrak{sl}_2\mathbb{C}$ and construct a subalgebra of $\mathfrak{sl}_3\mathbb{C}$ isomorphic to $\mathfrak{sl}_2\mathbb{C}$. This is done by the identification

$$E_\alpha \mapsto E \in \mathfrak{sl}_2\mathbb{C}, \qquad E_{-\alpha} \mapsto F \in \mathfrak{sl}_2\mathbb{C}, \qquad [E_\alpha, E_{-\alpha}] = H_{12} \mapsto H \in \mathfrak{sl}_2\mathbb{C}$$

Calculating the commutators we see that this is indeed a representation of $\mathfrak{sl}_2\mathbb{C}$ with the right eigenvalues (± 2). Let us call this subalgebra $\mathfrak{s}_{\omega_{12}}$. This string of eigenspaces, $W = \bigoplus_k \mathfrak{g}_{\alpha+k\omega_{21}}$ will be preserved by the action of $\mathfrak{s}_{\omega_{12}}$. In other words, this W is a representation of $\mathfrak{sl}_2\mathbb{C}$ (by restriction of the action on V to this particular subalgebra).

We can now use our knowledge of representations of $\mathfrak{sl}_2\mathbb{C}$ to deduce what this string looks like. We know the eigenvalues are integers and integral, and symmetric with respect to zero. That is, symmetric about the eigenspace V_ω corresponding to eigenvalue ω, such that that $\omega(H_{12}) = 0$.

Using this first step (generating a string that is a $\mathfrak{sl}_2\mathbb{C}$ representation), we can fully generalize to find what the weight diagram looks like. We introduce the *Weyl group.*

Definition 9. *The Weyl group $\mathfrak{W}$ of a Lie algebra $\mathfrak{g}$ is the group generated by the involutions $W_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha$ on $\mathfrak{h}^*$ where α is a root and $\beta \in \mathfrak{h}^*$*

These are exactly the reflections in the lines $\Omega_\alpha = \{\beta \in \mathfrak{h}^* :< H_\alpha, \beta >= 0\}$. Such Weyl groups may be constructed for any semisimple Lie algebra. Here we see the Weyl group is a group isomorphic to S_3, the permutations on three elements, as there are three possible reflections.

Let us draw such a line, for example $\Omega_{\omega_{12}}$. We take this as an illustration of what happens around one such line of reflection pertaining to the Weyl group. The same story holds for $\Omega_{\omega_{23}}$ and $\Omega_{\omega_{13}}$

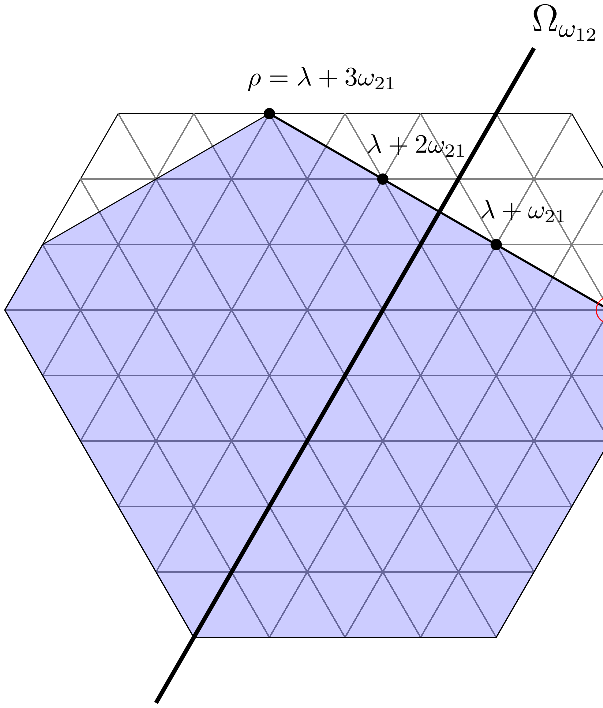
This line consists of alll $L \in \mathfrak{h}^*$ such that $\omega(H_{12}) = 0$, in the plane $\mathfrak{h}^*$. We conclude that the string generated by applying $E_{-\alpha}$ to the highest weight is symmetric about this line by the $\mathfrak{sl}_2\mathbb{C}$ properties, and hence is conserved under reflection in this line.
Let us draw this line in our picture of $\mathfrak{h}^*$. We know the functionals ω_1, ω_2 and ω_3 span $\mathfrak{h}^*$, so we write down $\Omega_{\omega_{12}}$ in terms of these. As we want zero eigenvalue on H_{12} for every functional on the line, it is easy to see that

$$\Omega_{\omega_{12}} = \{a\omega_1 + b\omega_2 + c\omega_3 : a = b\}$$

Putting this into a picture we immediatly see that this line is orthogonal to the string generated by applying $E_{-\alpha}$ to the highest weight space.

Let us now look at the left end of the string W generated by E_α from the highest weight vector. At some point we must have smallest integer m such that $E_\alpha^m(v) = 0$, by finite dimensionality of the representation. Let us look at the eigenspace V_ρ where $\rho = \omega + (m-1)\omega_{12}$. We see that $v' \in V_\rho$ is also a highest weight vector had we chosen a different functional, because it is a vector that is killed by the root spaces $E_{-\alpha}$, E_β and E_γ. Choosing a functional that maximizes the coefficients in front of these root spaces will have maximal weight ρ.

Then the statement after this, however, would not have changed. The entire representation would still be generated by the action of the now negative root spaces. We could again shade the area spanned by the negative root spaces and see we have confined our weight diagram to a smaller area in $\mathfrak{h}^*$. This is what happens around the line $\Omega_{\omega_{12}}$. We have the unbroken string of eigenspaces from λ to ρ which is symmtric about the line $\Omega_{\omega_{12}}$.

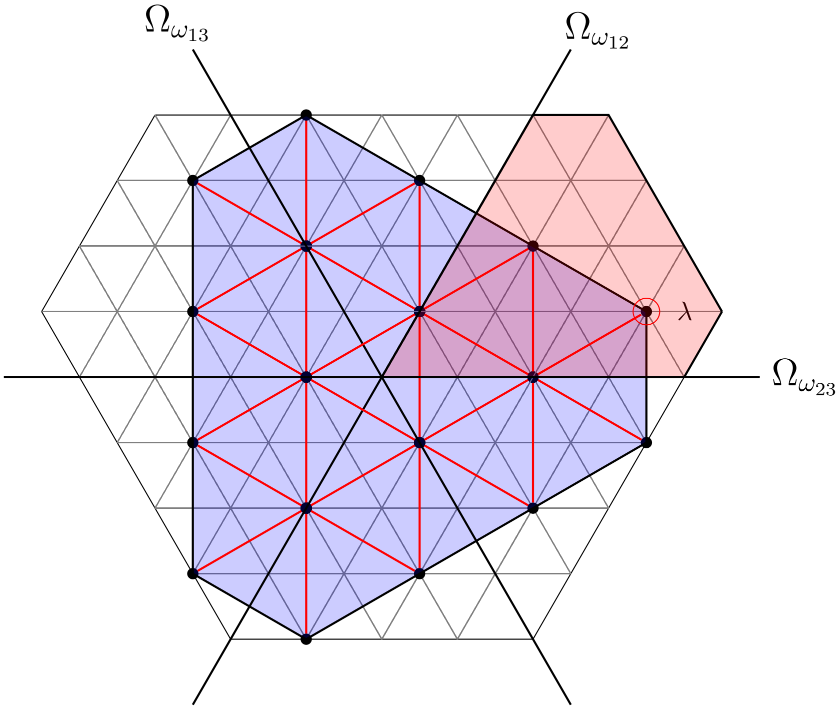


We can keep on repeating this process. Generate a string along the edge and find a subalgebra isomorphic to $\mathfrak{sl}_2\mathbb{C}$, by sending the root space generating the string, its negative and their commutators to the elements of $\mathfrak{sl}_2\mathbb{C}$. In this way we find again a line of reflection $\Omega_{\omega_{ij}}$ under which the string is invariant.
In this way, the set of α for which $V_\alpha \neq 0$ is symmetric with respect to the reflections of $\mathfrak{W}$. We have constructed a so called hull in which all the weights occuring the representation are contained. Note that by definition of the $\Omega_{\omega_{ij}}$ they intersect at $0 \in \mathfrak{h}^*$ as all coefficients must be equal, but $\omega_1 + \omega_2 + \omega_3 = 0$.

Now as the eigenvalues of the H_{ij} must be integers by virtue of the $\mathfrak{sl}_2\mathbb{C}$ subrepresentations, we know that the weights occuring will be integral linear combinations of the ω_1, ω_2 and ω_3. This is the case because any weight can be expressed as $\omega = a\omega_1 + b\omega_2 + c\omega_3$. It must have integer values on H_{12} and H_{23}, forcing $a, b, c \in \mathbb{Z}$. That is, the weights lie on the weight lattice

$$\Lambda_W = \{a\omega_1 + b\omega_2 + c\omega_3 : a, b, c \in \mathbb{Z}\}$$

We therefore find that the occuring weights must lie in the lattice generated by the weights Λ_W, as well as being congruent modulo the action of the root spaces, so congruent modulo Λ_R. Summarizing the entire preceding construction we can improve our diagram and find



Here, we see the root lattice drawn in red. From the highest weight λ, all weights within the hull are congruent to each other modulo Λ_R, whereas the weights themselves lie on the triangular grey grid, within the blue convex hull spanned by the image of λ under the Weyl group. This in fact is almost all there is to know about representations of $\mathfrak{sl}_3\mathbb{C}$. Given a finite dimensional irreducible representation of $\mathfrak{sl}_3\mathbb{C}$, we can search for the highest weight vector $\lambda \in \Lambda_W$. The set of eigenvalues occuring in V will the the functionals congruent to λ modulo Λ_R, lying inside the convex hull (a hexagon) generated by the images of λ under the Weyl group $\mathfrak{W}$.

Note however, that we are not entirely done. The multiplicities of the weights are something we have not calculated. This may be calculated, but the calculation is quite long so we just state the result. Going one shell inwards (every weight space reached by acting with one generating root space) has multiplicity 2 all around. As long as the shells have a hexogonal shape, the multiplicity increases by one on each shell going inwards. Once the shells become triangles, the multiplicity becomes constant among these triangles.

Finally we briefly introduce the concept of the Weyl chamber. At this time, it is not that useful, but for different semisimple algebras we can reduce the amount of calculation greatly as we need to only focus on this Weyl chamber. When we found a highest weight, we know we can generate the convex hull that contains all the weights of the irreducible representation associated with this highest weight. However, by acting with the Weyl group, we need very little information actually. We only need the cone between the lines $\Omega_{\omega_{12}}$ and $\Omega_{\omega_{23}}$, as all other weights will be conjugate to weights between these lines under the Weyl group.
We call this cone the Weyl chamber $\mathscr{W}$. In the above diagram we filled it purple. Now this subset of $\mathfrak{h}^*$ is convex. In fact, any weight inside is a linear combination of ω_1 and $-\omega_3$. Therefore, the highest weight will be within the Weyl chamber. Indeed, there is a bijection between the elements of $\Lambda_W \cap \mathscr{W}$ and irreducible representations of $\mathfrak{sl}_3\mathbb{C}$, as any weight $a\omega_1 - b\omega_3$ in this set can be associated to an irreducible representation $D^{(a,b)}$.

To summarize, given that we want to construct some irreducible representation, we know that this is equivalent to giving two integers m_1 and m_2 and constructing $D^{(m_1,m_2)}$. This construction is then done by what we calculated before. Let the Weyl group act on the highest weight $m_1\omega_1 - m_2\omega_3$, and fill out the rest of the diagram with weights congruent the lattice Λ_R on the lattice Λ_W. The multiplicities of on the inner shells are described by what we had before. Therefore, we have fully characterized all irreducible representations of $\mathfrak{sl}_3\mathbb{C}$.

3 Elementary particles

As presented in the introduction, we now apply our knowledge of representation theory to an important application in physics. Of great interest will be the algebras $\mathfrak{sl}_2\mathbb{C}$ and $\mathfrak{sl}_3\mathbb{C}$ which we have discussed in great detail.

3.1 Interaction in particle physics

As technology advanced, physicists were able to construct high energy particle accelerators. Experiments conducted with these accelerators produced a wealth of new particles that could be studied. From many types of collisions one may produce many types of particles. However, in this particle pandemonium many structures were observed. We can even classify them in a rather beautiful way using representation theory inspired by experimental data. By carefully studying all particles that appear, we may group them. From this experimental observation, a theory was developed which explained this phenomomen. This is the quark model. It is a great example of the way in which theory and experiment come together to produce a powerful model.
What we are concerned with is the so called *strong force*. Of course, other forces might be of interest, like electromagnetic forces in the charges or weak interaction in decay. However, we will concentrate only on the strong interaction, which is several orders of magnitude stronger than the other interactions. The strong interaction is what binds the nucleus together. To further explain why we need representation theory, let us explain how symmetry plays a role in physics.

3.2 Symmetry groups and physics

Many things about physical systems can be learned from their symmetries. In 1918, E. Noether derived conservation laws using the symmetries of physical systems. In this thesis we will focus on more restricted cases, and simplify some theory in order to remain concise.
Let us first state what we mean by a symmetry. Physical systems have an energy which in quantum mechanics is given by the Hamilton operator $\bar{H}$. The particles are states $|\psi\rangle$ which are vectors in some vector space.
A physical system is said to have a symmetry group G_α if the Hamiltonian describing the system commutes with the operators $\hat{U}(\alpha) \in G_\alpha$. These are matrix operators assigned to the group elements, so it is a representation of the group G_α on the vector space of states $|\psi\rangle$. Because the Hamiltonian governs time evolution of our system this is indeed very important. Let us make this statement more concrete.
Physical quantum systems will be described by wave functions governed by the Schrödinger equation. The interpetation of such wavefunctions is as vectors in a complex Hilbert space. They are functions of space and time, together with an inner product $< \ | \ >$. Then the inner product of a state with itself squared is interpeted as the probability of finding a particle at a position $\vec{r}$ at time t.
The Hamiltonian $\bar{H}$ is a way to measure the energy of a given state. Given that the Schrödinger equation holds for some initial state we have

$$i\hbar \frac{\partial \psi(r,t)}{\partial t} = \hat{H}\psi(r,t)$$

Applying the symmetry operator $\hat{U}(\alpha)$ we have

$$i\hbar\frac{\partial\hat{U}(\alpha)\psi(r,t)}{\partial t} = \hat{U}(\alpha)\hat{H}\hat{U}^{-1}(\alpha)\hat{U}(\alpha)\psi(r,t)$$

Then the displaced wave function $\hat{U}(\alpha)\psi(r,t)$ obeys the same Schrdinger equation as $\psi(r,t)$ if we have $[\hat{H},\hat{U}(\alpha)] = 0$ which is then what it means to be a symmetry of the Hamiltonian.
From this we can also make a statement about the infinitesimal operators associated to G_α. That is, the operators which generate the group. Note that in the case G_α is a Lie group this will be its associated Lie algebra. From the previous calculation, we then have for the infinitesimal operators that $[\hat{H}, e^{-i\alpha_k\hat{L}_k}] = 0$, where $\hat{L_k}$ are the infinitesimal operators. Then for small displacements we have

$$[\hat{H},\hat{L}_k] = 0$$

So for a symmetry group G_α, the infinitesimal generators $\hat{L}_k$ will also commute with the Hamiltonian. In physics, the irreducible representations of G_α on the vector space of states is called a *multiplet.* We will actually use the representations of the Lie algebra, as there is a 1-1 correspondence and they are easier to work with.
Now we wish that symmetry operators $\hat{U}_\alpha$ conserve the probability interpetation. That is, they preserve the inner product on the vector space of states. One may calculate that this implies that the representations are unitary. We will only be looking at the case of $SU(2)$ and $SU(3)$. We know that any representation of semisimple Lie algebras is completely reducible, so the action of G_α can be decomposed into multiplets.

Now we come to an important property of such multiplets. If a state $|\psi\rangle$ has an energy E, we mean that it has eigenvalue E for the Hamiltonian $\hat{H}$. That is

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

From commutativity of the Hamiltonian with group operators $\hat{U}(\alpha)$, we see that the displaced wave function $|\psi'\rangle = \hat{U}(\alpha)|\psi\rangle$ also has eigenvalue E for the Hamiltonian operator $\bar{H}$

$$\begin{aligned}\hat{H}|\psi'\rangle &= \hat{H}\hat{U}_\alpha|\psi\rangle \\ &= \hat{U}_\alpha\hat{H}|\psi\rangle \\ &= \hat{U}_\alpha E|\psi\rangle \\ &= E|\psi'\rangle\end{aligned}$$

Therefore, we can conclude the eigenvalue of the Hamiltonian operator for all these particles is the same; all states inside a multiplet have the same energy. Because the Hamiltonian also commutes with infinitisimal generators $\hat{L}_k$ also commute with $\hat{H}$ this is also true. However, there is another class of operators which will prove interesting.

We have introduced Casimir operators, which we shall denote $\hat{C}_\lambda$. They are operators that commute with all the infinitesimal generators and for every two $\hat{C}_\lambda$ and $\hat{C}_\mu$ we have $[\hat{C}_\lambda,\hat{C}_\mu] = 0$. One may deduce that for symmetry groups $SU(n)$ these operators will always be homogeneous polynomials in the infinitesimal generators: $\hat{C}_\lambda(\hat{L}_1,\hat{L}_2,\ldots,\hat{L}_k)$. This is referenced on page 75 of [5] (Biedenharn, 1963).
The number of independent Casimir operators depends on the rank of the Lie algebra (the dimension of $\mathfrak{h}$). In fact, the number of Casimir operators equals the rank of the Lie algebra (theorem by Chevalley).

Consider a rank l Lie algebra. The Casimir operators also commute with $\bar{H}$, and therefore, can be simultaneously diagonalized. In quantum mechanics this means that their eigenvalues can be simultaneously observed, together with the energy as this is the eigenvalue of the commuting operator $\bar{H}$. In some sense these operators are more important than the infinitesimal generators. By Schur's lemma the Casimir operators act as scalars on the multiplets. Then any multiplet can be characterized by the l eigenvalues of these l independent Casimir operators $\bar{C}_\lambda$, with $\lambda = 1, 2, \ldots, l$.

Before we start with our first part, the theory of isospin which arises from $SU(2)$ symmetry, let us introduce some more terminology. Particles that interact strongly are called *hadrons*. Note that this does not mean they cannot interact via electromagnetic or weak forces. However, we pretend that the only force is the strong interaction.
One may divide the hadrons into two groups called *mesons* and *baryons*. Mesons are particles with half-integer spins (fermions), whereas baryons have whole integer spins (bosons). Spin is an intrinsic way to characterize the internal angular momentum of elementary particles. The theory for isospin will be very similar to that of spin. From the many high energy collision experiments, very many mesons and baryons are known. We shall focus on these baryons and mesons.

3.3 Isospin and $\mathfrak{sl}_2\mathbb{C}$

After the discovery of the neutron it was noted that this particle was very similar to the proton. Of course, their charge differs. However, we will only be concerned with the strong forces that act on the particles, so we disregard this for the time being. Comparing some empirical properties of the neutron and proton yields

	Mass (in [MeV])	Spin	Lifetime (in [s])
p	983.213	$\frac{1}{2}$	stable
n	939.507	$\frac{1}{2}$	918 ± 14

We see that their spin is the same, and their masses are nearly the same. If we assume that the small mass difference is due to electromagnetic interactions, their masses are approximately equal with respect to the strong interaction. In 1932 Heisenberg proposed that the proton and neutron were two states of one and the same particle, the nucleon. This is reflected by taking the group $SU(2)$ as the symmetry group of the strong interaction. This means we are only concerned with the Hamiltonian of strong forces. Note that the Lie algebra that has representations corresponding to those of $SU(2)$ that we will use is $\mathfrak{sl}_2\mathbb{C}$. We know how to classify multiplets in this case. We need only one number $n \in \mathbb{N}$. In physics however, it is custom to normalize differently (the 2 arised as eigenvalue of H), such that the number n can now be half-integer: $n = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Then so are the representations of $\mathfrak{su}(2)$ classified, so also $SU(2)$.
The theory of spin was already present, which makes this a somewhat natural extension. Spin arises from interactions invariant under rotations due to symmetric potentials. Heisenberg proposed that likewise, the proton and neutron were invariant under isospin rotations.
If we see the proton and neutron as two states of one particle, we may denote them as two states in isospin space, that is two basis vectors

$$|p\rangle = e_1 \qquad |n\rangle = e_2$$

Now one wants to mix these components by means of $\mathfrak{sl}_2\mathbb{C}$ transformations. Now we group the neutron and proton under a isospin doublet, a multiplet consisting of two states. The smallest nontrivial multiplet we can take is the one classified by the number $\frac{1}{2}$. We call this number T, the *total isospin*. Now we can write these particles in the basis associated to this classification.

We take our standard basis with a little modification to the diagonal operator such that they are classified by half-integer numbers. That is the basis $\{E, F, \frac{1}{2}H\}$, which we call $\{\hat{T}_1, \hat{T}_2, \hat{T}_3\}$. Then we can denote the proton and neutron as $|T, T_3\rangle$, which is $|p\rangle = \left|\frac{1}{2}\frac{1}{2}\right\rangle$ and $|n\rangle = \left|\frac{1}{2} - \frac{1}{2}\right\rangle$.

Looking at other particles, we can again find three very similar particles. These are the pions, π^-, π^0 and π^+. Let us again look at their properties.

	Mass (in [MeV])	Spin	Lifetime (in [s])	Charge (in e)
π^-	139.59	0	$(2.55 \pm 0.03) \times 10^{-8}$	-1
π^0	135.00	0	0.83×10^{-16}	0
π^+	139.59	0	$(2.55 \pm 0.03) \times 10^{-8}$	0

Again, we see three very similar particles. Distinguish them again by isospin we can take the next smallest representation with three eigenspaces, whih has $T = 1$. Then their eigenvalues on the diagonal operator T_3 will be $\{1, 0, -1\}$. We can group them into an isotriplet with $T = 1$ as it is a three dimensional representation of $\mathfrak{sl}_2\mathbb{C}$

$$\left|\pi^+\right\rangle = |1,1\rangle$$
$$\left|\pi^0\right\rangle = |1,0\rangle$$
$$\left|\pi^-\right\rangle = |1,-1\rangle$$

This process seems somewhat arbitrary in the sense that we have stated whether this property of these particles, their isospin, can be observed. Spin for example can be demonstrated (Stern-Gerlach experiment). However, it turns out we can indeed see isospin as something more than bookkeeping.

In quantum mechanics, when decay takes place a certain isospin state decays into two or more different isospin states. The decay that can take place must satisfy certain rules. The way in which to calculate how a particle with a certain state will decay is done by seeing it as a composition of states of the end products with coefficients. These coefficients are called *Clebsch-Gordan coefficients*. They also arise from the representation theory but we will not calculate them. They can be interpeted as a chance of the particle decaying to those particular states. This theory may be found in 5.5 of [5]. Let us take as an example proton-deutron scattering as in [5] page 115-116. We consider two possibilities that can occur

$$p + d \to \pi^0 +^3 He$$
$$p + d \to \pi^+ +^3 H$$

One may calculate that the initial state in isospin space is $\left|\frac{1}{2}, \frac{1}{2}\right\rangle |0,0\rangle$, whereas the final states are $|1,1\rangle \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ and $|1,0\rangle \left|\frac{1}{2}, \frac{1}{2}\right\rangle$. The likelyhood of either of these interactions occuring may be calculated from the Clebsch-Gordan coefficients. It turns experimentally that the results are very similar to the results aquired from the mathematical theory. Over time, many such experiments have been conducted, giving credibility to isospin invariance.
We can therefore say that isospin is indeed a real property of particles.

3.4 Hypercharge and $\mathfrak{sl}_3\mathbb{C}$

As we saw, we can classify particles using isospin and thus classify them into multiplets (i.e. doublets, triplets). However, physicists wished to group these particles in bigger multiplets, thereby extending the symmetry group.

Historically, we need to start with the notion of strangeness. This notion was introduced to describe unexplained phenomena. For example, in proton-proton collisions there were single π-mesons being produced. However, K-meson production only occured in pairs. Furthermore, these K-mesons appeared to not be able to decay in a strong fashion. There was some mechanism preventing this.
This strange behaviour prompted the proposition of a new quantity called *strangeness*. The relative strangeness of particles can be determined by looking at production and decay reactions. To normalize, the proton and electron were assigned zero strangeness. In this way, we can assign to all particles their strangeness S. This strangeness was assumed to be a conserved quantity for strong interactions. Note that this means it can still be violated by, for example, a K^0 meson ($S = 1$) decaying to two π^0 mesons ($S = 0$). However, the decay is much slower than one would expect from strong interaction, so this is weak decay. So, strong interactions conserve strangeness, whereas weak interactions don't.

Empirically, a rule for the charge was derived using the concepts we saw so far. This is known as the *Gell-Mann - Nishijima rule* (introduced independently in 1953) which says

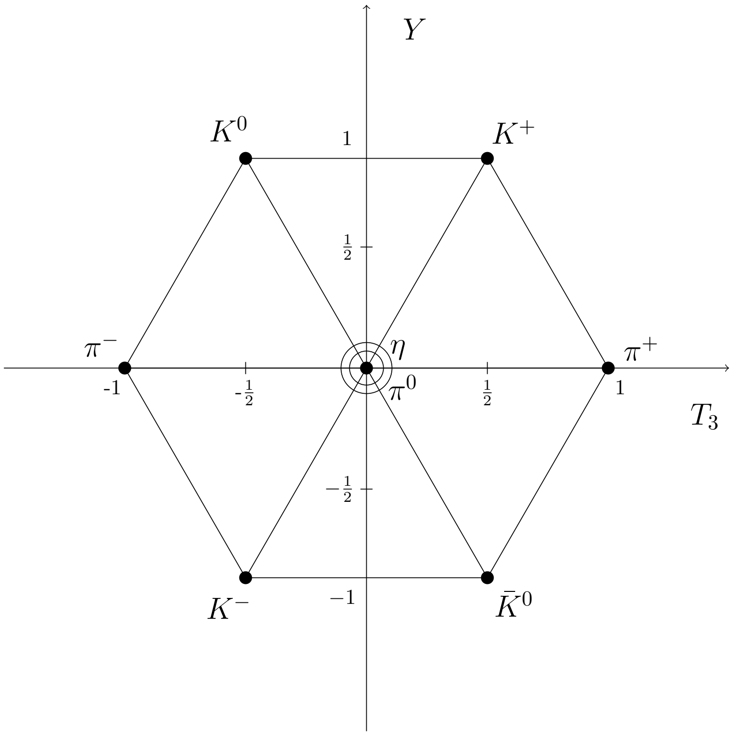
$$Q = \frac{1}{2}Y + T_3 \tag{9}$$

Here, T_3 is the three-component of the isospin. We are dealing with a representation of $\mathfrak{sl}_2\mathbb{C}$. There we had two operators whose actions shifted between eigenspaces and one diagonal operator determining which eigenspace one is looking at. The operator T_3 is that operator. The operator T is the operator which determines what representation we are dealing with. The eigenvalue for T for all particles within a multiplet is the same. The particles within the mulitplet are distinguished by their values of T_3, as T and T_3 are diagonalized simultaneously.
In terms of operators we had $Y = S + B$ is the *hypercharge*, where S is the strangeness and B the *baryon number*. The baryon number is a way to distinguish baryons and mesons among the hadrons. Baryon number 1 is associated to baryons, baryon number-1 to antibaryons and baryon number 0 to (anti)mesons. This new quantum number Y suggests we should enlarge the symmetry group of strong interactions. For isospin we had $SU(2)$ symmetry. A new theory was introduced, called $SU(3)$ *flavour symmetry*. The idea was, like for isospin, to group all hadrons into multiplets of this $SU(3)$ flavour symmetry.

Again, we can construct multiplets which will now be two dimensional. However, the masses thoughout the multiplet will be very different, with an, as we will see, up to 40% mass difference. The idea was that the lightest mesons should form a multiplet. These are the following particles. They are all mesons with zero spin.

Particle	Mass (in [MeV])	Isospin T	Isospin component T_3	Hypercharge Y
$\pi^{\pm}$	139.6	1	±1	0
π^0	135.0	1	0	0
$K^{\pm}$	493.7	$\frac{1}{2}$	$\pm\frac{1}{2}$	±1
K^0	497.7	$\frac{1}{2}$	$-\frac{1}{2}$	1
$\bar{K}^0$	497.7	$\frac{1}{2}$	$\frac{1}{2}$	-1
η	549.0	0	0	0

We can now draw these particles in the $Y - T_3$-plane using these properties. This yields the following picture



We can do this for other groups of particles in this empirically based way. However, a model was devoloped in which one may assign the known particles to composite states of a new type of particle. In this way, all emperical data coincides with the results of this model, proving its usefulness. The model is the quark model.

3.5 Quark model

As we have seen we can group particles in multiplets in an empirical way. What we look for is a model that explains this behaviour. Such a model is the quark model, in which we introduce new particles that build the known particles and derive their properties using this model.
Historically, empirical data suggested that indeed particles could be grouped up into multiplets as we saw in the previous section. Looking closely at this picture, it has the shape of the $D^{(1,1)}$ representation of $\mathfrak{sl}_3\mathbb{C}$. Inspired by the succesful idea of isospin, let us extend this to the representation theory of $SU(3)$, where the multiplets correspond to irreducible representations of $\mathfrak{sl}_3\mathbb{C}$.
Now the particles from the previous section form a multiplet. The question was what the most fundamental representations, out of which one can build all others, should correspond to.
These are new particles called quarks, and the model they arise in is called the quark model. From a standpoint of group representation, let us assign operators to the quantum numbers like we did in the case of isospin. The fundamental representations will then arise as the multiplets belonging to quarks. We will now drop the numbers corresponding to the Casimir operators that distinguish between the different multiplets and just look at the multiplets on their own. We now have two quantum numbers to distinguish particles by; the hypercharge Y and the third component of the isospin T_3. Both Y and T_3 should be assigned diagonal operators. For the rest, we take the basis we are familiar with. Only, we have now taken the diagonal elements differently

(that is, a different basis for $\mathfrak{h}$), so some normalization slightly changes. This is the custom way to represent these operators in physics. We take the following basis of $\mathfrak{sl}_3\mathbb{C}$

$$Y = \frac{1}{3}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -2\end{pmatrix} \qquad T_3 = \begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix} \qquad T_\pm = E_{\pm\alpha}, \qquad U_\pm = E_{\pm\beta}, \qquad V_\pm = E_{\pm\gamma}$$

The theory remains quite the same, only the eigenvalues will be somewhat different. Let us again calculate some commutation relations. We already know the commutation relations on $T_\pm$, $U_\pm$ and $V_\pm$. However, commutators between these operators and the new basis of the Cartan subalgebra will give a somewhat different numbers. We will still have that $[E_\delta, E_{-\delta}] \in \mathfrak{h}$ for $\delta \in \{\pm\alpha, \pm\beta, \pm\gamma\}$, but it will be a different linear combination of the operators. For completeness let us here show all commutation relations.

$$\begin{array}{cccc}
[T_3, T_\pm] = \pm T_\pm & [T_3, U_\pm] = \mp\frac{1}{2}U_\pm & [T_3, V_\pm] = \frac{1}{2}V_\pm & [T_3, Y] = 0\\
[Y, T_\pm] = 0 & [Y, U_\pm] = \pm U_\pm & [Y, V_\pm] = \pm V_\pm & \\
[T_+, T_-] = 2T_3 & [U_+, U_-] = \frac{2}{3}Y - T_3 & [V_+, V_-] = \frac{2}{3}Y + T_3 & \\
[T_+, U_+] = V_+ & [T_+, U_-] = 0 & [T_+, V_+] = 0 & [T_+, V_-] = -U_-\\
[T_-, U_+] = 0 & [T_-, U_-] = -V_- & [T_-, V_+] = U_+ & [T_-, V_-] = 0\\
[U_+, V_+] = 0 & [U_+, V_-] = T_- & [U_-, V_+] = -T_+ & [U_-, V_-] = -U_-
\end{array}$$

These are all the commutation relations (apart from, of course, the ones aquired by the skew symmetry of $[\ ,\ ]$). We see that these operators shift eigenspaces in six directions as before. Also note again the positive and negative versions of the operators T, U and V together with their commutator each form a $\mathfrak{sl}_2\mathbb{C}$ subrepresentation within our representation of $\mathfrak{sl}_3\mathbb{C}$. In this case, the $T_\pm$ operators form the $\mathfrak{sl}_2\mathbb{C}$ representations associated with the isospin. This means that taking the directions of $T_\pm$ horizontal, in our representations isospin multiplets can be found along a horizontal line.

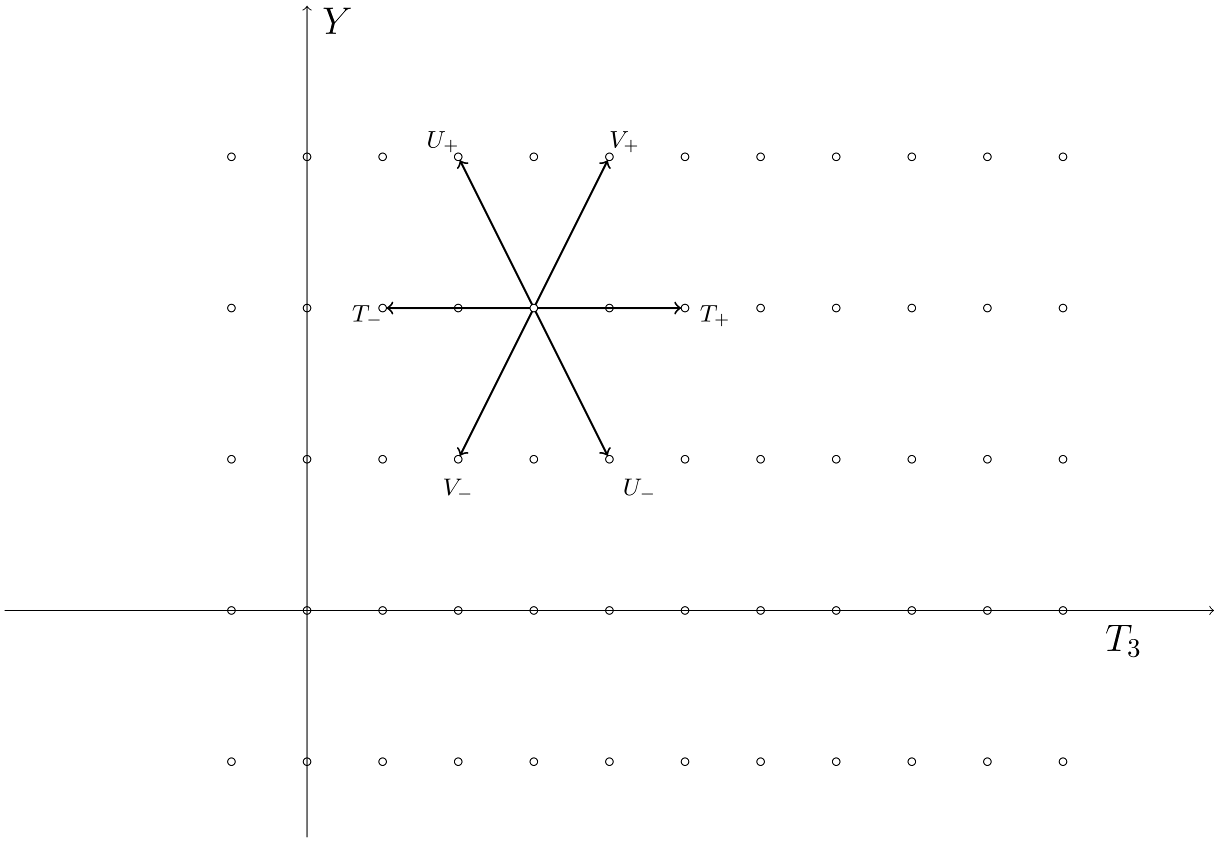
We now introduce the quark model. This is a model that explains the empirical data about particles regarding their isospin and hypercharge. It arises as the empiral data suggests a multiplet structure, which by using representation theory can be built using fundamental representations. The quarks and antiquarks then correspond to these fundamental representations (there are two). At first, the idea of quarks was merely regarded as mathematical bookkeeping. Only after many years, they were accepted as being "real" particles.
We will only discuss the quark model with three quarks. There are more flavours of quarks, but these three are historically the first ones proposed. Furthermore, the three flavour model is quite accurate as the weights of the quarks are very close together. The other quarks have very different weights, which is why it is hard to fit them into multiplets.

We begin by constructing the smallest possible non-trivial representation of $SU(3)$, the fundamental ones. In the basis we have picked, we can first look at the smallest isospin doublet we can aquire. Because the normalization chosen, the smallest representation has two eigenspaces with eigevalues $\frac{1}{2}$ and $-\frac{1}{2}$ for T_3. This is what we have seen before: it is exactly the property of $\mathfrak{sl}_2\mathbb{C}$, only now we have normalized the diagonal element associated with the $\mathfrak{sl}_2\mathbb{C}$ subalgebra differently, such that the eigenvalues differ by a factor $\frac{1}{2}$. However, also U-spin and V-spin ($U_\pm$ and $V_\pm$ with their commutatosr, as before) form $\mathfrak{sl}_2\mathbb{C}$ subalgebras. From our theory of $\mathfrak{sl}_3\mathbb{C}$ we know what the smallest non-trivial representations look like: these are $D^{(1,0)}$ and $D^{(0,1)}$.

However, like the $\mathfrak{sl}_2\mathbb{C}$ subalgebras, the normalization is different. Let us calculate what the eigenvalues corresponding to this representation are. We can again calculate how the operators shift the eigenvalues of Y and T_3 using the commutation relations. Let us calculate the action of V_+ as an example, and summarize the rest of the operators in a picture. Given a vector $v \in V_\alpha$ in some eigenspace, we see

$$\begin{aligned} T_3(V_+(v)) &= [T_3, V_+](v) + V_+(T_3(v)) \\ &= \frac{1}{2}V_+(v) + \alpha(T_3)V_+(v) \\ &= (\alpha(T_3) + \frac{1}{2})V_+(v) \end{aligned} \qquad \begin{aligned} Y(V_+(v)) &= [Y, V_+](v) + V_+(Y(v)) \\ &= V_+(v) + \alpha(Y)V_+(v) \\ &= (\alpha(Y) + 1)V_+(v) \end{aligned}$$

As we see, V_+ increases the value of T_3 by $\frac{1}{2}$ and the value of Y by 1. We can make the same calculation for all operators, and get the following picture.



From this picture one may easily read what the action of the different operators does to the eigenvalues of T_3 and Y. Along the T, U and V lines we have subrepresentations that are isomorphic to representations of $\mathfrak{sl}_2\mathbb{C}$, like we saw in the classification of the irreducible representations of $\mathfrak{sl}_3\mathbb{C}$.

Let us now look at what happens for the smallest representation. We have the three basis vectors e_1, e_2 and e_3, which we shall denote u, d and s respectively. Then the fundamental representation $D^{(1,0)}$ is just given by letting the matrices of $\mathfrak{sl}_3\mathbb{C}$ act on these basis vectors, as before. We get three eigenspaces corresponding the the span of each basis vector. Let us denote the vectors in a more physical notation as a ket where the first number indicates the eigenvalue of T_3 and the second one the eigenvalue of Y. Unlike before we do no denote the eigenvalues of the Casimir as we just calculate within multiplets and this would be cumbersome. Furthermore, Casimir operators are difficult to find so we just keep in mind their usefulness for keeping track of multiplets. We then have

$$|u\rangle = \left|\frac{1}{2}, \frac{1}{3}\right\rangle \qquad |d\rangle = \left|-\frac{1}{2}, \frac{1}{3}\right\rangle \qquad |s\rangle = \left|0, -\frac{2}{3}\right\rangle$$

Note that by a historical coincidence, the strange quark has negative strangeness $S = -1$. We also know how to construct the other fundamental representation, by acting with a factor -1 on the transpose. We denote e_1^*, e_2^* and e_3^* by $\bar{u}$,$\bar{d}$ and $\bar{s}$ respectively. We then have in the same way

$$|\bar{u}\rangle = \left|-\frac{1}{2}, -\frac{1}{3}\right\rangle \qquad |\bar{d}\rangle = \left|\frac{1}{2}, -\frac{1}{3}\right\rangle \qquad |\bar{s}\rangle = \left|0, \frac{2}{3}\right\rangle$$

What we do now is construct particles that are composed of these so called quarks. In quantum mechanics particles are composed as tensor products. That is, the composed particle is in a state that is the tensor product of its constituents. In this case, we can also take tensor products; we have seen before how this works. Let us compose compose an $|u\rangle$ and $|\bar{s}\rangle$. We calculate T_3 on $|u\rangle \otimes |\bar{s}\rangle$, as we are interested in the eigenvalues. This will give the values of T_3 and Y of the composit particle. We repress the $\otimes$-sign for cleaner notation.

$$\begin{aligned}
T_3(|u\rangle\,|\bar{s}\rangle) &= (T_3\,|u\rangle)\,|\bar{s}\rangle + |u\rangle\,(T_3\,|\bar{s}\rangle) \\
&= \frac{1}{2}\,|u\rangle\,|\bar{s}\rangle + |u\rangle\,0\,|\bar{s}\rangle \\
&= \frac{1}{2}\,|u\rangle\,|\bar{s}\rangle \\
Y(|u\rangle\,|\bar{s}\rangle) &= (Y\,|u\rangle)\,|\bar{s}\rangle + |u\rangle\,(Y\,|\bar{s}\rangle) \\
&= \frac{1}{3}\,|u\rangle\,|\bar{s}\rangle + |u\rangle\,\frac{2}{3}\,|\bar{s}\rangle \\
&= |u\rangle\,|\bar{s}\rangle
\end{aligned}$$

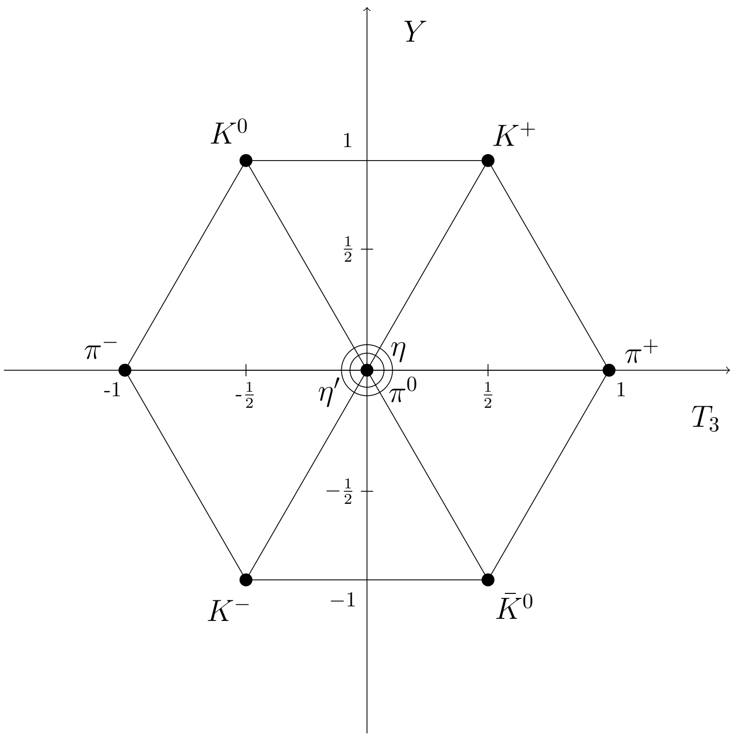
So we see the composite particle $|u\rangle \otimes |\bar{s}\rangle$ is the state $\left|\frac{1}{2}, 1\right\rangle$. We already know a particle with $T_3 = \frac{1}{2}$ and $Y = 1$; the kaon K^+. In this way we can associate to the quark content $|u\rangle \otimes |\bar{s}\rangle$ the particle K^+, such that we can say the particle K^+ is made up of one up quark and one anti-strange quark. One can now repeat this same computation for all combinations of quarks.

Mesons are made up of a quark-antiquark pair. The baryons will be a composite particle of three quarks (or three antiquarks). We will not look at any other possibilities. This has to do with the theory of quantumchromodynamics (QCD), in which an extra symmetry plays a role. This is associated to another charge possessed by quarks, the colour charge. Every quark also has a colour, wheras antiquarks have anticolours (red, blue, green and antired, antiblue, antigreen). The theory then says that hardrons must have neutral colour charge. This means a colour and its anticolour, or three different colours. Then we can only construct particles consisting of a quark and an antiquark, or of three quarks or three antiquarks. We shall however not concern ourselves with the theory of QCD and just restrict ourselves to the situation of q and $\bar{q}$ or three (anti-)quarks.

If we make the calculations, we see there are several composites with quantum numbers $T_3 = 0$ and $Y = 0$. Let us denote the fundamental representations $D^{(1,0)}$ and $D^{(0,1)}$ by their dimensions,

as this pertains to the number of particles in the multiplet (often used notation). Then we have [3] and $[\bar{3}]$ respectively.
Then what we are looking at for the mesons is $[3] \otimes [\bar{3}]$. Multiplets that arise as composites of these quarks are irreducible representations in this representation $[3] \otimes [\bar{3}]$. This representation is not irreducible, as it has multiplicity three for eigenvalue 0. We can see for example that the linear combination $\frac{u\bar{u}+d\bar{d}+s\bar{s}}{\sqrt{3}}$ has $T_3 = 0$ and $Y = 0$, and is killed by all the operators $T_\pm$, $U_\pm$ and $V_\pm$. This means it is an trivial irreducible subrepresentation $[1] \subset [3] \otimes [\bar{3}]$. In fact we can compute how $[3] \otimes [\bar{3}]$ decomposes in irreducible components. We have $[3] \otimes [\bar{3}] = [8] \oplus [1]$. By associating particles we have $\eta' = \frac{u\bar{u}+d\bar{d}+s\bar{s}}{\sqrt{3}}$. This particle η' is called the singlet, whereas the other eight particles belonging to the other multiplet are knows as the octet. Together they form a meson nonet. The normalization of η' is picked because of the probability interpetation of quantum mechanics. This means that when looking at the quark content of a η' particle, there is a chance of $\frac{1}{3}$ to find $u\bar{u}$, $d\bar{d}$ or $s\bar{s}$ as its content.

We get the following diagram. These are the so called pseudoscalar mesons. The nomenclature is of too much theoretical depth but we briefly give some properties. First of all they are all particles with 0 spin, which is integer as they are mesons. Secondly they are particles with odd parity. Parity is a concept that requires some knowledge of quantum field theory to explain, so we shall not delve into it here. Let us just say it is a property of particles that can be observed by how they interact with specific other particles, which is useful for classifying particles. These properties of the pseudoscalar mesons can be summarized in notation as $J^P = 0^-$ (here J stands for spin and P stands for parity).
In a diagram with Y and T_3 on the axis, the different pseudoscalar mesons may be depicted as such



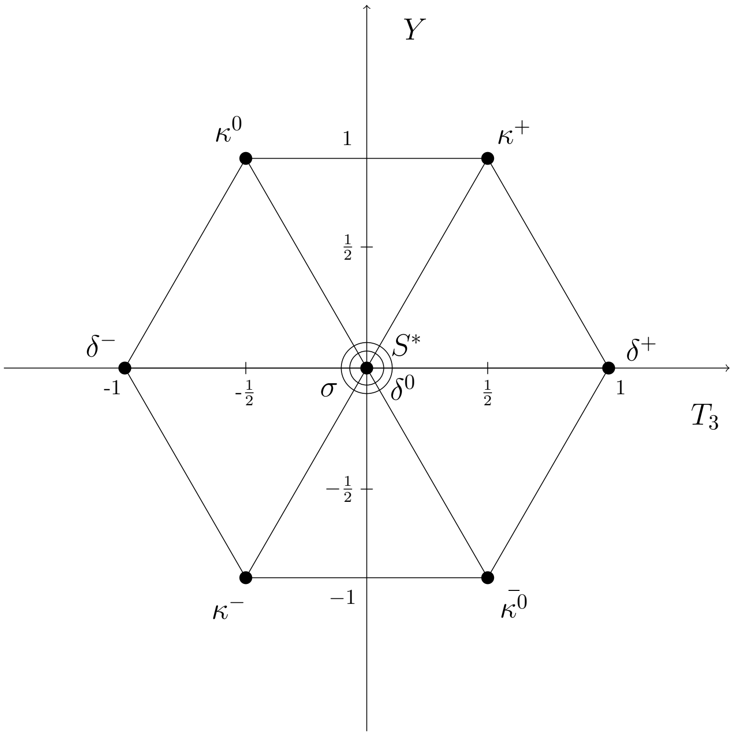
Note we can also put into this diagram the charges of the particles using the Gell-Mann Nishijima rule, where lines of constant charge turn out to be lines in the diagonal direction (left bottom to top right). We do not plot these as it is just a coincidence derived from the Gell-Mann Nishijima rule that one might notice. We then get the following quark content

Name	Multiplet	Quark content
π^+	[8]	$u\bar{d}$
π^-	[8]	$d\bar{u}$
π^0	[8]	$\frac{1}{\sqrt{2}}u\bar{u} - d\bar{d}$
K^+	[8]	$u\bar{s}$
K^-	[8]	$s\bar{u}$
K^0	[8]	$d\bar{s}$
$\bar{K}^0$	[8]	$s\bar{d}$
η	[8]	$\frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s}$
η'	[1]	$\frac{1}{\sqrt{3}}(u\bar{u} + d\bar{d} + s\bar{s}$

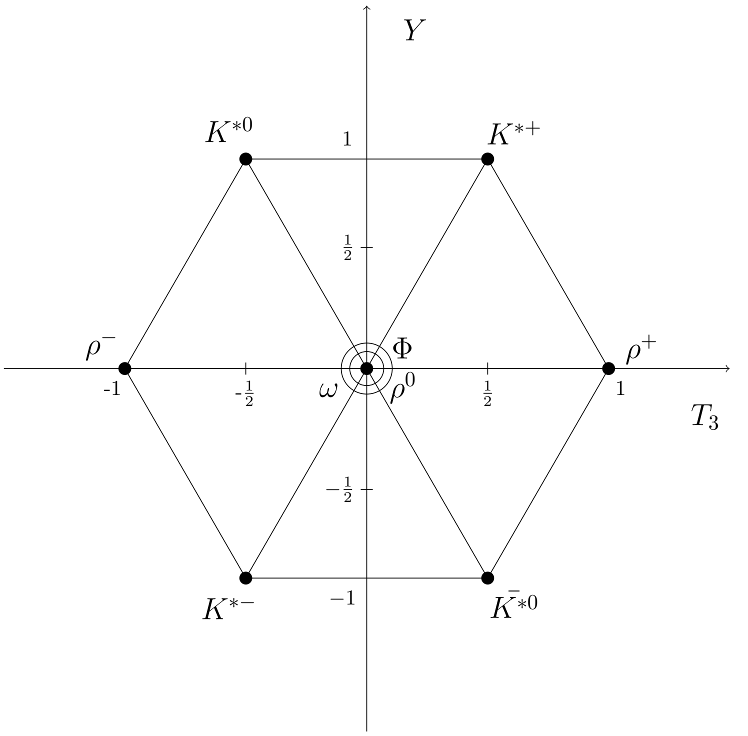
Note that the quark content of the particles with $T_3 = Y = 0$ may seem strange. The normalization factor is chosen such that the square sum of the coefficients equals 1. However, which linear combinations to pick for which particle follows from the Clebsch-Gordan coefficients as these follow from reactions within the multiplet. It is in fact quite involved, but a full calculation may be found in [5] 8.10. This always happens in the middle of the meson multiplets. For baryons it will be much easier as there this problem does not occur; no such fractional contents arise.

3.5.1 Mesons

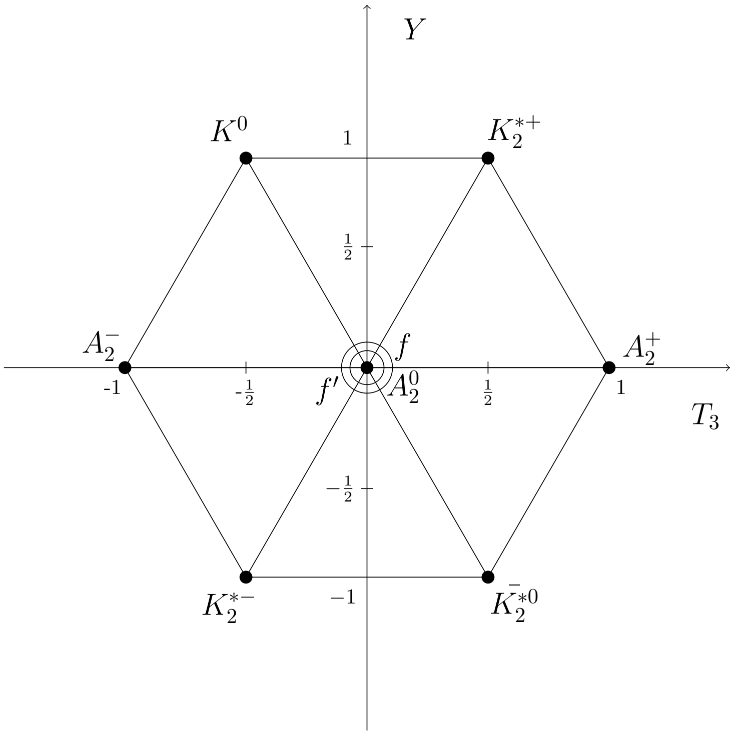
We encountered a meson multiplet before. However, there are more mesons with different spins and parities. The point though is essentially the same. From the particles known, empirical data suggests which particles form the multiplets. The different meson multiplets differ in spin and parity. This might seem strange as this means quark content does not fully determine the properties of a particle. Particles with the same quark content may have properties such as spin. However, the way these particles interact cannot be seen from this model which only takes into account a single symmetry and no interactions.
We will merely connect the empirical data to the multiplets with the corresponding values of Y and T as we did before. The pictures will be similar to the one of pseudoscalar mesons, only the actual particals are different. We list here several types of mesons, from which their properties can be deduced from the pictures. Let us explore other mesons. Another multiplet of mesons is the scalar mesons, which can be summerized by denoting them $J^P = 0^+$. They too belong to a meson nonet which looks like this



Then there are several more. We also have the vector mesons $J^P = 1^-$, which look like



and lastly the tensor mesons, the $J^P = 2^+$ mesons.



These are all mesons. We now wish to construct the particles that are composed of three quarks, which will be the baryons.

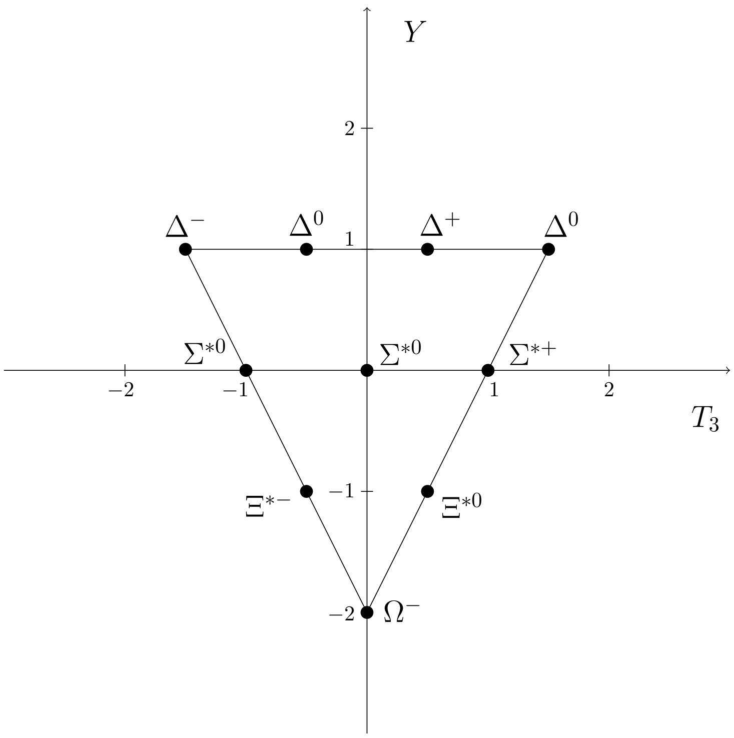
3.5.2 Baryons

We can also construct the baryons as $SU(3)$ multiplets like we did for the mesons. This means we need to decompose the representation [3]$\otimes$[3]$\otimes$[3]. We quickly invoke the previous notation, which means we wish to find the decomposition of $V \otimes V \otimes V$ in terms of the irreducible representations $D^{(m_1,m_2)}$. One may calculate that this can be decomposed as

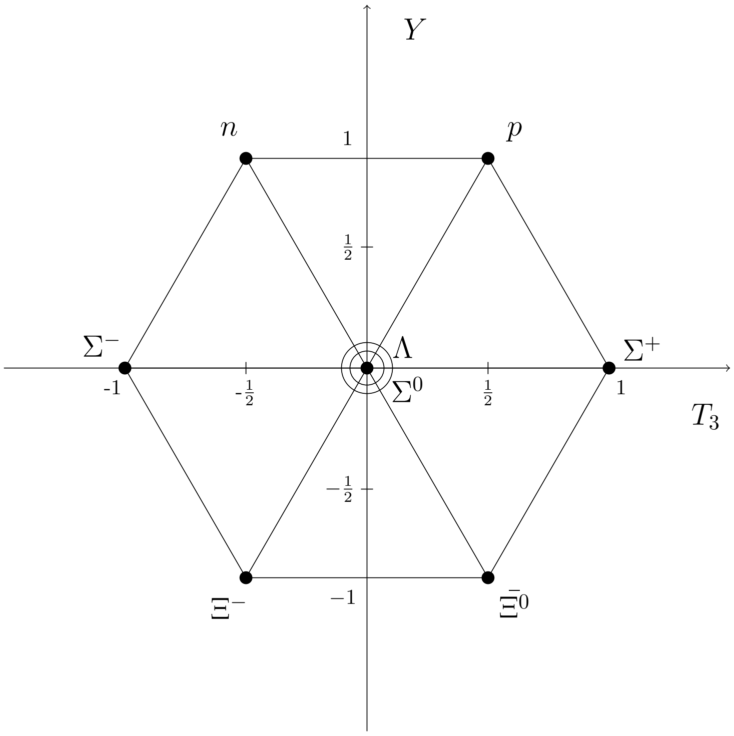
$$V \otimes V \otimes V = D^{(3,0)} \oplus D^{(1,1)} \oplus D^{(1,1)} \oplus D^{(0,0)}$$

That is, $[3] \otimes [3] \otimes [3] = [10] + [8] + [8] + [1]$. We already saw the proton and neutron, which were baryons. As they are made of only up and down quarks, knowing that their baryon number should be 1 implies that the hypercharge of the proton and neutron should be 1. We can do something similar for the Δ baryons. These are baryons with strangeness $S = 0$. Only this time, there are four particles in a isospin multiplet with $Y = 1$. This is exactly the shape associated to [10], which is a triangle with four weight spaces at $Y = 1$.
Now we can continue this process of classifying. For strangeness $S = -1$ we have the Σ baryons, as well as the Λ^0-baryon. Now the Σ baryons are Σ^-, Σ^0 and Σ^+ and their resonances Σ^{*-}, Σ^{*+} and Σ^{*0}. Resonances are very short lived particles that are excited states of other particles. However, the interaction does not concern us and their property are exactly what we want, so we just view them as particles in this model.
The particles and their resonances are distinguished by their spin which is $\frac{1}{2}$ for the first triplet,

and $\frac{3}{2}$ for the second triplet. Hence, we group the first ones with the proton and neutron in the [8] representation, and the second ones in the [10] representation where we had the Δ-baryons, also with spin $\frac{3}{2}$. The same story holds for the Ξ particles, which will be in the [8] representation with $Y = -1$, whereas their resonances will be in the [10] representation with $Y = -1$. Resonances of particles are excited states of these particles. They have different properties and are only very short lifed as they nearly immediatly decay. However, they are particles satisfying this multiplet structure in the sense that they have the required properties and therefore fit into the model.
To fill the multiplet [10] we need one more particle, which should be a $S = -3$ (or $Y = -2$) baryon with spin $\frac{3}{2}$. At the time the quark model was being developed, a particle with these properties had not been found. However, in 1964 the Ω^- particle as discovered. This particle had the right properties and the right decay products predicted by the quark model. This was a great triumph for the quark model, showing its usefulness. So far it seems as if one might only use it for classification purposesess. The discovery of the Ω^--baryon, however, shows that this also works the other way around in which the theory predicts the existence of real particles.
Summarizing the previous findings about the baryons, we draw similar pictures to the ones drawn for the mesons. We begin with the baryons associated to the [10]-representation, the baryon decuplet. By knowing their place in the diagram, we know their place in the representation and hence we are able to calculate the quark content of all these particles. The diagram for the baryon decouplet looks like this



Secondly, there are the baryons corresponding to the [8] representation, which is the baryon octet


Y
n
p
1
$\frac{1}{2}$
Σ^-
Λ
Σ^+
-1
-$\frac{1}{2}$
Σ^0
$\frac{1}{2}$
1
T_3
$-\frac{1}{2}$
Ξ^-
-1
Ξ^0

4 Conclusion

We have seen that using the mathematical concept of representation theory, one may produce a classification of light hadrons that is quite accurate. It is a great example of how representation theory may be used in physics. Its theory is quite concise, whereas for full understanding of these particles one needs knowledge of quantum fieldtheory, with reasonably little knowledge we can set up such a classification scheme. Though the quark model was very succesful, the small differences between model and reality become large for heavier particles by introducing more flavours of quarks. In due time, more quark flavors were postulated. Outside of the three mentioned in this thesis there are also the charm, top and bottom quark. For these higher numbers of quarks, the quark model becomes more and more inacurate. The deviations between masses within the multiplets become significantly large. This has to do with the fact the other quark flavours are extremely heavy. For three flavours, the quark model works very well, but it quickly breaks down for more flavours. The multiplet structure interpeted mathematically assumes that the weights within multiplets are exactly the same. That is, in a perfect world where flavour symmetry is exact, the proton and neutron have exactly the same weights. However, they don't, as this model does not take into consideration these weight differences as well as many other influences (such as magnetic and electric forces). When the quark masses are very different, the model becomes inaccurate and it is no longer sensible to fit the particles into multiplets. Furthermore, the model is not particularly useful for predicting particles. It primarily is a scheme by which to classify particles.

However, the concept of using such representation theory is present in many areas of physics. We have alluded to one particular use, that of quantum chromodynamics. Here, particles are assigned colour charge, and the operators of $SU(3)$ mix components of these colours. The same theory applies, but this time, QCD appears to be an exact symmetry of the strong force, in contrast with the flavour symmetry. The concepts of representation theory however, apply in the same way as the underlying $SU(3)$ symmetry is still the same. In this way, knowledge of representation theory of more Lie groups can be very useful in very many areas of physics. The theory covered in this thesis might serve as an introduction or priliminary to more complicated applications of representation theory in physics.

The symmetry of some physical systems may be more extensive, for example bigger groups. However, the concepts presented in this thesis may readily be extended. Objects like Weyl groups and Weyl chambers are very general in the sense of semisimple Lie algebras, and with the knowledge we have now we can easily extend to many more Lie algebras. Examples would be more general ones like $\mathfrak{sl}_n\mathbb{C}$, $\mathfrak{sp}_n\mathbb{C}$ and $\mathfrak{so}_n\mathbb{C}$.

The geometric interpetation of the weight diagrams can also be extended, and they can be beautifully summarized and classified using so called Dynkin diagrams. One may summarize all important information of a Lie algebra into a single small drawing. In this way, all Lie algebras may be classified in classes of classical ones for which the structure is similar throughout higher dimensions and five exceptional ones. The theory of $\mathfrak{sl}_2\mathbb{C}$ and $\mathfrak{sl}_3\mathbb{C}$ is only a small part of this larger context.

5 Acknowlegdements

I would like to thank my supervisor Andr Henriques for taking the time to answer my many questions. I appreciate his patience and many helpful tips to get through the theory and his help with scrutinizing the proofs. I would also like to thank Stefan Vandoren for answering the questions that were raised when studying the quark model and flavour symmetry in clear terms and explanations.

6 References

References

[1] W. Fulton, J. Harris *Representation Theory* 1991: Springer-Verlag New York, Inc.

[2] J. E. Humphreys *Introduction to Lie algebras and Representation Theory* 1972: Springer-Verlag New York, Inc.

[3] D. H. Sattinger, O.L. Weaver *Lie Groups and Algebras with Applications to Physics, Geometry and Mechanics* 1986: Springer-Verlag New York, Inc.

[4] J.P. Elliott, P.G. Dawber *Symmetry in Physics* 1979: The Macmillan Press Ltd.

[5] W. Greiner, B. Mller *Quantum Mechanics: Symmetries* 1989: Springer-Verlag Berlin Heidelberg.

[6] B. C. Hall *Lie Groups, Lie Algebras, and Representations* 2003: Springer-Verlag New York, Inc.

[7] B. Weber *Course on Lie algebras, Lecture 7*
$http://www.math.upenn.edu/ brweber/Courses/2012/Math650/Notes/L7_CompleteReducibility.pdf$