P420M: Groups and Symmetries (PHYS5007)

Dr Christoph Englert

Particle Theory Group Glasgow, Rm 534
Email: christoph.englert@glasgow.ac.uk

version of Sunday 19th March, 2017, 21:12

Contents

1 Introductory remarks

Symmetries are the cornerstones of our modern understanding of physics. Their mathematical realisation runs like a golden thread across the boundaries of classical mechanics, electrodynamics, quantum mechanics and quantum field theory. Today we accept that symmetries can be thought of as the fundamental entities which should be used to construct theories of nature. For instance the standard model of particle physics, which is a theoretical model capable of making the most precise predictions that have been confirmed by measurement, is a theory that is entirely based on the concept of symmetries.

If we want to solve e.g. the two body problem that underpins the planetary motion in our solar system, we will look for solutions of fixed energy, fixed angular momentum and fixed momentum of the centre-of-mass systems. In a way, this might seem trivial to you, mostly because statements like "no energy can be destroyed" are the doctrines of undergraduate teaching classes. While energy conservation is something that we can measure in a variety of experiments, the fundamental question we should ask really is "*Why is energy conserved?*". This question becomes even more interesting given the abundant experimental evidence of conservation laws related to energy, momentum and angular momentum. We see energy conservation at large distances, at small distances, in systems with few or many degrees of freedom, although we typically use different theoretical approaches to analyse the physics of these diverse systems. It goes without saying that we would miss a crucial point that underpins physics as a whole if we elevated the conservation laws to axioms.

As we will see, the mathematical reason behind these conservation laws is spelled out by Noether's theorem (which we will discuss in Sec. 2.1.2), namely that there are associated symmetries of space and time that will have to be reflected by the dynamics of a system that lives in space and time. This has to be true irrespective of it being small or large, quantum mechanical or classical. Once we have identified the symmetries that are intrinsic to a particular physical system, we can use their mathematical formulation to simplify the solution of its equations of motion: Noether's theorem connects symmetry transformations that leave physics invariant directly with conserved physical quantities:

$$\begin{aligned} \text{time translation invariance} &\implies \text{energy conservation} \\ \text{translation invariance} &\implies \text{momentum conservation} \\ \text{rotational invariance} &\implies \text{angular momentum conservation}\,. \end{aligned} \tag{1.0.1}$$

The crucial parameter for the potential of a two-body problem (after transforming into the centre-of-mass system) is the difference between the two bodies $|\vec{x}|$ which enters in the $1/|\vec{x}|$ dependence of the gravity potential. This difference is invariant under rotations and constant translations in three dimensions, and does not depend on time *explicitly*. The conserved quantities of Eq. (1.0.1) emerge from the invariance of the potential (the kinetic energy is trivial). For the $1/|\vec{x}|$ potential there is even an additional symmetry connecting momentum $\vec{p}$ and angular momentum $\vec{L}$, which is maybe less familiar from your other courses: The Lenz-Runge vector $\vec{A} = \vec{p} \times \vec{L} - m\alpha\vec{e}_r$, with $\alpha = GmM$ for gravity in the two-body problem. Labeling the solutions of the two-body problem (e.g. the planetary orbits around the sun) according to specific values of the conserved quantities like energy, momentum, Lenz-Runge vector and angular momentum, makes finding the solutions almost trivial. For example, using $\vec{L} = \vec{x} \times m\dot{\vec{x}} = \text{const.}$ it follows immediately

$$\vec{L} \cdot \vec{x} = \vec{L} \cdot \dot{\vec{x}} = \vec{L} \cdot \ddot{\vec{x}} = 0\,,$$

i.e. the orbit is localised in a plane orthogonal to $\vec{L}$, and we effectively deal with a problem in two dimensions instead of three, which simplifies the algebra considerably.

We need to learn to walk before we can run, so let's step back and retrace the above steps and develop the formalism that will allow us to express symmetries for classical and quantum systems bit by bit. We will first revisit classical physics and quantum mechanics with a particular emphasis on the concept of symmetries in these frameworks. We will see how symmetries can arise as global symmetries such as the Galilean transformations of classical mechanics, or as local symmetries as gauge symmetries (such as in electrodynamics). Common to all these continuous symmetries is that they are not only groups in an algebraic sense, but they are also finite dimensional differential manifolds, e.g. one can define continuously differentiable paths in the group space. Such objects are well-studied in maths literature - they are Lie groups and their representation theory is completely known and particularly transparent.

A note on textbooks and these notes

There are many textbooks on symmetries and there is no recommendation that I can give that would reflect this course completely, or best suited for every learner.

These notes are self-contained and all necessary concepts will be introduced and you can find a number of exercises to challenge your understanding of the material. This does not mean that you shouldn't explore other approaches to groups and symmetries in physics. There is an enormous amount of information available on the web or in the library and I encourage you to find your own, personal angle on this subject.

Please help me to improve these notes for future generations by informing me of any typos you may find. Your help is very much appreciated.

2 Symmetries and Physics*

2.1 Continuous Symmetries in Classical Physics

2.1.1 Lagrange Formalism and the Principle of Least Action

You are all familiar with Newton's axioms which state that the dynamics of a classical mechanics system is entirely determined by knowing its initial conditions, as well as all forces that act it. Solving the associated second order differential equations will give you the classical trajectories. However, this can become quite tedious as often a system is subject to constraints that are not always expressed straightforwardly when using Newton's axioms. For instance, a pendulum of length ℓ when moving in the $y-z$ direction is always subject to the constraint $A(\vec{x}) = \ell^2 - x(t)^2 - z^2(t) = 0$. The way this is accounted for in Newton's formalism is by introducing a normal force onto mass that compensates the gravitational pull, or put differently, a constraint $A(\vec{x}) = 0$ introduces a force into Newton's equations that needs to be included in finding the solution via $\vec{F}^{\text{const}} = \lambda \nabla \cdot A$, with λ a constant that needs to be calculated. So effectively we need to consider a sum of two forces to obtain the correct motion. As you can imagine, when adding more particles

and constraints (e.g. considering a rigid body), the bookkeeping of these constraints becomes increasingly tedious. Hence, there is motivation to find a version of Newton's laws that directly expresses these constraints such that we do not need to worry about those anymore when trying to find the solution.

This is precisely achievd by introducing generalised coordinates q_α (α labelling the new coordinates) for which the constraints become trivial. You all have seen how this works for the pendulum: You write everything in terms of the angle $x = \ell \sin\theta(t)$, $z = \ell\cos\theta(t)$, and of course

$$A(\vec{x}) = \ell^2 - \ell^2\cos^2\theta(t) - \ell^2\sin^2(t) = 0 \tag{2.1.1}$$

is trivially fulfilled at all times and you obtain a differential equation only for $\theta(t)$: We started with two variables and one constraint, leaving one degree of freedom.

In general the number of degrees of freedom f is given by the number of degrees of freedom of N point masses minus the number of constraints R: $f = 3N - R$. For the general case we simply follow the same strategy as for the pendulum, i.e. we re-write $\vec{x}_i(t) = \vec{x}_i(\{q_\nu(t)\}, t)$ where we have already generalised to a number of point masses labelled by i. With this we have

$$\dot{\vec{x}}_i = \frac{\mathrm{d}\vec{x}_i}{\mathrm{d}t} = \frac{\partial\vec{x}_i}{\partial t} + \sum_{\alpha=1}^{f}\frac{\partial\vec{x}_i}{\partial q_\alpha}\frac{\mathrm{d}q_\alpha}{\mathrm{d}t} = \frac{\partial\vec{x}_i}{\partial t} + \sum_{\alpha=1}^{f}\frac{\partial\vec{x}_i}{\partial q_\alpha}\dot{q}_\alpha\,, \tag{2.1.2}$$

which leads directly to

$$\frac{\partial\dot{\vec{x}}_i}{\partial\dot{q}_\alpha} = \frac{\partial\vec{x}_i}{\partial q_\alpha}\,. \tag{2.1.3}$$

With this we can re-write Newton's force law $m_i\ddot{\vec{x}}_i = \sum F_i + \sum F_i^{\mathrm{const}}$:

$$\sum_{i=1}^{N} m_i\ddot{\vec{x}}_i\frac{\partial\vec{x}_i}{\partial q_\alpha} - \sum_{i=1}^{N}\vec{F}_i\frac{\partial\vec{x}_i}{\partial q_\alpha} - \sum_{i=1}^{N}\sum_{\nu=1}^{R}\lambda_\nu\frac{\partial}{\partial\vec{x}_i}A_\nu(\{\vec{x}_k\},t)\frac{\partial\vec{x}_i}{\partial q_\alpha} = 0\,. \tag{2.1.4}$$

We limit ourselves to conservative forces for which we can find a potential V: $\vec{F} = -\nabla V$ and obtain

$$\sum_{i=1}^{N}\vec{F}_i\frac{\partial\vec{x}_i}{\partial q_\alpha} = -\sum_{i=1}^{N}\frac{\partial}{\partial\vec{x}_i}V(\{\vec{x}_k\},t)\frac{\partial\vec{x}_i}{\partial q_\alpha} = -\frac{\partial V}{\partial q_\alpha}\,. \tag{2.1.5}$$

The term related to the constraints vanishes

$$\sum_{i=1}^{N}\sum_{\nu=1}^{R}\lambda_\nu\frac{\partial A_\nu}{\partial\vec{x}_i}(\{\vec{x}_k\},t)\frac{\partial\vec{x}_i}{\partial q_\alpha} = \sum_{\nu=1}^{R}\lambda_\nu\frac{\partial A_\nu}{\partial q_\alpha} = 0 \tag{2.1.6}$$

as expected (since $A_\nu(\{q_\alpha\},t) = 0$).

Using Eq. (2.1.3), we can write the first term as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial\dot{q}_\alpha}\sum_{i=1}^{N}\frac{1}{2}m_i\dot{\vec{x}}_i^2\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{N}m_i\dot{\vec{x}}_i\frac{\partial\dot{\vec{x}}_i}{\partial\dot{q}_\alpha}\right) = \sum_{i=1}^{N}m_i\ddot{\vec{x}}_i\frac{\partial\vec{x}_i}{\partial q_\alpha} + m_i\dot{\vec{x}}_i\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\vec{x}_i}{\partial q_\alpha} \tag{2.1.7}$$

$$= \sum_{i=1}^{N}m_i\ddot{\vec{x}}_i\frac{\partial\vec{x}_i}{\partial q_\alpha} + \frac{\partial}{\partial q_\alpha}\sum_{i=1}^{N}\frac{1}{2}m_i\dot{\vec{x}}_i^2 \tag{2.1.8}$$

where we have used

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \vec{x}_i}{\partial q_\alpha} = \sum_{\beta=1}^{f} \frac{\partial^2 \vec{x}_i}{\partial q_\beta \partial q_\alpha}\frac{\mathrm{d}q_\beta}{\mathrm{d}t} + \frac{\partial^2 \vec{x}_i}{\partial t \partial q_\alpha} = \frac{\partial}{\partial q_\alpha}\left(\sum_{\beta=1}^{f}\frac{\partial \vec{x}_i}{\partial q_\beta}\dot{q}_\beta + \frac{\partial \vec{x}_i}{\partial t}\right) = \frac{\partial}{\partial q_\alpha}\frac{\mathrm{d}\vec{x}_i}{\mathrm{d}t}\,, \tag{2.1.9}$$

i.e. we are allowed to swap the order of the partial derivatives for sufficiently well-behaved functions (which we always assume in physics).

This finally yields with the total kinetic energy $T := \sum_{i=1}^{N} m_i \dot{\vec{x}}_i^2/2$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}_\alpha}T - \frac{\partial}{\partial q_\alpha}(T - V) = 0 \tag{2.1.10}$$

and since V is conservative it makes sense to introduce the *Lagrange function*

$$\mathcal{L}(\{q_\alpha, \dot{q}_\alpha\}) = T(\{q_\alpha, \dot{q}_\alpha\}) - V(\{q_\alpha\})\,. \tag{2.1.11}$$

and the "new" equations of motion are (the *Lagrange equation*)

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0, \quad \alpha = 1, \ldots, f\,. \tag{2.1.12}$$

What have we gained? Of course, there is not more information here than there is in Newton's axioms, but it is more distilled. Newton's laws are vector equations, whereas the Lagrange equation is derived from a single scalar quantity. The main difference is that we are able to simplify the equations of motion by injecting more information, i.e. we choose a "good" coordinate system right from the start. This reduces the complexity of finding the differential equations. Finding their solutions can still be complicated.

The main advantage of Eq. (2.1.12) is a more abstract and hence more general picture that emerges from the Lagrange equation. Just by counting the degrees of freedom of a system, we can see that we cannot simplify the physics any further; alternative approaches can be mapped onto Eq. (2.1.12) (such as the Hamilton formalism, which is the Legendre-transformation if Eq. (2.1.12), you will learn more about this in your Dynamics course). In this sense the "true" trajectory $\{q_\alpha(t)\}$ is fully encoded in the Lagrange equation. We can show this explicitly by defining the action functional (we'll drop the indices in the following, they carry through straightforwardly)

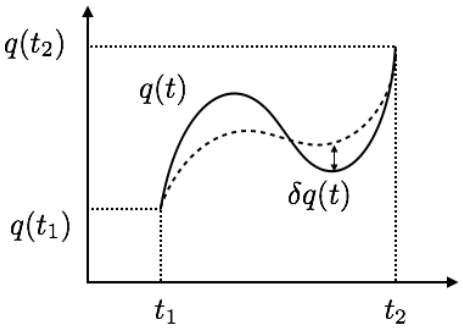


Figure 1: Illustration for the principle of least action.

$$S[q] = \int_{t_1}^{t_2} \mathrm{d}t\, \mathcal{L}(q, \dot{q}, t)\,, \tag{2.1.13}$$

which assigns a scalar number to every function $q(t)$ (not necessarily the one that is picked out by Eq. (2.1.12)).

What is special about the solution of Eq. (2.1.12) in this context? To answer this question we study how the functional responds to a variation of the function $q(t) \to q(t) + \delta q(t)$, for functions

that connect given start and endpoints $q(t_{1,2})$, i.e. the variation has to vanish in these points $\delta q(t_1) = \delta q(t_2) = 0$ (Fig. 1). This gives

$$\begin{aligned} S[q+\delta q] &= \int_{t_1}^{t_2} \mathrm{d}t\, \mathcal{L}(q+\delta q, \dot{q}+\delta\dot{q}, t) \simeq \int_{t_1}^{t_2} \mathrm{d}t\, \left(\mathcal{L}(q,\dot{q},t) + \frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta\dot{q} \right) \\ &= \int_{t_1}^{t_2} \mathrm{d}t\, \left(\mathcal{L}(q,\dot{q},t) + \frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{\mathrm{d}}{\mathrm{d}t}\delta q \right) \\ &= \int_{t_1}^{t_2} \mathrm{d}t\, \left(\mathcal{L}(q,\dot{q},t) + \left[\frac{\partial \mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} \right] \delta q \right) + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta q \Big|_{t_1}^{t_2} . \end{aligned} \tag{2.1.14}$$

The surface term vanishes vanishes by construction and we have

$$S[q+\delta q] = S[q] + \int_{t_1}^{t_2} \mathrm{d}t\, \left[\frac{\partial \mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} \right] \delta q + \mathcal{O}(\delta q^2) \,. \tag{2.1.15}$$

This means that the correct trajectory is picked out for a stationary action; this is the *principle of least action*

$$\frac{\delta S}{\delta q} = 0 \,. \tag{2.1.16}$$

This brings us closer to what symmetries mean for the mathematical description of a physical system: Given a solution of Eq. (2.1.16), a continuous symmetry transformation characterised by a parameter s (e.g. the angle of a 2d rotation) will map an allowed trajectory

$$h_s : q \longmapsto h_s(q) \,. \tag{2.1.17}$$

The image of the symmetry transformed path $h_s(q)$ still needs to be a physically allowed path that has to obey Eq. (2.1.16)! Any such function h_s will induce another map that operates on the velocities

$$\hat{h}_s : \dot{q} \longmapsto \frac{\partial h_s}{\partial q}\dot{q} \,. \tag{2.1.18}$$

The only way this can be implemented is if $(h_s, \hat{h}_s)$ do not change the action except for a surface term, which does not modify the Lagrange equation, see Eq. (2.1.14). At the integrand level of the action this means that the Lagrange functions need to be related by a total time derivative

$$\mathcal{L}(h_s(q), \hat{h}_s(\dot{q})) = \mathcal{L}(q,\dot{q}) + \frac{\mathrm{d}}{\mathrm{d}t} F_s(q,\dot{q}) \tag{2.1.19}$$

where F_s is an arbitrary function which can still depend on the symmetry parameter s.

2.1.2 Noether's Theorem

The spectacular consequence first derived by Emmy Noether in the 1910s is that an invariance of type Eq. (2.1.19) leads to a conserved quantity. This is *Noether's Theorem*.

To see this we investigate all allowed paths in the sense of the Lagrange equations connected by symmetry transformations h_s. To this end we define

$$Q(s,t) = h_s(q(t))\,. \tag{2.1.20}$$

We can rewrite Eq. (2.1.19)

$$\mathcal{L}(q,\dot{q}) = \mathcal{L}\left(Q(s,t), \frac{\partial}{\partial t}Q(s,t)\right) - \frac{\mathrm{d}}{\mathrm{d}t}F_s\,. \tag{2.1.21}$$

The left hand side does not know anything about the s-dependence of the right hand side and we have

$$0 = \frac{\partial}{\partial s}\left(\mathcal{L}\left(Q(s,t), \frac{\partial}{\partial t}Q(s,t)\right) - \frac{\mathrm{d}}{\mathrm{d}t}F_s\right) = \frac{\partial\mathcal{L}(q,\dot{q})}{\partial q}\frac{\partial Q}{\partial s} + \frac{\partial\mathcal{L}(q,\dot{q})}{\partial\dot{q}}\frac{\partial^2 Q}{\partial t\partial s} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F_s}{\partial s}\,, \tag{2.1.22}$$

which becomes with the equations of motion Eq. (2.1.12)

$$0 = \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}(q,\dot{q})}{\partial\dot{q}}\right)\frac{\partial Q}{\partial s} + \frac{\partial\mathcal{L}(q,\dot{q})}{\partial\dot{q}}\frac{\partial^2 Q}{\partial t\partial s} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F_s}{\partial s} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}(q,\dot{q})}{\partial\dot{q}}\frac{\partial Q}{\partial s} - \frac{\partial F_s}{\partial s}\right)\,, \tag{2.1.23}$$

and we have identified a conserved quantity

$$\mathcal{Q} = \frac{\partial\mathcal{L}(q,\dot{q})}{\partial\dot{q}}\frac{\partial Q}{\partial s} - \frac{\partial F_s}{\partial s}\,. \tag{2.1.24}$$

There are a few important remarks:

- We have used the equations of motion to arrive at this result. This means that $\mathcal{Q}$ is constant for any solution of the Lagrange equation and allows us to relate different trajectories straightforwardly with physically observable consequences. This is known as *mechanical similarity*, see exercise 1.
- The conserved quantity $\mathcal{Q}$ *is entirely determined by the infinitesimal symmetry transformations*, i.e. only $\partial Q/\partial s$ enters. If a coordinate is *cyclic*

  $$\frac{\partial\mathcal{L}}{\partial q} = 0 \tag{2.1.25}$$

  we can identify the conserved quantity directly from the Lagrange equation

  $$\mathcal{Q} = \frac{\partial\mathcal{L}}{\partial\dot{q}}\,. \tag{2.1.26}$$

Let us look at the example of a particle in a time-independent potential described by $\mathcal{L} = m\dot{\vec{x}}^2/2 - V(\vec{x})$ such that $V(\vec{x})$ is invariant under shifts $\vec{x} \to \vec{x} + s\vec{a}$ for a vector $\vec{a}$. We check the symmetry transformation

$$h_s : \vec{x} \longmapsto \vec{x} + s\vec{a}\,, \tag{2.1.27}$$

for which

$$\hat{h}_s(\dot{\vec{x}}) = 1 \tag{2.1.28}$$

and therefore

$$\mathcal{L}(h_s(\vec{x}), \hat{h}_s(\dot{\vec{x}})) = \frac{m}{2}\dot{\vec{x}}^2 - V(\vec{x}) = \mathcal{L}(\vec{x}, \dot{\vec{x}}) \tag{2.1.29}$$

and $\mathcal{L}$ is invariant, i.e. $F_s \equiv \text{const}$. By computing

$$\frac{\partial Q}{\partial s} = \frac{\mathrm{d}h_s(\vec{x})}{\mathrm{d}s} = \vec{a} \tag{2.1.30}$$

we obtain the conserved quantity

$$\mathcal{Q} = m\dot{\vec{x}} \cdot \vec{a}\,, \tag{2.1.31}$$

which is just the projection of the momentum onto the $\vec{a}$-direction. This is the proof of the second point of Eq. (1.0.1) that translation invariance leads to momentum conservation. I leave energy and angular momentum conservation as exercises 2 and 3.

These symmetries exhibit important algebraic structures, which will eventually help us to classify them: they form continuous groups. The concepts of groups, rings and fields have been introduced in Mathematical Methods 2, but I list the basic axioms in the appendix, which can act as a reference for the remainder of this course. For, e.g., the shift symmetries of Eq. (2.1.27), checking the group structure is almost trivial.

- If h_s and $h_{s'}$ are symmetry transformations, so is $h_s \circ h_{s'} : \vec{x} \to \vec{x} + (s + s')\vec{a} = h_{s+s'}(\vec{x})$ and the algebraic structure $(\{h_s\}, \circ)$ is well-defined,
- since $(\mathbb{R}^3, +)$ is associative, $(\{h_s\}, \circ)$ is associative,
- $h_0 = \mathbf{1}$ is the neutral element of $(\{h_s\}, \circ)$,
- h_{-s} is the inverse element of every h_s,
- the group is abelian.

Obviously, the group structure for the shift symmetry in space and time are inherited from the properties of the real numbers, which becomes clear in the above proof. While these symmetries follow directly from continuous $\mathbb{R}$ symmetries, angular momentum is slightly more complex and deserves a bit more attention (see below).

We can summarise our discussion of Noether's theorem from a physics angle. All "traditional" conserved quantities such as energy, momentum and angular momentum originate from symmetries in the underlying formalism. While we would find these conservation laws also by solving Newton's laws, we have developed a formalism to find the equations of motion of a mechanical system that also allows us to access and derive potential conservation laws directly from the outset. The price we pay for this conceptual clarity is an increased level of abstraction. But this is a small price to pay for what we gain. We can now forget Newton's laws, which are empirical, ad-hoc, and too clumsy in their engineer-like angle to mechanics. The principle of least action tells us that finding the solution to a mechanical simply amounts to finding the function that renders the action functional extremal. Continuous invariances of these solutions have direct physical consequences, e.g. we can look for solutions of fixed energy, angular momentum, etc. and by expressing the solution of the differential equations in terms of these quantities, we can find *all* solutions. Even more, the principle of least action is one way of seeing how classical mechanics arises as limit of quantum mechanics using the path integral formalism - ask your (relativistic) quantum mechanics lecturer about this!

We have seen that conserved quantities follow from symmetries of a Lagrange function. We can turn this argument around and elevate symmetries to the central building principle of the dynamics of a system. In the particular case of classical mechanics, these symmetries are the Galilean transformations:

$$t' = t + a \tag{2.1.32}$$

$$\vec{x}' = \mathcal{M}\vec{x} + \vec{b} \tag{2.1.33}$$

where $\mathcal{M}$ is a matrix we'll investigate shortly, and $\vec{b}$ is an arbitrary but constant vector. These transformations leave the line element

$$\mathrm{d}s^2 = \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{2.1.34}$$

invariant. δ_{ij} is the metric which is a rank (0,2) tensor. The three dimensional Euclidean space together with the metric (or scalar product in this case) is a canonical example of a Riemannian manifold; velocities are tangent space vectors, accelerations are cotangent space vectors. The symmetry of the manifold is therefore reflected in the equations of motion; Newton's axioms are form-invariant under these transformations that connect different inertial frames.

2.1.3 Rotational Symmetries

Rotations are isometries of 3 dimensional Euclidean space, i.e. the leave angles between vectors and their norm invariant. If we consider an infinitesimal change of a vector $\vec{x}$ which is represented by its coordinates x^i under the Galilean transformations

$$\mathrm{d}x^{i\prime} = \mathcal{M}^i{}_j\mathrm{d}x^j \tag{2.1.35}$$

The invariance of the line element therefore means

$$\mathrm{d}s^2 = \delta_{ij}\mathrm{d}x^{i\prime}\mathrm{d}x^{j\prime} = \delta_{ij}\mathcal{M}^i{}_k\mathcal{M}^j{}_s\mathrm{d}x^k\mathrm{d}x^s = \delta_{ks}\mathrm{d}x^k\mathrm{d}x^s \tag{2.1.36}$$

or simply

$$\mathcal{M}^T\mathcal{M} = \mathcal{M}\mathcal{M}^T = \mathbf{1}\,, \tag{2.1.37}$$

i.e. the matrices are orthogonal and have $\det\mathcal{M} = \pm 1$. Usually we are not interested in discrete transformations $x^i \to -x^i$ and keep the orientation when discussion rotations and therefore limit ourselves to positive determinants. Also, as we have seen in the previous section, the conserved quantities follow from the infinitesimal symmetry transformations, close to the identity. These conserved quantities will therefore follow from matrices with $\det\mathcal{M} = +1$, the special orthogonal transformations These matrices together with matrix multiplication form the continuous group $SO(3)$

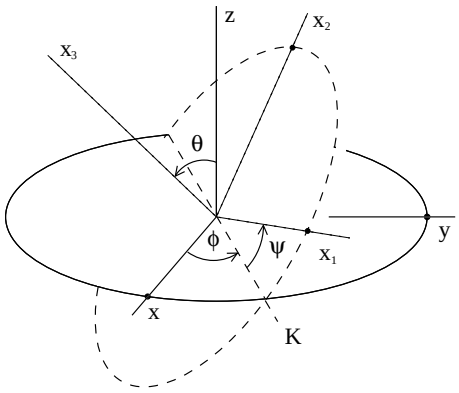


Figure 2: Euler angles for a general rotation $x, y, z \to x_1, x_2, x_3$.

$$SO(3) \simeq \{\mathcal{M} \in \mathbb{R}^{3\times 3} | \mathcal{M}^T\mathcal{M} = \mathbf{1}, \det\mathcal{M} = 1\}\,. \tag{2.1.38}$$

Interesting in the following will also be their complex counterpart, the special unitary 2×2 matrices $SU(2)$

$$SU(2) \simeq \{\mathcal{M} \in \mathbb{C}^{2\times 2} | \mathcal{M}^\dagger\mathcal{M} = \mathbf{1}, \det\mathcal{M} = 1\} \tag{2.1.39}$$

and you are asked to verify the group properties in exercise 4.

We can therefore think of each rotation as an abstract element of a group, which is represented by a particular matrix, i.e. by an element of the general group of linear transformations on a vector space or module of given dimension n. This point of view is the core of representation theory, which we will investigate later.

Any rotation can be built from the three basic rotations around the x, y, z axes using the Euler angles of Fig. 2

$$\mathcal{M}(\phi, \psi, \theta) = \mathcal{M}_z(\theta)\mathcal{M}_y(\psi)\mathcal{M}_x(\phi)\,. \tag{2.1.40}$$

This shows that the group of rotations in 3d Euclidean space depends on three continuous parameters.

The rotations in 3d Euclidean space around the x, y, z axes are given by

$$\mathcal{M}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad \mathcal{M}_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \quad \mathcal{M}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.1.41}$$

Together with the group structure of shift symmetries, the Galilean transformations form a group — the Galilean group.

2.1.4 Electrodynamics and Special Relativity

In the previous section we have seen that the symmetries largely determine the dynamics of a classical system and concluded that the conserved quantities we observe in classical physics are a consequence of the Galilean group symmetry.

That this is not the end of all wisdom became clear when Einstein tried to make meaning of the electromagnetism of moving objects in 1905, which led to the introduction of special relativity theory. A way of seeing how special relativity is built into Maxwell's equations is by looking for their no source-solutions

$$\begin{aligned} \nabla \cdot \vec{E} &= 0 \\ \nabla \cdot \vec{B} &= 0 \\ \nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \nabla \times \vec{B} &= \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}\,. \end{aligned} \tag{2.1.42}$$

With a little bit of vector analysis, which you have worked through in Electromagnetic Theory one can find the solutions to, e.g., the electric field to be a wave equation

$$\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0 \tag{2.1.43}$$

and since $\mu_0 \varepsilon_0 c^2 = 1$, we see the speed of light entering the equation as a characteristic speed of a solution of Maxwell's equation. This violates Galilean invariance as velocities add up when

transforming between frames, but Maxwell's equations pick out a unique speed in *all* frames for the classical interpretation of light, the wave solutions of Eq. (2.1.42).

If we want to hold on to the idea of symmetry dictating the dynamics, we are force to give up the Galilean transformations. We can do this by exploiting a constant speed of light in all frames that should be connected by these new transformations. This is motivation enough to set $c = 1$ (we work with natural units). We then have to ask ourselves whether we can reproduce the successful theory of classical mechanics in some limit of the new theory.

For a propagating light wave we have

$$\vec{1} = \frac{\Delta\vec{x}}{\Delta t} \implies 0 = (\Delta t)^2 - (\Delta\vec{x})^2 \,. \tag{2.1.44}$$

In general we can therefore consider the line element

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^{\,2} = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \,, \tag{2.1.45}$$

which is equal to zero for light-like solutions, and look for all transformations

$$(t, x) \mapsto \Lambda(t, x) + \text{const.} \tag{2.1.46}$$

that leave the line element invariant. We have collected the time and space coordinates into a contravariant four vector $x^\mu = (t, \vec{x})$ and introduced the (pseudo-euclidean) metric

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{2.1.47}$$

We can use the metric to arrive at a covariant four vector $x_\mu = g_{\mu\nu}x^\nu$ (there are a few exercises to understand the difference between covariant and contravariant vectors better, exercises 5, 6) The transformations Eq. (2.1.46) are called *Poincaré transformations* (the special case where we choose const. $= 0$ are the *Lorentz transformations*).

The invariance of the line element already gives rise to physical consequences, namely that space and time can transform into each other, i.e. we can find a reference frame $\mathrm{d}s'^2 = \mathrm{d}\tau^2$ and connect it to another one with $\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^{\,2}$ and since $\mathrm{d}s'^2 = \mathrm{d}s^2$ we have

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 \left(1 - \frac{\mathrm{d}\vec{x}^{\,2}}{\mathrm{d}t^2}\right) = \mathrm{d}t^2 \left(1 - \vec{v}^{\,2}\right) \,, \tag{2.1.48}$$

i.e. clocks in $(t, \vec{x} = \vec{v}t)$ show time intervals

$$\mathrm{d}t = \frac{\mathrm{d}\tau}{\sqrt{1 - \vec{v}^{\,2}}} = \gamma \mathrm{d}\tau \geq \mathrm{d}\tau \tag{2.1.49}$$

i.e. time dilation is a consequence of the invariance of the line element. For speeds much smaller than 1 (remember we are working in natural units) we have $\gamma = 1 + \mathcal{O}(\vec{v}^{\,2})$ and we recover the Galilean result as a limit.

Let us look at the general coordinate transformations that leave the line element invariant $x^\mu \mapsto x'^\mu = \Lambda^\mu{}_\nu x^\nu + b^\mu$, where b^μ is constant four vector to derive some properties of the Lorentz and Poincaré transformations (see exercises 7, 8 and 9). The invariance of the line element becomes

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = g_{\delta\epsilon}\mathrm{d}x'^\delta \mathrm{d}x'^\epsilon = g_{\delta\epsilon}\Lambda^\delta{}_\mu \Lambda^\epsilon{}_\nu \mathrm{d}x^\mu \mathrm{d}x^\nu \,. \tag{2.1.50}$$

So we have

$$g_{\mu\nu} = g_{\delta\epsilon}\Lambda^\delta{}_\mu \Lambda^\epsilon{}_\nu \quad \text{or} \quad g = \Lambda^T g \Lambda \tag{2.1.51}$$

This allows us to derive some properties of Λ:

(i) $\det(\Lambda) = \pm 1$,

(ii) $1 = g_{00} = \Lambda^\mu{}_0 \Lambda^\nu{}_0 g_{\mu\nu} = (\Lambda^0{}_0)^2 - (\Lambda^i{}_0)^2$.

A negative determinant means that we would again allow discrete symmetries $x^i \mapsto -x^i$, i.e. parity transformations, or $t \mapsto -t$, i.e. time inversions. Since this is not a continuous operation we exclude those and only consider positive determinants (similar to our discussion of rotations in three dimensions). The second condition can be interpreted as

$$(\Lambda^0{}_0)^2 = 1 + (\Lambda^i{}_0)^2 \geq 1 \,, \tag{2.1.52}$$

from which we choose $\Lambda^0{}_0 \geq 1$ i.e. we only allow transformations that are similar to Eq. (2.1.49) and preserve the direction of time (they are *orthochroneous*).

It is not surprising that the Poincaré transformations forms a group with these restrictions and the multiplication law

$$(\Lambda_2, b_2) \times (\Lambda_1, b_1) = (\Lambda_2\Lambda_1, b_2 + \Lambda_2 b_1) \,, \tag{2.1.53}$$

from which follows that the $b^\mu = 0$ transformations form a subgroup, the group of the special orthochroneous Lorentz transformations $SO^\uparrow_+(1,3)$ ("+" and "↑" correspond to the properties (i) and (ii) respectively) which is the connected part the group space that contains $\mathbf{1}$ and is therefore important for conserved quantities. The $(1,3)$ indicates that we deal with a metric which has signature $(+1,-1,-1,-1)$. $SO(3)$ is a subgroup of $SO^\uparrow_+(1,3)$ and implements the ordinary three-dimensional rotations.

Similar to the Euler angles, it can be shown that any Λ can be decomposed

$$\Lambda = R\Lambda_x(\psi)\tilde{R} \tag{2.1.54}$$

with

$$R = \begin{pmatrix} 1 & 0 \\ 0 & \mathcal{M} \end{pmatrix} \,, \quad \mathcal{M} \in SO(3) \tag{2.1.55}$$

and similar for $\tilde{R}$, and

$$\Lambda_x(\psi) = \begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \,. \tag{2.1.56}$$

$\Lambda_x(\psi)$ is called a *Lorentz boost* and $\psi \in \mathbb{R}$ is the *rapidity*

$$\cosh\psi = \gamma = \frac{1}{\sqrt{1-v^2}} \tag{2.1.57}$$

(see exercises 10 and 11). Alternatively, a general Lorentz transformation can be represented as a combination of a general boost and a general rotation $\bar{R}$

$$\Lambda = \Lambda_{\vec{n}}(\psi)\bar{R} \tag{2.1.58}$$

which shows that the Lorentz group has six continuous parameters: 3 boosts and 3 rotation angles.

Here we already see a qualitative difference between the Lorentz transformation and ordinary rotations. While the Euler angles are defined on a compact interval, the rapidity is not. We will see that this has important consequences when we try to build a quantum theory that respects Lorentz symmetry.

We can also look at momenta and energies to arrive at important kinematical implications of the invariance of the line element. Defining a four velocity by differentiating with respect to eigenzeit τ to make the transformation properties of the vector well defined, we have

$$v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\,. \tag{2.1.59}$$

We can define a four momentum by multiplying the rest mass[1]

$$p^\mu = m_0 v^\mu = \left(m_0\gamma, m_0\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau}\right) = (m_0\gamma, m_0\gamma\vec{v})\ . \tag{2.1.60}$$

An immediate implication is that the contracted version (which by construction is invariant similar to the line element under Eq. (2.1.46))

$$p^\mu p_\mu = m_0^2\gamma^2 - m_0^2\gamma^2\vec{v}^2 = m_0^2\,. \tag{2.1.61}$$

If we look at the spacial components of the four momentum we have

$$p^i = m_0\gamma v^i = m_0 v^i + \mathcal{O}(\vec{v}^2) \tag{2.1.62}$$

for velocities smaller than 1 in units of the speed of light. This is just the non-relativistic three momentum. Performing the same Taylor expansion with

$$p^0 = m_0 + \frac{m_0}{2}\vec{v}^2 + \mathcal{O}(\vec{v}^4) \tag{2.1.63}$$

we see that p^0 in the non-relativistic limit is just a constant in addition to the non-relativistic kinetic energy. This means that we have to understand p^0 as the relativistic energy and p^i as the relativistic three momentum. Then we have

$$p^\mu p_\mu = E^2 - \vec{p}^2 = m_0^2 \quad\Longleftrightarrow\quad E^2 = m_0^2 + \vec{p}^2 \tag{2.1.64}$$

which is the dispersion relation for a relativistic. One implication of special relativity is then the famous Einstein relation $E = m_0$ for particles at rest (in natural units).

[1] We will see that this quantity is well-defined in special relativity later.

2.2 Symmetries in Quantum Mechanics

2.2.1 Continuous Symmetries: Degeneracies of States

So far we have limited ourselves to classical physics but symmetries have important consequence for quantum mechanics as well. We start with Schrödinger's equation, the differential equation that describes the time evolution of a non-relativistic quantum system of states reads

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{x},t) = \hat{H}\psi(\vec{x},t)\,, \tag{2.2.1}$$

where $\psi(x) = \langle x|\psi\rangle$ describes the state a system $|\psi(t)\rangle$ in position space and $\hat{H}$ is the Hamilton operator, again in position space. In the simple case of a particle in a time-independent potential $V(\vec{x})$ we have

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\hat{\vec{x}})\,, \tag{2.2.2}$$

and the time evolution of a state is given by (now in natural units $\hbar = c = 1$)

$$\psi(\vec{x},t) = \exp\{-iH(x)t\}\psi(\vec{x},0)\,. \tag{2.2.3}$$

An observable $O(t)$ (which is represented by a unitary operator) when measured on the quantum state yields

$$O(t) = \langle\psi(t)|\hat{O}|\psi(t)\rangle = \langle\psi(0)|e^{i\hat{H}t}\hat{O}e^{-i\hat{H}t}|\psi(0)\rangle = \langle\psi(0)|\hat{O}_H(t)|\psi(0)\rangle \tag{2.2.4}$$

where we have introduced the Heisenberg picture which describes a basis transformation after which the operators carry the time dependence and states are time-independent. We can now study the time dependence of an operator in this basis straightforwardly (exercise 12)

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{O}_H(t) = \frac{\partial}{\partial t}\hat{O}_H(t) + i[\hat{H},\hat{O}_H(t)]\,. \tag{2.2.5}$$

In the Heisenberg picture we can directly sandwich this equation to study the time dependence of observables (this leads to a generalisation of Ehrenfest's theorem). If there is no explicit dependence on time, a conserved quantity in quantum mechanics is therefore characterised by

$$O(t) = \text{const.} \quad\Longleftrightarrow\quad [\hat{H},\hat{O}_H(t)] = 0\,, \tag{2.2.6}$$

i.e. the Hamilton operator commutes with this particular operator. If we have a symmetry in the system, such as a rotational symmetry in a $1/|\vec{x}|$ Coulomb potential of the hydrogen atom, angular momentum conservation therefore implies

$$[\hat{H},\hat{\vec{L}}] = \hat{H}\hat{\vec{L}} - \hat{\vec{L}}\hat{H} = 0 \tag{2.2.7}$$

so for the eigenfunctions of the Hamilton operator (the states characterised by a fixed energy)

$$\hat{H}\phi_n(\vec{x},t) = E_n\phi_n(\vec{x},t) \tag{2.2.8}$$

we also have

$$0 = [\hat{H},\hat{L}]\phi_n(\vec{x},t) = \hat{H}\hat{L}\phi_n(\vec{x},t) - \hat{L}\hat{H}\phi_n(\vec{x},t) = \hat{H}(\hat{L}\phi_n(\vec{x},t)) - E_n(\hat{\vec{L}}\phi_n(\vec{x},t))\,. \tag{2.2.9}$$

This means that $\hat{\vec{L}}\phi_n(\vec{x},t)$ is also an eigenstate of the Hamilton operator with identical eigenvalue E_n. We can label the degeneracy of energy levels due to the symmetry of the Hamiltonian with the values of the angular momentum. For instance, the non-relativistic hydrogen atom is therefore characterised by two additional quantum numbers in addition to energy: total (quantised) angular momentum and its z-component:

$$\psi_{nlm}(r,\theta,\phi) \sim R_n(r)Y_l^m(\theta,\phi), \text{ with } \hat{H}\psi_{nlm} = -\frac{1\text{Ry}}{n^2}\psi_{nlm}\,, \tag{2.2.10}$$

for discrete n, l, m, where Y_l^m are the spherical harmonics, the solutions of the angular part of the Laplacian operator. Every energy level n, hence, is degenerate by n^2.

If the rotational symmetry is broken, e.g. by applying an external magnetic or electric field, the degeneracy is lifted and the energy values are split according to their angular momentum quantum numbers (the Zeeman and Stark effects).

In general, commutation relations with the Hamilton operator point us towards characterising solutions with specific energies according to the global symmetries a particular physics problem entails.

2.2.2 Electrodynamics and Gauge Symmetries

All of the symmetries we have discussed in this section so far are cases of "external" and global symmetries, i.e. symmetries that tell us that it does not matter how we look at a system (rotational invariance, angular momentum conservation) or when we observe it (time translational invariance, energy conservation). But there can be "internal" symmetries too, i.e. when the degrees freedom are transformed amongst each other. A prominent example of this is the "eightfold" way of particle physics, which postulates that up, down and strange quarks are states belonging to the same symmetry multiplet governed by a $SU(3)$ symmetry:

$$\begin{pmatrix}|u\rangle\\|d\rangle\\|s\rangle\end{pmatrix} \longmapsto \mathcal{M}\begin{pmatrix}|u\rangle\\|d\rangle\\|s\rangle\end{pmatrix}, \quad \mathcal{M}\in SU(3)\,. \tag{2.2.11}$$

Gell-Mann (and independently Zweig) saw this as a solution to make a meaning of the zoology of particles discovered in the early 1970's. Postulating that these particles are built from (back then) hypothetical objects that transform under an approximate $SU(3)$ as in Eq. (2.2.11), with the additional assumption that the dynamics that binds them to composite objects respects this symmetry, the composite spectrum will inherit an approximate $SU(3)$ symmetry pattern as well.[2] This was successfully confirmed by investigating the mass spectrum and finding particles that were predicted by the symmetry pattern.

Later this decade we came to realise that quarks are indeed the fundamental building blocks of strongly interacting matter and that their interaction is mediated by a "gauge symmetry" which is $SU(3)$ (which has nothing to do with Gell-Mann's $SU(3)$). To understand the basic principle of a

[2]The quantum mechanical Hamiltonian will commute with the symmetry operations and the energy eigenstates, which are connected by $E = mc^2$ can be categorised according to their $SU(3)$ quantum number just like the hydrogen atom solutions are labelled with angular momentum quantum numbers.

gauge symmetry concept it is enough to combine electrodynamics with quantum mechanics, which we sketch in the following.

In electrodynamics electric and magnetic fields are manifestations of the same phenomenon, connected by the underlying Poincaré symmetry of the theory. Take Maxwell's equations

$$\begin{aligned} \nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \cdot \vec{B} &= 0 \\ \nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \nabla \times \vec{B} &= \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \end{aligned} \tag{2.2.12}$$

on the entire vector space $\mathbb{R}^3$. In this case the second equation tells us that we can write $\vec{B} = \nabla \times \vec{A}$ since $\nabla \cdot (\nabla \times \vec{A}) \equiv 0$. The third equation now becomes (assuming a sufficiently well-behaved vector field $\vec{A}$)

$$\nabla \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0 \tag{2.2.13}$$

and again we can write

$$\vec{E} + \frac{\partial \vec{A}}{\partial t} = -\nabla \phi \tag{2.2.14}$$

for a scalar potential ϕ since $\nabla \times (\nabla \phi) \equiv 0$. Effectively we have re-expressed electric field and magnetic flux density

$$\begin{aligned} \vec{E} &= -\nabla \phi - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \nabla \times \vec{A} \end{aligned} \tag{2.2.15}$$

but we can also identify a symmetry in these definitions. If $\vec{A} \mapsto \vec{A} + \nabla \alpha$ we have $\vec{B} \mapsto \vec{B}$ for $\alpha(\vec{x}, t)$ double-continuously differentiable. The electric field is invariant if $\phi \mapsto \phi - \partial \alpha / \partial t$. This function α is not related to a symmetry as in Eq. (1.0.1) but describes a "redundancy" in the definition of the electromagnetic potentials.

This freedom is termed a "gauge transformation". Once we combine electrodynamics with quantum mechanics via "minimal substitution"[3] of the momentum operator with a static background potential

$$\hat{\vec{p}} \to \hat{\vec{\Pi}} = \hat{\vec{p}} - q\vec{A}\,, \tag{2.2.16}$$

(the dimensions of q do not matter at this stage, but it is merely a coupling strength of the particle described by a wave function to the background field $\vec{A}$) a gauge transformation changes

$$\hat{\vec{\Pi}} \mapsto \hat{\vec{\Pi}} - q\nabla\alpha\,. \tag{2.2.17}$$

A static background field does not change the normalisation of a state, it can only change the quantisation of the energy levels (as encountered in the Stark and Zeeman effects). Hence, if

[3] It can be shown that this substitution describes the motion of a particle in a magnetic field.

physics, i.e. expectation values of a quantum mechanical measurements, are unaffected by this gauge symmetry transformation we can only allow for a phase change of the wave function. Trying an ansatz (keeping in mind $\hat{\vec{p}} = -i\hbar\,\partial/\partial\vec{x}$)

$$\hat{\vec{p}}\left(e^{iq\alpha(x,t)/\hbar}|\vec{k}\rangle\right) = \left(q\nabla\alpha + \hbar\vec{k}\right)\left(e^{iq\alpha(x,t)/\hbar}|\vec{k}\rangle\right) \tag{2.2.18}$$

we see that if

$$\varphi_{\vec{k}}(\vec{x},t) \mapsto e^{iq\alpha(x,t)/\hbar}\varphi_{\vec{k}}(\vec{x},t) \tag{2.2.19}$$

under a gauge transformation, we have indeed

$$\hat{\vec{\Pi}}\varphi_{\vec{k}}(\vec{x},t) \mapsto (\hat{\vec{\Pi}} - q\nabla\alpha)e^{iq\alpha(x,t)/\hbar}\varphi_{\vec{k}}(\vec{x},t) = \hat{\vec{\Pi}}\varphi_{\vec{k}}(\vec{x},t)\,. \tag{2.2.20}$$

Note that the new phase is not a constant but depends on space and time, i.e. in principle we can choose it differently at any point in space or time! It is not a global symmetry (this would be the case if $\alpha = \text{const.}$), but it is called a local symmetry. Since phases are not observable (they drop out of calculated expectation values which itself is a symmetry leading to a conservation law, exercise 13), this gauge invariance has no obvious physical implications, but expresses a freedom in defining electromagnetism. If you attend lectures on quantum field theory, you will learn that it is exactly this freedom that guarantees a vanishing photon mass.

This phase is also another example of a "representation" of a symmetry group. Since it is periodic, it is the one-dimensional unitary group rather than $\mathbb{R}$ itself. The phase transformations define a circle with radius one in the complex plane. Each point on this circle is a different group element, and again we deal with a continuous group since $\alpha \in \mathbb{R}$. $U(1)$ is the simplest case of a gauge symmetry, but taking the example of electrodynamics, we can generalise this concept to different continuous symmetry groups. But already $U(1)$ makes the concept of transforming internal degrees of freedom clear: real and imaginary parts of the wave function are transformed amongst each other in a space- and time-dependent way.

2.2.3 Discrete Symmetries

So far we have only dealt with continuous symmetries as those are directly motivated from the symmetry group of classical mechanics, which is defined on a continuous space. Everything else (i.e. special relativity and the concept of gauge invariance) evolved from tracing this concept through to electromagnetism. In developing the continuous symmetry concept with the aim to apply Noether's theorem, we've focused on parts of the group space that is connect to the identity. Other group transformations, such as space reflections were the first examples of discrete symmetry representations, i.e. transformations $\vec{x} \to -\vec{x}$ form a group with only two elements, which is isomorphic to $\mathbb{Z}_2 = \{+1, -1\}$.

A more abstract, yet still familiar discrete group is the group of permutations of a discrete finite set, the *symmetric group*. Writing $M = \{1, 2, \ldots, m\}$ we can represent a permutation $\pi : M \to M$, by

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & m \\ \pi(1) & \pi(2) & \ldots & \pi(m) \end{pmatrix} \tag{2.2.21}$$

where π is an automorphism on M. Using induction, one can show that there $m!$ such automorphisms if the domain M has m elements and one refers to them as S_m. For instance, for $m = 3$

we have S_3 :

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$
$$\pi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \pi_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

It can be shown that these permutations obey the group axioms and form another example of a discrete group (which is also very important for the classification of continuous group representations as we will find out).

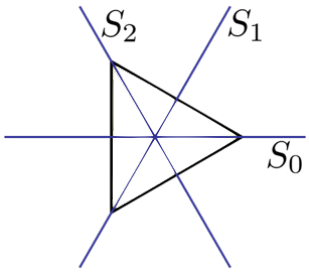


Figure 3: Reflexion symmetry axes of an equilateral triangle

If we relax the continuum point of view we are quickly drawn to more complicated discrete symmetries and richer finite groups. These can be motivated from studying the symmetry transformations of simplices. For example, an equilateral triangle is invariant under rotations about $2\pi/3$ and $4\pi/3$, as well as under reflexions about the centroid axes, see Fig. 3.

Cayley tables are a powerful tool to investigate the structure of discrete groups (exercise 14). If we denote $r_0 = \mathbf{1}$, and $r_{1,2}$ the rotations of the equilateral triangle by $2\pi/3$ and $4\pi/3$, respectively, and throw in the reflexions about the axes in Fig. 3, $s_{0,1,2}$ we can deduce the following group structure

	r_0	r_1	r_2	s_0	s_1	s_2
r_0	r_0	r_1	r_2	s_0	s_1	s_2
r_1	r_1	r_2	r_0	s_1	s_2	s_0
r_2	r_2	r_0	r_1	s_2	s_0	s_1
s_0	s_0	s_2	s_1	r_0	r_2	r_1
s_1	s_1	s_0	s_2	r_1	r_0	r_2
s_2	s_2	s_1	s_0	r_2	r_1	r_0

(2.2.22)

i.e. a rotation about $2\pi/3$ followed by another rotation about $2\pi/3$ equals a rotation about $4\pi/3$: $r_1 r_1 = r_2$. By knowing the Cayley table of a certain operation we can directly deduce a lot of information. For instance, every combination of symmetry operations leads to another symmetry operation which is again part of $\{r_i, s_i | i = 0, 1, 2\}$; the algebraic structure induced by successive symmetry operations is well-defined. Even more: *Cayley's theorem* states that any finite group is isomorphic to a subgroup of the symmetric group S_m. This means that every row or column of a group has to be a permutation of the base elements. This is indeed the case for the symmetries of the equilateral triangle. In particular, we have r_0 appearing just once, which means there exists a unique inverse element for every element representing a discrete symmetry transformation; for all elements r_0 acts (trivially) as the identity. Indeed the symmetry operations form a group, which, however, is not abelian as the Cayley table is not symmetric with respect to the diagonal. This group is called the *dihedral group* D_3.

D_3 is an example of a *point group*, i.e. isometries (or symmetries) $\vec{x} \mapsto \mathcal{U}\vec{x}$ that leave at least one point invariant. They exist in Euclidean spaces of any dimension but particularly relevant for, e.g., condensed matter physics is the three dimensional case, where there are 32 distinct families of point groups. Point groups, being non-translative by definition can be extended by including translational symmetries with a given bounded fundamental domain. This adds operations like

The 32 Point Groups and Their Symbols				
System	Schoenflies symbol	Hermann-Mauguin symbol		Examples
		Full	Abbreviated	
Triclinic	C_1	1	1	
	$C_i, (S_2)$	$\bar{1}$	$\bar{1}$	Al_2SiO_5
Monoclinic	$C_{2v}, (C_{1h}), (S_1)$	m	m	KNO_2
	C_2	2	2	
	C_{2h}	$2/m$	$2/m$	
Orthorhombic	C_{2v}	$2mm$	mm	
	$D_2, (V)$	222	222	
	$D_{2h}, (V_h)$	$2/m\ 2/m\ 2/m$	mmm	I, Ga
Tetragonal	S_4	$\bar{4}$	$\bar{4}$	
	C_4	4	4	
	C_{4h}	$4/m$	$4/m$	$CaWO_4$
	$D_{2d}, (V_d)$	$\bar{4}2m$	$\bar{4}2m$	
	C_{4v}	$4mm$	$4mm$	
	D_4	422	42	
	D_{4h}	$4/m\ 2/m\ 2/m$	$4/mmm$	TiO_2, In, $\beta - Sn$
Rhombohedral	C_3	3	3	AsI_3
	$C_{3i}, (S_6)$	$\bar{3}$	$\bar{3}$	$FeTiO_3$
	C_{3v}	$3m$	$3m$	
	D_3	32	32	Se
	D_{3d}	$\bar{3}2/m$	$\bar{3}m$	Bi, As, Sb, Al_2O_3
Hexagonal	$C_{3h}, (S_3)$	$\bar{6}$	$\bar{6}$	
	C_6	6	6	
	C_{6h}	$6/m$	$6/m$	
	D_{3h}	$\bar{6}2m$	$\bar{6}2m$	
	C_{6v}	$6mm$	$6mm$	ZnO, NiAs
	D_6	622	62	CeF_3
	D_{6h}	$6/m\ 2/m\ 2/m$	$6/mmm$	Mg, Zn, graphite
Cubic	T	23	23	$NaClO_3$
	T_h	$2/m\bar{3}$	$m3$	FeS_2
	T_d	$\bar{4}3m$	$\bar{4}3m$	ZnS
	O	432	43	β-Mn
	O_h	$4/m\ \bar{3}\ 2/m$	$m3m$	NaCl, diamond, Cu

Figure 4: The 32 point groups and their symbols in Schoenflies and Hermann-Mauguin conventions, and some examples of condensed matter systems that exhibit a specific symmetry (taken from Dresselhaus, "*Applications of Group Theory to the Physics of Solids*". In Schoenflies convention C_n stands for n-fold rotation axis, the addition of h, v mean the addition of a mirror plane perpendicular to the axis of rotation. S_{2n} refer to a $2n$-fold rotation axis. $C_n i$ has only a rotoinversion axis. D_n is dihedral (as already encountered in our example) and represents an n-fold rotation axis and n two-fold axes perpendicular to the rotation axis. The addition of h denotes an additional horizontal mirror, v represents n vertical mirror planes. T is the tetrahedral group and has three 2-fold and four 3-fold axes, d, h means addition of a diagonal or horizontal mirror planes. O, I refer to the octahedral and icosahedral groups, with h meaning again the inclusion of horizontal mirror planes.

glide reflections to the the discrete isometries of euclidean space, and there 230 such *space groups* in three dimensions. In solid state physics the 32 point groups these are communicated by using the Schoenflies or Hermann-Maugin conventions, Fig. 4.

From your solid state physics lectures you know that crystalographic patterns (accessed with e.g. the Debye-Scherrer method) is completely determined by the short-range symmetry encountered in a particular crystal.[4] Such a symmetry is reflected in the Hamiltonian that describes a the quantum behavior of a crystal:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\hat{\vec{x}})\,, \tag{2.2.23}$$

where the potential is periodic, reflecting a discrete symmetry:

$$\hat{V}(\hat{\vec{x}}) = \hat{V}(\hat{\vec{x}} + \vec{R})\,, \tag{2.2.24}$$

and where R is an infinite set of vectors defined from a *Bravais lattice*

$$\vec{R} = \sum_{i=1}^{3} n_i \vec{a}_i,\ n_i \in \mathbb{Z}\,, \tag{2.2.25}$$

with *primitive vectors* $\vec{a}_{1,2,3}$. Note that $\hat{V}$ can summarise a plethora of effects, the periodicity, however, remains as a crucial property of a solid state with short range order (like a crystal).

We can introduce the translation operator

$$\hat{T}_{\vec{R}} = \exp\left(\frac{i}{\hbar}\vec{R}\cdot\hat{\vec{p}}\right)\,. \tag{2.2.26}$$

In exercise 15 you can show that this indeed acts as a translation by a vector $\vec{R}$. With this result we find straightforwardly

$$\forall \vec{R}_1, \vec{R}_2 :\ \hat{T}_{\vec{R}_1}\hat{T}_{\vec{R}_2} = \hat{T}_{\vec{R}_1+\vec{R}_2}\,, \quad [T_{\vec{R}_1}, T_{\vec{R}_2}] = 0 \tag{2.2.27}$$

which tells us that $\hat{T}_{\vec{R}}$ are forming an abelian group. Solutions with momentum $\hbar\vec{k}$, denoted $|\vec{k}\rangle$ (but not necessarily plane waves!), that describe the system will transform,

$$\hat{T}_{\vec{R}}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{R}}|\vec{k}\rangle\,, \tag{2.2.28}$$

i.e. the wave function is only dressed with a phase under the translation symmetry, which does not enter physical observables. With Eq. (2.2.28) we can find the matrix elements of the translation operator

$$\langle\vec{k}'|\hat{T}_{\vec{R}}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{R}}\,\delta(\vec{k}-\vec{k}') \tag{2.2.29}$$

and therefore we have a wave function in position space

$$\psi_{\vec{k}}(\vec{x}+\vec{R}) = \langle\vec{x}|\hat{T}_{\vec{R}}|\vec{k}\rangle = \int \mathrm{d}^3\vec{k}'\langle\vec{x}|\vec{k}'\rangle\langle\vec{k}'|\hat{T}_{\vec{R}}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{R}}\psi_{\vec{k}}(\vec{x})\,. \tag{2.2.30}$$

[4] Similar techniques to the ones discussed below can also be applied to materials without short-ranged but long-ranged order like amorphous materials such as glasses.

This is the famous *Bloch-theorem*. While the details of the wave function depend on the specific form of the potential, which can be hard to compute even numerically, periodicity of the potential needs to be reflected by the Bloch condition.

We can ask the question when are two solutions $|\vec{k}\rangle, |\vec{k}'\rangle$ equivalent in the sense that we resolve degeneracies $\vec{k} \neq \vec{k}'$. This means we look for solutions $\psi_{\vec{k}}(\vec{x}) = \psi_{\vec{k}'}(\vec{x})$ for all $\vec{x}$. This implies

$$e^{i\vec{k}\cdot\vec{R}}\psi_{\vec{k}}(\vec{x}) = \psi_{\vec{k}}(\vec{x}+\vec{R}) = \psi_{\vec{k}'}(\vec{x}+\vec{R}) = e^{i\vec{k}'\cdot R}\psi_{\vec{k}'}(\vec{x}) = e^{i\vec{k}'\cdot R}\psi_{\vec{k}}(\vec{x}) \tag{2.2.31}$$

and therefore

$$(\vec{k}-\vec{k}')\cdot\vec{R} = 2\pi\mathbb{Z}\,. \tag{2.2.32}$$

Since the $\vec{R}$ vectors define a lattice, the solutions of $\vec{G}\cdot\vec{R} = 2\pi$ form again a lattice, *the reciprocal lattice*. Writing $\vec{G} = \sum_{i=1}^{3} m_i\vec{b}_i$ we can infer the lattice through the solutions of

$$\vec{b}_i \cdot \vec{a}_j = 2\pi\delta_{ij}\,. \tag{2.2.33}$$

Solutions are equivalent if they are momenta differ by a reciprocal lattice vector. This means we can characterise all solutions in a fundamental domain of a single representative of the symmetry pattern of the reciprocal lattice, giving rise to the notion of the *first Brioullin-zone*.

The meaning of "momentum" can therefore be different compared to ordinary plane wave solutions $\psi_{\vec{k}}(\vec{x}) \sim \exp(i\vec{k}\cdot\vec{x}/\hbar)$[5] and one typically speaks of *quasi-momentum* to highlight the differences. Performing an elastic scattering experiment of photons off a system described by a periodic potential therefore implies selection rules $\vec{k}-\vec{k}' = \vec{G}$, which is also known as the *Bragg-condition.* Static solutions of the Hamilton operator are therefore superpositions of solutions with momenta that differ by reciprocal lattice vectors and the physical significance of the shift-symmetry invariance is the conservation of the solutions' *quasi-momentum*.

The lattice is typically not static, but its dynamics are determined by the particular mechanism that creates a short-range order (such as the interactions of ions governed by the Coulomb force). Typically it is enough to investigating nearest neighbour-interactions to obtain a qualitative picture, which you have probably done in your solid state lectures. One finds that lattice excitations have a dispersion relation $E'(\vec{q})$ similar to free particles $E(\vec{k})$, and that this coherent phenomenon can therefore be interpreted as a quasi-particle (a *phonon*) with a given three momentum $\hbar\vec{q}$. Probing inelastic properties of a lattice, the Bragg-condition is modified by exciting quasi-particles

$$E(\vec{k}) - E(\vec{k}') = \pm E'(\vec{q})\,, \quad \vec{k}-\vec{k}' = \pm\vec{q}+\vec{G} \tag{2.2.34}$$

which can be used to investigate the dispersion relation of phonons experimentally.

In summary, starting from the point groups, we can find the corresponding Bravais lattice and for a specified basis of the lattice we can also deduce the phononic properties of a solid (at least numerically). Moving to the reciprocal lattice follows from a specified lattice via Eq. (2.2.33) and we can formulate predictions for the elastic and inelastic properties of photon (X-ray) scattering particular material which in turn allows us to infer the crystallography of a material experimentally. Similarly, the phonon and wave function solutions do impact key parameters, such as the specific heat capacity etc. and we can deduce these properties from the symmetry structure. This way many physical implications are determined by the specific symmetry pattern of a space group that describes an idealised solid state physics system.

[5]Note that the physical observable $\vec{p}$ is still the same for both systems.

2.3 *Exercises for Section 2*

Exercise 1

Consider transformations $\mathcal{L} \mapsto \kappa\mathcal{L}$ for some constant κ. How do the equations of motion change under such a transformation? Consider now transformations $\vec{x} \mapsto \alpha\vec{x}, t \mapsto \beta t$ for potentials which are homogeneous functions of degree k

$$V(\alpha\vec{x}) = \alpha^k V(\vec{x})\,.$$

Find $\beta = \beta(\alpha)$ such that $\mathcal{L} \mapsto \kappa\mathcal{L}$. Derive Kepler's 3rd law as well as independence of the period of the harmonic oscillator on the amplitude.

Exercise 2

Show that energy conservation follows from symmetry transformations $h_s : t \mapsto t + s$ for a system described by $\mathcal{L} = m\dot{q}^2/2 - V(q)$.

Exercise 3

Show that angular momentum conservation follows from symmetry transformations $h_s : \vec{x} \mapsto \vec{x} + s\vec{\omega} \times \vec{x}$ for a system described by $\mathcal{L} = m\dot{\vec{x}}^{\,2}/2 - V(|\vec{x}|)$.
No solution provided

Exercise 4

By checking the group axioms explicitly, verify that $SU(2)$ and $SO(3)$ form groups.

Exercise 5

In Euclidean space, we do not usually distinguish between covariant and contravariant indices, whereas in Minkowski space this is mandatory. Why?

Exercise 6

What is the inverse matrix g^{-1} of the metric g? Since g is used to lower a Lorentz index, then any tensor $T^{\mu\nu}$ would have both its indices lowered through

$$T_{\rho\sigma} = g_{\rho\mu}\, g_{\sigma\nu}\, T^{\mu\nu},$$

where the right-hand side can be written in matrix form as $g\,T\,g^T = g\,T\,g$, as g is symmetric. Hence, choosing $T_{\rho\sigma} = g_{\rho\sigma}$, find the matrix corresponding to $T^{\mu\nu}$, i.e. solve $g = gTg$ for T. This is clearly the definition of $g^{\mu\nu}$, which is used to raise an index.

Exercise 7

If $x^\mu = (t, x, y, z)$ is the position four-vector, the general Lorentz transformation matrix ${\Lambda^\mu}_\nu$ gives

$$x'^\mu = {\Lambda^\mu}_\nu\, x^\nu.$$

Take ${\Lambda^\mu}_\nu$ from the definition of a Lorentz transformation leading to Eq. (2.1.56). How do $\Lambda_{\mu\nu}$, ${\Lambda_\mu}^\nu$ and $\Lambda^{\mu\nu}$ differ from this? In particular, how is ${\Lambda_\mu}^\nu$ related to ${\Lambda^\mu}_\nu$?

Exercise 8

Show, by reconsidering the invariance of the line element that

$$x'_\mu = x_\nu {(\Lambda^{-1})^\nu}_\mu\ .$$

Show that $\partial_\mu \phi$ (for arbitrary *scalar* functions $\phi(x)$) does transform like x_μ and not x^μ (i.e. it is a contravariant vector). **Exercise 9**

Show that rapidities add for subsequent boosts along the same axis. How do velocities add and what happens in the limits $v \to 1$ as well as $v \to 0$? Can you explain these limits from symmetry point of view? **Exercise 10**

Show that lengths are contracted $L' = L\gamma$, where L is a known length scale at rest. **Exercise 11**

Show that in the Heisenberg picture **Exercise 12**

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{O}_H(t) = \frac{\partial}{\partial t}\hat{O}_H(t) + i[\hat{H}, \hat{O}_H(t)]\,.$$

Starting from the Schrödinger equation for the wave function $\psi(\vec{x}, t)$, show that the probability density $\rho = \psi^*\psi$ satisfies the continuity equation **Exercise 13**

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \ ,$$

where

$$\vec{J} = \frac{1}{2im}\left[\psi^*(\vec{\nabla}\psi) - (\vec{\nabla}\psi^*)\psi\right] \ .$$

What is the interpretation of $\vec{J}$?

Using the Cayley table for S_3, show explicitly that S_3 is a group and identify all subgroups. Is S_3 or any of the subgroups abelian? Can you see another discrete structure emerging from the S_3 Cayley table? **Exercise 14**

By studying the behavior of the position operator $\hat{\vec{x}}$, show that **Exercise 15**

$$\hat{T}_{\vec{R}} = \exp\left(\frac{i}{\hbar}\vec{R} \cdot \hat{\vec{p}}\right)$$

induces a translation by a vector $\vec{R}$ in operator space.

3 Lie Groups, Lie Algebras and their connection

Most of the symmetry concepts we have introduced so far remained vague. For instance how can we derive the number of states predicted by the eightfold way?

3.1 Definitions*

The two most famous symmetry groups, which you have already encountered in this course are $SO(3)$ and $SU(2)$. $SO(3)$ is the group of the *special orthogonal transformations in 3 dimensions*. $SU(2)$ is the group of *special unitary transformations in 2 dimensions*. Using homomorphisms they can be linearly represented by matrices. In this case we speak of a *(linear) representation* $\mathcal{R}$ *of a group* $(G,\cdot)$: $\mathcal{R}:\ G\to GL(V,n),\ G\ni g\to\mathcal{R}(g)\in GL(V,n)$ and

$$GL(V,n)\ni\mathcal{R}(g_1g_2)=\mathcal{R}(g_1)\mathcal{R}(g_2)\ \forall g_{1,2}\in G\,. \tag{3.1.1}$$

Two representations $\mathcal{R},\mathcal{R}'$ are called *equivalent*, if there is a transformation $\mathcal{R}'(g)=U^{-1}\mathcal{R}(g)U$ for all $g\in G$ (they are in the same *conjugacy class*). A representation is called *faithful* if it is injective: $\forall g_1,g_2\in G,\ g_1\neq g_2\Longrightarrow\mathcal{R}(g_1)\neq\mathcal{R}(g_2)$.

It is easy to verify that $SU(2)$ and $SO(3)$ are groups, we have seen this in an earlier exercise. But there is more to these groups. $SO(3)$ represent all rotations in three dimensions (i.e. all basis transformations in 3d that leave distances invariant) and is characterised by three angles. The same is true for $SU(2)$ in the complex sense, so both groups depend on continuous parameters. It can be shown that elements of $SU(2)$ can be written as (*the Cayley-Klein representation*):

$$\mathcal{M}=\begin{pmatrix}a & b\\ -b^\star & a^\star\end{pmatrix} \tag{3.1.2}$$

with $a,b\in\mathbb{C}$ and $aa^\star+bb^\star=1$, or expressed in real and imaginary parts

$$\Re\{a\}^2+\Im\{a\}^2+\Re\{b\}^2+\Im\{b\}^2=1\,. \tag{3.1.3}$$

This means that all elements of the group $SU(2)$ can be identified with a point on a three sphere S^3. On this sphere we can define differentiable paths (S^3 is a differentiable manifold) and also the space of group elements is compact (closed and bounded) in the language of topology. We can also contract any closed path continuously to a point: S^3 is simply connected (or null-homotopic). Groups like $SU(2)$ and $SO(3)$ which are also differentiable manifolds are called *Lie groups* G.

Particularly relevant in physics is not the Lie group but the Lie algebra. From our discussion of Noether's theorem you know that the elements of symmetry transformations that lead to conserved currents and charges are given by the infinitesimal ones. By studying the group elements close to the unity group element we will be able to introduce a new structure - the *Lie algebra* $\mathfrak{g}$.

Let's have a look again at $SU(2)$ and let's study elements $\mathcal{M}=\mathbf{1}+i\alpha t$ where α is a small parameter that measures the "distance" from the one-element and t is 2×2 complex matrix. We want to find out how these t matrices look like. Since $\mathcal{M}\mathcal{M}^\dagger=\mathbf{1}$ we have

$$\mathcal{M}\mathcal{M}^\dagger=(\mathbf{1}+i\alpha t)(\mathbf{1}-i\alpha t^\dagger)=\mathbf{1}+i\alpha(t-t^\dagger)+\mathcal{O}(\alpha^2)\overset{!}{=}\mathbf{1} \tag{3.1.4}$$

so this means that the t matrices are hermitian, $t = t^\dagger$. To use the determinant condition we need take a detour via the eigenvalues λ_i (we assume a diagonal form of $\mathcal{M}$; the equation we'll derive is valid for Lie groups)

$$\det \mathcal{M} = \prod_i \lambda_i = \prod_i \exp\{\log \lambda_i\} = \exp\left\{\sum_i \log \lambda_i\right\} = \exp\left\{\mathrm{Tr}(\log \mathcal{M})\right\} . \tag{3.1.5}$$

Using $\mathcal{M} = \mathbf{1} + i\alpha t$ we can Taylor-expand the $\log(\mathbf{1} + i\alpha t) = i\alpha t$ and we have

$$1 = \exp\left\{i\alpha \mathrm{Tr}(t)\right\} \tag{3.1.6}$$

so the t matrices have vanishing trace.

In summary, we deal with hermitian and traceless matrices in $\mathbb{C}^{2\times 2}$. These form a three-dimensional (unitary) vector space, but using matrix multiplication, this vector space becomes and algebra (we will drop the underlined notation for the algebra elements for convenience in the following). It also a special matrix algebra, a linear representation of the so-called the *Lie algebra* $\mathfrak{su}(2)$.[6] The set of matrices t^a, $a = 1, 2, 3$ are connected to the Pauli matrices

$$t^a = \frac{\sigma^a}{2} \quad \text{with } \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \ \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \ \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{3.1.7}$$

The algebra is special because on top of the matrix multiplication we have

$$[t^a, t^b] = t^a t^b - t^b t^a = i f^{ab}{}_c t^c \quad (f^{ab}{}_c = \varepsilon^{ab}{}_c \text{ for } \mathfrak{su}(2)) \tag{3.1.8}$$

where $\varepsilon^{ab}{}_c$ is the Levi-Civita tensor in 3d. The t^a are called the *generators* of the Lie algebra. You can show that quantum mechanical angular momentum $\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$ obeys this algebra in exercise 16. We can find an inner product via the trace

$$\mathrm{Tr}(t^a t^b) = T_r \delta^{ab} , \tag{3.1.9}$$

i.e. the $\{t^a\}$ form an orthonormal bases with respect to this scalar product. The scalar T_r will depend on the particular representation, for the fundamental representation we have $T_r = T_F = 1/2$. This is true for all the Lie algebras we study in this course.[7]

For a general Lie algebra the Levi-Civita tensor will be replaced by general totally anti-symmetric if^{abc}, which are called the *structure constants of the Lie group*. Furthermore, we also have a *Jacobi identity*: for all $x, y, z \in \mathfrak{su}(2)$

$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 \tag{3.1.10}$$

and

$$[x, x] = 0 , \text{ and } [x, y] = -[y, x] \tag{3.1.11}$$

[6] We will not make a difference between complex and real Lie algebras, unless stated otherwise.

[7] A note for completeness: For the Lie algebras we consider in this course, all symmetric bilinear forms are scalar multiples of the Killing form $\kappa^{ab} = \frac{1}{2}\mathrm{Tr}(\mathrm{ad}_{t^a} \circ \mathrm{ad}_{t^b})$, which defines a proper inner product in our cases. This was shown by Cartan.

For our construction bilinearity of $[.,.]$ was automatic, but say we start with a vector space and introduce an operation $[.,.]$ for which

$$[ax + by, z] = a[x, y] + b[y, z] \tag{3.1.12}$$

in addition to Eqs. (3.1.8)-(3.1.11) then we call such an algebraic object a Lie algebra.

So we've seen that by studying the infinitesimal elements of a Lie group, we obtain a Lie algebra. This raises the question whether we can invert this statement: Is there a way to express the Lie group elements in terms of Lie algebra ones? The answer is yes with a twist (without proof):

Theorem The image of

$$\exp : \begin{cases} \mathfrak{g} \to G \\ t \mapsto \exp t = \sum_0^\infty t^n/n! \end{cases} \tag{3.1.13}$$

is dense in G for the (matrix) Lie groups we discuss in this course. This means that almost all element of the Lie group can be expressed by the exponential map from the Lie algebra (we used this step to for the determinant condition earlier).

From your previous exposure to the Pauli Matrices you probably also now that we have an operation $\{.,.\}$ such that

$$\{\sigma^a, \sigma^b\} = \sigma^a\sigma^b + \sigma^b\sigma^a = 2\delta^{ab}\mathbf{1}\,. \tag{3.1.14}$$

An associative algebra (in this case $V = \text{span}\,[\mathbf{1}, \sigma^1, \sigma^2, \sigma^3]$) endowed with an operation $\{.,.\}$ like this is called a *Clifford algebra* (more later).

In exercise 16 you can show that angular momentum in quantum mechanics obeys an algebra that is similar to $\mathfrak{su}(2)$. Although the structure constants are identical, the algebra is not exactly $\mathfrak{su}(2)$ (which we have defined over $\mathbb{C}$) because rotations refer to $SO(3)$, i.e. they are defined over $\mathbb{R}$. Therefore, even though the structure constants are identical, the Lie algebra $\mathfrak{so}(3)$ is intrinsically defined over $\mathbb{R}$! In physics we typically do not worry about these things as we implicitly assume that we do quantum mechanics on a complex Hilbert space. As we will see shortly, this ignorance involves some pitfalls.

If we look at the complexified version of $\mathfrak{so}(3)$, we indeed have $\mathbb{C} \times \mathfrak{so}(3) = \mathfrak{su}(2)$, i.e. we can redefine the generators using complex scalar prefactors, which takes us away from $\mathbb{R}$. We say that $\mathfrak{so}(3)$ is a real form of $\mathfrak{su}(2)$. Depending on how we choose the generators and whether we allow complex coefficients, we can have a series of real lie algebras. We will not make a distinction here in the following, but understand this implicitly, i.e. we take $\mathfrak{su}(2) = \mathfrak{so}(3) = \mathfrak{sl}(2, \mathbb{C})$ etc. but will point out issues when they arise.

3.2 (Ir)Reducible Representations*

In the introduction we motivated, using quantum mechanics as an example, that we can categorise the quantum mechanical energy eigenstates according to their symmetry quantum numbers. If we have a symmetry of the Hamiltonian (i.e. the Hamiltonian commutes with the symmetry operators) the energy levels are degenerate but the time evolution given by the Hamiltonian will only mix those states characterised by the set of states that describes a specific energy level.

How do we count the states that characterise such a symmetry multiplet? Since the symmetry is expressed by the algebra, we have to find the different dimensionalities of the vector spaces on which the representation of the algebra acts (the module). In the previous section we encountered a 2-dimensional complex representation of $\mathfrak{su}(2)$ (called the *fundamental representation*). But one can show straightforwardly that there is another representation (exercise 17), called the *adjoint representation* which can also be written in a general fashion as

$$\forall x, y \in \mathfrak{g} : \mathrm{ad}_x(y) = [x, y] \,. \tag{3.2.1}$$

This representation considers linear maps on the algebra itself rather than the module. Therefore it always exists, which together with a smart choice of basis will allow the complete classification of finite dimensional Lie Algebras relevant for physics.

The adjoint representation therefore has dimension three in case of $\mathfrak{su}(2)$. In general, if we have a Lie algebra like $\mathfrak{su}(n)$ we can directly formulate a $n^2 - 1$-dimensional adjoint representation in addition to the fundamental representation of dimension n.

Let us continue building representations for the example of $\mathfrak{su}(2)$, keeping in mind that $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are related by complex extension of the real vector space. Consider the object (which is formally not(!) part of the Lie algebra)

$$t_a t^a = (t^1)^2 + (t^2)^2 + (t^3)^2 \,. \tag{3.2.2}$$

As part of the exercises we can show that for all $b = 1, 2, 3$ (and for general $\mathfrak{su}(n)$)

$$[t_a t^a, t^b] = 0 \,, \tag{3.2.3}$$

i.e. $t_a t^a$ commutes with all generators (exercise 18).

Such an object is called a *Casimir operator* of the algebra. $t^a t^a$, although not part of the algebra, still defines a linear map that intertwines the same representation. Schur's Lemma (unfortunately we do not have time to discuss this in detail) therefore guarantees that $t_a t^a$ is proportional to the unit matrix and we can use the value of $t_a t^a$ calculated on a particular module as a characteristic quantity[8] of the particular representation r

$$t_{r\,a} t_r^a = C_r \mathbf{1}_r \,, \tag{3.2.4}$$

see exercises 19 and 20. In case of $\mathfrak{su}(2)$, we can also label a state of the multiplet as an eigenvalue of one of the t^i, and conventionally one chooses t^3.

In $\mathfrak{su}(2)$ we can introduce raising and lowering operators

$$t^\pm = t^1 \pm i t^2 \tag{3.2.5}$$

and

$$t^0 = 2t^3 \,. \tag{3.2.6}$$

The elements of $\mathfrak{su}(2)$ after this transformation now obey the commutation rule

$$[t^0, t^\pm] = \pm 2 t^\pm \,, \quad [t^+, t^-] = t^0 \,, \tag{3.2.7}$$

[8] Making contact with what you already know from quantum mechanics, the Casimir corresponds to the total angular momentum $\vec{L}^2 = L_1^2 + L_2^2 + L_3^2$.

and we have also

$$(t^+)^\dagger = t^- \,, \tag{3.2.8}$$

i.e. the generators are not hermitian anymore.

In this basis, we have

$$t_a t^a = \frac{1}{4}(t^0)^2 + \frac{1}{2}(t^+t^- + t^-t^+) \,. \tag{3.2.9}$$

If $\underline{v}_\lambda$ is an eigenvector of t^0, $t^0\underline{v}_\lambda = \lambda\underline{v}_\lambda$ we find

$$t^0(t^\pm\underline{v}_\lambda) = (\lambda \pm 2)(t^\pm\underline{v}_\lambda) \,. \tag{3.2.10}$$

This means that $t^\pm\underline{v}_\lambda$ is an eigenvector $\underline{v}_{\lambda\pm2}$, i.e. raises or lowers the eigenvectors.

This paves the way to *irreducible representations* – minimal spaces on which the symmetry transformation acts. The fundamental and adjoint representations we have encountered so far are examples of irreducible representations. "Irreducible" means that we cannot rearrange the module to break the representation into smaller subspaces. If we take an n-dimensional irreducible representation V^n, this means we cannot decompose $V^n = V_1^r \oplus V_2^s$ with $s + r = n$. If a representation is not irreducible it is *reducible*. In practice we can connect reducible n-dimensional representations $\mathcal{R}(g)$ with irreducible ones by a basis transformation expressed through a matrix $\mathcal{U}$ such that

$$\mathcal{R}(g) \longmapsto \mathcal{R}'(g) = \mathcal{U}^{-1}\mathcal{R}\mathcal{U} = \begin{pmatrix} \mathcal{R}_1(g) & 0 & & \cdots & 0 \\ 0 & \mathcal{R}_2(g) & 0 & \cdots & 0 \\ 0 & 0 & \mathcal{R}_3(g) & \cdots & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \mathcal{R}_n(g) \end{pmatrix} . \tag{3.2.11}$$

with $\sum_i \dim \mathcal{R}_i(g) = n$. This means by a suitable choice of basis the representation takes a block-diagonal form and we can identify the block matrices $\mathcal{R}_i(g)$ as the irreducible representations that the original one can be decomposed into.

By definition an irreducible representation is finite-dimensional, so there has to be a *highest weight* Λ of the representation. By applying the lowering operator to the vector with highest weight we have to obtain zero at some point. So, let $\underline{v}_{\Lambda-2N}$ be the lowest non-vanishing vector

$$t^-\underline{v}_{\Lambda-2N} = \underline{o} \tag{3.2.12}$$

Let's consider t^+ now. Using that $t^+\underline{v}_\Lambda = \underline{o}$ we have

$$t^+\underline{v}_{\Lambda-2} = t^+t^-\underline{v}_\Lambda = [t^+, t^-]\underline{v}_\Lambda = t^0\underline{v}_\Lambda = \Lambda\underline{v}_\Lambda \tag{3.2.13}$$

i.e. $t^+\underline{v}_{\Lambda-2}$ has to be a scalar multiple of $\underline{v}_\Lambda$. By induction, we can show this for all other $\lambda \neq \Lambda$. Indeed, we have

$$t^+\underline{v}_{\Lambda-2n} = t^+t^-\underline{v}_{\Lambda-2n+2} = (t^-t^+ + t^0)\underline{v}_{\Lambda-2n+2} = (r_{n-1} + \Lambda - 2n + 2)\underline{v}_{\Lambda-2n+2} \tag{3.2.14}$$

where we have introduced $t^+\underline{v}_{\Lambda-2n} = r_n\underline{v}_{\Lambda-2n+2}$. This implies that there is a recursion relation

$$r_n = r_{n-1} + \Lambda - 2n + 2 \tag{3.2.15}$$

and if $t^+\underline{v}_\Lambda = \underline{o}$ (since Λ is the highest eigenvalue) we have $r_0 = 0$. Using this as initial condition we can show that

$$r_n = n(\Lambda - n + 1)\,. \tag{3.2.16}$$

Using that $t^-\underline{v}_{\Lambda-2N}$ vanishes we find

$$\underline{o} = t^+t^-\underline{v}_{\Lambda-2N} = (t^-t^+ + t^0)\underline{v}_{\Lambda-2N} = (r_N + \Lambda - 2N)\underline{v}_{\Lambda-2N}\,. \tag{3.2.17}$$

Substituting $r_N = N(\Lambda - N + 1)$ we obtain a quadratic equation $N^2 + (1-\Lambda)N - \Lambda = 0$ and since $N > 0$ we have $N = \Lambda$. The highest weight has to be a non-negative integer.

We can also read off the dimension of the module of the representation with highest weight Λ

$$\dim r_\Lambda = N + 1 = \Lambda + 1\,. \tag{3.2.18}$$

and value of the Casimir operator is

$$t_a t^a \underline{v}_\Lambda = \frac{1}{4}\Lambda(\Lambda+2)\underline{v}_\Lambda\,. \tag{3.2.19}$$

In physical applications we usually consider eigenvalues of $j_3 = \lambda/2$ of $L^3 = t^0/2$. The highest weight is then called *spin* $j = \Lambda/2$ and the dimension of the irreducible representation is $\dim r_j = 2j + 1$ and value of the Casimir $j(j+1)$.

3.3 The relation of $SO(3)$ and $SU(2)$

Already from the relation between the algebras $\mathbb{C} \times \mathfrak{so}(3) \simeq \mathfrak{su}(2)$ it becomes clear that the special orthogonal groups and the special unitary groups are related. From the last section (and your quantum mechanics lectures) you know that $j = 1/2$ is an allowed spin value. Spin half systems are modules of $SU(2)$ but not of $SO(3)$: they corresponds to the fundamental module of $SU(2)$ but the fundamental module of $SO(3)$ is three-dimensional (vectors in $\mathbb{R}^3$). So we seem to miss something here.

Let us analyse things at the group level and construct a homomorphism $SO(3) \to SU(2)$. Let $U \in SU(2)$. Let $\vec{x}$ be a component vector in $\mathbb{R}^3$, we can construct

$$\underline{x} = x_i\sigma^i \tag{3.3.1}$$

(note that although we contract the index, we understand this as identifying a vector with a hermitian and traceless matrix). from which we see that $\det \underline{x} = -|\vec{x}|^2$. Now, for $\mathcal{U} \in SU(2)$, we can define

$$\underline{x}_\mathcal{U} = \mathcal{U}\underline{x}\mathcal{U}^\dagger\,. \tag{3.3.2}$$

Since $\det(\mathcal{U}_1\mathcal{U}_2) = \det(\mathcal{U}_1)\det(\mathcal{U}_2)$ we have

$$|\vec{x}_\mathcal{U}|^2 = -\det x_\mathcal{U} = -\det x = |\vec{x}|^2 \tag{3.3.3}$$

But obviously $\vec{x}_\mathcal{U} \neq \vec{x}$ unless $\mathcal{U} = \mathbf{1}$ but $\underline{x}_\mathcal{U}$ still lies in the vector space defined by Eq. (3.3.1). So there has to be a $\Lambda \in SO(3)$: $\vec{x}_\mathcal{U} = \Lambda(\mathcal{U})\vec{x}$ (see exercise 21). In total we therefore have constructed

a prescription that maps elements of $SU(2)$ onto elements of $SO(3)$. But since this construction is bilinear in $\mathcal{U} \in SU(2)$ we have

$$\Lambda(\mathcal{U}) = \Lambda(-\mathcal{U}) \tag{3.3.4}$$

so in general we cannot invert this linear map to obtain a homomorphism $SU(2) \to SO(3)$. $SU(2)$ "double covers" $SO(3)$, and to highlight this we say that $SO(3)$ *is locally isomorphic* to $SU(2)$:

$$SO(3) \simeq SU(2)/\mathbb{Z}_2\,. \tag{3.3.5}$$

We have seen earlier that $SU(2)$ is compact and simply connected, and now we can see that $SU(2)$ covers $SO(3)$ in the sense of Eq. (3.3.5). This combination of properties is unique and $SU(2)$ is the *universal covering group (or spin group) of* $SO(3)$.

Cutting some mathematical corners we can say that actually we are not interested in linear representations but representations up to a phase if we are doing quantum mechanics. Such a phase will drop out from the physical observable expectation values. So in principle we can allow representations for group elements $x \in G$ which are of the form

$$\mathcal{R}(x) = e^{i\alpha(x)}\mathcal{U}(x)\,. \tag{3.3.6}$$

which is called a *projective representation*. In order to have a group we must then have

$$\mathcal{U}(x)\mathcal{U}(y) = e^{i\phi(x,y)}\mathcal{U}(x,y)\,. \tag{3.3.7}$$

The phase is usually an unwanted feature (see Eq. (3.1.1)) and we normally want to get rid of it. Hence, we need to assure that we are allowed to consistently choose $\phi(x,y) = 0$ for all $x, y \in G$. This is possible if G is simply connected (without proof), so ϕ can be removed for $SU(2)$. Now there is a theorem in representation theory that states that all projective representations of a group can be lifted to an ordinary representation of an extended group, which in our case is the universal covering group. This tells us that indeed we have to look for representations of $SU(2)$ to look for all representations of $SO(3)$ in quantum mechanics, even if the spin 1/2 representations are not transparent from $SO(3)$. We will run into a similar problem when we will discuss the representation theory of the Lorentz and Poincaré groups.

3.4 Constructing irreducible representations and modules*

Since we have found the minimal building block of $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$, we can construct irreducible representations by "adding spins". Consider the product of two $\Lambda = 1$ ($j = 1/2$) modules (or multiplets in the language of physics). Since this module is 2 dimensional, we introduce $|1, \pm 1\rangle$ as a shorthand notation (meaning $|\{\text{highest weight, } t^0 \text{ eigenvalue}\}\rangle$) inspired by quantum mechanics. The tensor product vector space is given by

$$|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2 \tag{3.4.1}$$

i.e. it is four dimensional. We define the generators of the composite system

$$t^a = (t_1^a \otimes \mathbf{1}_2) \oplus (\mathbf{1}_1 \otimes t_2^a) \tag{3.4.2}$$

and in particular

$$\begin{aligned} t^0(|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2) &= \left[(t_1^0 \otimes \mathbf{1}_2) \oplus (\mathbf{1}_1 \otimes t_2^0)\right](|1,\pm 1\rangle_1 \otimes |1,\pm 1\rangle_2) \\ &= \left[(t_1^0 \otimes \mathbf{1}_2)(|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2)\right] \oplus \left[(\mathbf{1}_1 \otimes t_2^0)(|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2)\right] \\ &= \lambda_1(|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2) + \lambda_2(|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2) \\ &= (\lambda_1+\lambda_2)(|1,\lambda\rangle_1 \otimes |1,\lambda\rangle_2) \end{aligned} \tag{3.4.3}$$

so we have (at least buried in the tensored module) a module with $\Lambda = 2$. This module, however, is only three-dimensional, so we miss one degree of freedom. This is exactly the $\Lambda = 0$ module on which the $\Lambda = 0$ representation acts.

The tensored module is an example of a reducible representation. In fact we can rearrange the basis of the tensor module in a simple way to make the decomposition in terms of reducible modules explicit. We can define raising and lowering operators $t^{\pm} = t^1 \pm it^2$ for the tensor generators and given a unique

$$|2,2\rangle = |1,1\rangle_1 \otimes |1,1\rangle_2 \tag{3.4.4}$$

we obtain the basis transformation

$$\begin{aligned} |2,0\rangle =& \frac{1}{\sqrt{2}}\left[(|1,1\rangle_1 \otimes |1,-1\rangle_2) \oplus (|1,-1\rangle_1 \otimes |1,1\rangle_2)\right] \\ |2,-2\rangle =& |1,-1\rangle_1 \otimes |1,-1\rangle_2 \,. \end{aligned} \tag{3.4.5}$$

We can get the remaining basis element by looking for an orthogonal combination in the tensor product, which is

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left[(|1,1\rangle_1 \otimes |1,-1\rangle_2) \ominus (|1,-1\rangle_1 \otimes |1,1\rangle_2)\right] . \tag{3.4.6}$$

There is an important thing to observe here: The irreducible representations emerge from how we can combine the two modules in terms of permutations. There is the symmetric combination (leading to $\Lambda = 2$) and the anti-symmetric combination (leading to $\Lambda = 0$). If we had known that before, we could have written down the module right away, and the representation would follow straightforwardly. That the classification of the tensor product in terms of the permutation group leads to irreducible representations is true for general $SU(n)$ and $SO(n)$ groups. And finding the right representations boils down to finding a representation of the permutation group on the tensor module. This will lead to a powerful tool called *Young Tableaux*, which we will discuss later.

But for now let's continue with constructing irreducible $SU(2)$ representations and modules. The elements that transform the tensored basis to the basis in terms of permutations are called *Clebsch Gordon coefficients* and usually in physics we express them in terms of the total angular momentum and the z component m. Fig. 5 shows a subset of these coefficients and there is the related exercise 22.

Since we have now found the basis transformation via the Clebsch Gordon coefficients, we can also express the generators as matrices in this irreducible space. For the $j = 1$ module we have a representation of e.g.

$$t^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{3.4.7}$$

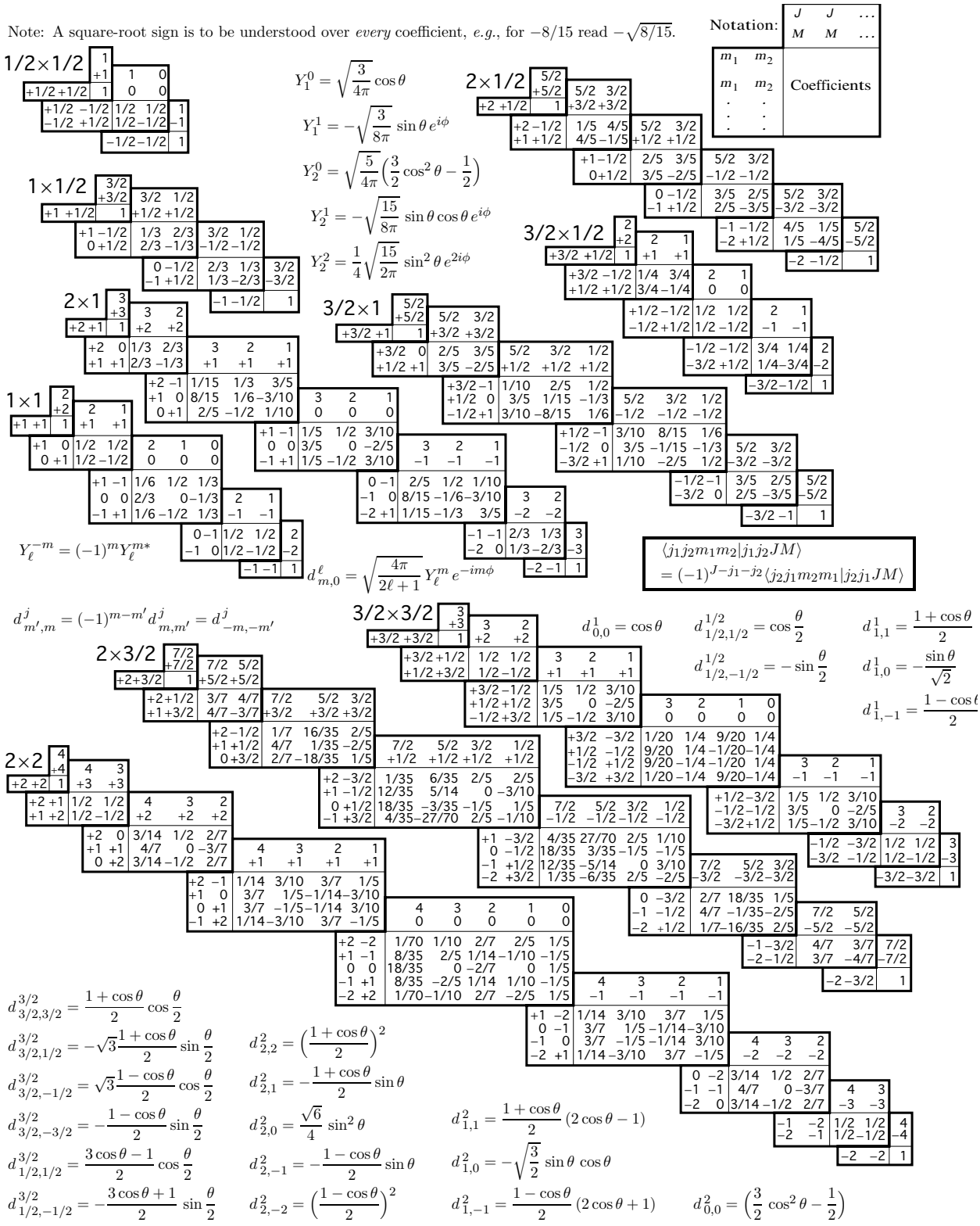


Figure 5: Some Clebsch Gordon coefficients (taken from The Review of Particle Physics).

From our discussion of the relation of $SU(2)$ and $SO(3)$, we anticipate that this representation $j = 1$ should be the fundamental representation of $SO(3)$. Let's verify this.

Using the exponential map, we show that the adjoint representation of t^3 and the exponential map can be interpreted as a rotation around the z axis. First we have

$$-it^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{3.4.8}$$

We include the $-i$ here because we want to obtain a real representation instead of a complex representation. Both representations are isomorphic in $\mathbb{C}$ sense but not if we understand them as maps on $\mathbb{R}$ vector spaces.

Let's look at the upper 2×2 block. We can see that $(-it^3)^{2n} = (-1)^n \tilde{\mathbf{1}}$, where the tilde indicates that there is a zero in the $3,3$ component of the matrix if $n \neq 0$. We also have $(-it^3)^{2n+1} = (-1)^n(-it^3)$, so

$$\begin{aligned} \exp(-i\phi t^3) = \sum_{n=0}^{\infty} \frac{(-i\phi t^3)^n}{n!} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \tilde{\mathbf{1}} \sum_{n=0}^{\infty} (-1)^n \frac{\phi^{2n}}{(2n)!} - it^3 \sum_{n=0}^{\infty} (-1)^n \frac{\phi^{2n+1}}{(2n+1)!} \\ &= \mathbf{1} \cos\phi - it^3 \sin\phi \\ &= \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{3.4.9}$$

which is indeed the standard form of 3×3 rotation matrix around the z axis. In a similar way, we can find other spin representations of $\mathfrak{so}(3)$ (exercises 23 and 24).

3.5 $SU(3)$

In the previous sections we have discussed strategies to obtain irreducible representations of $\mathfrak{su}(2)$ and of $SU(2)$ and $SO(3)$ in consequence. Instead of building representations with higher spin, we can also ask whether we can combine two $\mathfrak{su}(2)$ algebras non-trivially thus allowing for more quantum numbers (e.g. to describe the hadron spectrum). This will lead us to $\mathfrak{su}(n)$ algebras and the general representation theory of $SU(n)$.

3.5.1 Combining $\mathfrak{su}(2)$s

With a bit of hindsight we give the generators of two $\mathfrak{su}(2)$ algebras slightly different names and choose a basis in terms of raising and lowering operators

$$[H^1, E^1_\pm] = \pm 2E^1_\pm, \quad [E^1_+, E^1_-] = H^1 . \tag{3.5.1}$$

If we want to introduce a new quantum number with the second $\mathfrak{su}(2)$ algebra we need to require that both quantum numbers can be measured simultaneously

$$[H^1, H^2] = 0 . \tag{3.5.2}$$

We could also allow commutation relations of H^2 with $E^1_\pm$, but this would mean that we add another quantum number instead of enlarging the symmetry as is required from the observations in hadron physics. Instead we include a complete $\mathfrak{su}(2)$ algebra

$$[H^2, E^2_\pm] = \pm 2E^2_\pm, \quad [E^2_+, E^2_-] = H^2\,. \tag{3.5.3}$$

The next step involves how the two $\mathfrak{su}(2)$ algebras "interact". Requiring that if we apply a step operator to the mutual eigenvectors of H^1, H^2 gives back another eigenvector leads to

$$[H^i, E^j_\pm] = \pm A^{ji} E^j_\pm \quad i,j = 1,2 \tag{3.5.4}$$

with some numbers A^{ji}, which, however, have some restrictions.

First, because of the $\mathfrak{su}(2)$ relationships we have $A^{ii} = 2$, which leaves A^{12} and A^{21} unconstrained. We could choose zero values for them, which would mean that a the index 2 step operator would leave the H^1 eigenvalue unchanged. But this is not a situation of enlarged symmetry. If we have, say, a triplet under H^1 and a doublet H^2, then a larger symmetry seeks to organise the modules into a single irreducible module instead of tensoring the "old" modules. This can only happen if also the step operators of the two $\mathfrak{su}(2)$ algebras talk to each other, so $A^{12}, A^{21} \neq 0$.

To proceed we consider the commutators of $E^1_\pm$ and $E^2_\pm$ and give it name

$$E^\theta_\pm = [E^1_\pm, E^2_\pm]\,. \tag{3.5.5}$$

(it will become clear why we choose θ as a name). Because of the Jacobi identity

$$\begin{aligned}
[H^i, E^\theta_\pm] = [H^i, [E^1_\pm, E^2_\pm]] &= -\,[E^1_\pm, [E^2_\pm, H^i]] + [E^2_\pm, [E^1_\pm, H^i]] \\
&= \pm\, A^{2i}[E^1_\pm, E^2_\pm] \mp A^{1i}[E^2_\pm, E^1_\pm] = \pm(A^{1i} + A^{2i})[E^1_\pm, E^2_\pm] \\
&= \pm\,(A^{1i} + A^{2i}) E^\theta_\pm
\end{aligned} \tag{3.5.6}$$

we see that $E^\theta_\pm$ are again step operators. Similarly we have

$$[H^i, [E^1_\pm, E^2_\mp]] = \pm(A^{1i} - A^{2i})[E^1_\pm, E^2_\mp]\,. \tag{3.5.7}$$

In principle $[E^1_\pm, E^2_\pm]$ or $[E^1_\mp, E^2_\pm]$ could vanish. Not both of them can because the Jacobi identity implies

$$[E^1_+, [E^1_-, E^2_+]] - [E^1_-, \underbrace{[E^1_+, E^2_+]}_{=E^\theta_+}] = -[E^2_+, [E^1_+, E^1_-]] = -[E^2_+, H^1] = A^{21} E^2_+\,. \tag{3.5.8}$$

This shows for $A^{21} \neq 0$ not both E^θ_+ and $[E^1_-, E^2_+]$ can be zero (similar calculations apply for other combinations of $+$ and $-$. In summary there are *at least two additional independent step operators* and without loss of generality we choose these to be $E^\theta_\pm$.

The commutators of the step operators with H^i give again step operators, and the $E^1_\pm, H^1$ generate a $\mathfrak{su}(2)$ subalgebra. From our earlier discussion this means that the eigenvalues of H^1 are integers, and so must be the A^{ji} (see Eq. (3.5.7)).

The minimal possibility is to require vanishing of the other two potential step operators

$$[E^1_\pm, E^2_\mp] = 0 \tag{3.5.9}$$

and also require vanishing commutators with the other step operators

$$[E^j_\pm, [E^1_\pm, E^2_\pm]] = 0\,. \tag{3.5.10}$$

By considering identities like

$$0 = [E^1_-, [E^1_+, E^\theta_+]] = -(A^{11} + 2A^{21})E^\theta_+ \tag{3.5.11}$$

we can show that this corresponds to

$$A^{21} = A^{12} = -1\,, \tag{3.5.12}$$

which means that the eigenvalue of H^1 is changed by ∓ 1 if we act on the eigenvector with $E^2_\pm$. This is exactly the kind of cross talk between the $\mathfrak{su}(2)$s that is necessary to enlarge the symmetry.

This leaves us with two genuine quantum numbers H^1, H^2 – we say the resulting Lie algebra has rank 2. This expresses that there are two simultaneously diagonal basis vectors or the maximal abelian subalgebra, the *Cartan subalgebra*, spanned by H^1, H^2 has dimension two. Our good old $\mathfrak{su}(2)$ has rank 1. Additionally, there are six step operators $E^1_\pm, E^2_\pm, E^\theta_\pm$ and we have the following commutation relations

$$[H^1, H^2] = 0\,, \qquad [E^1_+, E^1_-] = H^1\,, \qquad [E^2_+, E^2_-] = H^2\,, \tag{3.5.13}$$

$$[H^1, E^1_\pm] = \pm 2E^1_\pm\,, \qquad [H^1, E^2_\pm] = \mp E^2_\pm\,, \quad [H^1, E^\theta_\pm] = \pm E^\theta_\pm\,, \tag{3.5.14}$$

$$[H^2, E^1_\pm] = \mp E^1_\pm\,, \qquad [H^2, E^2_\pm] = \pm 2E^2\pm\,, \quad [H^2, E^\theta_\pm] = \pm E^\theta_\pm\,, \tag{3.5.15}$$

$$[E^1_\pm, E^2_\pm] = \pm E^\theta_\pm\,, \qquad [E^1_\pm, E^\theta_\mp] = \mp E^2_\mp\,, \quad [E^2_\pm, E^\theta_\mp] = \pm E^1_\mp\,, \tag{3.5.16}$$

$$[E^1_\pm, E^2_\mp] = 0\,, \qquad [E^1_\pm, E^\theta_\pm] = 0\,, \qquad [E^2_\pm, E^\theta_\pm] = 0\,, \tag{3.5.17}$$

$$[E^\theta_+, E^\theta_-] = H^1 + H^2\,. \tag{3.5.18}$$

This structure becomes clearer if we define

$$E^3_+ = E^\theta_-\,, \quad E^3_- = E^\theta_+\,, \quad H^3 = -H^1 - H^2\,. \tag{3.5.19}$$

These operations satisfy

$$[H^3, E^3_\pm] = \pm 2E^3_\pm\,. \tag{3.5.20}$$

and they generate another $\mathfrak{su}(2)$ algebra. If we take this notation at face value, the result is completely invariant under permutations of the indices $i = 1, 2, 3$, which means that the vectors of eigenvalues $(\alpha^1, \alpha^2) = ([H^1, E] \propto E, [H^2, E] \propto E)$, with $E \in \{E^1_+, E^2_+, E^\theta_+\}$ are treated on an equal footing. We can read off those eigenvectors from the above system of commutation relations and find

$$\alpha^{(1)} = (2, -1)\,, \quad \alpha^{(2)} = (-1, 2)\,, \quad -\theta = (-1, -1)\,. \tag{3.5.21}$$

Interestingly, we have $\theta = \alpha^{(1)} + \alpha^{(2)}$. This will become important later. We also these vectors with negative signs when we consider the lowering operators.

Matrices that satisfy the commutation relations are given by

$$H^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{3.5.22}$$

$$E^1_+ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E^1_- = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$E^2_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad E^2_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$E^\theta_+ = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E^\theta_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

This is is the algebra $\mathfrak{su}(3)$ (again we do not discriminate between $\mathbb{C}$ or $\mathbb{R}$ vector spaces).

3.5.2 Orthogonal Basis

Let us make contact with the redefined version of the basis of $\mathfrak{su}(3)$ matrices. We look at different linear combinations

$$H^I = \frac{1}{2}H^1, \quad H^Y = \frac{1}{2\sqrt{3}}H^1 + \frac{1}{\sqrt{3}}H^2. \tag{3.5.23}$$

Using the original representation we have

$$H^Y = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{3.5.24}$$

The main advantage is that in this basis we have $\mathrm{Tr}\,(H^I H^I) = \mathrm{Tr}\,(H^Y H^Y) = 1/2$ and $\mathrm{Tr}\,(H^I H^Y) = \mathrm{Tr}\,(H^Y H^I) = 0$ (while in the old basis we have $\mathrm{Tr}(H^i H^j) = A^{ij}$). This tells us that the H_I and H_Y are orthogonal with respect to the scalar product that is induced by the trace. Writing $\vec{H} = (H^I, H^Y)$ we also obtain

$$[\vec{H}, E^1_\pm] = \pm(1,0)E^1_\pm, \quad [\vec{H}, E^2_\pm] = \pm\frac{1}{2}(-1,\sqrt{3})E^2_\pm, \quad [\vec{H}, E^\theta_\pm] = \pm\frac{1}{2}(1,\sqrt{3})E^\theta_\pm, \tag{3.5.25}$$

and

$$[E^1_+, E^1_-] = 2H^I, \quad [E^2_+, E^2_-] = -H_I + \sqrt{3}H_Y, \quad [E^\theta_+, E^\theta_-] = H_I + \sqrt{3}H_Y. \tag{3.5.26}$$

and we can read off the components of the vectors in this basis

$$\alpha^{(1)} = (1,0), \quad \alpha^{(1)} = \frac{1}{2}(-1,\sqrt{3}), \quad \theta = \frac{1}{2}(1,\sqrt{3}), \tag{3.5.27}$$

and again $\theta = \alpha^{(1)} + \alpha^{(2)}$ and there are additional vectors with minus signs.

3.5.3 Gell-Mann representation

A common representation of the $\mathfrak{su}(3)$ used in physics are the Gell-Mann matrices

$$\lambda^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.5.28}$$

$$\lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{3.5.29}$$

$$\lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{3.5.30}$$

$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \tag{3.5.31}$$

which are clearly reminiscent of the Pauli matrices (see e.g. $\lambda^{1,2,3}$). The Lie algebra $\mathfrak{su}(3)$ is then given by

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = i f^{ab}{}_c \frac{\lambda^c}{2} \tag{3.5.32}$$

with structure constants

$$f^{12}{}_3 = 1, \quad f^{14}{}_7 = f^{16}{}_5 = f^{24}{}_6 = f^{25}{}_7 = f^{34}{}_5 = f^{37}{}_6 = \frac{1}{2}, \quad f^{45}{}_8 = f^{67}{}_8 = \frac{\sqrt{3}}{2}. \tag{3.5.33}$$

3.6 Roots and Weights

We have already encountered the Cartan subalgebra as the subspace of simultaneously diagonalisable Lie algebra elements, as well as the eigenvectors of the step operators for these algebra elements, e.g. in the orthogonal basis we found

$$\alpha^{(1)} = (1, 0), \quad \alpha^{(2)} = \frac{1}{2}(-1, \sqrt{3}), \quad \theta = \frac{1}{2}(1, \sqrt{3}), \tag{3.6.1}$$

but they were not independent $\theta = \alpha^{(1)} + \alpha^{(2)}$.

These vectors $\alpha^{(i)}, \theta$ are called *roots* and as we have seen they are a direct reflection of the algebra (even though they do depend on the choice of basis). If $\alpha^{(i)}$ is a root, so is $-\alpha^{(i)}$. The roots, which cannot be represented by a linear combination of other roots are called *simple*. We call them *positive* when the first non-zero entry is positive (so $\alpha^{(1)}, -\alpha^{(2)}, \theta$ are positive). For $\mathfrak{su}(3)$ there are two simple roots – *exactly the dimension of the Cartan subalgebra or the rank of the Lie algebra*. This is true in general and extends beyond $\mathfrak{su}(3)$.

The root system is completely determined by the simple roots or by the angle between the simple roots. The latter for $\mathfrak{su}(3)$ is 120°, and cannot depend on our orthogonal basis choice. Hence,

this will give us a very compact and unique way of writing Lie algebras as the roots completely determine the algebra. Let's see why:

First of all since the generators of the Cartan subalgebra commute, they are simultaneously diagonalisable. We'll call the Cartan subalgebra $\mathfrak{g}_0$. But since this subalgebra is a vector space, *all elements h of the Cartan subalgebra are simultaneously diagonalisable*! From Eq. (3.2.1) we see that for all $y \in \mathfrak{g}$

$$\mathrm{ad}_h(y) = [h, y] = \alpha_y(h) y \,. \tag{3.6.2}$$

For a fixed $y \in \mathfrak{g}$, the eigenvalue $\alpha_h(y)$ is some number that will linearly depend on $h \in \mathfrak{g}_0$. The eigenvalues are the roots of the characteristic equation of ad_h, hence the name *root*. Different from normal eigenvalue problems, however, we only call it a root if it is non-zero. As $\mathfrak{g}$ is spanned by elements Eq. (3.6.2) we can decompose the algebra

$$\mathfrak{g} = \mathfrak{g}_0 \bigoplus_{\alpha>0} \mathfrak{g}_\alpha \,, \quad \mathfrak{g}_\alpha = \{x \in \mathfrak{g} | [h, x] = \alpha(h) x \text{ for all } h \in \mathfrak{g}_0\} \,. \tag{3.6.3}$$

This *root space decomposition* of $\mathfrak{g}$ relative to the Cartan subalgebra $\mathfrak{g}_0$ means that we can choose a basis, which in addition to $\{H^i\}$ is given by

$$[H^i, E^\alpha] = \alpha^{(i)} E^\alpha \,, \quad \text{for } i = 1, 2, ..., r = \dim \mathfrak{g} - \mathrm{rank}\, \mathfrak{g} \,, \tag{3.6.4}$$

which looks like the step operator commutation relations! In the example of $\mathfrak{su}(3)$ we also saw that each α-eigenspace is one-dimensional. This is always the case for semi-simple Lie algebras, i.e. such algebras which are direct groups of simple Lie algebras, where simple means the non-existence of a non-trivial subspace $\mathfrak{h} \subset \mathfrak{g}$ with $[\mathfrak{h}, \mathfrak{g}] \subset \mathfrak{h}$ (called an *ideal*).[9]

In summary there is a basis of a semi-simple Lie algebra

$$B = \{H^i | i = 1, \ldots, \mathrm{rank}\, \mathfrak{g}\} \cup \{E^\alpha | \alpha \text{ is a root of } \mathfrak{g}\} \,. \tag{3.6.5}$$

This basis is called the *Cartan-Weyl basis* of $\mathfrak{g}$ which is the standard way of writing a basis because it makes the step operators associated with a root explicit.

There is a way to find all roots of an Lie algebra. Consider again the Cartan subalgebra of the fundamental representation of $\mathfrak{su}(3)$ in the orthogonal basis.

$$H^I = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \,, \quad H^Y = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \,. \tag{3.6.6}$$

We can find three (simultaneous) eigenvectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \,, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \,, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \,, \tag{3.6.7}$$

with vectors of eigenvalues (called *weights* – exactly the weights we used to classify representations)

$$\nu^{(1)} = \frac{1}{2} \left(1, \frac{1}{\sqrt{3}}\right) \,, \quad \nu^{(2)} = \frac{1}{2} \left(-1, \frac{1}{\sqrt{3}}\right) \,, \quad \nu^{(3)} = \left(0, -\frac{1}{\sqrt{3}}\right) \,. \tag{3.6.8}$$

[9] $\mathfrak{h} = \{o\}$ and $\mathfrak{h} = \mathfrak{g}$ are trivial ideals.

We can now find the roots as the differences of weights $\nu^{(i)} - \nu^{(j)}, i \neq j$. In our case there are six differences which exactly give the roots as we obtained them in the orthogonal basis.

Dynkin Theorem. If α and β are two roots of a Lie algebra then

(i) $2(\alpha \cdot \beta)/|\alpha|^2$ is an integer

(ii) $\beta - [2(\alpha \cdot \beta)/|\alpha|^2]\alpha$ is also a root (somewhat similar to the Gram-Schmidt process)

If ν is a weight of some representation of a Lie algebra and α is a root then

(i) $2(\alpha \cdot \nu)/|\alpha|^2$ is an integer

(ii) $\nu - [2(\alpha \cdot \nu)/|\alpha|^2]\alpha$ is also a weight of the considered representation

As a consequence we can write all weights

$$\nu = \sum_i \mu_i \alpha^{(i)} \,. \tag{3.6.9}$$

with integer μ_i. Also, the highest weight Λ of a representation is the weight for which $\Lambda + \alpha^{(i)}$ is not a weight for all positive roots. E.g. for the fundamental representation of $\mathfrak{su}(3)$, the highest weight is $\Lambda = \nu^{(1)}$ because

$$\nu^{(2)} + \alpha^{(1)} = \nu^{(1)} \,, \quad \nu^{(3)} + \theta = \nu^{(1)} \,. \tag{3.6.10}$$

We can also count the number of irreducible representations for a given highest weight with the following

Highest Weight Theorem. If $\nu \cdot \alpha^{(i)} \geq 0$ for all simple roots i, then there is unique, irreducible, finite-dimensional representation of $\mathfrak{g}$ with highest weight $\nu = \Lambda$.

Indeed for $\mathfrak{su}(3)$ we have $\nu^{(1)} \cdot \alpha^{(i)} \geq 0$.

3.7 Classification of Lie algebras and Dynkin Diagrams*

All of what we have discussed in the previous section allows us to find the representation of a Lie algebra in a generic way. The root sector decomposition allows us to gather all knowledge about a Lie algebra from the simple roots and we can reconstruct the algebra in the *Cartan-Weyl basis*

$$\begin{aligned} [H^i, H^j] &= 0 \,, \\ [H^i, E^{\pm\alpha^{(p)}}] &= \pm\alpha^{(p)} E^{\pm\alpha^{(p)}} \,, \\ [E^{+\alpha^{(p)}}, E^{-\alpha^{(p)}}] &= \alpha_i^{(p)} H^i \,, \\ [E^{\alpha^{(p)}}, E^{\alpha^{(q)}}] &= N_{\alpha^{(p)}\alpha^{(q)}} E^{\alpha^{(p)}+\alpha^{(q)}} \\ &\qquad\qquad (\alpha^{(p)} + \alpha^{(q)} \neq 0) \,, \end{aligned} \tag{3.7.1}$$

with $i = 1, \ldots, \operatorname{rank}\mathfrak{g}$, $p = 1, \ldots, \dim\mathfrak{g} - \operatorname{rank}\mathfrak{g}$. We have chosen the orthogonal bases $\operatorname{Tr}(H^i H^j) = \delta^{ij}/2$ and $\operatorname{Tr}(E^{\alpha^{(p)}} E^{\alpha^{(q)}}) = \delta^{\alpha^{(p)}+\alpha^{(q)},0}/2$ (in particular the ladder operators are normalised by a factor $\sqrt{2}$ different than in Sec. 3.5.2). The $N_{\alpha^{(p)}\alpha^{(q)}}$ are zero if $\alpha^{(p)} + \alpha^{(q)}$ is not a root of $\mathfrak{g}$.

We also saw that while the roots are basis dependent, angles between the roots are not. This allows us to represent a particular Lie algebra via the angles between the simple roots. We saw that the simple roots are in general not orthonormal. This non-orthonormality is encoded in the *Cartan matrix* $A(\mathfrak{g})$ which we have found as the rank $\mathfrak{g}\times$rank $\mathfrak{g}$ matrix

$$A^{ij} = 2\,\frac{\alpha^{(i)}\cdot\alpha^{(j)}}{\alpha^{(j)}\cdot\alpha^{(j)}}\,. \tag{3.7.2}$$

This summarises the Lie algebra completely. So surveying all restrictions of the Cartan matrix allows us to classify all Lie algebras. We will not do this here but collect the results directly in terms of Dynkin diagrams.

Carrying out the classification programme one can show that there is only a handful of different finite dimensional algebras that can be built. They have very insightful names ($r = \text{rank}\,\mathfrak{g}$)

$$A_{\rm r}\ ({\rm r}\geq 1)\,,\ B_{\rm r}\ ({\rm r}\geq 3)\,,\ C_{\rm r}\ ({\rm r}\geq 2)\,,\ D_{\rm r}\ ({\rm r}\geq 4)\,, \tag{3.7.3}$$

(the cases are introduced to avoid double counting) in addition to five isolated cases

$$E_6\,,\quad E_7\,,\quad E_8\,,\quad G_2\,,\quad F_4\,. \tag{3.7.4}$$

The algebras of the infinite series are called the *classical Lie algebras*, and they are isomorphic to some of the matrix algebras that we have encountered, but they also include matrix algebras we have not discussed so far:

$$A_{\rm r}\simeq\mathfrak{su}(r+1)\,,\qquad B_{\rm r}\simeq\mathfrak{so}(2r+1)\,, \tag{3.7.5}$$

$$C_{\rm r}\simeq\mathfrak{sp}(r)\,,\qquad D_{\rm r}\simeq\mathfrak{so}(2r)\,. \tag{3.7.6}$$

The five isolated cases are called *exceptional Lie algebras* and they do not have a matrix representation.

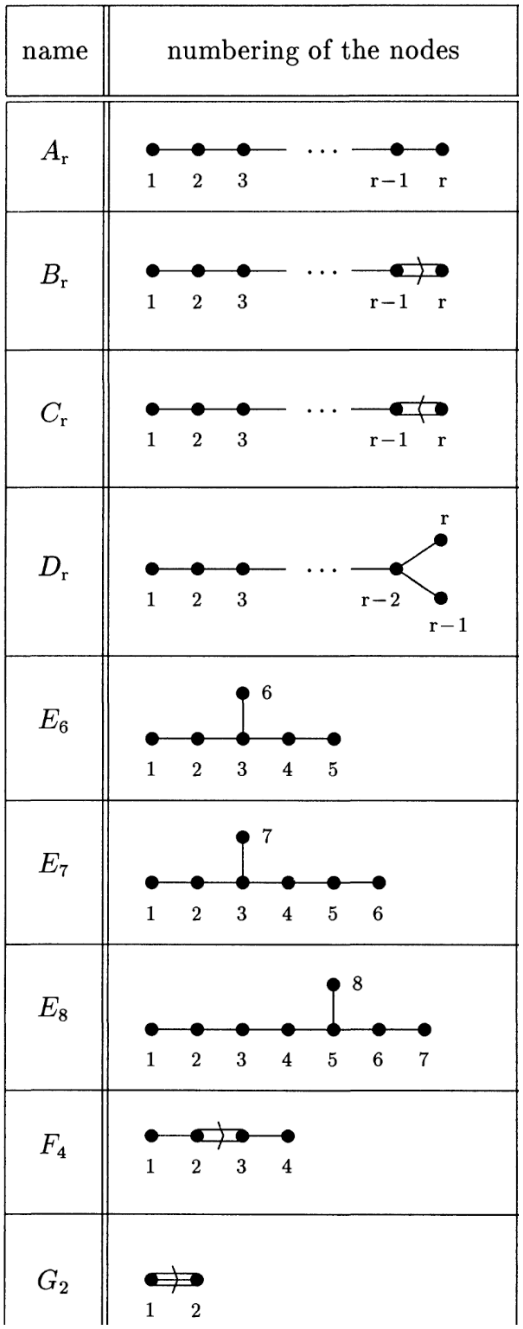


Figure 6: Dynkin diagrams of the finite-dimensional simple Lie algebras.

For any of these Lie algebras the angle between any two simple roots can only take one of the values

$$90^\circ\,,\quad 120^\circ\,,\quad 135^\circ\,,\quad 150^\circ\,. \tag{3.7.7}$$

By representing all angles between the simple roots in a diagrammatic way we can gather an enormous amount of information about the algebra. We can reproduce all roots, and from this the

algebra itself, and start building irreducible representations. In Dynkin diagrams we represent each simple root with a circle or node. If some of the simple roots are longer than others then this is represented by an arrowhead between the nodes. If the angle between simple roots is larger than 90° we connect the circle with links. The number of links between the simple roots indicates the angle between them

- one link corresponds to an angle of 120°
- two links represent 135°
- three links indicate an angle of 150° .

The resulting Dynkin diagrams are shown in Fig. 6. See also exercise 25.

3.8 Young tableaux*

After classifying all classic Lie algebras let us come back to building irreducible representation of $\mathfrak{su}(n)$. We have seen that for $\mathfrak{su}(2)$ we could obtain all irreducible representation by tensoring the fundamental representation building blocks and categorise the tensor product basis using the permutation (i.e. create a irreducible representation of the symmetric group) to decompose the tensor product space into irreducible modules. Young tableaux are a very helpful tool to do this as they bookkeep the permutations.

Let us have a look at $\mathfrak{su}(2)$ for simplicity and refresh our memories. We found that when combining two spin-1/2 systems

$$\begin{aligned} |2,2\rangle =& |1,1\rangle_1 \otimes |1,1\rangle_2 \\ |2,0\rangle =& \frac{1}{\sqrt{2}} \left[(|1,1\rangle_1 \otimes |1,-1\rangle_2) \oplus (|1,-1\rangle_1 \otimes |1,1\rangle_2) \right] \\ |2,-2\rangle =& |1,-1\rangle_1 \otimes |1,-1\rangle_2 \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ |0,0\rangle =& \frac{1}{\sqrt{2}} \left[(|1,1\rangle_1 \otimes |1,-1\rangle_2) \ominus (|1,-1\rangle_1 \otimes |1,1\rangle_2) \right] . \end{aligned} \tag{3.8.1}$$

we ended up with a spin 1 system and a spin 0 system. From the definition of the basis elements we see that the spin 1 system is symmetric under $1 \leftrightarrow 2$, the singlet is antisymmetric.

If we represent each fundamental representation with a box and understand the horizontal arrangement of boxes as symmetrising indices and the vertical arrangement as anti-symmetrising, we obtain a pictorial representation like

$$\square_1 \otimes \square_2 = \square\square_{\{12\}} \oplus \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}_{[12]} \tag{3.8.2}$$

(in physics we usually call the box or fundamental representation of $\mathfrak{su}(2)$ also just a "2" referring to the dimension of the fundamental representation, $2 \sim \square$). But since we know that the antisymmetric combination is only a singlet we have to identify two vertical boxes with a singlet

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} = \bullet \quad [\text{for } \mathfrak{su}(2)] . \tag{3.8.3}$$

The tensor space dimension is 4; there is 1 antisymmetric, and 3 symmetric index combinations.

Let's move to $\mathfrak{su}(3)$, which is a tad more complicated. We saw that we have a unique three-dimensional representation, and similar to $\mathfrak{su}(2)$ we get

$$\square_1 \otimes \square_2 = \square\square_{\{12\}} \oplus \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}_{[12]} . \tag{3.8.4}$$

but this time we have

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \neq \bullet \quad [\text{for } \mathfrak{su}(3)] \tag{3.8.5}$$

because we only stack two boxes instead of three (cf. the case of $\mathfrak{su}(2)$). The tensor space has dimension nine; the anti-symmetric combination has dimension 3, and the symmetric index combination amounts to 6 dimensions. This already shows that there is no singlet involved here this time. What does this mean?

Let's approach this from a slightly different angle, namely that from the group $SU(3)$. If we have a fundamental representation acting on a module with coordinates v_1^i, v_2^i we define a tensor $v_1^i v_2^j$. We can also have lower indices v_i which signalises that these objects transform under the complex conjugate representation, which is not necessarily identical to the fundamental representation (see later).

We now proceed to map out the symmetric group. Writing

$$v_1^i v_2^j = \frac{1}{2}\left\{v_1^i v_2^j + v_2^i v_1^j\right\} + \frac{1}{2}\left[v_1^i v_2^j - v_2^i v_1^j\right] \tag{3.8.6}$$

we have made (anti)symmetry explicit.

There is no anti-symmetric rank 2 tensor which is invariant under $SU(3)$. However, we have the rank three ϵ_{ijk} which transforms as $\varepsilon \mapsto \det(\mathcal{U})\varepsilon$ under $\mathcal{U} \in SU(3)$ and hence ε_{ijk} is invariant. Also if we write $\varepsilon_{ijk} v_1^i v_2^j = \varepsilon_{ij3} v_1^i v_2^j + \ldots$ we recover the antisymmetric part. However, there is a problem because we need a tensor of rank two and not of rank one. The only way to get back two indices whilst keeping antisymmetry is by glueing together two ε tensors

$$v_1^i v_2^j = \frac{1}{2}\left\{v_1^i v_2^j + v_2^i v_1^j\right\} + \frac{1}{2}\varepsilon^{ijk}\varepsilon_{klm} v_1^l v_2^m . \tag{3.8.7}$$

The object $\varepsilon_{klm} v_1^l v_2^m$ has one index and can be identified with a v_k, i.e. another three dimensional representation which is not the fundamental one. Since it is antisymmetric, we call it a $\bar{3}$ instead of $3 \sim \square$

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \sim \bar{3} \quad [\text{for } \mathfrak{su}(3)] \tag{3.8.8}$$

as it can be shown that it refers to the complex conjugate representation of $\mathfrak{su}(3)$ (exercise 26). For a general $\mathfrak{su}(n)$ we can show that

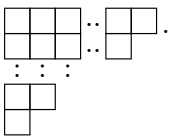
$$n \sim \square\,, \quad \bar{n} \sim \left.\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}\right\} n-1 \text{ times}\,, \quad 1 \sim \bullet = \left.\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}\right\} n \text{ times}\,. \tag{3.8.9}$$

The funny thing to realise is that a special property of $\mathfrak{su}(2)$ is encoded in this language:

$$\square \sim 2 \simeq \bar{2} \sim \square \tag{3.8.10}$$

the fundamental representation of $\mathfrak{su}(2)$ is *symplectic*, $\mathfrak{su}(2) \simeq \mathfrak{sp}(2)$ (for the fundamental representation). i.e. v^i and $v_i = \varepsilon_{ij} v^{*j}$ transform under the same representation and there is an automorphism that renders the upper and lower case representations equivalent.

We also need a prescription to combine more than two boxes, which boils down to an exercise in combinatorics. At the end of the day one can see that going from left to right, the number of rows has to decrease, so when building modules that host irreducible representations we need to produce tableaux that have the following shape



(3.8.11)

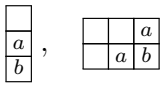
The way we build this structure is by the following rules. For $V^{(i)} \times V^{(j)}$

1. label each box of the j Young tableau with $a, b, c, \ldots$ according to their row from top to bottom.
2. attach boxes $a, b, c, \ldots$ from of the j tableau to i in all possible ways, one at a time, starting with a. The resulting diagram should be a Young tableau with no two $a, b, c, \ldots$ being in the same column (due to antisymmetrisation).
3. for any box in the diagram, there should be no more b's than a's to the right and above. Likewise no more c's than b's etc. A diagram like

   (3.8.12)

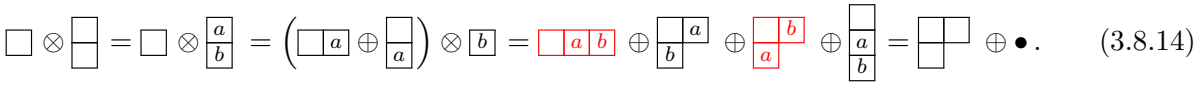
   is not allowed but

   (3.8.13)

   are allowed.
4. two tableaux of identical shape are different if the labels are distributed in a different way.

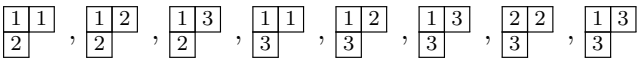
With this we are equipped to build irreducible representations. For instance we can look at the hadron spectrum which is described by a $\mathfrak{su}(3)$ symmetry. If quarks form 3 and the antiquarks form a $\bar{3}$ we have mesons (quark—anti–quark bound states) forming



(3.8.14)

The red Young tableaux have to be dropped according to the above rules.

Let us count how many different possibilities we have for the non-singlet component (remember indices in rows are symmetrised and indices in columns are anti-symmetrised. Making the indices explicit we have



(3.8.15)

so the dimension is eight (and we also see why we have to arrange the boxes the way introduced in Eq. (3.8.11)). From the form of the octet Young tableau we can see that the index combinations are neither completely symmetric nor completely anti-symmetric, the have a mixed symmetry. Since the adjoint representation of $\mathfrak{su}(3)$ has $d = 8$, we have found its representation as a Young tableau. The prediction of the "eightfold way" (i.e. three species of degenerate quarks and antiquarks that transform under $SU(3)$) is then the existence of eight degenerate mesons, which is observed in nature. I leave the baryons as an exercise below.

A convenient way to find the dimension of a $\mathfrak{su}(n)$ module is the *hook formula*

$$d = \prod_{(i,j)\in\text{tableau}} \frac{n-i+j}{h_{ij}} \,. \tag{3.8.16}$$

Here (i, j) denote the coordinates of the boxes of the tableau with i numbering the rows from top to bottom, and j numbering the columns from left to right. The h_{ij} is the hook length, which is the number of boxes that belong to the hook that has (i, j) as the box in the upper left corner, i.e. those boxes (i', j') with $i' = i, j' \geq j$ or $i' \geq i, j' = j$. In terms of the lengths a_j of the jth column and b_i of the ith row, we therefore have

$$h_{ij} = a_j + b_i - i - j + 1 \,. \tag{3.8.17}$$

In practice this works as follows. We assign a ratio to each box. We start in the upper left part of the tableau and assign a numerator n. The denominator is given by the sum of the boxes we find to the right and below this box $+1$ for the box itself. Moving to the next box on the right we assign a numerator $n + 1$ and the denominator, which is again given by the number of boxes to the right and below . When moving to boxes below the first one, we assign a numerator $n - 1$, the denominator is analogous to how we handle boxes in rows. At the end we multiply all ratios.

Here are some examples for what we have discussed for $\mathfrak{su}(3)$.

$$\begin{array}{|c|}\hline \frac{3}{2} \\ \hline \frac{2}{1} \\ \hline \end{array} = \frac{3}{2}\cdot\frac{2}{1} = 3\,, \qquad \begin{array}{|c|}\hline \frac{3}{3} \\ \hline \frac{2}{2} \\ \hline \frac{1}{1} \\ \hline \end{array} = \frac{3}{3}\cdot\frac{2}{2}\cdot\frac{1}{1} = 1\,, \qquad \begin{array}{|c|c|}\hline \frac{3}{3} & \frac{4}{1} \\ \hline \frac{2}{1} \\ \cline{1-1} \end{array} = \frac{3}{3}\cdot\frac{4}{1}\cdot\frac{2}{1} = 8\,. \tag{3.8.18}$$

Coming back to the eightfold way, we can construct baryons, which are bound states of three quarks.

3.9 Clifford Algebras and $SO(n)$

It can be shown that the formalism of the previous section can be extended to other Lie algebras and groups. The details become very technical and unfortunately we do not have the time to cover all this as part of this lecture course.[10]

[10] If you really want to get your hands dirty you can have a look at chapter 10 of Hamermesh's "*Group Theory and its Application to Physical Problems*".

Let us repeat the programme of $\mathfrak{su}(2)$ for $\mathfrak{so}(n)$. Writing a matrix $\mathcal{U} \in SO(n)$ as $\mathcal{U} = \exp(i\alpha^a T^a)$ where the T^a are the generators of the Lie algebra $\mathfrak{so}(n)$, we see that

$$T^a = -(T^a)^T \text{ and } \mathrm{Tr}\,(T^a) = 0\,. \tag{3.9.1}$$

It is customary to use a double index ij instead of a; one can show that the generators of $\mathfrak{so}(n)$ are given by

$$t_{ij} = e_{i,j} - e_{j,i} \tag{3.9.2}$$

with $i < j \leq n$ (which leads to $n(n-1)/2$ index combinations) where the matrices $(e_{i,j})_{kl} = \delta_{ik}\delta_{jl}$. The Lie algebra reads

$$[t_{ij}, t_{kl}] = \delta_{jk}t_{il} + \delta_{il}t_{jk} - \delta_{jl}t_{ik} - \delta_{ik}t_{jl}\,. \tag{3.9.3}$$

We come now back to Clifford algebras that we introduced earlier. Let's refresh our memory: An associative algebra with an operation $\{.,.\}$ and a non-degenerate, symmetric bi-linear form $g(.,.)$ such that

$$\{e_a, e_b\} = e_a e_b + e_b e_a = 2g(e_a, e_b)\mathbf{1} \tag{3.9.4}$$

is called a *Clifford algebra* $\mathcal{C}$. A Clifford algebra is generated by all elements $e_{i_1}^{m_1} \cdots e_{i_n}^{m_n}$ for which $i_1 \leq i_2 \leq \cdots \leq i_n$ and $m_i = 0, 1$ due to the anticommutation relations. These elements can be shown to be linearly independent and the dimension, hence, is 2^n.

If we choose $\mathbb{C}$ as a base field we can always find a basis such that $g(e_a, e_b) = \delta_{ab}$, whereas in the $\mathbb{R}$ case we can have different signatures $g(e_a, e_b) = \eta_a\delta_{ab}$ ($\eta_a = \pm 1$). We will mostly work in the complex case in the following though. The simplest case you are familiar with are the Pauli matrices $e^a = \sigma^a$, but we take this as a general definition.

We can make contact between the Clifford algebra $\mathcal{C}^n(\mathbb{C})$ and $\mathfrak{so}(n+1)$! Consider the elements

$$\begin{aligned} S_{0,a} = -S_{a,0} = e_a \quad & \text{for } a = 1, 2, \ldots, n \\ S_{a,b} = -S_{b,a} = [e_a, e_b] \quad & \text{for } a, b = 1, 2, \ldots, n,\ a \neq b, \end{aligned} \tag{3.9.5}$$

we can see that the $S_{a,b}$ obey a $\mathfrak{so}(n+1)$ algebra so $\mathcal{C}^n \supset \mathfrak{so}(n+1)$. This means that if we find a representation of $\mathcal{C}^n$, we have also found a representation of $\mathfrak{so}(n+1)$. Earlier we have seen that the Pauli matrices generate indeed a Clifford algebra, and we have seen that $SU(2)$ is the spin group of $SO(3)$. This suggests indeed that by finding a representation of the complex Clifford algebra we can find a representation of the universal covering group of $SO(n)$.

So let us quickly discuss the representation theory of Clifford algebras $\mathcal{C}^n(\mathbb{C})$. First we start with $n = 1$. The algebra $\mathcal{C}^1$ is generated by two elements $\mathbf{1}$ and e_1, with $e_1^2 = \mathbf{1}$. Hence it is isomorphic to the discrete group $\mathbb{Z}_2$. We can use a different basis $(\mathbf{1} \pm e_1)/2$ and see that this algebra is isomorphic to two copies of $\mathbb{C}$:

$$\mathcal{C}^1(\mathbb{C}) \simeq \mathbb{C} \oplus \mathbb{C} \tag{3.9.6}$$

The discrete group $\mathbb{Z}_2$ has two representations, the trivial one, where both elements are represented by 1, and a "faithful" (i.e. injective) one where the elements are represented by ± 1. The two representations are directly related to the two parts of the direct product above.

Next we consider $n = 2$. We have encountered this Clifford algebra as the Pauli matrices, the dimension of this Clifford algebra is four (remember that $\mathbf{1}$ is part of the Clifford algebra). We

choose $\mathcal{R}(e_1) = \sigma^1$ and $\mathcal{R}(e_2) = \sigma^2$, and we have realised the $n = 2$ Clifford algebra. Since $\sigma^1\sigma^2 = i\sigma^3$ is part of the algebra, the Clifford algebra for $n = 2$ is represented by all complex 2×2 matrices

$$\mathcal{C}^2(\mathbb{C}) \simeq \mathcal{M}_2(\mathbb{C})\,. \tag{3.9.7}$$

This representation in faithful.

Next $n = 3$. This Clifford algebra has dimension eight. We can show that $\mathcal{R}(e_a) = \sigma^a$ is again a representation. Actually, for $\mathcal{R}(e^a) = \sigma^a$ we have seen the Clifford algebra already in Eq.(3.1.14):

$$\{\sigma^a, \sigma^b\} = \sigma^a\sigma^b + \sigma^b\sigma^a = 2\delta^{ab}\mathbf{1}\,. \tag{3.9.8}$$

The representation $\mathcal{R}(e_a) = \sigma^a$ is however not faithful:

$$\mathcal{R}(e_1e_2e_3) = \mathcal{R}(e_1)\mathcal{R}(e_2)\mathcal{R}(e_3) = i\mathbf{1} = \mathcal{R}(i\mathbf{1}) \tag{3.9.9}$$

We could have started with $\mathcal{R}'(e_a) = -\sigma^a$ and would find another (unfaithful) representation of the $n = 3$ Clifford algebra. This representation is not equivalent as can be seen from the following exercise

So we have found a second representation, and the combination of R and R' exhausts all possible irreducible representations. Each individual representation is equivalent to a complex 2×2 matrix algebra and therefore

$$\mathcal{C}^3(\mathbb{C}) \simeq \mathcal{M}_2(\mathbb{C}) \oplus \mathcal{M}_2(\mathbb{C})\,. \tag{3.9.10}$$

This reveals the general recurrence relation

$$\mathcal{C}^{n+2}(\mathbb{C}) = \mathcal{C}^n(\mathbb{C}) \otimes \mathcal{C}^2(\mathbb{C}) \tag{3.9.11}$$

which can be proven by induction. This leads to

$$\mathcal{C}^n \simeq \begin{cases} \mathcal{M}_{2^{n/2}}(\mathbb{C}) & \text{for } n = 2\mathbb{Z} \\ \mathcal{M}_{2^{(n-1)/2}}(\mathbb{C}) \oplus \mathcal{M}_{2^{(n-1)/2}}(\mathbb{C}) & \text{for } n = 2\mathbb{Z}+1 \end{cases}\,. \tag{3.9.12}$$

3.10 $SO(3) \simeq SU(2)/\mathbb{Z}_2$ revisited

Equipped with this knowledge about the representation of $\mathfrak{so}(n)$ algebras as Clifford algebras we can come back to $\mathfrak{so}(3)$. We understand $\mathfrak{so}(3)$ as a complex algebra, so we directly have from Eq. (3.9.5),

$$\mathfrak{so}(3) \subset \mathcal{C}^2(\mathbb{C}) \tag{3.10.1}$$

and, using Eq. (3.9.12)

$$\mathfrak{so}(3) \subset \mathcal{M}_2(\mathbb{C}) \tag{3.10.2}$$

which is pretty much the result that we have encountered earlier, i.e. there is two-dimensional complex module that hosts a (projective) representation of $\mathfrak{so}(3)$. Such a module is called *spinor module* (similar to the name of the universal covering or spin group).

So far we have only been concerned with representations of the algebra. How about the group? Consider matrices $\mathcal{U} \in SO(n)$, we denote the entries $\mathcal{U}_a{}^b$, and a representation of the Clifford algebra $\mathcal{C}^{n-1}(\mathbb{C}) \supset \mathfrak{so}(n)$, $\rho_a = \mathcal{R}(e_a)$. We can now find a different representation

$$\rho'_a = \mathcal{U}_a{}^b \, \rho_b \tag{3.10.3}$$

then

$$\{\rho'_a, \rho'_b\} = \mathcal{U}_a{}^c \mathcal{U}_b{}^d \, \{\rho_c, \rho_d\} = 2\,\delta_{cd} \mathcal{U}_a{}^c \mathcal{U}_b{}^d = 2\,\mathcal{U}_a{}^c \mathcal{U}^T{}_{cb} = 2\,\delta_{ab}\,, \tag{3.10.4}$$

i.e. ρ' are again a representation.

For an even $n-1$, i.e. odd n, we have only a single representation Eq. (3.9.12) (up to an isomorphism), so both representations must be equivalent. In other words, for every $\mathcal{U} \in SO(n)$ there must be an invertible matrix $\mathcal{M}(\mathcal{U})$ such that

$$\rho' = \mathcal{M}(\mathcal{U})^{-1} \rho \, \mathcal{M}(\mathcal{U})\,. \tag{3.10.5}$$

So $\mathcal{M}(\mathcal{U})$ is uniquely defined up to a scalar multiple (we only need $\det[\mathcal{M}(\mathcal{U})] \neq 0$ to be able to find an invertible matrix $\mathcal{M}$). $\mathcal{M}$ needs to implement the $SO(n)$ group action. So for $\mathcal{U}_1\mathcal{U}_2$ we need

$$\mathcal{M}(\mathcal{U}_1\mathcal{U}_2) = \omega(\mathcal{U}_1,\mathcal{U}_2)\mathcal{M}(\mathcal{U}_1)\mathcal{M}(\mathcal{U}_2) \tag{3.10.6}$$

where ω is a scalar. We have seen this earlier as an example of a projective representation. For $SO(n)$, n odd, we cannot remove[11] ω, but we have found earlier that, e.g. $SO(3) \simeq SU(2)/\mathbb{Z}_2$. The discrete $\mathbb{Z}_2 = \{+1, -1\}$ states that $SU(2)$ is the two-fold cover of $SO(3)$, as you have seen from the last exercise in Sec. 3.4.

3.11 Representation Theory of the Lorentz Group

3.11.1 The Universal Covering Group of $SO^\uparrow_+(1,3)$ and Spinors

Let us first construct the universal covering group of $SO^\uparrow_+(1,3)$. We follow the strategy that we used to show $SO(3) \simeq SU(2)/\mathbb{Z}_2$. We extend the set of Pauli matrices by the 2×2 unity matrix

$$\sigma^0 = \mathbf{1}\,,\ \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,,\ \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,,\ \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \tag{3.11.1}$$

and associate with a contravariant four vector the matrix

$$\underline{x} = x_\mu \sigma^\mu = \begin{pmatrix} x^0 - x^3 & ix^2 - x^1 \\ -ix^2 - x^1 & x^0 + x^3 \end{pmatrix}\,. \tag{3.11.2}$$

Note that this is not really a Lorentz contraction (i.e. it is not invariant under Lorentz transformations), but we *identify* Lorentz vectors with hermitian 2×2 matrices. It is easy to see that $\underline{x}^\dagger = \underline{x}$ and $\det \underline{x} = g_{\mu\nu} x^\mu x^\nu$. These form a *real* vector space of linear maps that operate on a two-dimensional complex vector space. If we transform this complex vector space with an unimodular complex matrix $\mathcal{U} \in SL(2,\mathbb{C})$ (understood as a *real* group) we can write

$$\underline{x}' = \mathcal{U}\,\underline{x}\,\mathcal{U}^\dagger\,. \tag{3.11.3}$$

[11] For even n the implications are similar, we will see an example of this below.

and we still have $\underline{x}'^{\dagger} = \underline{x}'$ and $\det \underline{x}' = \det \underline{x}$ due to $\det \mathcal{U} = \det \mathcal{U}^{\dagger} = 1$. Therefore, for such a $\mathcal{U} \in SL(2,\mathbb{C})$ there is a transformation $\Lambda \in SO^{\uparrow}_{+}(1,3)$ with $x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$. This means we have constructed a homomorphism $SL(2,\mathbb{C}) \to SO^{\uparrow}_{+}(1,3)$, which, again is not injective (exercise 32). $\pm\mathcal{U}$ are mapped to the same element $\Lambda(\mathcal{U}) = \Lambda(-\mathcal{U})$, we have

$$SO^{\uparrow}_{+}(1,3) \simeq SL(2,\mathbb{C})/\mathbb{Z}_2 \,. \tag{3.11.4}$$

Note again that we understand $SL(2,\mathbb{C})$ as a real Lie group in this context. Similar to constructing the universal covering group of $SO(3)$ we can again use the trace to find the explicit representation of the local isomorphism.

It is worthwhile to pause again at this stage and re-evaluate everything from the perspective of the algebra. Upon complexification of the real algebra we obtain the relation[12]

$$\mathbb{C} \times \mathfrak{so}^{\uparrow}_{+}(1,3)_{\mathbb{R}} = \mathfrak{so}(4)_{\mathbb{C}} \subset \mathcal{C}^3(\mathbb{C}) \simeq \mathcal{M}_2(\mathbb{C}) \oplus \mathcal{M}_2(\mathbb{C}) \,. \tag{3.11.5}$$

Hence we expect two spinor modules to from a faithful representation also for the universal covering group $SL(2,\mathbb{C}) = \mathrm{Spin}(1,3)$. This is indeed the case: If $\mathcal{U}$ is a representation of $SL(2,\mathbb{C})$, then $\mathcal{U}^{\star}$ is an non-equivalent one!

Let's start by constructing irreducible representations of the Lorentz group and irreducible modules by using what is called *the Weyl-van der Waerden* or *two-component spinor* formalism. We start with a vector of the complex two-dimensional space and denote it by ψ_α (also called *a left-handed Weyl spinor*), which gives rise to the first $SL(2,\mathbb{C})$ module:

$$\psi_\alpha \longmapsto \psi'_\alpha = \mathcal{U}_\alpha{}^\beta \psi_\beta \,. \tag{3.11.6}$$

A second module is given by the action of the complex conjugate matrices, and we denote an element of the module by $\psi_{\dot\alpha}$ (also called *a right-handed Weyl spinor*):

$$\psi_{\dot\alpha} \longmapsto \psi'_{\dot\alpha} = \mathcal{U}^{\star}_{\dot\alpha}{}^{\dot\beta} \psi_{\dot\beta} \,. \tag{3.11.7}$$

This also means that complex conjugation changes from one module to another, i.e.

$$\psi^{\star}_{\alpha} = \psi_{\dot\alpha} \,, \quad \psi^{\star}_{\dot\alpha} = \psi_{\alpha} \,. \tag{3.11.8}$$

From Eq. (3.11.3) we see that the σ^μ need to have and index structure $\sigma^{\mu}_{\alpha\dot\alpha}$. Furthermore, we have invariant tensors in each of the modules

$$\varepsilon_{\alpha\beta} \longmapsto \varepsilon_{\delta\epsilon} \mathcal{U}_\alpha{}^\delta \mathcal{U}_\beta{}^\epsilon = \det(\mathcal{U}) \varepsilon_{\alpha\beta} = \varepsilon_{\alpha\beta} \tag{3.11.9}$$

and its inverse $\varepsilon^{\alpha\beta}\varepsilon_{\beta\delta} = \delta^\alpha{}_\delta$ with identical relations for the dotted module. We can use this epsilon tensor to raise and lower indices in the corresponding modules

$$\psi^\alpha = \varepsilon^{\alpha\beta}\psi_\beta \,, \quad \psi_\alpha = \varepsilon_{\alpha\beta}\psi^\beta \tag{3.11.10}$$

and define an invariant

$$\psi^\alpha \phi_\alpha = \varepsilon^{\alpha\beta}\psi_\beta\phi_\alpha = \varepsilon^{\beta\alpha}\psi_\alpha\phi_\beta = -\varepsilon^{\alpha\beta}\phi_\beta\psi_\alpha = -\phi^\alpha\psi_\alpha \,. \tag{3.11.11}$$

[12] We can also show that $\mathbb{C} \times \mathfrak{so}^{\uparrow}_{+}(1,3)_{\mathbb{R}} = \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}}$.

So we have to pay attention how we contract indices! By convention we contract undotted indices top-left to bottom-right, and dotted indices from bottom-left to top-right:

$$\psi_{\dot\alpha}\phi^{\dot\alpha} = -\phi_{\dot\alpha}\psi^{\dot\alpha}\,. \tag{3.11.12}$$

Equipped with these epsilon tensors we can also define a new set of Pauli matrices

$$(\bar\sigma^\mu)^{\dot\alpha\alpha} := \varepsilon^{\dot\alpha\dot\beta}\varepsilon^{\alpha\beta}\sigma^\mu_{\beta\dot\beta} \quad [\bar\sigma^\mu = (\mathbf{1}, -\sigma^i)]\,, \tag{3.11.13}$$

Note that in particular

$$(i\bar\sigma^2)^{\dot\alpha\alpha}(i\sigma^2)_{\alpha\dot\beta} = \delta^{\dot\alpha}{}_{\dot\beta}\,, \quad (i\sigma^2)_{\alpha\dot\alpha}(i\bar\sigma^2)^{\dot\alpha\beta} = \delta_\alpha{}^\beta\,. \tag{3.11.14}$$

From our discussion of $\mathcal{C}^3(\mathbb{C})$ we know that we need both $\bar\sigma$ and σ for the representation *of the associated complex Clifford algebra to be faithful*. With these matrices we can now write down an explicit representation of $\mathcal{C}^3(\mathbb{C})$. By defining

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar\sigma^\mu & 0 \end{pmatrix} \tag{3.11.15}$$

we can check that

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbf{1}\,. \tag{3.11.16}$$

We can now make contact between the conjugate and the ordinary representations by looking at

$$\psi^\star_\alpha \longmapsto \mathcal{U}^\star{}_\alpha{}^\beta\psi^\star_\beta\,, \tag{3.11.17}$$

and we can use the sigma matrices to move to dotted indices

$$\psi^\star_\alpha \longmapsto \mathcal{U}^\star{}_\alpha{}^\beta\delta_\beta{}^\epsilon\psi^\star_\epsilon = \mathcal{U}^\star{}_\alpha{}^\beta(i\sigma^2)_{\beta\dot\gamma}(i\bar\sigma^2)^{\dot\gamma\epsilon}\psi^\star_\epsilon \tag{3.11.18}$$

and therefore

$$\begin{aligned}(i\bar\sigma^2)^{\dot\alpha\alpha}\psi^\star_\alpha \longmapsto &\left[(i\bar\sigma^2)^{\dot\alpha\alpha}\mathcal{U}^\star{}_\alpha{}^\beta(i\sigma^2)_{\beta\dot\gamma}\right](i\bar\sigma^2)^{\dot\gamma\epsilon}\psi^\star_\epsilon \\ &= \mathcal{U}^{\star\dot\alpha}{}_{\dot\gamma}(i\bar\sigma^2)^{\dot\gamma\epsilon}\psi^\star_\epsilon\,.\end{aligned} \tag{3.11.19}$$

The representation $\mathcal{U}^{\star\dot\alpha}{}_{\dot\gamma}$ is contragredient to $\mathcal{U}^\star{}_{\dot\alpha}{}^{\dot\gamma}$ and equivalent as can be seen from the properties of $\sigma^2, \bar\sigma^2$ or

$$\mathcal{U}^{\star\dot\alpha}{}_{\dot\beta} = \varepsilon^{\dot\alpha\dot\gamma}\varepsilon_{\dot\beta\dot\delta}\mathcal{U}^\star{}_{\dot\gamma}{}^{\dot\delta}\,. \tag{3.11.20}$$

Eq. (3.11.19) means that we can define a dotted spinor, which is slightly different from the "ordinary" complex conjugation[13] of Eq. (3.11.8),

$$\bar\psi^{\dot\alpha} := (i\bar\sigma^2)^{\dot\alpha\beta}\psi^\star_\beta\,. \tag{3.11.21}$$

[13]The alert reader might point out that indeed $\varepsilon = i\sigma^2$, so we could have just lifted the dotted index with an epsilon tensor. The choice of σ^2, however, is purely conventional.

With this, the γ matrices represent $SL(2,\mathbb{C})$ on the combined module

$$\Psi = \begin{pmatrix} \psi_\alpha \\ \bar{\phi}^{\dot{\alpha}} \end{pmatrix} . \tag{3.11.22}$$

Ψ is called a *Dirac spinor*[14] and is the central object of relativistic quantum mechanics. From a representation theory point of view, the Dirac spinor is slightly cumbersome as we build it from two undotted modules and take a detour instead of working directly with dotted and undotted indices.

Ultimately, we want to find a Lorentz- (or even Poincaré-)covariant tensor equation, which would allow us to formulate a differential equation (similar to our discussions of tensors in Sec. A.7), which after we finding a solution, can be compared to experiments. Using the two $SL(2,\mathbb{C})$ modules, together with $\partial_\mu = \partial/\partial x^\mu$, we can write down two differential equations for the left- and right-handed spinors

$$\sigma^\mu_{\alpha\dot{\alpha}} \partial_\mu \bar{\phi}^{\dot{\alpha}} = 0 \,, \quad (\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} \partial_\mu \psi_\alpha = 0 \,, \tag{3.11.23}$$

which, by construction, are form-invariant under Lorentz transformations. In this case both left- and right-handed modules (in this notation) decouple. But we can couple both modules too. From the index structure, we can couple the above equations, at the price of an additional Lorentz scalar m

$$i\sigma^\mu_{\alpha\dot{\alpha}} \partial_\mu \bar{\phi}^{\dot{\alpha}} - m\psi_\alpha = 0 \,, \quad i(\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} \partial_\mu \psi_\alpha - m\bar{\phi}^{\dot{\alpha}} = 0 \,. \tag{3.11.24}$$

In the Dirac notation this becomes

$$(i\gamma^\mu \partial_\mu - m)\Psi = 0 \,. \tag{3.11.25}$$

This is the famous Dirac equation that describes the relativistic dynamics[15] of a spin-1/2 particle with mass m. Already in our discussion of $\mathfrak{su}(2)$ we have seen that the spin-1/2 representation is two dimensional, the two-dimensional module refers to the different spin states. The Dirac equation can be solved by looking at the Fourier transformed Ψ in the rest frame where $p^\mu = (m, \vec{o})$, and then use Lorentz transformations to find the solution for general p^μ.

Denoting left- and right handed modules with $(.,.)$ and using the spin language instead of the highest weight notation we can see that

$$\text{left-handed Weyl spinor :} \quad \left(\frac{1}{2}, 0\right) \tag{3.11.26}$$

$$\text{right-handed Weyl spinor :} \quad \left(0, \frac{1}{2}\right) \tag{3.11.27}$$

$$\text{Dirac spinor :} \quad \left(\frac{1}{2}, 0\right) \oplus \left(0, \frac{1}{2}\right) . \tag{3.11.28}$$

[14] By definition a Dirac spinor is an element of the complex vector space that carries an irreducible representation of the complex Clifford algebra.

[15] In SI units m would be replaced by $mc/\hbar$, which is just the inverse (reduced) Compton wavelength of a particle with mass m. Hence the Dirac equation contains a characteristic length scale which is introduced from a relativistic particle-wave matching: $E = mc^2 = hc/\lambda = \hbar c/(\lambda/2\pi)$.

From the mapping $SL(2,\mathbb{C}) \ni \mathcal{U} \mapsto \Lambda(\mathcal{U}) \in SO^{\uparrow}_{+}(1,3)$ (the index structure of σ^{μ}) we furthermore see that

$$\text{Lorentz vector}: \quad \left(\frac{1}{2},\frac{1}{2}\right). \tag{3.11.29}$$

We have quietly extended the base field to the complex numbers to look for the representations of the Lorentz group, making contact with the complex irreducible representations of the Clifford algebra in three dimensions. Since we deal, however, with a (pseudo)metric $g_{\mu\nu}$ we can't be sure whether the representations that we have constructed are still irreducible in their real sense, see Eq. (3.11.5). In fact, they are not. Looking at

$$i\gamma^2\Psi^{\star} = \begin{pmatrix} 0 & (i\sigma^2)_{\alpha\dot{\alpha}} \\ (i\bar{\sigma}^2)^{\dot{\alpha}\alpha} & 0 \end{pmatrix} \begin{pmatrix} \psi_{\alpha} \\ \bar{\phi}^{\dot{\alpha}} \end{pmatrix}^{\star} = \begin{pmatrix} [(i\sigma^2)_{\alpha\dot{\alpha}}\bar{\phi}^{\dot{\alpha}}]^{\star} \\ (i\bar{\sigma}^2)^{\dot{\alpha}\alpha}\psi^{\star}_{\alpha} \end{pmatrix} = \begin{pmatrix} \phi_{\alpha} \\ \bar{\psi}^{\dot{\alpha}} \end{pmatrix} \tag{3.11.30}$$

we can reduce the number of degrees of freedom by requiring $i\gamma^2\Psi^{\star} \stackrel{!}{=} \Psi$ (which can be though of as the analogue of the identification of complex and complex conjugate numbers to get the real numbers). This is tantamount to identifying the dotted and undotted modules by imposing the condition

$$\phi = \psi\,. \tag{3.11.31}$$

The resulting spinor

$$\Psi_M = \begin{pmatrix} \psi_{\alpha} \\ \bar{\psi}^{\dot{\alpha}} \end{pmatrix} \tag{3.11.32}$$

is called *Majorana spinor* and acts as representation space of the *real* Clifford algebra. The dotted and undotted representations can be thought of as aligned in the sense that solving the Dirac equation reduces to finding a solution in one of the two modules. This means we can't really distinguish between complex conjugate and ordinary transformations anymore and we have identified complex dotted and undotted module although they are still inequivalent. We can understand $i\gamma^2$ together with complex conjugation as charge conjugation; the Majorana spinor is "real", which suggests that the Dirac spinor collects particle and anti-particle, whereas the particle and anti-particle are identical in the Majorana case.

From the definition of the γ matrices we can also introduce two special elements of the Clifford algebra

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -\mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}. \tag{3.11.33}$$

This is again reminiscent of our discussion of Clifford algebras, and indeed one can show that

$$\{\gamma^{\mu},\gamma^5\} = 0, \quad (\gamma^5)^2 = \mathbf{1}\,. \tag{3.11.34}$$

With this we can define two projectors in the Clifford algebra

$$P_{L,R} = \frac{1}{2}\left(1 \mp \gamma^5\right), \quad P^2_{L,R} = P_{L,R} \tag{3.11.35}$$

which projects out the left- and right-handed spinor components. In a slightly fancier language, we have identified the irreducible representations of the *even subalgebra* of the Clifford algebra. As a by-product we have found a representation of $\mathcal{C}^4(\mathbb{C})$.

3.12 The Lorentz and Poincaré Algebras

We turn to a more careful treatment of the Lorentz algebra. By studying the Lorentz transformations of Eq. (2.1.56) close to the one element we can find the generators and the associated Lie algebra $\mathfrak{so}(1,3)$

$$[t_{\iota\rho}, t_{\kappa\lambda}] = -i(g_{\rho\kappa}t_{\iota\lambda} + g_{\iota\lambda}t_{\rho\kappa} - g_{\rho\lambda}t_{\iota\kappa} - g_{\iota\kappa}t_{\rho\lambda})\,, \tag{3.12.1}$$

which is a real form of the analytically continued $\mathfrak{so}(4)$.

Equipped with our knowledge that we can give this algebra (in the complex sense) a meaning in terms of a Clifford algebra $\mathcal{C}^3$, we can indeed show that

$$\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu] \tag{3.12.2}$$

is representation of the Lie algebra Eq. (3.12.1) and we can use again the exponential map to describe nearly all elements of the Lie group, just like we did with the $\mathfrak{su}(2)$ or $\mathfrak{so}(3)$.

In particular, we have $[\sigma^{\mu\nu}, \gamma^5] = 0$ so the projectors of the even subalgebra are Lorentz-invariant. From a physics point of view this means that if we prepare a fermion in a left-handed state it will look left-handed in all systems that are connected by Lorentz transformations.

If we include the translations to get a representation of the Poincaré algebra we need to extend Eq. (3.12.1) by four generators of time and space translations p_δ, which can again be found by looking at the respective transformation close to unity[16] and according to Eq. (2.1.53) will have non-trivial commutation relations:

$$[t_{\mu\nu}, p_\delta] = i(g_{\mu\nu}p_\delta - g_{\nu\delta}p_\mu) \quad \text{and} \quad [p_\mu, p_\nu] = 0\,. \tag{3.12.3}$$

With this we can see that there are two Casimirs

$$C_1 = p^\delta p_\delta \tag{3.12.4}$$

which we have already encountered (however this time it refers to generators and is an a posterio justification of choosing the rest mass earlier in Eq. (2.1.60)) and

$$C_2 = -w^\mu w_\mu \quad \text{with} \quad w_\mu = \frac{1}{2}\varepsilon_{\mu\nu\rho\delta}t^{\nu\rho}p^\delta\,. \tag{3.12.5}$$

W^μ is called the *Pauli-Lubanski vector*. Using the definition and the Poincaré algebra we can see that

$$w^\mu p_\mu = 0\,, \quad [w_\mu, p_\nu] = 0\,, \tag{3.12.6}$$

and

$$[t_{\mu\nu}, w_\delta] = i(g_{\mu\delta}w_\nu - g_{\nu\delta}w_\mu) \quad \text{and} \quad [w_\mu, w_\nu] = i\varepsilon_{\mu\nu\rho\delta}w^\rho p^\delta\,. \tag{3.12.7}$$

[16] These are obviously related to four momenta and a representation in terms of four derivatives is $p^\mu = i\partial/\partial x_\mu$.

3.12.1 The Stability Subgroup

The Casimir C_1 tells us that we can find irreducible representations by first looking at all states with a given value of C_1, i.e. we are interested in the transformations that transform a four momentum q^α onto q'^α leaving $q^\alpha q_\alpha = m^2$ invariant. But even more, we can ask for all automorphic transformations of the sub-vector space characterised by a given four momentum q. E.g. a Dirac spinor can be characterised by a given four momentum (through solving Eq. (3.11.25) in Fourier space), and we can ask for the number of degrees of freedom that specify the spinor on top of its four momentum. The Poincaré transformations which do exactly that form a subgroup of all Poincaré transformations and are called the *stability subgroup for q*. Using the exponential map we can write a Poincaré transformation as

$$U(\Lambda, b) = \exp\left[i\left(b^\mu p_\mu + K^{\mu\nu} t_{\mu\nu}\right)\right] \tag{3.12.8}$$

where the $K^{\mu\nu}$ are the six "angles" of the Lorentz transformation.

According to Eq. (3.12.3) the map $\exp(iK^{\mu\nu}t_{\mu\nu})$ transforms a four momentum vector q^μ into

$$q'^\mu = \exp(iK^{\mu\nu}t_{\mu\nu})^\mu{}_\nu q^\nu =: \exp(K)^\mu{}_\nu q^\nu \tag{3.12.9}$$

So we can find the stability subgroup by demanding that $q' = q$ and see that $K^{\mu\nu} = \varepsilon^{\mu\nu\rho\delta} q_\rho n_\delta$ for an arbitrary vector n. So the stability subgroup for q consists of elements

$$\exp\left[i\left(b^\mu p_\mu + \varepsilon^{\mu\nu\rho\delta} q_\rho n_\delta t_{\mu\nu}\right)\right] \tag{3.12.10}$$

for arbitrary b and n. The way we have constructed the elements of the stability subgroup of q look like

$$\exp\left[i\left(b^\mu p_\mu + \varepsilon^{\mu\nu\rho\delta} q_\rho n_\delta t_{\mu\nu}\right)\right] = \exp(\alpha)\exp(-in_\mu w^\mu)\,. \tag{3.12.11}$$

The irreducible representations of the stability subgroup are, hence, given by irreducible representations of the algebra of the Pauli-Lubanski vector. We can show that for $C_1 = 0$ (i.e. a massless representation) the only finite-dimensional representation which is relevant for physics is a discrete quantum number $\lambda = 0, \pm 1/2, \pm 1, \pm 3/2$ called *helicity*. In the massive case $C_1 > 0$ the Pauli-Lubanski vector describes a $\mathfrak{su}(2)$ algebra, and the finite dimensional representations can be labelled with a Casimir $m^2 s(s+1)$, where $s = 0, 1/2, 1, 3/2$ is the spin. This again tells us that a Dirac fermion has spin 1/2.

3.13 *Exercises for Section 3*

Exercise 16

Using

$$\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$$

show that the rotation group forms a Lie algebra similar to $\mathfrak{su}(2)$ by verifying the commutation relations.

Exercise 17

Verify that the adjoint representation is a representation of a Lie algebra. Find a representation of the adjoint representation of $\mathfrak{su}(2)$ and verify the Lie algebra relations. Does this representation give rise to a Clifford algebra?

Show that $[t_a t^a, t^b] = 0$ for general $\mathfrak{su}(n)$ Lie algebras.

Exercise 18

Calculate the Casimir operator explicitly for the fundamental and the adjoint representations of $\mathfrak{su}(2)$.

Exercise 19

Show that the adjoint representation has dimension $n^2 - 1$ for $\mathfrak{su}(n)$. Use this result to show that the Casimir value the fundamental representation of $\mathfrak{su}(n)$ is

$$C_F = \frac{n^2 - 1}{2n} .$$

Exercise 20

Formulate the relation of $SO(3) \simeq SU(2)/\mathbb{Z}_2$ explicitly, i.e. find the expression for the matrix $\Lambda(\mathcal{U})$

$$\Lambda^{ab}(\mathcal{U}) = \frac{1}{2}\mathrm{Tr}\,\left\{\mathcal{U}^\dagger \sigma^a \mathcal{U} \sigma^b\right\} .$$

Exercise 21

Find the irreducible representations' modules from combining two spin 1 systems.

Exercise 22

Find the matrix representations for spin-1/2 representations of $SU(2)$ transformations that correspond to rotation around the x, y, z axes respectively.

Exercise 23

Use the result of the previous exercise and the identification $SU(2) \to SO(3)$ to derive representations of the elements as three-dimensional rotations.

Exercise 24

Write down the Cartan matrix for the following Dynkin diagram:

Exercise 25

Show for $SU(3)$ that $v_i = \varepsilon_{ijk} v^j v^k$ does transform under the complex conjugate representation.

Exercise 26

Decompose the $\mathfrak{su}(3)$ module

$$\square \otimes \square \otimes \square$$

into irreducible modules. Determine the modules' dimensions using the hook formula.

Exercise 27

Exercise 28

Show for $\mathfrak{su}(3)$ that

$$6 \otimes 3 = 10 \oplus 8\,, \quad 6 \otimes \bar{3} = 15 \oplus 3\,, \quad 8 \otimes 8 = 27 \oplus 10 \oplus \overline{10} \oplus 8 \oplus 8 \oplus 1\,.$$

Show for $\mathfrak{su}(5)$ that

$$5 \otimes 5 = 15 \oplus \overline{10}\,, \quad \overline{10} \otimes \overline{10} = 45 + \overline{50} + \bar{5}\,.$$

Exercise 29

Show that the (anti)symmetric combination of two n-dimensional vectors has dimension

$$d_s = \frac{n(n+1)}{2}\,, \quad d_a = \frac{n(n-1)}{2}$$

such that $d_s + d_a = d = n^2$.

Exercise 30

Derive Eq. (3.9.1) explicitly and show that there are $n(n-1)/2$ generators for $\mathfrak{so}(n)$. Is this what you would have expected for the case $n = 3$?

Exercise 31

Show that the representations $\mathcal{R}(e_a) = \sigma^a$ and $\mathcal{R}'(e_a) = -\sigma^a$ are not equivalent.

Exercise 32

Show explicitly that $\underline{x}'^{\dagger} = \underline{x}'$ in Eq. (3.11.2) and derive an expression for the associated $\Lambda(\mathcal{U})$.

Exercise 33

Show explicitly that the gamma matrices as defined in Eq. (3.11.15) form the Clifford algebra Eq. (3.11.16).

Exercise 34

In String Theory we deal with a fundamental space-time which is 10 or 11 dimensional. By consistently compactifying 6 or 7 dimensions, we obtain the "normal" pseudo-euclidean four dimensions. A potentially important intermediate step is a theory defined in 1+4 dimensions instead of 1+3 (our world). How many components does a spinor of a 5d Lorentz-invariant theory have? Can we make a distinction between left- and right-handed spinors?

Appendix*

A Elements of Mathematical Methods*

Before we go into more details we have to remind ourselves of the central algebraic structures that play a role in group theory. We repeat this part of MM2 and lecture course's part 1 to make these notes self-contained.

A.1 Groups

We collect the axioms that give rise to a group:

(i) $(A, \star)$ is an algebraic structure (well-definition)

(ii) $\forall a, b, c \in A : \ (a \star b) \star c = a \star (b \star c)$ (associativity)

(iii) $\exists n \in A \, \forall a \in A : \ a \star n = n \star a = a$ (existence of a neutral element)

(iv) $\forall a \in A \, \exists \bar{a} \in A : \ a \star \bar{a} = \bar{a} \star a = n$ (existence of inverse elements)

Furthermore, if

(v) $\forall a, b \in A : \ a \star b = b \star a$ (commutativity)

the group is called *abelian*.

A.2 Rings

A ring is an algebraic structure $(R, +, \cdot)$ with two operations

(i) $(R, +)$ is an abelian group

(ii) $(R, \cdot)$ is associative

(iii) The following *distributive properties* are fulfilled $\forall a, b, c \in R$

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c \,, \\ (b + c) \cdot a &= b \cdot a + c \cdot a \,. \end{aligned} \tag{A.2.1}$$

If $(R, \cdot)$ is commutative, then both distributive properties are identical and we can drop the second property from the definition.

Some textbooks introduce a Ring together with the existence of a one element of $(R, \cdot)$.

A.3 Fields

There is a special class of rings, which justify the introduction of yet another name for this algebraic object - a *field*. A field is a ring $(R, +, \cdot)$ for which $(R \backslash \{0\}, \cdot)$ is an abelian group.

A.4 Vector Spaces

Let F be a field and V a set. Let there be a structure $V \times V \to V$ with $(\underline{a}, \underline{b}) \mapsto \underline{a} + \underline{b}$ (summation) such that $(V, +)$ is an abelian group. Furthermore let there be an "external" operation $F \times V \to V$ with $(\alpha, \underline{a}) \mapsto \alpha\underline{a}$ (F-multiplication) such that for all $\alpha, \beta \in F$ and $\underline{a}, \underline{b} \in V$

(i) $1\underline{a} = \underline{a}$

(ii) $\alpha(\beta\underline{a}) = (\alpha\beta)\underline{a}$

(iii) $(\alpha + \beta)\underline{a} = \alpha\underline{a} + \beta\underline{a}$

(iv) $\alpha(\underline{a} + \underline{b}) = \alpha\underline{a} + \alpha\underline{b}$

Then V is called a *vector space over the field* F. The elements of V are called *vectors*, the elements of F are called *scalars*. In this course the scalars are elements of the real or complex numbers.

A.5 Modules

A module is the generalisation of a vector space, replacing fields with rings. As such this algebraic structure has a broader application and provides the best playground to study representation theory. For instance, if M is a set, R is a ring with one element and we have an operation "+" such that $(M, +)$ is an abelian group, and we define a map $(R, M) \to M$ such that for $\alpha, \beta \in R$ and $\underline{a}, \underline{b} \in M$ we have (i),(ii),(iii),(iv) from the previous section, then we call M a left R-module. Since there is not necessarily a concept of commutativity in R we need to make a distinction between left and right modules.

For the remainder of the course this will not make a difference between vector spaces and modules, for our studies of classical Lie algebras we will not need an algebraic object that is so broadly defined. Although we will refer to the underlying space as a module, you can always think of it as a vector space, however, it is important to keep in mind that we could also operate on a more general space.

A.6 Algebras

Let A be a set and F a field. Let there be two algebraic structures $+, \otimes$ on A: $A \times A \to A$ and an external operation $\cdot : F \times A \to A$ such that $(A, +, \cdot)$ is a F-vector space and $(A, +, \otimes)$ is a ring with one-element and the condition

$$\forall \underline{a}, \underline{b} \in A \; \forall \lambda \in F : \; (\lambda \cdot \underline{a}) \otimes \underline{b} = \underline{a} \otimes (\lambda \cdot \underline{b}) = \lambda \cdot (\underline{a} \otimes \underline{b}) \tag{A.6.1}$$

is fulfilled. Then we call $(A, +, \cdot, \otimes)$ an *algebra*.

Particularly important for our purposes is the algebra of quadratic matrices with the matrix multiplication operation (you can easily show that this defines an algebra).

A.7 Tensors

Another important concept in representation theory that we will need later is the tensor. Let V be a n-dimensional F-vector space and $V^\star$ the corresponding dual vector space. A multi-linear form

$$T : \underbrace{V^\star \times \cdots \times V^\star}_{r \text{ times}} \times \underbrace{V \times \cdots \times V}_{s \text{ times}} \longrightarrow F \tag{A.7.1}$$

is called a tensor of rank (r, s) – s-times covariant (i.e. the indices transform co-variantly with a basis transformation on V) and r-times contravariant (the indices transform with the inverse of a basis transformation). For given bases of $V, V^\star$ we can express a tensor as F-scalars

$$T\left[(\underline{b}^{\star\, i_1}, \ldots, \underline{b}^{\star\, i_r}, \underline{b}_{j_1}, \ldots, \underline{b}_{j_s})\right] = T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \,. \tag{A.7.2}$$

According to this definition, a vector is, e.g., a rank $(1, 0)$ tensor.

Tensors are central objects in theoretical physics because they allow us to write equations in a form-invariant way. Formulating the equations of motion in terms of tensors, we can choose a coordinate system, in which the equations are easier to solve. After finding the solution we can use the transformation properties to find solutions in different coordinate systems. If we formulate the equations of motion in terms of covariant tensors, we can directly infer the transformation from the involved basis transformation. One of the most prominent examples of such a set of equations are the Einstein field equations of General Relativity

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \,. \tag{A.7.3}$$

In this equation, $R_{\mu\nu}$ is the rank $(0, 2)$ curvature (or Ricci) tensor of space-time, which is a contraction of a $(1, 3)$ tensor. R is the Ricci scalar (a rank $(0, 0)$ tensor), $g_{\mu\nu}$ is the metric tensor and $T_{\mu\nu}$ is the energy-momentum tensor. Transforming into a system at rest, the $T_{\mu\nu}$ typically simplify and finding the solution of ten coupled partial differential equations becomes easier.