The classification of bound quark states

Schijven, P.A.M.

P.A.M. Schijven

The classification of bound quark states

Bachelorscriptie, 15 januari 2009

Scriptiebegeleiders: dr. R.J. Kooman
prof. dr. A. Achucarro

Mathematisch Instituut, Universiteit Leiden

Contents

Introduction

Symmetry has always been an important concept in the history of science. For example, in the time of the ancient Greeks, Plato suggested in his dialogue *Phaedo* that all forms in nature try to be like their perfect symmetric forms. Like that a line in real nature always has a length, breadth and height, but that a perfect line only has length as a dimension. The Pythagorean school believed that some form of harmony or symmetry underlies all things in the universe. For instance, they applied this harmony to their theory of music.

In the 20th century the idea of symmetry arose once more among physicists. In the year 1915, the German mathematician Emmy Noether discoverd her famous theorem which relates continuous symmetries with conserved quantities in nature. This proved to be the solution to the problem of the failure of local energy conservation in the theory of general relativity. Some years later, after the advent of quantum mechanics and the discovery of a large number of different elementary particles, physicists needed some way to classify all these particles. Again, the solution was to look at the symmetries of nature. In particular, the physicist Gell-Mann reinvented the theory of some particular continuous groups to classify the particles in a scheme which he called *"The eightfold way"*. This classification scheme also led to the hypothesis that all the elementary particles are build up of quarks and antiquarks.

In this bachelor thesis we explore the ideas of Gell-Mann to classify the simple elementary particles. First we will look at the basic mathematics of Lie groups. These Lie groups are basically the result of the unification of the theory of abstract groups with the theory of manifolds. First, we will define what Lie groups are and then look at how we can study them by looking at their representations. We will see that to do this, we'll need to look at the Lie algebra that is associated to the Lie group. Finally we will prove the important Campbell-Baker-Haussdorf formula. This formula basically gives us the result of a product of 2 exponentials of matrices.

In the next chapter we will explore the idea of symmetry in physics. First, we'll look at how we can use symmetry to study the solutions of the Schrödinger equation. As a prime example we'll study the symmetry group $SU(2)$ to classify the lightweight elementary particles in terms of isospin. To do this we will first give a complete classification of the irreducible representations of $SU(2)$. Then we apply these results to construct the elementary particles out of the two light quarks, namely the up and down quarks.

In the last chapter we will study the quark model set up by Gell-Mann. First we will introduce some general quantities used in elementary particle physics, such as the baryon- and lepton number, strangeness and hypercharge. Then we introduce the $SU(3)$ symmetry to use the three types of light quarks to construct the multiplets for the mesons and baryons. Again, this will be done by looking at the irreducible representations of $SU(3)$. As a last application, we will derive a mass formula, by which we predict the mass of the Ω^- particle.

Chapter 1

Lie groups and Lie algebras

1.1 Lie groups

The theory of Lie groups is widely used in physics and in numerous parts of mathematics. The general idea is to consider a differentiable [1] manifold which is also a group and where the multiplication and inverse operations are differentiable.

Definition 1.1.1. A Lie group G is a differentiable manifold with a group structure defined on it, such that the maps $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are differentiable.

Definition 1.1.2. A Lie subgroup $H \subset G$ of G is both a subgroup and a submanifold of G.

Definition 1.1.3. A map between Lie groups G and H is a homomorphism $\rho : G \to H$ such that ρ is differentiable.

There are plenty of examples of Lie groups. One of the most simple examples is the real line $\mathbb{R}$ with addition along a line as the group operation. In this thesis we'll only look at matrix Lie groups. The most general matrix Lie group is of course the group of linear transformations of an n-dimensional vector space with nonzero determinant, which is denoted as $GL(n, K)$, where K is a field. Mostly we take $K = \mathbb{R}$ or $K = \mathbb{C}$. This group has many different subgroups, like the group $SL(n, K)$ which consist of all the linear maps with determinant equal to 1. Another example is the subgroup consisting of the upper triangular linear maps. If we view $GL(n, K)$ as the group of automorphisms of a n-dimensional K-vector space, we will denote it as $\mathrm{Aut}(V)$.

It is also possible to create subgroups of $GL(n, K)$ by looking at some bilinear form $Q : V \times V \to V$ defined on V. These subgroups consist of the matrices A which preserve Q, in the sense that $Q(Av, Aw) = Q(v, w)$ for all $v, w \in V$. If $K = \mathbb{R}$ we can write $Q(v, w) = v^T M w$ for some fixed matrix M. This definition is the most general bilinear form on $\mathbb{R}$. If a matrix A preserves Q then $v^T M w = Q(v, w) = Q(Av, Aw) = (Av)^T M (Aw) = v^T A^T M A w$. This means that $A^T M A = M$. If we now take $M = I$ we get the subgroup $O(n, \mathbb{R})$, the subgroup of orthogonal matrices. The subgroup containing the orthogonal matrices with determinant 1 is denoted as $SO(n, \mathbb{R})$. We can also do the same construction for complex Lie groups by using

[1] In this thesis differentiable always means infinitely differentiable.

the general symmetric hermitian form $H(v,w) = \bar{v}^T M w$. The subgroup of $GL(n,\mathbb{C})$ which preserves H for $M = I$ is the subgroup of the unitary matrices $U(n)$. The subgroup of unitary matrices with determinant 1 is denoted as $SU(n)$.

1.2 Representations

To study the properties of Lie groups, it may be helpful to see how they may be represented as a group of linear mappings acting on some Hilbert space or vector space. This has the great advantage that we can use the full machinery of linear algebra to find out numerous properties of Lie groups. Representation theory of Lie groups is also of great importance in physics, because by looking at the representations of continuous symmetry groups (which are always Lie groups) on some Hilbert space, we can understand how these symmetries act on the solutions of a physical system.

Definition 1.2.1. A representation of a group G on a vector space V is a homomorphism $\rho : G \to \mathrm{Aut}(V)$. A subrepresentation of a representation is a homomorphism $\rho' : G \to \mathrm{Aut}(W)$, $W \subset V$ a subspace of V, such that $\rho(G)(W) \subset W$. A representation ρ on a vector space V is called irreducible if it only contains $W = 0$ or $W = V$ as subrepresentations.

Often, when the context is clear, we refer to V as the representation instead of the homomorphism ρ. There are of course many possible ways to represent a group G on a vector space V. We therefore introduce some sense of equivalence. We call two representations ρ and ρ' equivalent if $\rho'(g) = U\rho(g)U^{-1}$ for all $g \in G$ and $U \in \mathrm{Aut}(V)$ fixed.
Given a vector space V, one can perform a number of operations on it, for example taking the direct sum or the tensor product with some other vector space. One can also look at the dual space of V. Tensor products are of importance in quantum mechanics, because if you try to combine two subsystems into one larger system, the total system is given by the tensor product of the two smaller ones. If we are given representations on the vector spaces V and W, we can easily define how the related representations on $V \oplus W$ and $V \otimes W$ look like:

Definition 1.2.2. Let $\rho : G \to GL_n(V)$ and $\rho' : G \to GL_n(W)$ be two representations on respectively the vector spaces V and W. We then define the direct sum and direct product representations as follows:

1. $\rho \oplus \rho'(g)(v + w) = \rho(g)(v) + \rho'(g)(w) \in V \oplus W$
2. $\rho \otimes \rho'(g)(v \otimes w) = \rho(g)(v) \otimes \rho'(g)(w) \in V \otimes W$

To define the dual representation of a representation ρ, we remember that for every linear map $\phi : V \to V$, there exists a map $\phi^T : V^* \to V^*$, which is called *the transpose map*. This map is defined as $\phi^T(f)(v) = f(\phi(v))$. With this transpose map one easily thinks the dual representation ρ^* is defined as $\rho^*(g) = \rho(g)^T$. There is however a slight problem with this definition because ρ^* isn't a homomorphism but an anti-homomorphism: $\rho^*(g_1g_2) = \rho(g_1g_2)^T = (\rho(g_1)\rho(g_2))^T = \rho(g_2)^T\rho(g_1)^T = \rho^*(g_2)\rho^*(g_1)$. We can fix this problem easily though, by noting that if f is a anti-homomorphism, we can define a homomorphism f' by setting $f'(x) = f(x^{-1})$.

Definition 1.2.3. Let ρ be a representation of a group G on a vector space V. The dual representation ρ^* is defined as $\rho^*(g) = \rho(g^{-1})^T$.

The following lemma due to Schur is of great use in identifying the irreducible representations of a group[2]:

Lemma 1.2.1. *(Schur) Let G be a Lie group.*
A representation $\rho : G \to GL(V)$ is irreducible $\Longleftrightarrow$ the only operators on V that commute with all $\rho(g)$ are of the form $\lambda \cdot \mathrm{id}_V$.

1.3 Lie algebras

1.3.1 Algebraic definitions

In the previous two sections we talked about Lie groups and defined the notion of a representation of a Lie group. In order to study these in more detail we are going to use another algebraic structure defined on a Lie group, namely the Lie algebra. In the algebraic sense, an algebra is a K-vector space together with a compatible K-bilinear map, mostly called the "multiplication". For a Lie algebra this map is called the Lie bracket which is denoted as $[,]$. In the case of a matrix Lie algebra this Lie bracket is just the commutator of two matrices: $[A, B] = AB - BA$.

Definition 1.3.1. A Lie algebra is a vector space V together with an antisymmetric bilinear map $[,] : V \times V \to V$ which satisfies the Jacobi identity $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in V$.

The most general example of a matrix Lie algebra is the Lie algebra $\mathfrak{gl}(V)$ which consists of all the linear mappings $V \to V$, where V is a K-vector space. We can also perform all the standard constructions with Lie algebras, like looking at subalgebras, at maps between Lie algebras and looking at representations of Lie algebras on some vector space. This gives us the following set of definitions:

Definition 1.3.2. Let $\mathfrak{g}$ be a Lie algebra.

1. A subspace $\mathfrak{h} \subset \mathfrak{g}$ is called a Lie subalgebra if $[X, Y] \in \mathfrak{h}$ for all $X, Y \in \mathfrak{h}$.

2. A linear map $\phi : \mathfrak{g} \to \mathfrak{s}$ from $\mathfrak{g}$ to a Lie algebra $\mathfrak{s}$ is called a Lie algebra map if $\phi([X, Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$.

3. A representation of $\mathfrak{g}$ on a vector space V is a Lie algebra map $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$, where the Lie bracket on $\mathfrak{gl}(V)$ is defined as the commutator of maps: $[f, g] = f \circ g - g \circ f$ for all $f, g \in \mathfrak{gl}(V)$. A subrepresentation ρ' of a representation ρ on V is the restriction of ρ to an invariant subspace of V. This means that ρ' is a map $\rho' : \mathfrak{g} \to \mathfrak{gl}(W)$, $W \subset V$, $\rho'(X)(Y) = \rho(X)(Y)$ for all $X \in \mathfrak{g}$ and for all $Y \in W$, and $\rho(X)(W) \subset W$ for all $X \in \mathfrak{g}$. A representation on V is called irreducible if it only contains the zero space and $W = V$ as subrepresentations.

[2] A proof of the lemma for general groups is given in lemma 1.7 on page 7 of [4]

4. Let ρ and ρ' be two representations of $\mathfrak{g}$ on respectively the vector spaces V and W. We then define the direct sum, tensor product and dual representations as follows:
   (a) $\rho \oplus \rho'(X)(Y+Z) = \rho(X)(Y) + \rho'(X)(Z)$
   (b) $\rho \otimes \rho'(X)(Y \otimes Z) = \rho(X)(Y) \otimes Z + Y \otimes \rho'(X)(Z)$
   (c) $\rho^*(X) = \rho(-X)^T = -\rho(X)^T : V^* \to V^*$

1.3.2 General construction of the associated Lie algebra to a Lie group

In order to show that a Lie group G admits the structure of a Lie algebra, we first define what the tangent space at a point $p \in G$ is. It's possible to define this tangent space in number of different ways, for example using the notion of a derivation. A derivation is a linear map f which satisfies the product rule $f(vw) = vf(w) + f(v)w$. A vector space which consists of derivations can easily be made into a Lie algebra by defining the Lie bracket to be the commutator of maps: $[f, g] = f \circ g - g \circ f$. One can now define T_pG by saying that T_pG consists of derivations $v : \mathcal{E}_p \to \mathbb{R}$ which satisfy $v(fg) = v(f)g(p) + f(p)v(g)$. $\mathcal{E}_p$ is called the set of germs at p. The germs at p are the equivalence classes of differentiable functions which agree on some open neighbourhood of p. This definition of the tangent space is often called the algebraic tangent space of G at p. But in the most straightforward way the tangent space T_pG is defined as the space of tangent vectors at the point p. More rigourously we say that T_pG is the space of equivalence classes of curves $\gamma : (-\epsilon, \epsilon) \to G$, with $\gamma(0) = p$ and $\epsilon > 0$ sufficiently small, under the equivalence relation $\gamma_1 \sim \gamma_2 \iff d/dt(h \circ \gamma_1)(0) = d/dt(h \circ \gamma_2)(0)$ for some chart[3] (U, h, U') around p. One can show that this definition is independent of the chosen chart. T_pG has the structure of a vector space and it has the same dimension as G. By using this definition one can also consider the differential of the map $\phi : G \to H$ between the Lie groups G and H. The differential $(d\phi)_p$ of ϕ at the point p is defined as a map $(d\phi)_p : T_pG \to T_{\phi(p)}H$ and is given by $(d\phi)_p(\frac{d\gamma}{dt}(0)) = \frac{d}{dt}(\phi \circ \gamma)(0)$ for a curve γ through p.

To discover more about the structure of a Lie group G, we can use a well known device in abstract algebra, namely to consider the action of G on some set X. If we take X to be a vector space, this group action is just a representation of G on X. But we have a very natural choice for this vector space X, namely one of the tangent spaces of G. Since all these tangent spaces are isomorphic to each other[4], it doesn't really matter which one we choose. To find such a representation of G on X it is useful to consider the action of a Lie group on itself which respects any Lie group map. By this we mean, that if G and H are Lie groups, $\rho : G \to H$ is a Lie group map and $\Phi_K : K \to \mathrm{Aut}(K) : k \mapsto \phi_k$ is an action for some arbitrary Lie group K

[3] A chart around a point $p \in G$ is a triple (U, h, U') with $U \subset G$ open and $p \in U$, $U' \subset \mathbb{R}^n$ and h a homeomorphism $U \to U'$.

[4] The isomorphism between T_vG and T_wG for any $v \in G$, $w = gv$ for some $g \in G$, is constructed via the differential of the map $m_g : G \to G : h \mapsto gh$ at v which is a map $T_vG \to T_{gv}G = T_wG$. This differential is an isomorphism since m_g is a diffeomorphism.

on itself, we have the following commutative diagram:

$$\begin{array}{ccc} G & \xrightarrow{\rho} & H \\ \phi_g \big\downarrow & & \big\downarrow \phi_{\rho(g)} \\ G & \xrightarrow[\rho]{} & H \end{array}$$

A natural candidate for this is to consider the action Ψ of G on itself by conjugation, and then to look at the differential at the identity e, because e is fixed ($\Psi_g(e) = e$). This gives us the following maps:

$$\begin{aligned} \Psi : \quad G \quad &\to \mathrm{Aut}(G) : g \mapsto \Psi_g(h) = ghg^{-1} \\ \mathrm{Ad} : \quad G \quad &\to \mathrm{Aut}(T_eG) : g \mapsto \mathrm{Ad}(g) = (d\Psi_g)_e : T_eG \to T_eG \end{aligned}$$

We will call this map Ad *the adjoint representation* of G. This map has the required nice property that every homomorphism $\rho : G \to H$ between the Lie groups G and H respects the adjoint representation of G. This is true because every homomorphism respects conjugation. We can summarize these results in a nice commutative diagram.

$$\begin{array}{ccc} T_eG & \xrightarrow{(d\rho)_e} & T_eH \\ \mathrm{Ad}(g) \big\downarrow & & \big\downarrow \mathrm{Ad}(\rho(g)) \\ T_eG & \xrightarrow[(d\rho)_e]{} & T_eH \end{array}$$

If we now take the differential at e of the adjoint representation we arrive at a map $\mathrm{ad} = (d\mathrm{Ad})_e : T_eG \to \mathrm{End}(T_eG)$. We can shuffle things a little bit and write it as $\mathrm{ad} : T_eG \times T_eG \to T_eG$. This map is a bilinear map on T_eG. Using the commutative diagram above we can construct a new commutative diagram for ad.

$$\begin{array}{ccc} T_eG & \xrightarrow{(d\rho)_e} & T_eH \\ \mathrm{ad}(X) \big\downarrow & & \big\downarrow \mathrm{ad}(d\rho_e(X)) \\ T_eG & \xrightarrow[(d\rho)_e]{} & T_eH \end{array}$$

We can use the map ad to define a Lie algebra structure on T_eG by setting $[X, Y] \equiv \mathrm{ad}(X)(Y)$. Using this definition and the commutative diagram above we find that for any homomorphism $\rho : G \to H$ between Lie groups we have the relation

$$d\rho_e([X, Y]) = [d\rho_e(X), d\rho_e(Y)] \tag{1.1}$$

so $d\rho_e$ is a map of Lie algebras. This whole construction is quite abstract, and we haven't even proved yet that the map ad actually satisfies the properties for a Lie algebra. We will not consider the general proof of this here, but instead only focus ourselves at matrix Lie groups and Lie algebras, where the proof is a lot easier to carry out.

1.4 Matrix Lie groups and Lie algebras

Like we have seen in first section, the most general matrix Lie group is the group $G = GL(n, K)$, and all other matrix Lie groups are subgroups of this group. To study the properties, and most importantly the representations of this group we saw that we could look at the Lie algebra associated to G. To find out what this Lie algebra looks like, we would like to have a nice formula for the Lie bracket. We can construct this quite easily by noting that in the case of $G = GL(n, K)$, we can extend the map $\Psi : G \to \mathrm{Aut}(G)$ to $\Psi' : G \to \mathrm{End}(K^n)$. Because the tangent space of $\mathrm{End}(K^n)$ at the unit matrix is just $\mathrm{End}(K^n)$ itself, this means that we now have a simple expression for the adjoint representation: $\mathrm{Ad}(g)(X) = gXg^{-1}$. We can now find ad by looking at the differential of Ad at the unit matrix e. To do this we take two vectors $X, Y \in T_eG$ and consider a curve $\gamma : I \to G$ trough the identity element $\gamma(0) = e$ and with tangent vector $\dot{\gamma}(0) = X$. Then by definition of the Lie bracket we have:

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

So the Lie bracket on T_eG is just the commutator of matrices. This is nice, because it coincides with the most natural choice of a Lie bracket on a vector space of matrices.

At a first glance, it seems that the move from looking at Lie groups to looking at Lie algebras is a bad move, because Lie algebras are purely algebraic, where instead Lie groups also have a topological and differential structure. The good news is however, that this is not true, because we can link the Lie algebra to the Lie group in a special way, namely via the exponential map. With the help of this map we can prove that we can get all the representations of a Lie group by looking at the representations of its Lie algebra and then using the exponential map to lift them to the Lie group.

To motivate the form and the name of the exponential map, we will first look at the form of the elements of a matrix Lie group. Let us for example take the Lie group $SO(2)$, the group of rotations of the plane. Abstractly this group is defined by the real invertible matrices X which satisfy $X^TX = I$ and $\det(X) = 1$. But we can also give a realization of these matrices X in terms of a parameter θ. So let's consider a passive rotation of the plane, that means a rotation of the x and y axis about an angle θ. In this new coordinate system the coordinates x' and y' are related to the old x and y by $x' = \cos\theta x + \sin\theta y$ and $y' = -\sin\theta x + \cos\theta y$. This means that we can realize every element R of SO(2) as:

$$R = R(\theta) = \begin{pmatrix} \cos\theta & +\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

The inverse matrix $R^{-1}(\theta)$ is just a rotation about an angle $-\theta$, so $R^{-1}(\theta) = R(-\theta)$. We can also realize $SO(2)$ as the unit circle [5] S^1 by letting $R(\theta) = e^{2\pi i\theta}$. We can conclude from this

[5] This basically gives an isomorphism between $U(1)$ and $SO(2)$.

that the group $SO(2)$ is described by a single parameter θ. In general the group $SO(n)$ can be described by $n(n-1)/2$ parameters. It is also possible to describe all the other matrix Lie groups in terms of a set of n continuous parameters $\alpha_1, \dots, \alpha_n$. So we can write every element $g \in G$ as $g = g(\alpha_1, \dots, \alpha_n)$. We therefore say that Lie groups are continuous groups, that is groups depending on a set of continuous parameters. To find a realization of the elements of such a continuous group we consider an infinitesimal transformation in the neighbourhood of the identity element, by using the Taylor expansion up to first order since we can neglect all higher order terms:

$$g(\delta\alpha_1, \dots, \delta\alpha_n) = g(0) + \frac{\partial g}{\partial \alpha_\mu}(0)\delta\alpha_\mu = I + \frac{\partial g}{\partial \alpha_\mu}(0)\delta\alpha_\mu$$

Note that we have used the Einstein summation convention. This convention says that when a index appears twice (once as lower index, and once as upper index) in a term, you should sum over all its values. In this case we sum over all the values of μ. Let us now denote $L^\mu = \frac{\partial g}{\partial \alpha_\mu}(0)$. We call these L^μ the infinitesimal generators of the group. They are also elements of T_eG. We can now obtain the form of the elements of G near e by applying an infinite amount of infinitesimal transformations:

$$g(\alpha_1, \dots, \alpha_n) = \lim_{k \to \infty} g\left(\frac{\alpha_1}{k}, \dots, \frac{\alpha_n}{k}\right)^k = \lim_{k \to \infty}\left(I + \frac{L^\mu \alpha_\mu}{k}\right)^k = \exp(L^\mu \alpha_\mu)$$

This means that all the elements of G inside an open neighbourhood of e can be written as a matrix exponential. Because G is a Lie group, any open neighbourhood of e is a generator for the connected component of e[6]. So we can write every element in the connected component of e as the exponential of a matrix. We can now define the exponential map as the map which sends an element X of T_eG to $\exp(X)$.

$$\exp : T_eG \to G : X \mapsto \exp(X) \tag{1.2}$$

So we can use the exponential map to relate the Lie algebra to its Lie group. If we now choose a basis $\{E_i\}_{i=1}^n$ for our Lie algebra we can characterize it by the so called *structure constants* of the Lie algebra. These constants C_{ij}^k define the commutation relations between the basis elements: $[E_i, E_j] = C_{ij}^k E_k$. They are of course antisymmetric in the lower indices because the commutator is antisymmetric. Using these basis elements, we can now construct the so-called one-parameter subgroups of G. A one-parameter subgroup of G is a homomorphism $\phi : \mathbb{R} \to G$. So effectively it is a parameterization of the elements in some subgroup of G. If we now consider the lines λE_i in T_eG, we obtain a one-parameter subgroup of G by using the exponential map. This one-parameter subgroup is the homomorphism $\phi(\lambda) = \exp(\lambda E_i)$. By the Campbell-Baker-Hausdorff formula, the product of these n one-parameter subgroups is again an exponential $\exp(C)$, where C depends on the λ_i, the infinitesimal generators E_i and the repeated commutators of the E_i. But we can write all these commutators in terms of the original basis elements E_i by using the structure constants C_{ij}^k. We can conclude from this that the Lie algebra 'contains' the algebraic structure of its Lie group. Now for completeness, we will prove the Campbell-Baker-Hausdorff formula since it is of such an importance to our discussion above.

[6] For a proof, see [1], theorem 4.11, page 75

Lemma 1.4.1. *Let A and B be two matrices in a Lie algebra $\mathfrak{g}$. Consider the function $\rho_\theta(B) = e^{\theta A}Be^{-\theta A}$. Then $\rho_\theta(B)$ can be written as $\rho_\theta(B) = \exp(\mathrm{ad}(A)\theta)(B) = I + \theta[A,B] + \frac{\theta^2}{2}[A,[A,B]] + O(\theta^3)$. So $\rho_\theta(B) \in \mathfrak{g}$.*

Proof. To prove this result we will use the 'commutator-derivative' trick. This means we will derive a differential equation for $\rho_\theta(B)$, for which the solution has the required form.

$$\begin{aligned}\frac{d}{d\theta}\rho_\theta(B) &= Ae^{\theta A}Be^{-\theta A} - e^{\theta A}BAe^{-\theta A}\\ &= Ae^{\theta A}Be^{-\theta A} - e^{\theta A}Be^{-\theta A}A\\ &= [A, \rho_\theta(B)] = \mathrm{ad}(A)(\rho_\theta(B))\end{aligned}$$

The solution to this differential equation is (if we assume that $\rho_0(B) = I$) just

$$\rho_\theta(B) = e^{\theta \mathrm{ad}(A)}(B) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!}(\mathrm{ad}(A))^n(B) = I + \theta[A,B] + \frac{\theta^2}{2}[A,[A,B]] + O(\theta^3)$$

□

Lemma 1.4.2. *Let $A(t)$ be a matrix valued function of t. Then it holds that $e^{A(t)}\frac{d}{dt}e^{-A(t)} = -f(\mathrm{ad}(A(t)))\dot{A}(t)$ where $f(x) = (e^x - 1)/x$.*

Proof. We will essentially use the same trick as in the previous lemma. Let us denote $B(s,t) = e^{sA(t)}\frac{d}{dt}e^{-sA(t)}$. Then we have that:

$$\begin{aligned}\frac{\partial B}{\partial s} &= A(t)e^{sA(t)}\frac{d}{dt}e^{-sA(t)} - e^{sA(t)}\frac{d}{dt}\left(e^{-sA(t)}A(t)\right)\\ &= A(t)e^{sA(t)}\frac{d}{dt}e^{-sA(t)} - e^{sA(t)}e^{-sA(t)}\dot{A}(t) - e^{sA(t)}\frac{d}{dt}\left(e^{-sA(t)}\right)A(t)\\ &= [A,B] - \dot{A}\end{aligned}$$

The solution to this differential equation is $B(s,t) = e^{s\mathrm{ad}(A)}(B(0,t)) - f(s, \mathrm{ad}(A))(\dot{A})$ where $f(s,X) = (e^{sX} - 1)/X$. Since $B(0,t) = 0$ and setting $s = 1$ we get that

$$e^A\frac{d}{dt}e^{-A} = -f(\mathrm{ad}(A))\dot{A}$$

□

Theorem 1.4.1. *(Campbell-Baker-Hausdorff) Let $\mathfrak{g}$ be a matrix Lie algebra, and $A, B \in \mathfrak{g}$ inside a sufficiently small open neighbourhoud of the origin. Then the matrix $C = \ln(e^A e^B)$ is uniquely defined and $C \in \mathfrak{g}$. C is expressed only in terms of A, B and the repeated commutators of $[A,B]$*

Proof. Let us denote $C(t) = \ln(e^{tA}e^B)$. Then we have that:

$$e^{C(t)}\frac{d}{dt}e^{-C(t)} = e^{tA}e^B\frac{d}{dt}\left(e^{-B}e^{-tA}\right) = -A$$

So by lemma 1.4.2 we have that $A = f(\text{ad}(C(t)))\dot{C}(t)$. We now want to solve this equation for $\dot{C}(t)$ to obtain a differential equation for $C(t)$. To do this, we must compute $f(\text{ad}(C(t)))$. We can do this by noting that by lemma 1.4.1 we have that $e^{\text{ad}C}(h) = e^{C}he^{-C} = e^{tA}e^{B}he^{-B}e^{-tA} = e^{t\text{ad}(A)}e^{\text{ad}(B)}(h)$, so for $\|\text{ad}(A)\| < \frac{\ln 2}{2t}$ and $\|\text{ad}(B)\| < \frac{\ln 2}{2}$[7] we have that

$$\text{ad}(C(t)) = \ln\left(e^{t\text{ad}(A)}e^{\text{ad}(B)}\right)$$

Now consider the matrix valued function $g(x) = \ln(x)/(x-1)$. This gives us that $f(\ln(x)) = g(x)^{-1}$, so $f(\text{ad}(C(t))) = f(\ln(e^{t\text{ad}(A)}e^{\text{ad}(B)})) = g(e^{t\text{ad}(A)}e^{\text{ad}(B)})^{-1}$. Using this in the above equation for A we have that

$$\dot{C}(t) = g(e^{t\text{ad}(A)}e^{\text{ad}(B)})A$$

If we integrate this differential equation for $t = 0$ to $t = 1$ we obtain a integral form for C:

$$C = B + \int_0^1 g(e^{t\text{ad}(A)}e^{\text{ad}(B)})Adt$$

To compute this integral we expand the function g into a power series. The power series for g is given by

$$g(x) = \frac{\ln(x)}{x-1} = \sum_{n=1}^{\infty}\frac{(-1)^n}{n}(x-1)^{n-1}$$

If we now compute the integral expression for C, by also using the power series expansion for $e^{t\text{ad}(A)}e^{\text{ad}(B)}$ we find that

$$C = B + E(\text{ad}(A), \text{ad}(B))(A)$$

where $E(\text{ad}(A), \text{ad}(B))$ is a convergent power series in $\text{ad}(A)$ and $\text{ad}(B)$. From this we can conclude that C is totally determined by the matrices A, B and their commutators. □

1.5 Representations of $U(1)$

Let us now consider the construction of the irreducible representations of the Lie group $U(1)$. As we have seen in section 4 of this chapter, it holds that $U(1)$ is isomorphic to $SO(2)$ and the elements of $U(1)$ can be parameterized by an angle θ. So $U(1)$ consists of the elements $e^{i\theta}$ where θ is bounded between 0 and 2π. It's Lie algebra is defined as the tangent space to the identity element. This means that it consists of the elements $\frac{d}{d\theta}e^{i\theta}|_{\theta=0} = i$. So it is a 1 dimensional Lie algebra isomorphic to $i\mathbb{R}$ and we will denote it as $\mathfrak{u}(1) = \{\alpha i | \alpha \in \mathbb{R}\}$. It is obvious that all the elements of this matrix Lie algebra commute with each other and that they are diagonal.

Let us now consider a representation ρ of $\mathfrak{u}(1)$ on some vector space V. Because $\mathfrak{u}(1)$ is commutative and consists of diagonal matrices, we have that $\rho(X)$ is diagonalizable and commutative for all $X \in \mathfrak{u}(1)$ because Jordan decomposition is preserved for all representations of $\mathfrak{u}(1)$. So from this we see that all the irreducible representations of $\mathfrak{u}(1)$ are the representations ρ on an one dimensional vector space V, and are given by $\rho(i\alpha) = pi\alpha$ for some real number p.

[7] The norm we use here is the L_2 norm

If we exponentiate this result we find the irreducible representations of $U(1)$ which are given by $\rho(e^{i\alpha}) = e^{pi\alpha}$. But from this formula we can now deduce what the possible values of p are:

$$\begin{aligned} e^{ip\alpha} &= \rho(e^{i\alpha}) = \rho(e^{i(\alpha+2\pi)}) = e^{ip(\alpha+2\pi)} \\ 1 &= e^{2\pi ip} \end{aligned}$$

From this equation we can see that p must be an integer number. So all the irreducible representations of $U(1)$ are classified by an integer number p.

Chapter 2

Isospin

2.1 Symmetries and multiplets

In the theory of quantum mechanics, states are described by the Schrödinger equation

$$\hat{H} |\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle \tag{2.1}$$

$\hat{H}$ is the Hamiltonian operator of the system which is a measure for the total energy of a system and $|\Psi\rangle$ is the wave function which describes the state of the system. These wave functions live inside a complex Hilbert space with a hermitian scalar product and they are normalized to unity. We interpret $\langle \Psi(\vec{r}, t) \,|\, \Psi(\vec{r}, t) \rangle$ as the probability of finding the particle at position $\vec{r}$ and at time t. One would of course also like the quantum versions of the important observable quantities which appear in classical mechanics. This is done by the process which goes under the name of *canonical quantization*. In the hamiltonian formulation of classical mechanics, we can describe every state by specifying its position and its momentum and all the important physical observables are functions of these momenta and positions. The idea of canonical quantization is to replace these momenta p_i and positions x_i by the corresponding hermitian operators $\hat{x}_i = x_i$ and $\hat{p}_i = -i\hbar \frac{d}{dx_i}$. These operators also satisfy the commutation relations $[\hat{x}_i, \hat{p}_j] = ih\delta_{ij}$. So the set of these operators form a Lie algebra called the Heisenberg algebra. These commutation relations are of great importance because it results in the Heisenberg uncertainty principle, which says that one cannot know the precise position and momentum of a state at any given time.

To study the behaviour of the solutions of the Schrödinger equation or other equations describing some part of nature, it's very useful to look at the symmetries of the physical system. For example, in the case of classical mechanics which is described by Newton's equations, it is assumed that space is homogenous and isotropic and that time is also homogenous. This means that space is invariant under translations and rotations, and that time is invariant under translations. These symmetries of space and time together with the transformations $\vec{x} \mapsto \vec{x}' = \vec{x} - \vec{v}t$ and $t \mapsto t' = t$, which describe the relative motion between coordinate systems, form the Galilei group. In other physical theories this group is no longer a good symmetry group of a physical system. For example, in the theory of special relativity the Galilei group is replaced by the Poincaré group, where the transformations between relatively moving coordinate systems are given by the Lorentz transformations. In 1915, Emmy Noether discovered a famous theorem, stating that every continuous symmetry of the Lagrangian which describes a physical system, corresponds to a conserved quantity. We will explain shortly what this Lagrangian actually is. To give some examples, this theorem implies that the rotational invariance corresponds to conservation of angular momentum, and the translational invariance of space corresponds to conservation of momentum. The translational invariance of time corresponds to the conservation of energy.

As we described in the first paragraph of this section, the Schrödinger equation is described in terms of a wave function and a Hamiltonian. This Hamiltonian is a measure for the total energy of the system, in the sense that the expectation value of $\hat{H}$ represents the energy. But there also exists an equivalent description in terms of the so called action. This action S is a functional defined as follows:

$$S[\phi] = \int_{t_1}^{t_2} dtL(\phi, \partial_\mu \phi, t) = \int_{t_1}^{t_2} d^n x dt \mathcal{L}(\phi, \partial_\mu \phi, t) \tag{2.2}$$

where we call L the Lagrangian of the physical system, $\mathcal{L}$ the Lagrangian density, and $\partial_0 = \partial_t$. Most modern theories work with this Lagrangian density since it appears to be far more useful in relativistic theories than the standard Lagrangian. In the theory of classical mechanics the Lagrangian is specified as the difference between the kinetic energy and the potential energy. We can derive from this action the equations of motion, by invoking the principle of least action, which states that the trajectory a particle takes is an extremum of the action. In more mathematical terms this means that the variation of the action should be zero, e.g. $\delta S = 0$. This condition leads to the Euler-Lagrange equations whose solutions describe the motion of the particle:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0 \tag{2.3}$$

One can also perform the same construction in quantum mechanics, which goes under the name of the path integral formalism, which was introduced by Richard Feynman. Let us make one more remark to clarify the use of the name symmetry. By a symmetry of a Lagrangian, we mean a infinitesimal transformation which leaves the form of the lagrangian the same. So under such an infinitesimal transformation the equations of motion stay invariant. Consider for example the Lagrangian density $\mathcal{L} = (\partial_\mu \phi)^\dagger \partial^\mu \phi$ with ϕ a complex field. This density is clearly invariant under the transformations $\phi \to \phi' = e^{i\theta}\phi$ with $0 \leq \theta < 2\pi$, which means that the Lie group $U(1)$ is a symmetry group for this particular density.

Theorem 2.1.1. *(Noether) Every continuous symmetry of a lagrangian $L(\phi, \partial_\mu \phi, t)$ corresponds to a conserved quantity*

To see how these symmetry groups G act on the Hilbert space of states $\mathcal{H}$ for a particular physical system, we consider the action of G on $\mathcal{H}$ by matrix multiplication. In other words, we consider the representations of G on $\mathcal{H}$. But let us first remark that not all representations of

G are possible: We need to have that the probability interpretation is preserved, which means that the representations must preserve the scalar product on $\mathcal{H}$. So the representations of G on $\mathcal{H}$ must be unitary! In this thesis we will only look at the symmetry group $SU(n)$. It is easy to see that all the representations ρ of this symmetry group are unitary since $\rho(g)\rho(g)^\dagger = \rho(g)\rho(g)^{-1} = \rho(gg^{-1}) = \rho(e) = I$. Another special property of this symmetry group is that it is a semi-simple group (even a simple group!). These semi-simple groups have the property that if $W \subset V$ is an invariant subspace under a representation ρ on V, then there exists a complementary invariant subspace $W' \subset V$ such that $V = W \oplus W'$. From this we can conclude that every representation of a semi-simple Lie group can be decomposed as a direct sum of irreducible representations. A common name for these irreducible representations is the name *multiplet*.

Proposition 2.1.1. *Let G be a symmetry group of a quantum mechanical system described by states $|\psi\rangle \in \mathcal{H}$. Then all the states inside a multiplet of G have the same energy.*

Proof. Let $g \in G$ and let $\hat{g}$ be the corresponding matrix representation of g in a multiplet. Consider a state $|\psi\rangle$ inside this multiplet of G with energy E. Because G is a symmetry group of the physical system, it must hold that the transformation

$$|\psi\rangle \to |\psi'\rangle = \hat{g}\,|\psi\rangle$$

leaves the Schrödinger equation invariant, so that $|\psi'\rangle$ is also a solution of the same Schrödinger equation. Because $|\psi\rangle$ was inside a multiplet of G, $|\psi'\rangle$ must also be inside the same multiplet. By comparing the S.E. for $|\psi'\rangle$ and the S.E. for $|\psi\rangle$ multiplied with $\hat{g}$ we get

$$\hat{H}\,|\psi'\rangle = i\hbar\frac{\partial\,|\psi'\rangle}{\partial t} = i\hbar\frac{\partial}{\partial t}\hat{g}\,|\psi\rangle = \hat{g}\hat{H}\,|\psi\rangle = \hat{g}\hat{H}\hat{g}^{-1}\hat{g}\,|\psi\rangle = \hat{g}\hat{H}\hat{g}^{-1}\,|\psi'\rangle$$

so $\hat{g}\hat{H}\hat{g}^{-1} = \hat{H}$, which means that $[H, \hat{g}] = 0$. The energy of the state $|\psi'\rangle$ is now calculated as follows:

$$\hat{H}\,|\psi'\rangle = \hat{H}\hat{g}\,|\psi\rangle = \hat{g}\hat{H}\,|\psi\rangle = \hat{g}E\,|\psi\rangle = E\hat{g}\,|\psi\rangle = E\,|\psi'\rangle$$

So the state $|\psi'\rangle$ also has energy E. $\square$

So from this theorem we can conclude that if G is a symmetry group of a physical system, all the elements in an irreducible representation of G commute with the Hamiltonian, and all the states inside this multiplet have the same energy. This also means that the representations of the infinitesimal generators all commute with the Hamiltonian. Therefore it seems worthwhile to look at certain operators which commute with the representations of all the infinitesimal generators. To study this somewhat more precisely we introduce the notion of an *universal enveloping algebra* for a particular representation ρ of the Lie algebra $\mathfrak{g}$:

Definition 2.1.1. The enveloping algebra for a faithful representation ρ of a Lie algebra $\mathfrak{g}$ consists of all the products and sums of the $\rho(L^\mu)$, where the L^μ are a basis of $\mathfrak{g}$.

The center of this enveloping algebra consists by definition of the matrices which commute with the representations of the infinitesimal generators L^μ. We also call these operators

Casimir operators. The number of these operators depends on the rank of the Lie algebra. The rank of a Lie algebra is defined as the maximum number of commuting basis elements. So if a Lie algebra has full rank, all the basis elements are Casimir operators. One can also show that for the Lie algebras $\mathfrak{su}(n)$ the Casimir operators are homogenous polynomials in the basis elements. For example, the Casimir operator for a faithful representation ρ of the Lie algebra $\mathfrak{su}(2)$ with basis $\{L_\mu\}_{\mu=1}^3$ is given by $C_1 = g_{\mu\nu}\rho(L_\mu)\rho(L_\nu)$, where the metric tensor $g_{\mu\nu}$ for a representation ρ is defined as $g_{\mu\nu} = \mathrm{Tr}\rho(L_\mu)\rho(L_\nu)$, and where we sum over the indices μ and ν.

Consider a faithful representation ρ of a Lie algebra $\mathfrak{g}$ of rank l on a vector space V. Let C be a Casimir operator for this Lie algebra. Schur's lemma implies that for an irreducible subrepresentation of ρ on $V' \subset V$, $C = \lambda \cdot \mathrm{id}_V$, since C commutes with all the elements of $\rho(\mathfrak{g})$. This means that V' lies in an eigenspace of C. By a theorem of Chevalley, it holds that it is possible to find exactly l indepedent Casimir operators, such that an irreducible subrepresentation of ρ corresponds to a unique common eigenspace of these Casimir operators. All the irreducible subrepresentations of ρ can therefore be classified by the eigenvalues of these l Casimir operators.

2.2 Representations of $SU(2)$

The group $SU(2)$ is a very important symmetry group in the theory of quantum mechanics. It occurs in the theory of angular momentum, as being the 'quantum version' of the rotational symmetry group $SO(3)$ which occurs as the symmetry group in classical mechanics that gives rise to conservation of angular momentum. The reason we use $SU(2)$ in quantum mechanics comes from the fact that $SU(2)$ is the double cover of $SO(3)$, which is essential in the theory of spin, where we encounter half-valued representations. This means that all the representations can be classified by a number n which can take the values $0, \frac{1}{2}, 1, \ldots$. The half-valued representations cannot occur if we just used $SO(3)$ as the symmetry group, because this symmetry group only gives integer valued representations (so these representations are classified by an integer n).
The reason why we explore the representations of $SU(2)$ in this thesis is that it also occurs as a intrinsic symmetry group in elementary particle physics, namely as *isospin*. This symmetry arose originally from the fact that the proton and the neutron have approximately the same mass. Therefore Heisenberg proposed that they should be regarded as different states of the same particle, called the nucleon. In a later stage, the isospin group was regarded as the flavor symmetry group of the up and down quarks. Since the masses of these two quarks are not exactly equal to each other this symmetry is only an approximate symmetry.

2.2.1 The Lie group $SU(2)$ and its Lie algebra

As we saw in section 1.1, the group $SU(2)$ is given by the set of 2x2 matrices over $\mathbb{C}$ and which preserve the standard hermitian scalar product and have determinant equal to 1. The defining relation for these matrices is that for all $A \in SU(2)$ it must hold that $A^\dagger A = I$. The Lie algebra of $SU(2)$ is given by the tangent space at the identity. By differentiating the defining identity of $SU(2)$ at $A = I$, we find that for every $X \in \mathfrak{su}(2)$ it must hold that $X^\dagger = -X$, so

$\mathfrak{su}(2)$ consists of anti-hermitian matrices. We haven't completely specified $\mathfrak{su}(2)$ yet because we haven't used the fact that the determinant of $A \in SU(2)$ should be equal to 1. Let the set $\{E_1, E_2, E_3\}$ be an arbitrary basis for $\mathfrak{su}(2)$. Since $SU(2)$ is connected and compact, we can exponentiate $\mathfrak{su}(2)$ to obtain $SU(2)$. This gives that every element $g \in SU(2)$ can be written as $g(\alpha) = \exp(\alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3)$. Because g must have determinant 1, we now find a condition on the trace of the generators E_i.

$$1 = \det e^{(\alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3)} = e^{\mathrm{Tr}(\alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3)} \Longrightarrow \mathrm{Tr} E_i = 0$$

So we can conclude that $\mathfrak{su}(2)$ consists of the 2x2 anti-hermitian matrices with trace zero. A very useful basis for this Lie algebra is given by means of the Pauli matrices. The Pauli matrices σ_i are given by:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.4}$$

The Pauli matrices are a basis for the hermitian trace zero matrices. In order to find a basis for the space of anti-hermitian trace zero matrices we just multiply the σ_j by i. To find the representations of this Lie algebra, which we will do in the next paragraph, it is helpful to consider the complexification of the Lie algebra $\mathfrak{su}_2$. This complexification is the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. It consists of the linear combinations of the matrices in $\mathfrak{su}_2$ with complex coefficients. It thus holds that the Pauli matrices are a basis of this Lie algebra. But we can now define another very useful basis, by taking two linear combinations of the first two Pauli matrices: $\sigma_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$. In matrix form, the new basis looks like:

$$\sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.5}$$

Their commutation relations are given by

$$[\sigma_+, \sigma_-] = \sigma_3 \qquad [\sigma_3, \sigma_\pm] = \pm 2\sigma_\pm \tag{2.6}$$

2.2.2 The irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$

By using the basis we derived in the last section for the complexified version of $\mathfrak{su}_2$ we are able to classify the irreducible representations of $\mathfrak{su}_2$. A known fact in the theory of Lie algebras which we do not prove here[1], is the preservation of the Jordan decomposition under a representation. The Jordan decomposition of a matrix $g \in \mathfrak{sl}_2(\mathbb{C})$ is the decomposition of g into $g_d + g_n$ where g_d is diagonalizable, g_n is nilpotent and $[g_d, g_n] = 0$. We can use this property to study the representations of $\mathfrak{sl}_2(\mathbb{C})$, since σ_3 is diagonalizable, so the preservation of Jordan decomposition means that $\rho(\sigma_3)$ is diagonalizable in $\mathrm{Aut}(V)$. Now because $H = \rho(\sigma_3)$ is diagonalizable, we can write V as $V = \bigoplus V_\alpha$, where the V_α are the eigenspaces of H corresponding to the eigenvalues α of H.

[1] See for example appendix C.2 in [4]

How do we go about constructing the irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$? To do this, we will use the method of symmetric tensor powers. This means that we will construct the m-th irreducible representation by taking the m-th symmetric power of the fundamental representation. The fundamental representation is the representation ρ on $V = \mathbb{C}^2$ with $\rho(X) = X$ for all $X \in \mathfrak{sl}_2(\mathbb{C})$. But before we proceed to perform this construction, I will first explain the notion of a symmetric power of a vector space. Consider a n-dimensional (complex) vector space V. Let us denote the basis vectors of V as $\{x_i\}_{i=1}^n$. Then we define the m-th symmetric power of V as the vector space $\mathrm{Sym}^m V$ which has a basis

$$\{e_{i_1} \cdot e_{i_2} \cdot \ldots \cdot e_{i_m} | i_1 \leq \ldots \leq i_m\}$$

with the property that every element $v_1 \cdot \ldots \cdot v_m$ of $\mathrm{Sym}^m V$ is invariant under permutation of the vectors v_i. One can also easily see that the dimension of $\mathrm{Sym}^m V$ is given by:

$$\dim \mathrm{Sym}^m V = \binom{n+m-1}{m} \tag{2.7}$$

To illustrate this, consider $\mathrm{Sym}^2\mathbb{C}^2$. Let $\{x, y\}$ be a basis of $\mathbb{C}^2$. Then the basis of $\mathrm{Sym}^2\mathbb{C}^2$ is given by $\{x^2, xy, y^2\}$. So basically $\mathrm{Sym}^2\mathbb{C}^2$ consists of the symmetric polynomials of degree 2 in the variables x and y. In general this also holds: We can view the m-th symmetric power $\mathrm{Sym}^m V$ of an n-dimensional vector space V as the vector space of symmetric polynomials of degree m in the n basis variables of V. One can also construct the m-th symmetric power of V from the m-th tensor product $V^{\otimes m}$ of V by taking the quotient of $V^{\otimes m}$ with the subspace generated by $v_1 \otimes \ldots \otimes v_n - v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}$ for all $\sigma \in S_n$, the group of permutations of the set $\{1, \ldots, n\}$, and $\{v_1, \ldots, v_n\}$ a basis for V. This basically means that you mod out all the non-commuting tensor products of the basis vectors. We now give the two main theorems of this section, which give us the classification of the irreducible representations and the decomposition of tensor products of them in irreducible representations.

Theorem 2.2.1. *A representation ρ of $\mathfrak{sl}_2(\mathbb{C})$ on a vector space V is irreducible if and only if $V = \mathrm{Sym}^n\mathbb{C}^2$ for some $n \geq 0$, with $\mathbb{C}^2$ the fundamental representation of $\mathfrak{sl}_2(\mathbb{C})$.*

Proof. The proof of this statement is somewhat lengthy. To carry it out we will first consider the action of $H = \rho(\sigma_3)$ on the vector space $\mathrm{Sym}^n\mathbb{C}^2$. This will give us a decomposition of V into a number of subspaces. To proof the irreducibility we will consider the action of $X_\pm = \rho(\sigma_\pm)$ on these subspaces. Let us now denote the standard basis of $\mathbb{C}^2$ as $\{x, y\}$. Then the basis of $V = \mathrm{Sym}^n\mathbb{C}^2$ is given by $\{x^n, x^{n-1}y, \ldots, y^n\}$. Since we consider $\mathbb{C}^2$ as the fundamental representation we have that $H(x) = x$ and $H(y) = -y$. This gives us that

$$\begin{aligned} H(x^{n-k}y^k) &= (n-k)H(x)x^{n-k-1}y^k + kH(y)x^{n-k}y^{k-1} \\ &= (n-k)x^{n-k}y^k - kx^{n-k}y^k \\ &= (n-2k)x^{n-k}y^k \end{aligned}$$

So the eigenvalues of H are the numbers $n, n-2, \ldots, -n$. This means that we can decompose V as

$$V = \bigoplus_{i=0}^{n} V_{n-2i} = V_n \oplus \ldots \oplus V_{-n}$$

where V_j is the eigen space corresponding to the eigenvalue j. Let us now consider a eigenvector v in one of the subspaces V_j with $j \neq \pm n$. We can use the operators $X_\pm$ on v to show that V must be irreducible. By using the commutation relations of $\mathfrak{sl}_2(\mathbb{C})$ we get that

$$HX_\pm(v) = X_\pm H(v) + [H, X_\pm](v) = (j \pm 2)X_\pm(v)$$

so $X_\pm(v)$ is also an eigenvector of H with eigenvalue $j \pm 2$, so $X_\pm(v) \in V_{j\pm2}$. This means basically that $X_\pm$ is a map $V_j \to V_{j\pm2}$, so V has no invariant subspace, since we can get all the vectors by repeatedly applying $X_\pm$ to an eigenvector $v \in V$. We can conclude that V is irreducible. Now we only need to proof that if we consider an arbitrary irreducible representation V, that $V = \mathrm{Sym}^n\mathbb{C}^2$ for some n. So let V be an arbitrary irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$. Let v be an eigenvector of H with eigenvalue α. We know from above that under the application of $X_\pm$ this gives us an eigenspace decomposition of V as $V = \bigoplus_n V_{\alpha+2n}$. Because V is finite dimensional by hypothesis, we must have an upper bound $n_{\max}$ and a lower bound $n_{\min}$ for the eigenvalues. Let v be an nonzero eigenvector in $V_{n_{\max}}$. It thus holds that $X_+(v) = 0$. We now show that the vectors $\{v, X_-(v), X_-^2(v), \ldots, X_-^m(v)\}$, where m is the smallest power of X_- which annihilates v, form a basis for V. Because V is irreducible it suffices to show that the subspace generated by these vectors is an invariant subspace. Because H and X_- obviously preserve this subspace, we only need to show that X_+ preserves it. For $k = 1, \ldots, m$ we have that

$$X_+(X_-^k(v)) = k(n_{\max} - k + 1)X_-^{k-1}(v)$$

so X_+ preserves the subspace too. This means that all the eigenspaces are one dimensional and that the number $n_{\max}$ must be real and equal to $m-1$ since $0 = X_+(X_-^m(v)) = m(n_{\max} - m+1)X_-^{m-1}$. We can conclude that V is $n_{\max}+1$ dimensional and is uniquely determined by the eigenvalues $n_{\max}, n_{\max}-2, \ldots, -n_{\max}$. But since $\mathrm{Sym}^{n_{\max}}\mathbb{C}^2$ satisfies the same properties as V, it must be that $V = \mathrm{Sym}^{n_{\max}}\mathbb{C}^2$. □

Theorem 2.2.2. *(Clebsch-Gordan) Let $a \geq b$ be integers and V be the fundamental representation of $\mathfrak{sl}_2(\mathbb{C})$. Then it holds that* $\mathrm{Sym}^aV \otimes \mathrm{Sym}^bV = \bigoplus_{i=0}^b \mathrm{Sym}^{a+b-2i}V$.

Proof. The idea behind this proof is to consider all the possible eigenvalues of H on $\mathrm{Sym}^aV \otimes \mathrm{Sym}^bV$. We can find all these eigenvalues by considering all the possible combinations of the eigenvalues of Sym^aV and Sym^bV, because the eigenvalues in a tensor product just add. We then proceed to identify all the possible eigenvalues with their multiplicities as vector spaces isomorphic to some symmetric power. In order to count all the multiplicities it is useful to consider a formal Laurent polynomial in one variable for the vector spaces Sym^nV, or products of it, where the powers are the eigenvalues of these vector spaces and the coefficients are the multiplicities. This gives us that the polynomial for the vector space $\mathrm{Sym}^aV \otimes \mathrm{Sym}^bV$ is just

$$\left(\sum_{i=0}^{a} x^{a-2i}\right) \cdot \left(\sum_{j=0}^{b} x^{b-2j}\right) = \sum_{k=0}^{a}\sum_{l=0}^{b} x^{a+b-2k-2l}$$

From this formula you can easily see that the multiplicity of the eigenvalue $a+b-2z$ for $z \leq (a+b)/2$ equals $z+1$. By symmetry this also determines the multiplicities of the negative eigenvalues: The multiplicity of eigenvalue α equals the multiplicity of the eigenvalue $-\alpha$. Now consider an eigenvector v with eigenvalue $a+b$. By the previous theorem, the action of $\mathfrak{sl}_2(\mathbb{C})$ on v generates an invariant subspace W of $\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V$ isomorphic to $\mathrm{Sym}^{a+b} V$. $\mathfrak{sl}_2\mathbb{C}$ is semi-simple [2], which means that there exists a complementary invariant subspace W' such that $W \oplus W' = V$. We can now find an invariant subspace W'' of W' isomorphic to $\mathrm{Sym}^{a+b-2} V$ because all the eigenvalues have dropped by one. By continuing this process until the eigenvalue 1 or 0 is reached, the only invariant subspace is the zero subspace. This exactly gives the required decomposition. □

So we see that all the irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$ (so also all the irreducible representations of $\mathfrak{su}_2$) are uniquely determined by some non-negative integer n and that the tensor products of these irreducible representations decompose as a direct sum of irreducible representations. This is what physicists like to call 'addition of angular momentum'. When physicists speak of the representations of $\mathfrak{su}_2$ they use a somewhat different convention from the one we used above. Their definition of the basis matrices of $\mathfrak{su}_2$ differ with a factor of one-half. So instead of the integers n they got the numbers j, where simply $2j = n$. They also call j the *angular momentum* quantum number.

In order to proceed with the calculations in the next section we first write down the Casimir operator for $\mathfrak{sl}_2\mathbb{C}$ using the basis most physicists use, namely the basis $\{\sigma_+, \sigma_-, \sigma_3/2\}$, where the $\sigma_\pm$ and σ_3 are defined by equation 2.5. Let us now consider an irreducible representation ρ of $\mathfrak{sl}_2\mathbb{C}$ on a vector space $V = \mathrm{Sym}^{2j}\mathbb{C}^2$. Like in the proof of theorem 2.2.1 we denote $H = \frac{1}{2}\rho(\sigma_3)$ and $X_\pm = \rho(\sigma_\pm)$. By using the identity $[AB, C] = A[B, C] + [A, C]B$ for arbitrary linear operators A, B, C, one can easily see that the operator

$$C = X_\pm X_\mp + H^2 \mp H \tag{2.8}$$

commutes with H and the $X_\pm$. In other words, C is a Casimir operator for the representation ρ of $\mathfrak{sl}_2\mathbb{C}$. Let us now consider an eigenvector v in the eigenspace V_l corresponding to the maximum eigenvalue l of H . Then we have that

$$Cv = (X_- X_+ + H^2 + H)v = (0 + l^2 + l)v = l(l+1)v \tag{2.9}$$

so C has the eigenvalue $l(l+1)$ on the whole representation since it commutes with all the basis elements. We can now also immediately write down the eigenvalues for the operators $X_\pm$ by using the fact that $X_\pm = (X_\mp)^\dagger$. So let v_m be an eigenvector in the eigenspace V_m. Then from equation 2.8 it follows that

$$\begin{aligned} X_\mp X_\pm v_m &= (l(l+1) - m(m \pm 1))v_m & (2.10)\\ \langle X_\pm v_m | X_\pm v_m \rangle &= \langle v_m | X_\mp X_\pm v_m \rangle = (l(l+1) - m(m \pm 1)) \langle v_m | v_m \rangle & (2.11)\\ X_\pm v_m &= \sqrt{l(l+1) - m(m \pm 1)}\, v_{m\pm 1} & (2.12) \end{aligned}$$

Note that we chose a real and positive phase here. This phase is also called the *Condon-Shortley phase*.

[2] See [4], paragraph 9.3, page 123,128, and appendix C

2.3 Lightweight elementary particles

As mentioned in the introductory paragraph in the preceding section, the proton and the neutron have approximately the same mass. Despite the fact that they do not have the same charge, the proton and neutron behave almost exactly the same in all other aspects. For example, the strong force does not differentiate between a neutron and a proton, and the strong force is charge independent. This gives a strong indication that there must be a symmetry group for the proton and the neutron which is invariant under the strong interaction. This symmetry group was called *isospin* since its mathematical properties are identical to those of ordinary spin. So this isospin symmetry group is just the group $SU(2)$ which we studied in the previous section. In this section however, we will not consider the isospin group as the symmetry group of the nucleon, but as a flavor symmetry group of the up and down quarks. There also exist other flavors of quarks, but they are not described by $SU(2)$, but by $SU(3)$ and higher. We will look at them in the next chapter.

As we have seen in the previous section, all the irreducible representations of $SU(2)$ can be obtained by taking the symmetric tensor product of the fundamental representation $\mathbb{C}^2$. Using the hypothesis that the up and down quarks are the smallest constituents of the elementary particles, we can therefore say that the up and down quarks must lie in the fundamental representation of $SU(2)$.

$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad d = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Because we know all the possible eigenvalues of the irreducible representations, we can also denote the states in the perhaps more familiar Dirac notation. From now on we will denote the I-th irreducible representation as D^I. We will label the states inside D^I by the quantum number I and by the eigenvalues I_3 of this representation. We will also follow the convention used by physicists to divide the eigenvalues by 2 (this means that we also divide I by two). So a general state will be written as $|I\,I_3\rangle$. In this notation the up quark becomes $|u\rangle = |\frac{1}{2}\,\frac{1}{2}\rangle$, and the down quark becomes $|d\rangle = |\frac{1}{2} - \frac{1}{2}\rangle$.

2.3.1 The pion triplet

By using the building blocks of the up and down quarks we can construct the lightweight elementary particles. Let us first look at the case where we consider a composite system of 2 quarks. This can either be a uu, ud or dd state or linear combinations of them. This means we have to consider the tensor product representation of two copies of the fundamental representation, so we consider the representation $D^{1/2} \otimes D^{1/2}$. By theorem 2.2.2 we know that this tensor product decomposes as $D^1 \oplus D^0$. Let us look some closer to the states in these 2 irreducible representations. Since the eigenvalues for D^1 are 1, 0 and -1, its states are given by $|1\,1\rangle$, $|1\,0\rangle$ and $|1-1\rangle$. The case of D^0 is simple, because it only allows for the state $|0\,0\rangle$. We can also write these 4 states in terms of the up and down quarks. To do this we use the fact that under the application of a tensor product the eigenvalues just add. The only possible way to get the state $|1\,1\rangle$ out of a tensor product is to tensor two states with eigenvalues 1/2. This means that $|1\,1\rangle = |u\rangle \otimes |u\rangle$. By using the same argument we also find that $|1-1\rangle = |d\rangle \otimes |d\rangle$. We can now get the state $|1\,0\rangle$ by either applying X_+ to $|1\,-1\rangle$ or X_- to $|1\,1\rangle$. This gives us that $|1\,0\rangle = X_-\,|1\,1\rangle = X_-(|u\rangle \otimes |u\rangle) = (X_-\,|u\rangle) \otimes |u\rangle + |u\rangle \otimes (X_-\,|u\rangle) = (|d\rangle \otimes |u\rangle + |u\rangle \otimes |d\rangle)/\sqrt{2}$. The only other remaining possibility for an other state with eigenvalue 0 and which is orthogonal[3] to the other three, is the combination $(|u\rangle \otimes |d\rangle - |d\rangle \otimes |u\rangle)/\sqrt{2}$. The factors of $1/\sqrt{2}$ come from normalizing the state. These factors also appear if we look at how to decompose a tensor product of angular momentum states in general. These decompositions of the direct product states go under the name of the *Clebsch-Gordan series* and the coefficients that appear in this series go under the name of the *Clebsch-Gordan coefficients*. We will not derive these coefficients since it is quite a laborious work. Instead we will just use them in the following calculations and the reader can refer to page 188 of Griffiths's quantum mechanics text which lists a table of the most common Clebsch-Gordan coefficients. Now by using these coefficients we find:

$$
\begin{aligned}
|1\,1\rangle &= |u\rangle \otimes |u\rangle \\
|1\,0\rangle &= (|d\rangle \otimes |u\rangle + |u\rangle \otimes |d\rangle)/\sqrt{2} \\
|1\,-1\rangle &= |d\rangle \otimes |d\rangle \\
|0\,0\rangle &= (|u\rangle \otimes |d\rangle - |d\rangle \otimes |u\rangle)/\sqrt{2}
\end{aligned}
$$

But how do we interpret these states as elementary particles? We can do this be looking at the experimental properties of the Pion triplet.

Properties of the Pions

pion	mass (MeV)	charge	lifetime (s)	spin
π^+	139.59	e	$(2.55 \pm 0.03) \cdot 10^{-8}$	0
π^0	135.00	0	$0.83 \cdot 10^{-16}$	0
π^-	139.59	$-e$	$(2.55 \pm 0.03) \cdot 10^{-8}$	0

In this table we see that the masses of the three different pions are almost equal to each other. The mass of the π^+ equals that of the π^- and the mass of the π^0 is just 4.59 MeV smaller. One can argue that this mass difference is caused by other interactions, like the electromagnetic interaction or the weak interaction. We can therefore regard these pions as members of an isospin triplet, since isospin is conserved under the strong interaction and because the strong interaction is charge-independent. By using the results we found above we find that

$$
\begin{aligned}
|1\,1\rangle &\longleftrightarrow |\pi^+\rangle \\
|1\,0\rangle &\longleftrightarrow |\pi^0\rangle \\
|1\,-1\rangle &\longleftrightarrow |\pi^-\rangle
\end{aligned}
$$

We also see that we can find an easy relation between the charge of the pions and the eigenvalues[4]:

$$Q = eI_3 \tag{2.13}$$

[3] Here we consider orthogonality with respect to the induced inner product on the tensor product space, which is defined by: $\langle v \otimes w, v' \otimes w' \rangle = \langle v, v' \rangle \langle w, w' \rangle$, where $\langle , \rangle$ is the standard inner product on $\mathbb{C}^2$.

[4] We will see this equation appear more generally in the next chapter as the Gell-Mann/Nishijima rule

But the construction we used isn't entirely justified. In the experiments we observe that a quark-quark bounded state isn't possible. On the other hand we observe lots of quark-antiquark bounded states. These quark antiquark states are called *mesons*. To explain why for example the quark-quark states cannot appear in nature, the notion of *color* was introduced. Each quark can carry one of the colors red, green or blue. The antiquarks carry the colors antired, antigreen and antiblue. This would suggest that there are numerous different particles which are composed of the same quarks. But there is a conjecture, called the *Color-singlet conjecture* which says that only the particles with neutral color can exist in nature. This conjecture immediately explains why there do not exist any quark-quark bounded states. These states do not have neutral color! Only a quark-antiquark state has neutral color.
In our isospin model however, the fundamental representation for the anti-particles is isomorphic to the standard fundamental representation, which means that in this model the antiquarks correspond to the quarks by $u \leftrightarrow -\bar{d}$ and $d \leftrightarrow \bar{u}$. The minus sign in front of the antidown quark comes from the fact that we want the charge conjugation operator (The operator which sends T_3 to $-T_3$) to commute with the rotations in $SU(2)$. This gives us the true composition of the three pions:

$$|\pi^+\rangle = -|u\rangle \otimes |\bar{d}\rangle \qquad |\pi^0\rangle = (|u\rangle \otimes |\bar{u}\rangle - |d\rangle \otimes |\bar{d}\rangle)/\sqrt{2} \qquad |\pi^-\rangle = |d\rangle \otimes |\bar{u}\rangle$$

2.3.2 The lightweight baryons

We now go one step further and consider the elementary particles which are made up of three quarks/antiquarks. We call these particles *Baryons*. Because the total combination of these 3 quarks must be color neutral, the only possible baryons are those which are made up entirely out of quarks or entirely out of antiquarks. To classify these particles we proceed in the same way as the previous section: we consider the threefold tensor product of the fundamental representation $D^{1/2}$. By theorem 2.2.2 this tensor product decomposes as $(D^{1/2})^{\otimes 3} = D^{3/2} \oplus D^{1/2} \oplus D^{1/2}$. Let us look some closer at the $D^{3/2}$ representation. The possible states in this representation are of course $|\frac{3}{2}\,\frac{3}{2}\rangle$, $|\frac{3}{2}\,\frac{1}{2}\rangle$, $|\frac{3}{2}-\frac{1}{2}\rangle$ and $|\frac{3}{2}-\frac{3}{2}\rangle$. To find out the composite structure of these baryon states we will use the same procedure as in the previous section. The only possible combination for the $|\frac{3}{2}\,\frac{3}{2}\rangle$ state is $|u\rangle \otimes |u\rangle \otimes |u\rangle$ since this is the only state with the maximum eigenvalue. The other states follow from applying the I_- operator to this state.

$$\begin{aligned}
|3/2\;3/2\rangle &= |u\rangle \otimes |u\rangle \otimes |u\rangle \\
|3/2\;1/2\rangle &= (|d\rangle \otimes |u\rangle \otimes |u\rangle + |u\rangle \otimes |d\rangle \otimes |u\rangle + |u\rangle \otimes |u\rangle \otimes |d\rangle)/\sqrt{3} \\
|3/2-1/2\rangle &= (|d\rangle \otimes |d\rangle \otimes |u\rangle + |d\rangle \otimes |u\rangle \otimes |d\rangle + |u\rangle \otimes |d\rangle \otimes |d\rangle)/\sqrt{3} \\
|3/2-3/2\rangle &= |d\rangle \otimes |d\rangle \otimes |d\rangle
\end{aligned}$$

Again we would like to find a type of elementary particles which fit in this isospin quartet. These particles are called the Δ baryons and they have the following properties:

Properties of the Δ baryons

Δ baryon	mass (MeV)	charge	lifetime (s)	spin
Δ^{++}	1232 ± 2	$2e$	$5.49 \cdot 10^{-24}$	$3/2$
Δ^{+}	1232 ± 2	e	$5.49 \cdot 10^{-24}$	$3/2$
Δ^{0}	1232 ± 2	0	$5.49 \cdot 10^{-24}$	$3/2$
Δ^{-}	1232 ± 2	$-e$	$5.49 \cdot 10^{-24}$	$3/2$

As you can see from this table, the lifetimes of these particles are very short. This comes from the fact that these baryons only exist as an intermediary product in some reactions. One can create these particles by scattering a proton off a positive pion. In this reaction the Δ^{++} particle[5] is formed, since it is the only Δ baryon with charge 2e. But this particle immediately decays into a proton and a positive pion again. If one looks at the differential cross-section of these pion-nucleon processes, one sees a sharp peak at the mass of the Δ baryons, so at 1232 MeV. We therefore also call these types of baryons *baryon resonances*. There exists a nice fitting formula for the cross-section of these resonances which is called the Breit-Wigner formula. This formula gives the energy dependence of the cross-section of a reaction between two particles x and y, for energies E close to the resonance energy E_0.

$$\sigma(E) = \frac{\lambda^2(2S+1)}{4\pi(2s_x+1)(2s_y+1)} \frac{\Gamma^2}{(E-E_0)^2+\Gamma^2/4} \tag{2.14}$$

Here λ is the wavelength in the center of mass system, s_x and s_y are the spins of the particles x and y, S is the total spin of the resonance and Γ is the width of the resonance which is related to the lifetime τ by $\tau = \hbar/\Gamma$. So using the lifetime which is given in the table above we obtain a resonance width of about 120 MeV. So for a reaction between a proton and a positive pion the cross-section for energies in MeV becomes

$$\sigma(E) \approx \frac{7200\lambda^2}{\pi((E-1232)^2+3600)}$$

Of course, most of the time people will use this formula the other way around. They use it for example to find the lifetime of a resonance by fitting this formula to the measured data to find the width of the peak. Or they use it to find the wavelength of the particle from which they deduce its momentum.

But let us now consider the other isospin multiplet which appeared in the tensor product of the three lightweight quarks, namely the $D^{1/2}$ representation. In this representation we have the states $|1/2\,1/2\rangle$ and $|1/2\,-1/2\rangle$. We already mentioned a candidate for 2 particles

[5]It's perhaps interesting to mention that the discovery of this particle is one of the reasons why physicists introduced the notion of color. They did it because without color this particle wasn't allowed to exist by the Pauli exclusion principle since it is made up of three up quarks which each have only two spin directions.

that fit in this isospin doublet, namely the proton and the neutron. These 2 particles are one of the most common particles in nature and all the atoms are composed of them. For completeness we give the properties of the proton and the neutron:

Properties of the nucleons

Nucleon	mass (MeV)	charge	lifetime (s)	spin	composition
p	938.27	e	stable	1/2	uud
n	939.57	0	886	1/2	udd

There also exist other types of lightweight baryons who fit inside an isospin doublet, namely the $N(M)$ baryons. These baryons also appear as resonances in nucleon-pion scattering processes and appear at 1515 MeV and at 1688 MeV. Just like in the case of the nucleons, there is one N-baryon state with positive charge of e and one neutral state. We denote the positive charged state as $N^+(M)$ and the neutral state as $N^0(M)$. With the same analysis as before we can conclude that the $N^+(M)$ state is build up from 2 up quarks and one down quark, and the $N^0(M)$ state is build up from 2 down quarks and one up quark. This is also obvious from the fact that the two isospin doublets are isomorphic to each other. The wavefunctions of these resonances only differ in the non-isospin part since they have different energies and different angular momentum.

Chapter 3

The quark model

3.1 Hypercharge and strangeness

In the last section of the previous chapter we looked at the elementary particles which are made up of the up and down quarks. We classified these particles into two groups, the mesons and the baryons. The mesons consist of a quark-antiquark pair, and the baryons consist of combinations of three quarks. In a larger picture, one can view the mesons and baryons as the two parts of the larger class of the *hadrons*. The hadrons are the particles which can interact through the strong interaction. There are also particles which never experience strong interactions. These particles are called *leptons*. To classify these 2 groups somewhat more mathematically, one can define the lepton number L and the baryon number B as follows:

$$L = \begin{cases} 1 & \text{if the particle is a lepton} \\ -1 & \text{if the particle is an antilepton} \\ 0 & \text{otherwise} \end{cases} \qquad B = \begin{cases} 1 & \text{if the particle is a baryon} \\ -1 & \text{if the particle is an antibaryon} \\ 0 & \text{otherwise} \end{cases}$$

It is generally believed that the lepton number and the baryon number are each additively conserved. This means for example that in a reaction between two different baryons, the resulting product must contain at least some other baryon. The same holds for the leptons. Because B and L are conserved, we know from Noether's theorem that there must be a symmetry group G that corresponds to them. To find a possible candidate for this symmetry group we can regard B as an eigenvalue of a hermitian operator $\hat{B}$. We now take $i\hat{B}$ to be the basis element of a Lie algebra $\mathfrak{g}$ isomorphic to $\mathfrak{u}(1)$. If we exponentiate $\mathfrak{g}$ we obtain a Lie group isomorphic to $U(1)$. The irreducible representations of $U(1)$ are classified by a number p which can take on the values $0, \pm 1, \pm 2, \ldots$. This also means that the irreducible representations of $\mathfrak{u}(1)$ and therefore the irreducible representations of $\mathfrak{g}$ are classified by the same number p. So from this we can conclude that the reason that B takes on integer values comes from the fact that it has a $U(1)$ symmetry. The same holds for the lepton number L.

In the section on the pion triplet we mentioned a relation between the charges of the pions and the eigenvalue I_3 (equation 2.13). But if we look at the tables for the lightweight baryons in the previous chapter we see that this relation does not hold for the baryons. We see that the difference between the true charge and the charge predicted by equation 2.13 is always equal to 1/2 for the normal baryons and equal to −1/2 for the antibaryons. This suggests that we use the baryon number B to modify equation 2.13:

$$Q = e(I_3 + \frac{1}{2}B) \tag{3.1}$$

For some more numerology, one can also see that for a given isospin multiplet the baryon number B equals the maximum charge Q_{max} plus the minimum charge Q_{min} in that multiplet divided by the electric charge e.
In later experiments it was discovered that there were certain particles which were produced through the strong interaction but had a lifetime of about 10^{-10} seconds. This is much longer than the expected lifetime of about 10^{-23} seconds. In order to try to quantify this, the quantum number strangeness was introduced. This quantum number S is assumed to be an integer and is conserved by the electromagnetic interaction and by the strong interaction. The reason why these particles had such a long lifetime is thus explained by the fact that strangeness is not conserved in the reactions which produce these particles. This means that these particles decay through the weak interaction. In a later stage, when the quark model was introduced, Gell-Mann postulated that these particles consist of a new type of quark, *the strange quark*. To make everything consistent with earlier classifications based on the strangeness property the strange quark was assigned the value $S = -1$ and the antistrange quark the value $S = 1$. This means that in general S is equal to the number of $\bar{s}$-quarks minus the number of s-quarks. People also found a generalization of equation 3.1 by including the number S:

$$Q = e(I_3 + \frac{1}{2}B + \frac{1}{2}S) \tag{3.2}$$

By looking at this equation it seems more useful to consider the quantum number Y defined by $Y = B + S$. We will call Y the *hypercharge*. In terms of this hypercharge equation 3.2 becomes $Q = e(I_3 + Y/2)$. We call this the *Gell-Mann/Nishijima rule*.

As in the case of the baryon number B it seems logical to assume that the possible values for Y are eigenvalues of some hermitian operator $\hat{Y}$ and that the underlying symmetry group is $U(1)$ since Y is integer valued. This means that the operator $i\hat{Y}$ should be a basis element of a Lie algebra isomorphic to $\mathfrak{u}(1)$. It should also hold that the operator $\hat{Y}$ commutes with the operator $\hat{I}_3$ since Y and I_3 are simultaneously measurable. This gives us a strong indication that the Lie algebra corresponding to the flavour symmetry group of the particles which consist of the up, down and strange quarks must be at least a rank[1] 2 Lie algebra and that it should have $\mathfrak{u}(1) \oplus \mathfrak{su}(2)$ as a Lie subalgebra. One of the simplest rank 2 Lie algebras is the Lie algebra $\mathfrak{su}(3)$. The classification based on this particular Lie algebra agrees quite well with experimental observations. In the next section we will derive the irreducible representations of this Lie algebra which we will then use to classify the particles which consist of the u,d and s quarks.

[1]For the definition, see the top of page 17

3.2 The Lie algebra $\mathfrak{sl}_3(\mathbb{C})$ and its representations

3.2.1 The Lie algebra

The arguments in section 2.2.1 to derive the general form of the elements of the Lie algebra $\mathfrak{su}(2)$ also applies to the general Lie algebra $\mathfrak{su}(n)$. This means that the elements of the Lie algebra $\mathfrak{su}(3)$ consist of the 3x3 anti-hermitian matrices with trace zero. As in the previous chapter we are going to consider the complexified version of $\mathfrak{su}(3)$, which is the Lie algebra $\mathfrak{sl}_3(\mathbb{C})$. One can choose a particular useful basis of matrices for this Lie algebra, which are called the Gell-Mann matrices. Because the dimension of $SU(3)$ is 8, there are 8 basis elements for this Lie algebra. Let us denote these basis elements as λ_i. Since $\mathfrak{sl}_2(\mathbb{C})$ is a subalgebra of $\mathfrak{sl}_3(\mathbb{C})$ we can define the first three λ_i as the original Pauli matrices extended to 3x3 matrices by just adding zeroes to the third row and column. We can now construct λ_4 and λ_6 by shifting the nonzero elements of λ_1 down the opposite diagonal. In the same way we find λ_5 and λ_7 by shifting down the nonzero elements of λ_2. Since $\mathfrak{sl}_3(\mathbb{C})$ has rank 2, λ_8 must be diagonal. It is chosen in such a way that the condition $\mathrm{Tr}\lambda_i\lambda_j = 2\delta_{ij}$ holds for all the basis elements λ_i.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

By looking at these eight matrices, it is almost immediately evident that they are linearly independent. Especially if you take the linearly independence of the Pauli matrices into account. These basis elements also satisfy the commutation relations

$$[\lambda_i, \lambda_j] = f_{ij}^k \lambda_k \tag{3.3}$$

where f_{ij}^k is completely antisymmetric in i,j and k. A table of all the values for f_{ij}^k is given on page 190 in [5]. Because λ_3 and λ_8 are diagonal we see that they commute with each other. This means that we can associate the hypercharge operator and the isospin operator to these two basis elements. So let us now define a somewhat more useful basis for $\mathfrak{sl}_3(\mathbb{C})$ like we did in the case of $\mathfrak{sl}_2(\mathbb{C})$: $T_\pm = (\lambda_1 \pm i\lambda_2)/2$, $V_\pm = (\lambda_4 \pm i\lambda_5)/2$, $U_\pm = (\lambda_6 \pm i\lambda_7)/2$, $T_3 = \lambda_3/2$ and $Y = \lambda_8/\sqrt{3}$. We can also define the operators U_3 and V_3 as $2U_3 = 3/2Y - T_3$ and $2V_3 = 3/2Y + T_3$. These definitions give us the following important commutation relations:

$$[T_+, T_-] = 2T_3, \qquad [T_3, T_\pm] = \pm T_\pm \tag{3.4}$$

$$[U_+, U_-] = 2U_3, \qquad [U_3, U_\pm] = \pm U_\pm \tag{3.5}$$

$$[V_+, V_-] = 2V_3, \qquad [V_3, V_\pm] = \pm V_\pm \tag{3.6}$$

From these commutation relations we see that the T operators generate a Lie subalgebra of $\mathfrak{sl}_3(\mathbb{C})$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. This also holds for the U and V operators. It also means that we can view the operators $T_\pm$, $U_\pm$ and $V_\pm$ as shifting operators. The rest of the commutators are given on page 191 of Greiner's text.

3.2.2 The representations

To find the irreducible representations of $\mathfrak{sl}_3(\mathbb{C})$ we would like to use the same approach as the one used in the previous chapter. There we looked for a diagonal matrix and classified the irreducible representations by the eigenvalues of this matrix. The case of $\mathfrak{sl}_3(\mathbb{C})$ is however slightly more complicated because it has two diagonal matrices in its basis. We denoted these basis elements by T_3 and Y to make the link with isospin and hypercharge. These two operators commute with each other, so they form a Lie subalgebra which we will denote as $\mathfrak{h}$.

Let us now consider a representation ρ of $\mathfrak{sl}_3(\mathbb{C})$ on some vector space V. We will now first look at what happens if we apply ρ to $\mathfrak{h} \subset \mathfrak{sl}_3(\mathbb{C})$. For clarity we shall from now on write $X(v)$ instead of $\rho(X)(v)$ for all $X \in \mathfrak{sl}_3(\mathbb{C})$ and $v \in V$. Because Jordan decomposition is preserved under ρ, $\rho(\mathfrak{h})$ also consists of diagonalizable matrices which commute with each other. Because they all commute with each other they have a common set of eigenvectors. Let us now consider an element $H \in \mathfrak{h}$ and let v be an eigenvector for the action of $\mathfrak{h}$ on V. We then have that $H(v) = \alpha(H)v$. Here $\alpha(H)$ is a complex eigenvalue which depends on H. So in fact we have that $\alpha \in \mathfrak{h}^*$[2]. We can now decompose V as

$$V = \bigoplus_\alpha V_\alpha$$

where V_α is the eigenspace for the action of $\mathfrak{h}$ with eigenvalue $\alpha \in \mathfrak{h}^*$. We already mentioned at the end of the previous subsection that we can regard the $T_\pm$, $U_\pm$ and $V_\pm$ operators as shifting operators. We based this on the fact that they generate subalgebras isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. Let us now compute how these operators change the eigenvalues of the members in $\mathfrak{h}$. As a basis for $\mathfrak{h}$ we take of course T_3 and Y. Let v be an eigenvector for $\mathfrak{h}$. Then we have that $T_3(v) = \alpha(T_3)v$ and $Y(v) = \alpha(Y)v$. Since $[T_3, V_\pm] = \pm V_\pm/2$, $[T_3, U_\pm] = \mp U_\pm/2$ and $[T_3, T_\pm] = \pm T_\pm$ we have that:

$$\begin{aligned}
T_3 V_\pm(v) &= [T_3, V_\pm](v) + V_\pm T_3(v) = (\alpha(T_3) \pm 1/2)V_\pm(v) \\
T_3 U_\pm(v) &= [T_3, U_\pm](v) + U_\pm T_3(v) = (\alpha(T_3) \mp 1/2)U_\pm(v) \\
T_3 T_\pm(v) &= [T_3, T_\pm](v) + T_\pm T_3(v) = (\alpha(T_3) \pm 1)T_\pm(v)
\end{aligned}$$

Also because $[Y, V_\pm] = \pm V_\pm$, $[Y, U_\pm] = \pm U_\pm$ and $[Y, T_\pm] = 0$ we have that:

$$\begin{aligned}
YV_\pm(v) &= [Y, V_\pm](v) + V_\pm Y(v) = (\alpha(Y) \pm 1)V_\pm(v) \\
YU_\pm(v) &= [Y, U_\pm](v) + U_\pm Y(v) = (\alpha(Y) \pm 1)U_\pm(v) \\
YT_\pm(v) &= [Y, T_\pm](v) + T_\pm Y(v) = \alpha(Y)T_\pm(v)
\end{aligned}$$

[2] $\mathfrak{h}^*$ is the dual vector space of $\mathfrak{h}$. The dual space V^* of a K-vector space V consists of all the linear maps $V \to K$

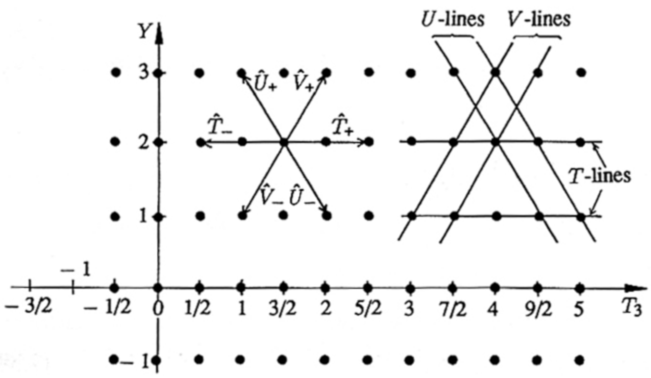

Figure 3.1: Example of a weight diagram and an illustration of the actions of the 6 shifting operators.

We can also illustrate these results in the so called *weight diagrams* for a given representation. This is a graph in which we plot all the possible eigenvalues for $\mathfrak{h}$. In this case it is a 2 dimensional graph since $\mathfrak{h}$ is 2 dimensional. For the x-axis we take the eigenvalues of T_3 and for the y-axis the eigenvalues of Y. Let us now restrict our attention to an irreducible representation of $\mathfrak{sl}_3(\mathbb{C})$ on a vector space V. We can now construct the weight diagram for this representation by using the shifting operators $T_\pm$, $U_\pm$ and $V_\pm$. Because the set of T operators is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ we know from chapter 2 that the possible eigenvalues $\alpha(T_3)$ for T_3 range from n to $-n$ by integer steps, where n is also an integer. This means that the weight diagram must be symmetric with respect to the Y axis. Since the sets of U and V operators are both also isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ we have that $U_3 = 0$ on the V-axis, V_3 is zero on the U-axis and that the weight diagram is symmetric with respect to the U and V axes. The U axis is defined by the line $\alpha(Y) = 2\alpha(T_3)/3$ and the V axis is defined by the line $\alpha(Y) = -2\alpha(T_3)/3$. All these four axes intersect each other in the origin, and the T, U and V axes form angles of 120 degrees between them, if we rescale the Y axis by $\sqrt{3}/2$.

Let us now consider an eigenvector v for which the pair of eigenvalues $(\alpha(T_3), \alpha(Y))$ is maximal[3]. We call this *the state with maximal weight*. Then we certainly have that $T_+(v) = U_-(v) = V_+(v) = 0$. It must also hold that there exists an integer a such that $V_-^{a+1}(v) = 0$ because the V-algebra is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. Similarly there must exist an integer b such that $T_-^{b+1}V_-^a(v) = 0$. Because of the symmetries of the multiplet these numbers a and b completely specify the multiplet. From now on we shall label the irreducible representations by these two numbers. So we write $\Gamma_{a,b}$ for the irreducible representation which is specified by a and b. Like in the previous chapter we can construct these irreducible representations from tensor products of the fundamental representation. One must be careful here however because there

[3]A pair $(\alpha(T_3), \alpha(Y))$ is called larger than $(\alpha(T_3)', \alpha(Y)')$ if $\alpha(T_3) \geq \alpha(T_3)'$ and $\alpha(Y) > \alpha(Y)'$.

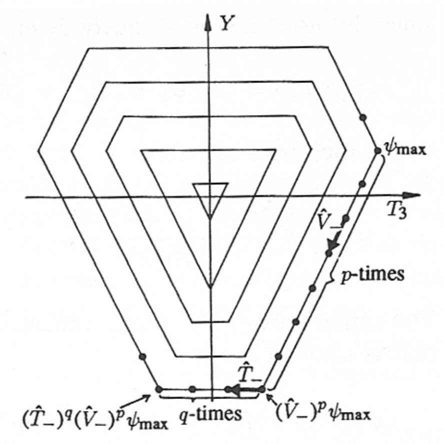

Figure 3.2: The construction of a $SU(3)$ multiplet

are two fundamental representations, since the dual representation is not isomorphic to the normal one. One can easily see this, by noting that the weights in the dual representation are just the negatives of the weights in the normal representation. This implies in turn that the weight diagrams of these 2 are different, and thus that the representations cannot be isomorphic (isomorphic representations have the same weights). There is however a automorphism of $\mathfrak{sl}_3(\mathbb{C})$ which carries the two weight diagrams into each other, namely the automorphism which sends $X \mapsto -X^T$, but this transformation does not satisfy the required relation for the equivalence of two representations. From now on we shall write $V = \mathbb{C}^3$ for the normal fundamental representation and $V^* = (\mathbb{C}^3)^*$ for the dual fundamental representation.

Lemma 3.2.1. *Let ρ be a representation of $\mathfrak{sl}_3(\mathbb{C})$ on the vector space* $\mathrm{Sym}^n V$*, where V is the fundamental representation of $\mathfrak{sl}_3(\mathbb{C})$. Then the weights in the weight diagram all have multiplicity one and the weight diagram has a triangular form*

Proof. We'll employ the same technique as in chapter two, where we considered the isomorphism of $\mathrm{Sym}^n V$ with the vector space of symmetric polynomials in $\dim(V)$ variables with total degree n. Let us take as a basis of V the three unit vectors, which we denote as x, y and z. To determine the eigenvalues of the diagonal operators in $\mathfrak{h}$ we first need to determine the actions of T_3 and Y on the three basis vectors. It's easy to see that $T_3(x) = x/2$, $T_3(y) = -y/2$ and $T_3(z) = 0$. Furthermore it holds that $Y(x) = x/3$, $Y(y) = y/3$ and $Y(z) = -2z/3$. The general form of an element of $\mathrm{Sym}^n V$ is of the form $x^{n-i-k}y^k z^i$, where $0 \leq k, i \leq n$ and $i + k \leq n$. We can now determine the actions of T_3 and Y on such a general element to determine all the possible eigenvalues:

$$\begin{aligned}
T_3(x^{n-i-k}y^k z^i) &= (n-i-k)T_3(x)x^{n-i-k-1}y^k z^i + kT_3(y)x^{n-i-k}y^{k-1}z^i + iT_3(z)x^{n-i-k}y^k z^{i-1} \\
&= \frac{1}{2}(n-i-2k)x^{n-i-k}y^k z^i \\
Y(x^{n-i-k}y^k z^i) &= (n-i-k)Y(x)x^{n-i-k-1}y^k z^i + kY(y)x^{n-i-k}y^{k-1}z^i + iY(z)x^{n-i-k}y^k z^{i-1} \\
&= (n/3 - i)x^{n-i-k}y^k z^i
\end{aligned}$$

So the general form of the weights for this representation is $((n-i-2k)/2, n/3-i)$. From this we can see that these weights have multiplicity one; Suppose that $x^{n-i-k}y^kz^i$ and $x^{n-i'-k'}y^{k'}z^{i'}$ are two different eigenvectors but have the same eigenvalue. This means in general that $i \neq i'$ and/or that $k \neq k'$. But from the general form of the eigenvalues we see that it must hold that $i = i'$ and $k = k'$. So the eigenvalues correspond to exactly one eigenvector from which we can conclude that the multiplicity of the eigenvalues is one. If you plot these weights in a weight diagram we see that the form of the diagram is triangular. One can also see this by noting that the weight diagram of V is triangular, so the weight diagram of $\text{Sym}^n V$ is also triangular, since it is obtained by doing n times a pairwise addition of all the weight vectors. □

Corollary 3.2.1. *The weight diagram of* $\text{Sym}^n V^*$ *has a triangular form and all the weights have multiplicity one*

Proof. We can relate the weight diagram of $\text{Sym}^n V$ to the weight diagram of $\text{Sym}^n V^*$ by the automorphism $X \mapsto -X^T$. This map only mirrors the weight diagram, so the multiplicities and geometric form of the diagram of $\text{Sym}^n V^*$ equal to those of $\text{Sym}^n V$. So by Lemma 3.2.1 we've proved the corollary. □

From these two results we can deduce that the symmetric powers of the two fundamental representations must be irreducible and that all the representations which possess a triangular weight diagram must be some symmetric power of either the normal or the dual fundamental representation. In the notation used before we have that $\text{Sym}^n V = \Gamma_{n0}$ and $\text{Sym}^n V^* = \Gamma_{0n}$.

Let us now denote the standard basis vectors of V as e_1,e_2 and e_3 so that the corresponding dual basis vectors of V^* are denoted by e_1^*, e_2^* and e_3^*. From the proof of the lemma above we can see that e_1 is a highest weight vector of V, which in turn implies that e_1^n is a highest weight vector of $\text{Sym}^n V$ with weight $(n/2, n/3)$. Since the weights in the dual representation are just the negatives of the weights in the normal representation, one easily sees that the highest weight of V^* is given by $(0, 1/3)$, which corresponds to the weight vector e_3^*. This, in turn, implies that $(e_3^*)^n$ is the highest weight vector of $\text{Sym}^n V^*$.

We can now also consider the representations on the vector space $\text{Sym}^a V \otimes \text{Sym}^b V^*$ with say $a \geq b$. Certainly, it does not a priori have to hold that this representation is irreducible. So in general we may write that $\text{Sym}^a V \otimes \text{Sym}^b V^* = \Gamma_{i,j} \oplus V'$ for some i, j and subrepresentation V'. One way to do this is to find a surjective map from $\text{Sym}^a V \otimes \text{Sym}^b V^*$ to some other vector space s.t. $\Gamma_{i,j}$ is in the kernel of this map. We have a natural choice for such a map, namely the tensor contraction map;

$$\begin{aligned} i_{a,b} : \text{Sym}^a V \otimes \text{Sym}^b V^* \to \quad & \text{Sym}^{a-1} V \otimes \text{Sym}^{b-1} V^* \\ (v_1 \cdot \ldots \cdot v_a) \otimes (w_1^* \cdot \ldots \cdot w_b^*) \mapsto \quad & \sum_{ij} \langle v_i, w_j^* \rangle \, (v_1 \cdot \ldots \cdot \hat{v}_i \cdot \ldots \cdot v_a) \otimes (w_1^* \cdot \ldots \cdot \hat{w}_j^* \cdot \ldots \cdot w_b^*) \end{aligned}$$

Theorem 3.2.1. *Let ρ be a representation of $\mathfrak{sl}_3(\mathbb{C})$ on the vector space* $\text{Sym}^a V \otimes \text{Sym}^b V^*$*, where V is the fundamental representation of $\mathfrak{sl}_3(\mathbb{C})$. Then it holds that $i_{a,b}$ is surjective and that* $\text{Ker}(i_{a,b}) = \Gamma_{ab}$.

Proof. It is evident from the description of the map $i_{a,b}$ that is a surjective map. Let us now consider the highest weight vector v of $\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^*$. From the discussion above we know that $v = e_1^a \otimes e_3^{*b}$. Let us now compute the action of $i_{a,b}$ on v:

$$i_{a,b}(v) = i_{a,b}(e_1^a \otimes e_3^{*b}) = ab \langle e_1, e_3^* \rangle\, e_1^{a-1} \otimes e_3^{*b-1} = 0$$

So we see that the highest weight vector v lies in the kernel of $i_{a,b}$. We know that every highest weight vector generates an irreducible representation, so the irreducible representation Γ_{ab} generated by v lies in the kernel of $i_{a,b}$. It now remains to show that Γ_{ab} is the entire kernel. We will do this by using a formula which gives the dimension of Γ_{ab}[4]:

$$\dim(\Gamma_{ab}) = (a+b+2)(a+1)(b+1)/2 \tag{3.7}$$

We can also easily calculate the dimension of the kernel of $i_{a,b}$ if we use formula (2.7):

$$\begin{aligned}
\dim\ker(i_{a,b}) &= \dim(\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^*) - \dim(\mathrm{Sym}^{a-1} V \otimes \mathrm{Sym}^{b-1} V^*) \\
&= \binom{a+2}{a}\binom{b+2}{b} - \binom{a+1}{a-1}\binom{b+1}{b-1} \\
&= ((a+2)(a+1)(b+2)(b+1) - ab(a+1)(b+1))/4 \\
&= (a+1)(b+1)(a+b+2)/2
\end{aligned}$$

which is the same as the dimension of Γ_{ab}. Because Γ_{ab} lies in the kernel of $i_{a,b}$ it must be that $\Gamma_{ab} = \ker(i_{a,b})$. □

Corollary 3.2.2. *Let ρ be a representation of $\mathfrak{sl}_3(\mathbb{C})$ on the vector space $\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^*$ where V is the fundamental representation. This representation decomposes as:*

$$\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^* = \bigoplus_{i=0}^{b} \Gamma_{a-i,b-i} \tag{3.8}$$

Proof. Because the map $i_{a,b}$ is surjective like we said before, we have by theorem (3.2.1) the following decomposition of $\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^*$ for $a \geq b$:

$$\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^* = \left(\mathrm{Sym}^{a-1} V \otimes \mathrm{Sym}^{b-1} V^*\right) \oplus \Gamma_{ab} \tag{3.9}$$

By continuing this process of contracting the tensor product we obtain the required decomposition □

3.3 The SU(3) classification of the mesons and baryons

3.3.1 The up, down and strange quarks

Let us now return to the classification of the mesons and baryons, but this time we will also consider mesons and baryons that may consist of one ore more (anti) strange quarks. By drawing parallels between the results of chapter 2, we proceed with this by assigning the elementary

[4] See formula 15.17 in [4]

particles to the irreducible representations of $\mathfrak{sl}_3(\mathbb{C})$. We have seen in the previous section that we can construct these irreducible representations from appropriate symmetric tensor products of the two fundamental representations. Since we use the hypothesis that all the baryons and mesons consist of quarks, we will assign the (anti) up, (anti) down and (anti) strange quarks to the (dual) fundamental representation.

Let us first consider how to assign the up, down and strange quark to the fundamental representation op $\mathfrak{sl}_3(\mathbb{C})$. From the formulas in the proof of lemma 3.2.1 we see that the possible eigenvalues for T_3 are $-1/2$, 0 and $1/2$. The possible eigenvalues for the hypercharge Y are respectively $1/3$, $-2/3$ and $1/3$. In chapter two we assigned the eigenvalue $\alpha(T_3) = 1/2$ to the up quark and $\alpha(T_3) = -1/2$ to the down quark. This means that if we want to be consistent to this convention, we should assign the eigenvalue $\alpha(T_3) = 0$ to the strange quark. If we put this information in the dirac notation, we get:

$$|u\rangle = |1/2, 1/3\rangle\,, \qquad |d\rangle = |-1/2, 1/3\rangle\,, \qquad |s\rangle = |0, -2/3\rangle \tag{3.10}$$

We can now easily determine the electrical charge of the three quarks by using the Gell-Mann/Nishijima formula since we now know the isospin and hypercharge of them.

$$Q_u = 2e/3, \qquad Q_d = -e/3, \qquad Q_s = -e/3 \tag{3.11}$$

However, nobody has actually been able to measure these fractional charges, since nobody has been able to observe a free quark. Physicists try to explain this by a model called quark-confinement. This model says for example that the strong force increases with distance, so that if you try to pull the quarks from each other they are going to resist more and more. This phenomenon is due to the inter gluonic interactions. In the theory of the strong force which is called quantum chromodynamics, the gluons are the exchange bosons that couple to the colour charge. This is analogous to the electromagnetic interaction in which the photons are the exchange bosons. The only difference however, is that the gluons can interact with themselves because they also carry a colour charge. So by increasing the distance between the quarks, you get more and more interacting gluons which causes the potential acting on the quark to rise.

3.3.2 Mesons

Let us now continue to use the quarks as building blocks to construct the elementary particles. We will begin by looking at various types of mesons. Recall from the previous chapter that mesons are elementary particles that consist of a quark-antiquark pair. From this definition we see that the mesons must lie in the $V \otimes V^*$ representation of $SU(3)$. This representation is not irreducible as we have seen in the previous section; we can decompose it into $\Gamma_{1,1} \oplus \mathbb{C}$.
Let us now look at the geometrical shape of the irreducible representation $\Gamma_{1,1}$. Since $V \otimes V^*$ is 9 dimensional, we see that $\Gamma_{1,1}$ must be 8 dimensional. In the previous section we've also seen that we can construct the outer border of Γ_{ab} by repeatedly applying the operator V_- to the highest weight vector v of Γ_{ab} until you get zero, and then by repeatedly applying the operator T_- to this vector $V_-^a v$, a total of b times, after which we get the entire border by using

the symmetries of the weight diagram. So in the case of $\Gamma_{1,1}$ this results in a hexagonal shape for the outer border of the multiplet. One can also easily see that the state in the center of the multiplet has multiplicity two. After a simple calculation one sees that the top two states have hypercharge $\alpha(Y) = 1$. The three center states have $\alpha(Y) = 0$ and that the two lower states have $\alpha(Y) = -1$.

There are however some minor difficulties with classifying the mesons into an octet and a singlet. If you look at the baryons for example, you see that the baryons and antibaryons lie in different multiplets. In the case of mesons however, you can immediately see that the mesons and their anti counterparts lie in the same multiplet. There is also another phenomenon that occurs in the meson multiplets, namely $SU(3)$ mixing; consider two $SU(3)$ meson multiplets with equal spin, parity and baryon number. It's now possible that for a state in the first multiplet there exists a state in the other multiplet with the same values for $\alpha(Y)$ and $\alpha(T_3)$. The physical states of the corresponding particles are now mixtures of these multiplet states. For example, there is a state in the octet which has the same eigenvalues of the hypercharge and isospin as the singlet state. Because of this mixing, we will only consider the meson nonet, instead of the octet and singlet separately. This also means that we can classify the mesons into different nonets, where every nonet corresponds to a different value for the spin and parity. To label these different meson nonets we introduce the notation J^P for them, where J stands for the spin, and P for the parity:

1. $J = 0$ and $P = \pm$: (pseudo)scalar mesons
2. $J = 1$, $P = -$: vector mesons
3. $J = 1$, $P = +$: axial vector mesons
4. $J = 2$, $P = \pm$: (pseudo)tensor mesons

Let us first look at the pseudoscalar mesons, because we also looked at these mesons in chapter 2. There we classified these mesons in the pion triplet. These three mesons are composed of various combinations of the up and down quarks. This means that the strangeness of the pions is zero which in turn implies that the hypercharge is zero since mesons have baryon number $B = 0$. We will call the pseudoscalar mesons with strangeness $S = \pm 1$ the kaons. The two kaons with $S = 1$ are denoted by K^0 and K^+, where K^0 has isospin $\alpha(T_3) = -1/2$ and K^+ has isospin $\alpha(T_3) = 1/2$. The antikaons K^- and $\bar{K}^0$ are placed in the corresponding diagonal mirror positions of the kaons with $S = 1$. Let us now display the nonets for the (pseudo)scalar, vector and tensor mesons together with the particle properties:

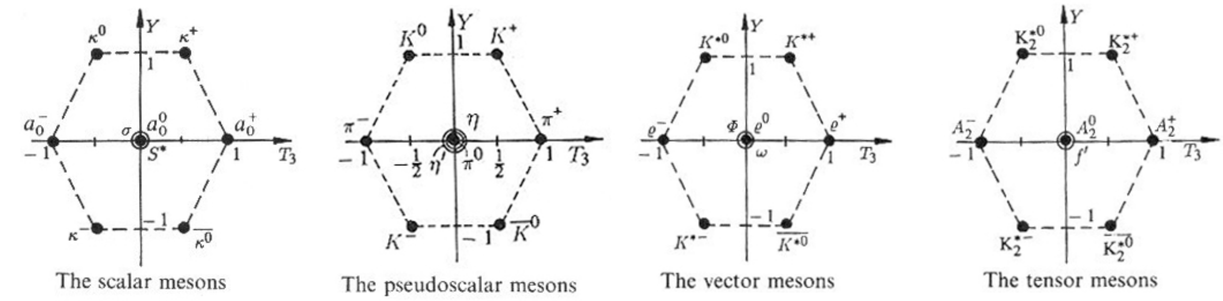

Figure 3.3: The meson nonets [5]

particle	mass (MeV)	charge (e)	width (MeV)	Y	T_3	spin
The scalar mesons						
a_0	976	± 1,0	50	0	± 1, 0	0
κ	1250	+1, 0	450	± 1	$\pm 1/2$	0
S^*	993	0	40	0	0	0
σ	750	0	600	0	0	0
The pseudoscalar mesons						
π^+	139.59	+1	$2.5 \cdot 10^{-13}$	0	1/2	0
π^0	135.00	0	$7.4 \cdot 10^{-6}$	0	0	0
π^-	139.59	-1	$2.5 \cdot 10^{-13}$	0	$-1/2$	0
K^+	493.82	+1	$5.4 \cdot 10^{-14}$	1	1/2	0
K^0	497.82	0	50% K_s / 50% K_l	1	$-1/2$	0
K^-	493.82	-1	$5.4 \cdot 10^{-14}$	-1	$-1/2$	0
$\bar{K}^0$	497.82	0	50% K_s / 50% K_l	-1	1/2	0
						0
η	548.6	0	0.07	0	0	0
η'	958	0	0.99	0	0	0
$K_s = 0.74 \cdot 10^{-11}$						
$K_l = 0.11 \cdot 10^{-13}$						

particle	mass (MeV)	charge (e)	width (MeV)	Y	T_3	spin
The vector mesons						
ρ^+	773	+1	154	0	1/2	1
ρ^0	773	0	154	0	0	1
ρ^-	773	−1	154	0	−1/2	1
ω	782.7	0	10	0	0	1
K^{*+}	892	+1	50	1	1/2	1
K^{*0}	898	0	50	1	−1/2	1
K^{*-}	892	−1	50	−1	−1/2	1
$\bar{K}^{*0}$	898	0	50	−1	1/2	1
						1
Φ	1019	0	4.2	0	0	1
						1
The tensor mesons						
A_2	1320	$\pm1, 0$	110	0	±1,0	2
K_2^*	1426	$\pm1, 0$	100 ± 3	±1	±1/2	2
f_2'	1525	0	76 ± 10	0	0	2
f_2	1274	0	185 ± 20	0	0	2
						1

3.3.3 Baryons

In this section we will proceed by looking at how to classify the baryons by the $SU(3)$ classification scheme. As we know from previous discussions, the baryons are (anti) quark triples. This means that we can classify the baryons by looking at the tensor product $V \otimes V \otimes V$ and the antibaryons by looking at the tensor product $V^* \otimes V^* \otimes V^*$. But we only need to look at the decomposition of $V \otimes V \otimes V$, since this immediately gives us the decomposition of $V^* \otimes V^* \otimes V^*$; the dual representation of $\Gamma_{a,b}$ is just $\Gamma_{b,a}$. There is however one slight problem: We can't use the decomposition formula we had derived in the previous section because our tensor product is not of the form $\mathrm{Sym}^a V \otimes \mathrm{Sym}^b V^*$. So to proceed, we need to find the decomposition by looking at the eigenvalues of the tensor product.

Let us first look at the tensor product $V \otimes V$. If we choose the standard basis e_i in V, we've seen that e_1 is the highest weight vector for V. So the highest weight vector for $V \otimes V$ is e_1^2. This weight vector generates the irreducible representation $\Gamma_{2,0} = \mathrm{Sym}^2 V$. To find the complement of this representation we can see by writing out all the 9 possibilities for the weights of $V \otimes V$, that the weights of the complement are $(0, 2/3)$ and $(\pm 1/2, -1/3)$. These are precisely the weights of the irreducible representation V^*. So in total we find the following decomposition of $V \otimes V$:

$$V \otimes V = \Gamma_{2,0} \oplus V^* \tag{3.12}$$

We can use this result to partially find a decomposition of $V \otimes V \otimes V$:

$$V \otimes V \otimes V = V \otimes (\Gamma_{2,0} \oplus V^*) = (V \otimes \Gamma_{2,0}) \oplus (V \otimes V^*) = (V \otimes \Gamma_{2,0}) \oplus \Gamma_{1,1} \oplus \mathbb{C} \tag{3.13}$$

This isn't the final decomposition, because the representation $V \otimes \Gamma_{2,0}$ is not irreducible: It has the highest weight vector e_1^3 which generates the irreducible subrepresentation $\Gamma_{3,0}$. This subrepresentation is 10 dimensional, so its complement must be 8 dimensional. This means that the complement must either be $\Gamma_{1,1}$ or the sum of irreducible representations for which the sum of their dimensions equals 8. The latter cannot be possible so the complement must be $\Gamma_{1,1}$. This gives us the final decomposition for $V \otimes V \otimes V$:

$$V \otimes V \otimes V = \Gamma_{3,0} \oplus \Gamma_{1,1} \oplus \Gamma_{1,1} \oplus \mathbb{C} \tag{3.14}$$

So we see that we can classify the baryons into a decuplet, an octet and into a singlet. The two octets that appear in this decomposition are isomorphic to each other.

In the previous chapter we already encountered some examples of baryons, namely the Δ-baryons and the two nuclei, the proton and the neutron. These baryons are composed of up and down quarks. This means that they have strangeness $S = 0$. Because baryons have baryon number $B = 1$, we see from the definition of the hypercharge, that these baryons have hypercharge $Y = 1$. We can easily see from this that we can assign the Δ-baryons to the representation $\Gamma_{3,0}$ since it has 4 weights with eigenvalue $\alpha(Y) = 1$ and we can assign the two nuclei to the representation $\Gamma_{1,1}$ because it has two weights with $\alpha(Y) = 1$. The four baryons that have strangeness $S = -1$ are called the three Σ-baryons and the Λ^0-baryon, and the two baryons that have $S = -2$ are called the Ξ-baryons. The baryon that has strangeness $S = -3$ is called the Ω^--baryon. The Σ-, Λ^0- and the Ξ-baryons are placed in the representation $\Gamma_{1,1}$ because they have spin 1/2. Their resonance counterparts Σ^* and Ξ^* are placed in the representation $\Gamma_{3,0}$ because they have spin 3/2. The Ω^- particle is placed in the $\Gamma_{3,0}$ representation because the $\Gamma_{1,1}$ representation doesn't contain the weight with eigenvalue $\alpha(Y) = -2$.

The singlet representation does not contain a physically possible particle due to the Pauli-exclusion principle. This state would have to consist of a uds quark combination, and also has to share the same relative spin as one of the other three uds-baryons (Σ^{*0},Σ^0 and Λ^0), which is not possible.

Let us now give a table with the properties of these elementary particles, like we did with the mesons in the previous section:

particle	mass (MeV)	charge (e)	lifetime (s)	spin	Y	T_3
The baryon decuplet						
Δ^{++}	1232	+2	$6 \cdot 10^{-24}$	3/2	1	3/2
Δ^{+}	1232	+1	$6 \cdot 10^{-24}$	3/2	1	1/2
Δ^{0}	1232	0	$6 \cdot 10^{-24}$	3/2	1	−1/2
Δ^{-}	1232	−1	$6 \cdot 10^{-24}$	3/2	1	−3/2
Σ^{*+}	1382	+1	$1.78 \cdot 10^{-25}$	3/2	0	1
Σ^{*0}	1382	0	$1.78 \cdot 10^{-25}$	3/2	0	0
Σ^{*-}	1387	−1	$1.78 \cdot 10^{-25}$	3/2	0	-1
Ξ^{*0}	1531.8	0	$9.4 \cdot 10^{-23}$	3/2	−1	1/2
Ξ^{*-}	1535.0	−1	$9.4 \cdot 10^{-23}$	3/2	−1	−1/2
Ω^{-}	1672	−1	$0.82 \cdot 10^{-10}$	3/2	−2	0
The baryon octet						
p	938.3	+2	stable	1/2	1	1/2
n	939.6	+1	886	1/2	1	−1/2
Σ^{+}	1189.4	+1	$0.8 \cdot 10^{-10}$	1/2	0	1
Σ^{0}	1192.5	0	$6 \cdot 10^{-20}$	1/2	0	0
Σ^{-}	1197.4	−1	$1.5 \cdot 10^{-10}$	1/2	0	-1
Λ^{0}	1115.6	0	$2.6 \cdot 10^{-10}$	1/2	0	0
Ξ^{0}	1315	0	$2.9 \cdot 10^{-10}$	1/2	−1	1/2
Ξ^{-}	1321	−1	$1.5 \cdot 10^{-10}$	1/2	−1	−1/2

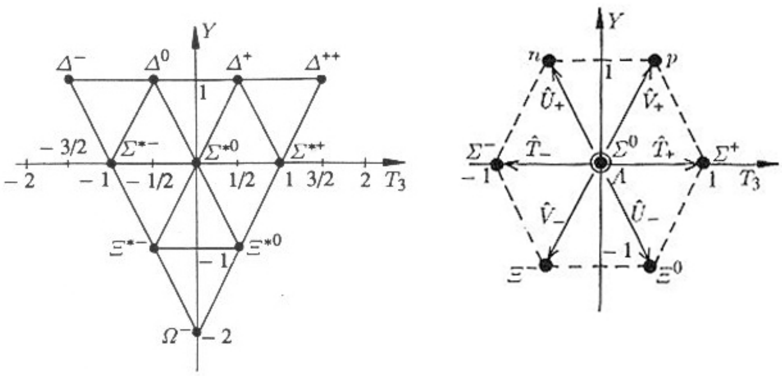

Figure 3.4: The baryon decuplet and octet [5]

3.3.4 Mass splitting in the $SU(3)$ multiplets

One thing that one can immediately conclude from the experimental measurements of the particle masses, which are given in the previous subsection, is that the masses are not constant on each multiplet. From this we can conclude that the $SU(3)$ symmetry can not be an exact symmetry of the strong interaction, because if it were we would have that the energies (and thus the masses) are constant on a multiplet (proposition 2.1.1). If we look at the masses of the baryon decuplet, we see that the mass splitting $\Delta M/M$ is of the order of 100 MeV, as is also the case for the baryon octet. We also see that the masses of the baryons increase by the about the same amount when $|S|$ increases. This suggests that the mass of the strange quark is somewhat larger than the masses of the up and down quarks and that we have a constant mass splitting in the multiplets.

Let us now try to find the form of the Hamiltonian for the simple case of the three flavor quarks q_i, where we assume that $m_u = m_d < m_s$. This assumption is justified if we ignore the electromagnetic mass splitting between the quarks, which implies that the Hamiltonian is invariant under the isospin operator T_3. We can now easily see that

$$(\langle q_i| H_{\text{strong}} |q_j\rangle) = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix} = \left(\frac{2m_u + m_s}{3}\right) I + \left(\frac{m_u - m_s}{\sqrt{3}}\right) \lambda_8$$

In this case, we see that the Hamiltonian consists of a $SU(3)$ invariant term, and a term that is proportional to the hypercharge. Because of our previous observations on the masses in the baryon multiplets, we can assume that a similar result holds for all the baryon representations. So the Hamiltonian H_{strong} can now be written as

$$H_{\text{strong}} = H_0 + H_{\text{ms}}$$

where H_0 is the $SU(3)$ invariant term, and H_{ms} the term responsible for the constant mass splitting. Since the mass splitting in the multiplets is of the order of 10%, we can assume that H_{ms} is small with respect to H_0. Perturbation theory can thus be used on H_{ms} to calculate the masses of the particles inside a baryon multiplet. Let us now consider a representation ρ on a baryon multiplet. Let $|\psi\rangle$ be a particle in this multiplet, and let us, for notational simplicity, write X instead of $\rho(X)$ for all $X \in \mathfrak{sl}_3(\mathbb{C})$. The expectation value of M of H_{strong} in the associated perturbed state $|\psi'\rangle$ can then written as

$$M = \langle \psi'| H_{\text{strong}} |\psi'\rangle = \langle \psi| H_0 |\psi\rangle + \langle \psi| H_{\text{ms}} |\psi\rangle = M_0 + \langle \psi| H_{\text{ms}} |\psi\rangle \tag{3.15}$$

We can evaluate $\langle \psi| H_{\text{ms}} |\psi\rangle$ with a result proved by S. Okubo[5]. This result states that H_{ms} can be written as

$$H_{\text{ms}} = a\lambda_8 + b\sum_{j,k} d_{8jk}\lambda_j\lambda_k \tag{3.16}$$

[5]Equation A8 in Progress of Theoretical Physics, Vol. 27, No. 5, May 1962

where a and b are constants on a multiplet,and where d_{ijk} is the totally symmetric symbol[6]. Writing out the second term in the expression, we find that

$$\begin{aligned}
\sum_{j,k} d_{8jk}\lambda_j\lambda_k &= \frac{1}{\sqrt{3}}(\lambda_1^2+\lambda_2^2+\lambda_3^2) - \frac{1}{2\sqrt{3}}(\lambda_4^2+\lambda_5^2+\lambda_6^2+\lambda_7^2) - \frac{1}{\sqrt{3}}\lambda_8^2 \\
&= -\frac{1}{2\sqrt{3}}\sum_i \lambda_i^2 + \frac{3}{2\sqrt{3}}(\lambda_1^2+\lambda_2^2+\lambda_3^2) - \frac{1}{2\sqrt{3}}\lambda_8^2
\end{aligned}$$

If we now use the identifications $\lambda_8 = \sqrt{3}Y$ and $\lambda_1^2+\lambda_2^2+\lambda_3^2 = 4T^2$, we find that

$$\begin{aligned}
H_{\mathrm{ms}} &= a\sqrt{3}Y + b\left(-\frac{1}{2\sqrt{3}}\sum_i \lambda_i^2 + \frac{3}{2\sqrt{3}}4T^2 - \frac{1}{2\sqrt{3}}3Y^2\right) \\
&= a\sqrt{3}Y + 2\sqrt{3}b\left(-\frac{1}{12}\sum_i \lambda_i^2 + T^2 - \frac{1}{4}Y^2\right)
\end{aligned}$$

We can now, without loss of generality, absorb the square roots into the constants a and b and write:

$$M = M_0 + \langle\psi|\, aY + b\left(-\frac{1}{12}\sum_i \lambda_i^2 + T^2 - \frac{1}{4}Y^2\right)|\psi\rangle$$

It can be easily seen that the operator $\sum_i \lambda_i^2$ is a Casimir operator for $\mathfrak{sl}_3(\mathbb{C})$, which means that is proportional to the identity on a multiplet, so this term can be absorbed into M_0. We now obtain the *Gell-Mann-Okubo* mass formula for the masses of the baryons[7].

$$M = M_0 + aY + b(T(T+1) - Y^2/4) \tag{3.17}$$

We can now try to use the mass formula to calculate the mass of the Ω^- particle. If we apply the formula to the decuplet we get the following four equations:

$$\begin{aligned}
M'_\Delta &= M_\Delta - M_0 = a + 7b/2 \\
M'_{\Sigma^*} &= M_{\Sigma^*} - M_0 = 2b \\
M'_{\Xi^*} &= M_{\Xi^*} - M_0 = -a + b/2 \\
M'_{\Omega^-} &= M_{\Omega^-} - M_0 = -2a - b
\end{aligned}$$

From this we can easily see that $M'_{\Omega^-} = 2M'_{\Xi^*} - M'_{\Sigma^*}$. This means in turn that:

$$M_{\Omega^-} = 2M_{\Xi^*} - M_{\Sigma^*} \tag{3.18}$$

If we now insert the experimental values of the masses of the Ξ^* and Σ^* into equation (3.18) we obtain a mass of 1683 MeV for the mass of the Ω^- particle. This predicted value deviates only by about 0.5% from the experimentally measured value. So we see from this that

[6] A table of all its values can be found in [7], page 483

[7] Note that we write Y instead of $\alpha(Y)$ to simplify the notation. Here, T stands for the maximal value of $\alpha(T_3)$ for the isospin multiplet associated to the hypercharge $\alpha(Y)$

the Gell-Mann-Okubo mass formula is quite well satisfied for the baryons. The mass formula however, does not work for mesons. This can be attributed to the fact that mesons satisfy the Klein-Gordon equation,

$$(\partial_\mu \partial^\mu + m^2)\phi = 0 \tag{3.19}$$

because they are bosons.[8]. We see that this equation contains quadratic terms in the mass. Baryons however, which are fermions, satisfy the Dirac equation

$$(i\gamma^\mu \partial_\mu + m)\phi = 0 \tag{3.20}$$

which contains linear terms in the mass. Therefore it seems plausible to substitute M^2 for M in equation (3.17) if we want to calculate the masses of the mesons. So for the mesons we obtain the following mass formula:

$$M^2 = M_0^2 + aY + b(T(T+1) - Y^2/4) \tag{3.21}$$

It is however somewhat more difficult to work with this mass formula due to the mixing of meson multiplets, which means that you have to take the mixing angles into account. Altough more difficult to work with, it resulted in a good prediction of the mass of the η meson. This in turn, was another victory for the $SU(3)$ classification scheme.

[8]Notice that we have given the Klein-Gordon equation for a free field. In our case there will appear interaction terms in the equation, but this doesn't change the argument

Conclusion

In this thesis, we saw that we could classify the elementary particles by only invoking the principles of symmetry and some simple quantum mechanics. This is in fact quite remarkable, since the full theory of elementary particles requires extensive use of quantum field theory. However, the standard model still relies a substantial amount on the principles of symmetry. So in a certain sense we can view the work we have done in this thesis as a sort of preliminary setup for all the gauge theories that have been constructed for the elementary particle interactions.

But the quark model which we explored, still has some major disadvantages. First of all, it has not much predictive power; it's only a classification scheme. Within this framework we derived the general properties of the particles, but as we have seen, we can not predict the masses of the particles inside a multiplet without knowing some of the masses in that multiplet! The quark model thus has some free parameters that we cannot predict without resorting to other theories. The second disadvantage of the quark model is that we cannot cover the whole hadron spectrum; the particles that are made up of non u, d and s quarks can't be sensibly fit into multiplets, because these other quarks differ too much in mass, especially if one looks at the particles that are made up of the top and bottom quarks. So we see that the predictive and the classificational power of the quark model is a bit limited, but that it still gives quite good results in the regime were we can apply it, and this gives us a real good intuition that the idea of symmetry is a very important idea in particle physics.

Bibliography

[1] M.A. Armstrong. *Basic Topology.* Springer-Verlag, New York, 1983.

[2] M. Chaichian and R. Hagedorn. *Symmetries in quantum mechanics: from angular momentum to supersymmetry.* Institute of Physics Publishing, Bristol, 1998.

[3] J.F. Cornwell. *Group Theory in Physics volume 2.* Academic Press, London, 1984.

[4] W. Fulton and J. Harris. *Representation theory: a first course.* Springer-Verlag, New York, Berlin, 1991.

[5] W. Greiner and B. Müller. *Quantum Mechanics - Symmetries.* Springer-Verlag, Berlin, Heidelberg, 1994.

[6] D.J. Griffiths. *Introduction to elementary particles.* Wiley, New York, 1987.

[7] O. Nachtmann. *Elementary Particle Physics: Concepts and Phenomena.* Springer-Verlag, New York, Berlin, 1990.

[8] D.H. Sattinger and O.L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry and Mechanics.* Springer-Verlag, New York Berlin Heidelberg Tokyo, 1986.