3

Lie Group Theory

Chapter Overview

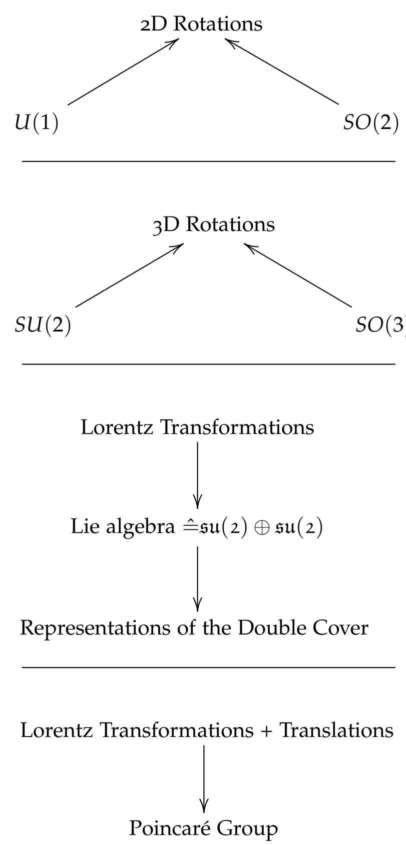
The final goal of this chapter is the derivation of the **fundamental representations of the double cover of the Poincaré group**, which we assume is the fundamental symmetry group of spacetime. These fundamental representations are the tools needed to describe all elementary particles, each representation for a different kind of elementary particle. The representations will tell us what types of elementary particles exist in nature.

We start with the definition of a **group**, which is motivated by two easy examples. Then, as a first step towards Lie theory we introduce two ways for describing rotations in two dimensions:

- 2×2 rotation matrix and
- unit complex numbers.

Then we will try to find a similar second description of rotations in three dimensions. This leads us to an extremely important group, called[1] **SU(2)**. After that, we will learn about **Lie algebras**, which enable us to learn a lot about something difficult (a Lie group) by using something simpler (the corresponding Lie algebra). There are in general many groups with the same Lie algebra, but only one of them is truly fundamental. We will use this knowledge to reveal the true fundamental symmetry group of nature, which doubly covers the Poincaré group. We usually start with some known transformations, derive the Lie algebra and use this Lie algebra to get different representations of the symmetry transformations. This will enable us to see that the representation we started with is just one special case out of many. This knowledge can then be used to learn something fundamental about the **Lorentz group**, which is an important part of the

This diagram explains the structure of this chapter. You should come back here whenever you feel lost. There is no need to spend much time here at a first encounter.



[1] The S stands for special, which means $\det(M) = 1$. U stands for unitary: $M^\dagger M = 1$ and the number 2 is used because the group is defined in the first place by 2×2 matrices.

J. Schwichtenberg, *Physics from Symmetry*, Undergraduate
Lecture Notes in Physics, https://doi.org/10.1007/978-3-319-66631-0_3

Poincaré group. We will see that the Lie algebra of the double cover of the Lorentz group consists of two copies of the $SU(2)$ Lie algebra. Therefore, we can directly use everything we learned about $SU(2)$. Finally, we include translations into the considerations, which leads us to the Poincaré group. The Poincaré group is the Lorentz group plus translations. At this point, we will have everything at hand to classify the fundamental representations of the double cover of the Poincaré group. We use these fundamental representations in later chapters to derive the fundamental laws of physics.

3.1 Groups

If we want to utilize the power of symmetry, we need a framework to deal with symmetries mathematically. The branch of mathematics that deals with symmetries is called **group theory**. A special branch of group theory that deals with continuous symmetries is **Lie Theory**.

Symmetry is defined as invariance under a set of of transformations and therefore, one defines a group as a collection of transformations. Let us get started with two easy examples to get a feel for what we want to do:

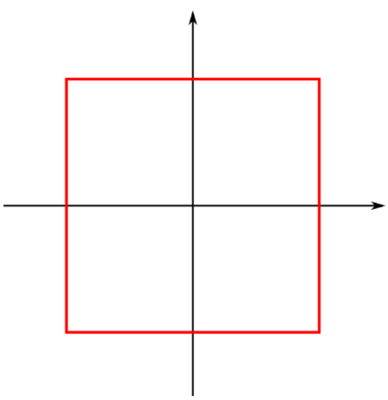


Fig. 3.1: Illustration of a square

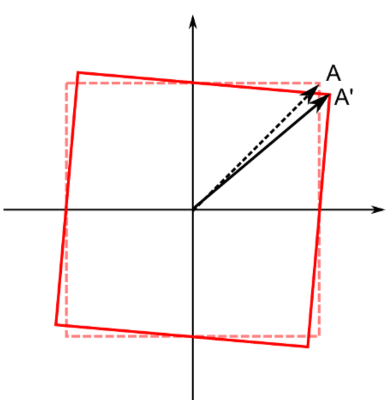


Fig. 3.2: Illustration of a square, rotated by 5°

1. A square is mathematically a set of points (for example, the four corner points are part of this set) and a symmetry of the square is a transformation that maps this set of points into itself.

   Examples of symmetries of the square are rotations about the origin by 90°, 180°, 270° or 0°. These rotations map the square into itself. This means they map every point of the set to a point that lies again in the set and one says the set is invariant under such transformations.

   Take note that not every rotation is a symmetry of the square. This becomes obvious when we focus on the corner points of the square. Transforming the set by a clockwise rotation by, say 5°, maps these points into points outside the original set that defines the square. For example, as shown in Figure 3.2, the corner point A is mapped to the point A', which is not found inside the set that defined the square in the first place. Therefore a rotation by 5° is not a symmetry of the square. Of course the rotated object is still a square, but a different square (=different set of points). Nevertheless, a rotation by 90° is a symmetry of the square because the point A is mapped to the point B, which lies again in the original set. This is shown in Fig. 3.3.

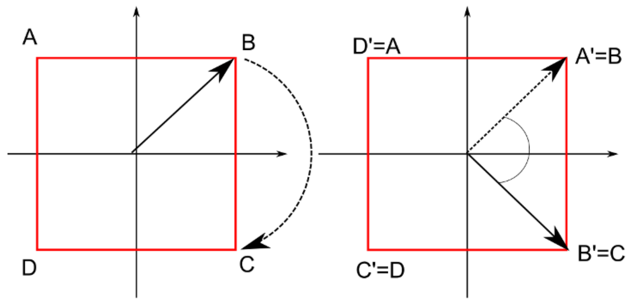


Fig. 3.3: Illustration of the square rotated by 90°

Here's another helpful perspective: Imagine you close your eyes for a moment, and someone transforms a square in front of you. If you can't tell after opening your eyes whether the person changed anything at all, then the transformation was a symmetry transformation.

The set of transformations that leave the square invariant is called a group. The transformation parameter, here the rotation angle, can't take on arbitrary values and the group is called a discrete group.

2. Another example is the set of transformations that leave the unit circle invariant. Again, the unit circle is defined as a set of points and a symmetry transformation is a map that maps this set into itself.

   The unit circle is invariant under all rotations about the origin, not just a few. In other words: the transformation parameter (the rotation angle) can take on arbitrary values, and the group is said to be a continuous group.

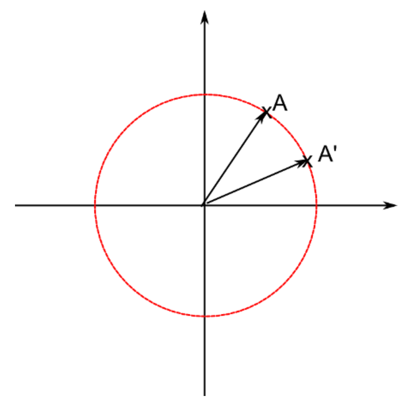


Fig. 3.4: Illustration of the rotation of the unit circle. For arbitrary rotations about the origin, all rotated points lie again in the initial set.

We are, of course, not only interested in symmetries of geometric shapes. For examples, considering vectors, we can look at the set of transformations that leave the length of any vector unchanged. For this reason, the definition of symmetry at the beginning of this chapter was very general: Symmetry means invariance under a transformation. Luckily, there is **one** mathematical theory, called group theory, that lets us work with **all** kinds of symmetries[2]

[2] As a side-note: Group theory was invented historically to describe symmetries of equations

To make the idea of a mathematical theory that lets us deal with symmetries precise, we need to distill the defining features of symmetries into mathematical terms:

- Leaving the object in question unchanged ("doing nothing") is always a symmetry and therefore, every group needs to contain an identity element. In the examples above, the identity element is the rotation by $0°$.

- Transforming something and afterwards doing the inverse transformation must be equivalent to doing nothing. Therefore, there must be, for every element in the set, an inverse element. A transformation followed by its inverse transformation is, by definition of the inverse transformation, the same as the identity transformation. In the above examples this means that the inverse transformation of a rotation by 90° is a rotation by -90°. A rotation by 90° followed by a rotation by -90° is the same as a rotation by 0°.

- Performing a symmetry transformation followed by a second symmetry transformation is again a symmetry transformation. A rotation by 90° followed by a rotation by 180° is a rotation by 270°, which is a symmetry transformation, too. This property of the set of transformations is called closure.

- The combination of transformations must be associative[3]. A rotation by 90° followed by a rotation by 40°, followed by a rotation by 110° is the same as a rotation by 130° followed by a rotation by 110°, which is the same as a rotation by 90° followed by a rotation by 150°. In a symbolic form:

$$R(110^\circ)R(40^\circ)R(90^\circ) = R(110^\circ)\big(R(40^\circ)R(90^\circ)\big) = R(110^\circ)R(130^\circ) \tag{3.1}$$

  and

$$R(110^\circ)R(40^\circ)R(90^\circ) = \big(R(110^\circ)R(40^\circ)\big)R(90^\circ) = R(150^\circ)R(90^\circ) \tag{3.2}$$

  This property is called **associativity**.

- To be able to talk about the things above, one needs a rule, to be precise: a **binary operation**, for the combination of group elements. In the above examples, the standard approach would be to use rotation matrices[4] and the rule for combining the group elements (the corresponding rotation matrices) would be ordinary matrix multiplication. Nevertheless, there are often different ways to describe the same thing[5] and group theory enables us to study this systematically. The branch of group theory that deals with different descriptions of the same transformations is called **representation theory**[6].

To work with ideas like these in a rigorous, mathematical way, one distils the defining features of such transformations and promotes them to axioms. All structures satisfying these axioms are then called groups. This paves the way for a whole new branch of mathematics, called group theory. It is possible to find very abstract structures

[3] But not commutative! For example rotations around different axes do not commute. This means in general: $R_x(\theta)R_z(\Phi) \neq R_z(\Phi)R_x(\theta)$

[4] If you want to know more about the derivation of rotation matrices have a look at Appendix A.2.

[5] For example, rotations in the plane can be described alternatively by multiplication with unit complex numbers. The rule for combining group elements is then complex number multiplication. This will be discussed later in this chapter.

[6] Representation theory is the topic of Section 3.5.

satisfying the group axioms, but we will stick with groups that are very similar to the rotations we explored above.

After the discussion above, we can see that the abstract definition of a group simply states (obvious) properties of symmetry transformations:

A group $(G, \circ)$ is a set G, together with a binary operation $\circ$ defined on G, that satisfies the following axioms[7]

[7] Do not worry too much about this. In practice one checks for some kind of transformation if they obey these axioms. If they do, the transformations form a group and one can use the results of group theory to learn more about the transformations in question.

- Closure: For all $g_1, g_2 \in G$, $g_1 \circ g_2 \in G$
- Identity element: There exists an identity element $e \in G$ such that for all $g \in G$, $g \circ e = g = e \circ g$
- Inverse element: For each $g \in G$, there exists an inverse element $g^{-1} \in G$ such that $g \circ g^{-1} = e = g^{-1} \circ g$.
- Associativity: For all $g_1, g_2, g_3 \in G$, $g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$.

To summarize: The set of all transformations that leave a given object invariant is called a symmetry group. For Minkowski spacetime, the object that is left invariant is the Minkowski metric[8] and the corresponding symmetry group is called the Poincaré group.

[8] Recall, this is the tool which we use to compute distances and lengths in Minkowski space.

Take note that the characteristic properties of a group are defined completely independent of the object the transformations act on. We can therefore study such symmetry transformations without making references to any object we extracted them from. This is very useful, because there can be many objects with the same symmetry or at least the same kind of symmetry. Using group theory we no longer have to inspect each object on its own, but are now able to study general properties of symmetry transformations.

3.2 Rotations in two Dimensions

As a first step, we start with an easy, but important, example. What transformations in two dimensions leave the length of any vector unchanged? After thinking about it for a while, we come up with[9] rotations and reflections. These transformations are of course the same ones that map the unit circle into the unit circle. This is an example of how one group may act on different kinds of objects: On the circle, which is a geometric shape, and on a vector. Considering vectors, one can represent these transformations by rotation matrices[10],

[9] Another kind of transformation that leaves the length of a vector unchanged are translations, which means we move every point a constant distance in a specified direction. These are described mathematically a bit different and we are going to talk about them later.

[10] For an explicit derivation of these matrices have a look a look at Appendix A.2.

which are of the form

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{3.3}$$

and describe two-dimensional rotations about the origin by angle θ. Reflections at the axes can be performed using the matrices:

$$P_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad P_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{3.4}$$

You can check that these matrices, together with the ordinary matrix multiplication as binary operation $\circ$, satisfy the group axioms and therefore these transformations form a group.

We can formulate the task of finding "all transformations in two dimensions that leave the length of any vector unchanged" in a more abstract way. The length of a vector is given by the scalar product of the vector with itself. If the length of the vector is the same after the transformation $a \to a'$, the equation

$$a' \cdot a' \stackrel{!}{=} a \cdot a \tag{3.5}$$

must hold. We denote the transformation with O and write the transformed vector as $a \to a' = Oa$. Thus

$$a \cdot a = a^T a \to a'^T a' = (Oa)^T Oa = a^T O^T Oa \stackrel{!}{=} a^T a = a \cdot a, \tag{3.6}$$

where we can see the condition a transformation must fulfil to leave the length of a vector unchanged is

$$O^T O = I, \tag{3.7}$$

where I denotes the unit matrix[11]. You can check that the well-known rotation and reflection matrices we cited above fulfil exactly this condition[12]. This condition for two dimensional matrices defines the group $O(2)$, which is the group of all[13] orthogonal 2×2 matrices. We can find a subgroup of this group that includes only rotations, by observing that it follows from the condition in Eq. 3.7 that

$$\begin{aligned} &\det(O^T O) \stackrel{!}{=} \det(I) = 1 \\ &\to \det(O^T O) = \det(O^T)\det(O) \stackrel{!}{=} \det(I) = 1 \\ &\to (\det(O))^2 \stackrel{!}{=} 1 \to \det(O) \stackrel{!}{=} \pm 1 . \end{aligned} \tag{3.8}$$

The transformations of the group with $\det(O) = 1$ are rotations[14] and the two conditions

$$O^T O = I \tag{3.9}$$

[11] $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

[12] With the matrix from Eq. 3.3 we have $R_\theta^T R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos^2(\theta) + \sin^2(\theta) & 0 \\ 0 & \sin^2(\theta) + \cos^2(\theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \checkmark$

[13] Every orthogonal 2×2 matrix can be written either in the form of Eq. 3.3, as in Eq. 3.4, or as a product of these matrices.

[14] As can be easily seen by looking at the matrices in Eq. 3.3 and Eq. 3.4. The matrices with $\det O = -1$ are reflections.

$$\det O = 1 \tag{3.10}$$

define the $SO(2)$ group, where the "S" denotes special and the "O" orthogonal. The special thing about $SO(2)$ is that we now restrict it to transformations that keep the system orientation, i.e., a right-handed[15] coordinate system stays right-handed. In the language of linear algebra this means that the determinant of our matrices must be $+1$.

[15] If you don't know the difference between a right-handed and a left-handed coordinate system have a look at the Appendix A.5.

3.2.1 Rotations with Unit Complex Numbers

There is a different way to describe rotations in two dimensions that makes use of complex numbers: rotations about the origin by angle θ can be described by multiplication with a unit complex number ($z = a + ib$ which fulfils the condition[16] $|z|^2 = z^\star z = 1$).

[16] The $\star$ symbol denotes complex conjugation: $z = a + ib \rightarrow z^\star = a - ib$

The unit complex numbers together with ordinary complex number multiplication are a group, called[17] $U(1)$, as you can check by looking at the group axioms. To establish the connection with the group definitions for $O(2)$ and $SO(2)$ introduced above, we write the defining condition for unit numbers as[18]

[17] The U stands for unitary, which means the condition $U^\dagger U = 1$, where the symbol "$\dagger$", called "dagger", denotes transposition plus complex conjugation: $U^\dagger = U^{T\star}$. For numbers, the dagger operation $\dagger$ is the same as taking only the complex conjugate, because every number satisfies trivially $z^T = z$. Hence the condition reduces here simply to $U^\star = U$.

[18] For more general information about the definition of groups involving a complex product, have a look at the appendix in Section 3.10.

$$U^\star U = 1\,. \tag{3.11}$$

Another way to write a unit complex number is[19]

[19] This is known as Euler's formula, which is derived in Appendix B.4.2. For a complex number $z = a + ib$, a is called the real part of z: $Re(z) = a$ and b the imaginary part: $Im(z) = b$. In Euler's formula $\cos(\theta)$ is the real part, and $\sin(\theta)$ the imaginary part of R_θ.

$$R_\theta = \mathrm{e}^{i\theta} = \cos(\theta) + i\sin(\theta), \tag{3.12}$$

because then

$$R_\theta^\star R_\theta = \mathrm{e}^{-i\theta}\mathrm{e}^{i\theta} = \big(\cos(\theta) - i\sin(\theta)\big)\big(\cos(\theta) + i\sin(\theta)\big) = 1 \tag{3.13}$$

Let's have a look at an example: we rotate the complex number $z = 3 + 5i$, by $90°$:

$$z \rightarrow z' = \mathrm{e}^{i90°} z = (\underbrace{\cos(90°)}_{=0} + i\underbrace{\sin(90°)}_{=1})(3 + 5i) = i(3 + 5i) = 3i - 5\,. \tag{3.14}$$

The two complex numbers are plotted in Fig. 3.6 and we see the multiplication with $\mathrm{e}^{i90°}$ does indeed rotate the complex number by $90°$. In this description, the rotation operator $\mathrm{e}^{i90°}$ acts on complex numbers instead of on vectors. To describe a rotation in two dimensions, one parameter is necessary: the angle of rotation θ. A complex number has two degrees of freedom and with the constraint to unit complex numbers $|z| = 1$, one degree of freedom is left as needed.

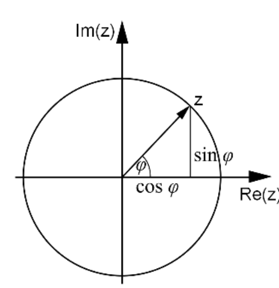


Fig. 3.5: The unit complex numbers lie on the unit circle in the complex plane.

We can connect this description to the previous one, where we used rotation matrices, by representing complex numbers by real 2×2 matrices. We define

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.15}$$

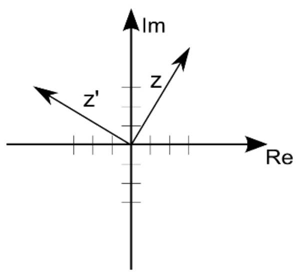


Fig. 3.6: Rotation of a complex number, by multiplication with a unit complex number

You can check that these matrices fulfil

$$1^2 = 1, \qquad i^2 = -1, \qquad 1i = i1 = i. \tag{3.16}$$

So now, the complex representation of rotations of the plane reads

$$R_\theta = \cos(\theta) + i\sin(\theta) = \cos(\theta)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin(\theta)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{3.17}$$

By making the map

$$i \to \text{real matrix},$$

we go back to the familiar representation of rotations of the plane. Maybe you have noticed a subtle point: The familiar rotation matrix in 2-dimensions acts on vectors, but here we identified the complex unit i with a real matrix (Eq. 3.15). Therefore, the rotation matrix will act on a 2×2 matrix, because the complex number we act on becomes a matrix, too.

A generic complex number in this description reads

$$z = a + ib = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \tag{3.18}$$

Let us take a look at how rotations act on such a matrix that represents a complex number:

$$z' = \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = R_\theta z = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$
$$= \begin{pmatrix} \cos(\theta)a - \sin(\theta)b & -\cos(\theta)b - \sin(\theta)a \\ \sin(\theta)a + \cos(\theta)b & -\sin(\theta)b + \cos(\theta)a \end{pmatrix}. \tag{3.19}$$

By comparing the left-hand side with the right-hand side, we get

$$\to a' = \cos(\theta)a - \sin(\theta)b$$
$$\to b' = \sin(\theta)a + \cos(\theta)b\,, \tag{3.20}$$

which is the same result that we get when we act with R_θ on a column vector

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos(\theta)a - \sin(\theta)b \\ \sin(\theta)a + \cos(\theta)b \end{pmatrix} = \begin{pmatrix} a' \\ b' \end{pmatrix}. \quad (3.21)$$

We see that both representations do exactly the same thing and mathematically speaking we have an isomorphism[20] between $SO(2)$ and $U(1)$. This is a very important discovery and we will elaborate on such lines of thought in the following chapters.

[20] An isomorphism is a one-to-one map Π that preserves the product structure $\Pi(g_1)\Pi(g_2) = \Pi(g_1 g_2)\ \forall g_1, g_2 \in G$.

Next we want to describe rotations in three dimensions and find similarly two descriptions for rotations in three dimensions[21].

[21] Things are about to get really interesting! Analogous to the two-dimensional case we discussed in the preceding section, we will find a second description of rotations in three dimensions and this alternative description will reveal something fundamental about nature.

3.3 Rotations in three Dimensions

The standard method to rotate vectors in three dimensions is to use 3×3 rotation matrices. The "basis rotations" around the three axes can be described by the following matrices:

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad R_y = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$

$$R_z = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.22)$$

If we want to rotate the vector

$$\vec{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

around the z-axis[22], we multiply it with the corresponding rotation matrix

[22] A general, rotated vector is derived explicitly in Appendix A.2.

$$R_z(\theta)\vec{v} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix}. \quad (3.23)$$

To get a second description for rotations in three dimensions, the first thing we have to do is find a generalisation of complex numbers in higher dimensions. A first guess may be to go from 2-dimensional complex numbers to 3-dimensional complex numbers, but it turns out that there are no 3-dimensional complex numbers. Instead, we can find 4-dimensional complex numbers, called quaternions. The quaternions will prove to be the correct second tool to describe rotations in 3-dimensions and the fact that this tool is 4-dimensional reveals something deep about the universe. We could have anticipated this result, because to describe an arbitrary rotation in 3-dimensions, 3 parameters are needed. Four dimensional complex numbers, with the constraint to unit quaternions[23], have exactly 3 degrees of freedom.

[23] Remember that we used the constraint to unit complex numbers in the two dimensional case, too.

3.3.1 Quaternions

The 4-dimensional complex numbers can be constructed analogous to the 2-dimensional complex numbers. Instead of just one complex "unit" we introduce three, named $\mathbf{i}, \mathbf{j}, \mathbf{k}$. These fulfil

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1. \tag{3.24}$$

Then a 4-dimensional complex number, called a quaternion, can be written as

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}. \tag{3.25}$$

We now need multiplication rules for $ij \; =?$ etc., because products like this will occur when one multiplies two quaternions. The extra condition

$$\mathbf{ijk} = -1 \tag{3.26}$$

suffices to compute all needed relations, for example **ij=k** follows from multiplying Eq. 3.26 with **k**:

$$\mathbf{ij}\underbrace{\mathbf{kk}}_{=-1} = -\mathbf{k} \rightarrow \mathbf{ij} = \mathbf{k}. \tag{3.27}$$

The set of *unit* quaternions $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ satisfy the condition[24]

$$q^\dagger q \stackrel{!}{=} 1$$
$$\rightarrow (a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k})(a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) = a^2 + b^2 + c^2 + d^2 \stackrel{!}{=} 1. \tag{3.28}$$

[24] The symbol †, here is called "*dagger*" and denotes transposition plus complex conjugation: $a^\dagger = (a^\star)^T$. The ordinary scalar product always includes a transposition $a \cdot b = a^T b$, because matrix multiplication requires that we multiply a row with a column. In addition, for complex entities we include complex conjugation that makes sure we get something real, which is important if we want to interpret things in terms of length.

Exactly as the unit complex numbers formed a group under complex number multiplication, the unit quaternions form a group under quaternion multiplication.

Analogous to what we did for two-dimensional complex numbers, we now represent each of the three complex units with a matrix. There are different ways of doing this, but one choice that does the job is as complex 2×2 matrices:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \mathbf{i} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\mathbf{j} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad , \quad \mathbf{k} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{3.29}$$

You can check that these matrices fulfil the defining conditions in Eq. 3.24 and Eq. 3.26. Using these matrices a generic quaternion can then be written

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = \begin{pmatrix} a+di & b+ci \\ -b+ci & a-di \end{pmatrix}. \tag{3.30}$$

Furthermore, we have

$$\det(q) = a^2 + b^2 + c^2 + d^2. \tag{3.31}$$

Comparing this with Eq. 3.28 tells us that the set of unit quaternions is given by matrices of the above form with unit determinant. The unit quaternions, written as complex 2×2 matrices therefore fulfil the conditions

$$U^\dagger U = 1 \quad \text{and} \quad \det(U) = 1. \tag{3.32}$$

Take note that the way we define $SU(2)$ here, is analogously to how we defined $SO(2)$. The S denotes special, which means $\det(U) = 1$ and U stands for unitary, which means the property[25] $U^\dagger U = 1$. Through the map in Eq. 3.29, every unit quaternion can be identified with an element of $SU(2)$.

[25] For some more information about this, have a look at the Appendix 3.10 at the end of this chapter.

Now, how is $SU(2)$ and with it the unit quaternions related to rotations? Unfortunately, the map between $SU(2)$ and[26] $SO(3)$ is not as simple as the one between $U(1)$ and $SO(2)$.

[26] Recall that $SO(3)$ is the set of the usual rotation matrices acting on 3 dimensional vectors.

In 2-dimensions the 2 parameters of a complex number $z = a + \mathbf{i}b$ could be easily identified with the two spatial axes, i.e. $v = x + \mathbf{i}y$. The restriction to unit complex numbers automatically makes sure that the resulting matrix preserves the length of any vector[27]

$$(Uz)^\star Uz = z^\star U^\star Uz = z^\star z.$$

[27] Recall that here R is a unit complex number, because complex numbers can be rotated by multiplication with unit complex numbers. Therefore we have $U^\star U = 1$, which is the defining condition for unit complex numbers.

The quaternions have 4 parameters, so the identification with the 3 coordinates of a three-dimensional vector is not obvious. When we define

$$v \equiv x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \tag{3.33}$$

and use the matrix representation of the quaternions (Eq. 3.29), we can compute

$$\det(v) = x^2 + y^2 + z^2. \tag{3.34}$$

Therefore, if we want to consider transformations that preserve the length of the vector (x, y, z), we must use matrix transformations that preserve determinants. Therefore the restriction to **unit** quaternions means that we must restrict to matrices with **unit** determinant[28]. Everything may now seem straight forward, but there is a subtle point. A first guess would be that a unit quaternion u induces a rotation on v simply by multiplication. This is not the case, because the product of u and v may not belong to $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$. Therefore, the transformed vector can have a component we are not able to interpret[29]. Instead the transformation that does the job is given by

$$v' = qvq^{-1}. \tag{3.35}$$

It turns out that by making this identification unit quaternions can describe rotations in 3-dimensions.

[28] This follows from the general rule $\det(BA) = \det(B)\det(A)$. Therefore, if B has determinant 1, the product matrix BA has the same determinant as A.

[29] We identify our spatial components x, y, z as components in the subspace $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$. If, as a result of a transformation, we end up with a coefficient that does not belong to this subspace, i.e. does not appear together with $\mathbf{i}$, $\mathbf{j}$ or $\mathbf{k}$, we can't interpret it.

Let's take a look at an explicit example: To make the connection to our example in two dimensions, we define u as a unit quaternion that only takes values in $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ and denote a general unit quaternion with

$$t = \cos(\theta) + \sin(\theta)u. \tag{3.36}$$

Using Eq. 3.33 a generic vector can be written

$$\vec{v} = (v_x, v_y, v_z)^T = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k} \underbrace{=}_{\text{Eq. 3.30}} \begin{pmatrix} iv_z & v_x + iv_y \\ -v_x + iv_y & -iv_z \end{pmatrix}. \tag{3.37}$$

With the identifications made above, we want to rotate, as an example, a vector $\vec{v} = (1, 0, 0)^T$ around the z-axis. We will make a particular choice for the vector and for the quaternion representing the rotation and show that it works. We write, using quaternions in their matrix representation (Eq. 3.29)

$$\vec{v} = (1, 0, 0)^T \rightarrow v = 1\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.38}$$

In addition, we use the following quaternion

$$R_z(\theta) = \cos(\theta)\mathbf{1} + \sin(\theta)\mathbf{k} = \begin{pmatrix} \cos(\theta) + i\sin(\theta) & 0 \\ 0 & \cos(\theta) - i\sin(\theta) \end{pmatrix}, \tag{3.39}$$

and then calculate that it is the correct quaternion that describes a rotation around the z-axis. We can rewrite R_z using Euler's formula[30] $\mathrm{e}^{ix} = \cos(x) + i\sin(x)$:

[30] For a derivation have a look at Appendix B.4.2.

$$\Rightarrow R_z(\theta) = \begin{pmatrix} \mathrm{e}^{i\theta} & 0 \\ 0 & \mathrm{e}^{-i\theta} \end{pmatrix}. \tag{3.40}$$

Inverting the quaternion rotation matrix yields

$$R_z(\theta)^{-1} = \begin{pmatrix} \cos(\theta) - i\sin(\theta) & 0 \\ 0 & \cos(\theta) + i\sin(\theta) \end{pmatrix} = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}. \tag{3.41}$$

Using Eq. 3.35 the rotated vector reads

$$v' = R_z(\theta) v R_z^{-1}(\theta) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & e^{i2\theta} \\ -e^{-2i\theta} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \cos(2\theta) + i\sin(2\theta) \\ -\cos(2\theta) + i\sin(2\theta) & 0 \end{pmatrix}. \tag{3.42}$$

On the other hand, an arbitrary vector can be written in our quaternion notation (Eq. 3.37)

$$v' = \begin{pmatrix} iv'_z & v'_x + iv'_y \\ -v'_x + iv'_y & -iv'_z \end{pmatrix}, \tag{3.43}$$

which we now compare with Eq. 3.42. This yields

$$v'_x = \cos(2\theta) \quad , \quad v'_y = \sin(2\theta) \quad , \quad v'_z = 0. \tag{3.44}$$

Therefore, written again in the conventional vector notation

$$\rightarrow \vec{v}' = (\cos(2\theta), \sin(2\theta), 0)^T. \tag{3.45}$$

Our identifications do indeed induce rotations[31], but something needs our attention. We haven't rotated $\vec{v}$ by θ, but by 2θ. Therefore, we define $\phi \equiv 2\theta$, because then ϕ really represents the angle we rotate. Using this definition we rewrite Eq. 3.36, which yields

$$t = \cos(\frac{\phi}{2}) + \sin(\frac{\phi}{2}) u\,. \tag{3.46}$$

We can now see that the identifications we made are not one-to-one, but rather we have two unit-quaternions describing the same rotation. For example[32]

$$t_{\phi=\pi} = u \qquad\qquad t_{\phi=3\pi} = -u$$

$$\searrow \qquad \swarrow$$

Vector Rotation by π

This is the reason $SU(2)$ is called the **double-cover** of $SO(3)$. It is always possible to go unambiguously from $SU(2)$ to $SO(3)$ but not

[31] See the example in Eq. 3.23 where we rotated the vector, using the conventional rotation matrix

[32] A rotation by π is the same as a rotation by $3\pi = 2\pi + \pi$ for ordinary vectors, because $2\pi = 360°$ is a full rotation. In other words: We can see that two quaternions u and $-u$ can be used to rotate a vector by π.

vice versa. One may think this is just a mathematical side-note, but we will understand later that groups which cover other groups are indeed more fundamental[33].

[33] To spoil the surprise: We will use the double cover of the Lorentz group, instead of the Lorentz group itself, because otherwise we miss something important: **Spin**. Spin is some kind of internal momentum and one of the most important particle labels. This is discussed in detail in Section 4.5.4 and Section 8.5.5.

To be able to discover the group that covers a given group, we need to introduce the most important tool of Lie theory: Lie algebras. This is the topic of the next section.

Take note that the fact we had one quaternion parameter too many, may be interpreted as a hint towards relativity. One may argue that a more natural identification would have been, as in the two-dimensional case, $v = t\mathbf{1} + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. We see that pure mathematics pushes us towards the idea of using a 4th component and what could it be, if not time? If we now want to describe rotations in 4-dimensions, because we know that the universe we live in is 4-dimensional, we have two choices:

- We could search for even higher dimensional complex numbers or
- we could again try to work with quaternions.

From the last paragraph it may seem that quaternions have something to say about rotations in 4 dimensions, too. An arbitrary rotation in 4 dimensions is described by[34] 6 parameters. There is no 7-dimensional generalisation of complex numbers, which together with the constraint to unit objects would have 6 free parameters. However, **two** unit quaternions have exactly 6 free parameters. Therefore, maybe it's possible to describe a rotation in 4-dimensions by two quaternions? We will learn later that there is indeed a close connection between two copies of $SU(2)$ and rotations in four dimensions.

[34] Using ordinary matrices, we need in four dimensions 4×4 matrices. The two conditions $O^T O = 1$ and $\det(O) = 1$ reduce the 16 components of an arbitrary 4×4 matrix to six independent components.

3.4 Lie Algebras

Lie theory is all about continuous symmetries. An example is the continuous symmetry of the unit circle we discussed at the beginning of this chapter. Continuous here means that there are infinitely many symmetry transformations, that can be parametrized continuously by one or several parameters. For example, the rotation angle ϕ in the circle example can be any value: $0.1°$, $0.11°$, $0.11991°$, ... In contrast, for discrete symmetries like a reflection symmetry there is no continuous transformation parameter.

An important observation in Lie theory is that for continuous symmetries there are elements of the group which are arbitrarily close to the identity transformation[35]. In contrast, for a discrete group there is no continuous transformation parameter and therefore no element arbitrarily close to the identity. Consider again the symmetries of a

[35] The identity transformation is the transformation that changes nothing at all. For example, a rotation by $0°$ is an identity transformation.

square. A rotation by 0,000001°, which is very close to the identity transformation (= a rotation by 0°), is not in the set of symmetry transformations of the square. In contrast, a rotation by 0,000001° is a symmetry of the circle. The symmetry group of a circle is continuous, because the rotation parameter (the rotation angle) can take on arbitrary (continuous) values. Mathematically, with the identity denoted I, an element g close to the identity is denoted

$$g(\epsilon) = I + \epsilon X \,, \tag{3.47}$$

where ϵ is, as always in mathematics, a really, really small number and X is an object, called generator, we will talk about in a moment. Such small transformations, when acting on some object change barely anything. In the smallest possible case such transformations are called infinitesimal transformation. Nevertheless, repeating such an infinitesimal transformation often, results in a finite transformation. Think about rotations: many small rotations in one direction are equivalent to one big rotation in the same direction. Mathematically, we can write the idea of repeating a small transformation many times

$$h(\theta) = (I + \epsilon X)(I + \epsilon X)(I + \epsilon X)... = (I + \epsilon X)^k, \tag{3.48}$$

where k denotes how often we repeat the small transformation. If θ denotes some finite transformation parameter, e.g. 50° or so, and N is some really big number which makes sure we are close to the identity, we can write the element close to the identity as

$$g(\theta) = I + \frac{\theta}{N} X. \tag{3.49}$$

The transformations we want to consider are the smallest possible, which means N must be the biggest possible number, i.e. $N \to \infty$. To get a finite transformation from such a infinitesimal transformation, one has to repeat the infinitesimal transformation infinitely. Mathematically

$$h(\theta) = \lim_{N \to \infty} (I + \frac{\theta}{N} X)^N, \tag{3.50}$$

which is in the limit just the exponential function[36]

$$h(\theta) = \lim_{N \to \infty} (I + \frac{\theta}{N} X)^N = e^{\theta X} \,. \tag{3.51}$$

In some sense the object X generates the finite transformation h, which is why it's called the **generator**.

If we want to calculate the generator X of a given transformation, we can differentiate the above formula

[36] This is often used as a definition of the exponential function. A proof, showing the equivalence of this limit and the exponential series we derive in Appendix B.4.1, can be found in most books about analysis.

$$\frac{d}{d\theta}h(\theta) = \frac{d}{d\theta}e^{\theta X} = Xe^{\theta X}\,. \tag{3.52}$$

Thus, if we evaluate this formula at $\theta = 0$, we get

$$X = \frac{dh(\theta)}{d\theta}\Big|_{\theta=0}. \tag{3.53}$$

The idea behind such lines of thought is that one can learn a lot about a group by looking at the important part of the infinitesimal elements (denoted X above): the **generators**. This will be made more precise in a moment, but first let's look at this from another perspective that makes it even clearer in what sense the generators generate a finite transformation:

If we consider a continuous group of transformations that are given by matrices, we can make a Taylor expansion[37] of an element of the group about the identity. The Taylor series is given by

[37] If you've never heard of the Taylor expansion, or Taylor series before, you are encouraged to have a look at Appendix B.3.

$$h(\theta) = I + \frac{dh}{d\theta}.|_{\theta=0}\theta + \frac{1}{2}\frac{d^2h}{d\theta^2}.|_{\theta=0}\theta^2 + ... = \sum_n \frac{1}{n!}\frac{d^nh}{d\theta^n}\Big|_{\theta=0}\theta^n. \tag{3.54}$$

This shows nicely how the generators generate transformations.

For matrix Lie groups one defines the corresponding Lie algebra as the collection of objects that give an element of the group when exponentiated. This is an easy definition one can use when restricting to matrix Lie groups. Later we will introduce a more general definition. In mathematical terms[38]

[38] The Lie algebra which belongs to a group G is conventionally denoted by the corresponding "Fraktur" letter $\mathfrak{g}$

For a Lie Group G (given by $n \times n$ matrices), the Lie algebra $\mathfrak{g}$ of G is given by those $n \times n$ matrices X such that $e^{tX} \in G$ for $t \in \mathbb{R}$, together with an operation, called the Lie bracket $[,]$ that tells us how we can combine these matrices.

The last part of this definition can be a bit confusing and thus we now spent some time discussing it.

We know from the definition of a group, that a group is more than just a collection of transformations. The definition of a group includes a binary operation $\circ$ that tells us how to combine group elements. For matrix Lie groups this is just ordinary matrix multiplication. Naively one may think that the same combination rule $\circ$ is valid for elements of the Lie algebra, but this is not the case! The elements of the Lie algebra are given by matrices[39], but the multiplication of two Lie algebra elements doesn't need to be an element of the Lie algebra. Instead there is another combination rule for the Lie

[39] A famous theorem of Lie group theory, called Ado's Theorem, states that every Lie algebra is isomorphic to a matrix Lie algebra.

algebra, already mentioned in the definition above, that is directly connected to the combination rule of the corresponding Lie group.

The connection between the combination rule of the Lie group and the combination rule of the Lie algebra is given by the famous Baker-Campbell-Hausdorff formula[40]

$$e^X \circ e^Y = e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\dots} \tag{3.55}$$

On the left hand side, we have the multiplication of two elements of the Lie group, let's name them g and h, which we can write in terms of the corresponding generators (=elements of the Lie algebra)

$$\underbrace{g}_{\in G} \circ \underbrace{h}_{\in G} = e^X \circ e^Y = \underbrace{e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\dots}}_{\in G} \tag{3.56}$$

with the generators[41] $X, Y \in \mathfrak{g}$. On the right-hand side we have a single object of the group and the multiplication of the group elements have been translated to a sum of Lie algebra elements. The new symbol in this sum [,] is called **Lie bracket** and for matrix Lie groups it is given by $[X,Y] = XY - YX$, which is called the **commutator** of X and Y. The elements XY and YX need not to be part of the Lie algebra, but their difference always is[42]!

We learn from the Baker-Campbell-Hausdorff-Formula that the natural product of the Lie algebra is, not as one would naively think, ordinary matrix multiplication, but the Lie bracket [,]. One says, the Lie algebra is **closed** under the Lie bracket, just as the group is closed under the corresponding composition rule $\circ$, e.g. matrix multiplication. Closure means that the composition of two elements lies again in the same set[43].

After looking at an example to illustrate these new notions, we will have a look at the modern definition of a Lie algebra. The main component of this definition is how the generators of a group behave when put into the Lie bracket. By using this general definition we will see that it is possible to say that different groups have the **same** Lie algebra. With the definition above saying something like this would make little sense. Nevertheless, this new way of thinking about Lie algebras will enable us to reveal the most fundamental description corresponding to a given transformation. This is possible because there is a theorem in Lie theory that tells us that there is exactly one **distinguished** Lie group for each Lie algebra. The Lie algebra however, according to the abstract definition, corresponds to many Lie groups. We will make this more concrete after introducing the modern definition of a Lie group.

[40] We will not talk about the proof of this formula in this book. Proofs can be found in most books about Lie theory, for example in William Fulton and Joe Harris. *Representation Theory: A First Course*. Springer, 1st corrected edition, 8 1999. ISBN 9780387974958

[41] The Lie algebra which belongs to a group G is conventionally denoted by the corresponding "Fraktur" letter $\mathfrak{g}$.

[42] A very illuminating proof of this fact can be found in John Stillwell. *Naive Lie Theory*. Springer, 1st edition, August 2008. ISBN 978-0387782140

[43] For group elements $g, h \in G$ we have $g \circ h \in G$. For elements of the Lie algebra $X, Y \in \mathfrak{g}$ we have $[X,Y] \in \mathfrak{g}$ and in general $X \circ Y \notin \mathfrak{g}$

Now we want to take a look at an explicit example of how we can derive the Lie algebra of a given group.

3.4.1 The Generators and Lie Algebra of $SO(3)$

The defining conditions of the group $SO(3)$ are (Eq. 3.10)

$$O^T O \stackrel{!}{=} I \qquad \text{and} \qquad \det(O) \stackrel{!}{=} 1. \tag{3.57}$$

We can write every group element O in terms of a generator J:

$$O = \mathrm{e}^{\theta J}. \tag{3.58}$$

Putting this into the first defining condition yields

$$O^T O = \mathrm{e}^{\theta J^T} \mathrm{e}^{\theta J} \stackrel{!}{=} 1 \to J^T + J \stackrel{!}{=} 0. \tag{3.59}$$

Using the second condition in Eq. 3.57 and the identity[44] $\det(\mathrm{e}^A) = \mathrm{e}^{\mathrm{tr}(A)}$ for the matrix exponential, we see

$$\begin{aligned}(O) \stackrel{!}{=} 1 \quad &\to \quad \det(\mathrm{e}^{\theta J}) = \mathrm{e}^{\theta \mathrm{tr}(J)} \stackrel{!}{=} 1 \\ &\to \quad \mathrm{tr}(J) \stackrel{!}{=} 0\,. \end{aligned} \tag{3.60}$$

[44] $\mathrm{tr}(A)$ denotes the trace of the matrix A, which means the sum of all elements on the main diagonal. For example for $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, we have $\mathrm{tr}(A) = A_{11} + A_{22}$.

Three linearly independent[45] matrices fulfilling both conditions (Eq. 3.59, Eq. 3.60) are

[45] This is explained in Appendix A.1.

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.61}$$

These matrices form a **basis** for the generators of the group $SO(3)$. This means any generator of the group can be written as a linear combination of these basis generators: $J = aJ_1 + bJ_2 + cJ_3$, where a, b, c denote real constants. These generators can be written more compactly by using the Levi-Civita symbol[46]

[46] The Levi-Civita symbol is explained in Appendix B.5.5.

$$(J_i)_{jk} = -\epsilon_{ijk}, \tag{3.62}$$

where j, k denote the components of the generator J_i. For example,

$$\begin{aligned}(J_1)_{jk} = -\epsilon_{1jk} \leftrightarrow \begin{pmatrix} (J_1)_{11} & (J_1)_{12} & (J_1)_{13} \\ (J_1)_{21} & (J_1)_{22} & (J_1)_{23} \\ (J_1)_{31} & (J_1)_{32} & (J_1)_{33} \end{pmatrix} &= -\begin{pmatrix} \epsilon_{111} & \epsilon_{112} & \epsilon_{113} \\ \epsilon_{121} & \epsilon_{122} & \epsilon_{123} \\ \epsilon_{131} & \epsilon_{132} & \epsilon_{133} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \end{aligned} \tag{3.63}$$

Let's see what finite transformation matrix we get from the first of these basis generators. The connection between the generators and the finite transformations is given by the exponential function: $R_{1\text{ fin}} = e^{\theta J_1}$. To calculate this, we can focus on the lower right 2×2 matrix[47] j_1 in J_1 and ignore the zeroes for a moment:

[47] This is exactly minus the two-dimensional Levi-Civita symbol $(j_1)_{ij} = -\epsilon_{ij}$ in matrix form (see Appendix B.5.5), which is the generator of rotations in two dimensions (of $SO(2)$).

$$J_1 = \begin{pmatrix} 0 & \\ & \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\equiv j_1} \end{pmatrix} . \tag{3.64}$$

We can immediately compute

$$(j_1)^2 = -1, \tag{3.65}$$

therefore

$$(j_1)^3 = \underbrace{(j_1)^2}_{=-1} j_1 = -j_1 , \quad (j_1)^4 = +1 , \quad (j_1)^5 = +j_1 . \tag{3.66}$$

In general, we have

$$(j_1)^{2n} = (-1)^n I \qquad \text{and} \qquad (j_1)^{2n+1} = (-1)^n j_1 , \tag{3.67}$$

which we can use if we evaluate the exponential function as series expansion[48]

[48] This is derived in Appendix B.4.1. The trick used here is explained in more detail in Appendix B.4.2 and the series expansions of sine and cosine are derived in Appendix B.4.1, too.

$$\begin{aligned} \tilde{R}_{1\text{ fin}} = e^{\theta j_1} &= \sum_{n=0}^{\infty} \frac{\theta^n j_1^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{\theta^{2n}}{(2n)!} \underbrace{(j_1)^{2n}}_{(-1)^n I} + \sum_{n=0}^{\infty} \frac{\theta^{2n+1}}{(2n+1)!} \underbrace{(j_1)^{2n+1}}_{(-1)^n j_1} \\ &= \underbrace{\left(\sum_{n=0}^{\infty} \frac{\theta^{2n}}{(2n)!} (-1)^n \right)}_{=\cos(\theta)} I + \underbrace{\left(\sum_{n=0}^{\infty} \frac{\theta^{2n+1}}{(2n+1)!} (-1)^n \right)}_{=\sin(\theta)} j_1 \\ &= \cos(\theta) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin(\theta) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} . \end{aligned} \tag{3.68}$$

Using $e^0 = 1$ for the upper-left component, the complete, finite transformation matrix therefore reads

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} , \tag{3.69}$$

which we can recognize as one of the well-known rotation matrices in 3-dimensions that were quoted at the beginning of this chapter

(Eq. 3.22). Following the same steps, we can derive the matrices for rotations around the other axes, too.

We now have the generators of the group in explicit matrix form (Eq. 3.61) and this allows us to compute the Lie bracket relations between the basis generators by brute force[49]. The result is[50]

$$[J_i, J_j] = \epsilon_{ijk} J_k, \tag{3.70}$$

where ϵ_{ijk} is again the Levi-Civita symbol. In physics it's conventional to define the generators of $SO(3)$ with an extra "i". Concretely this means that instead of $e^{\tilde{\phi} J}$, we write $e^{i\phi J}$ with $\phi = -\tilde{\phi}$. Our generators are then

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \qquad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.71}$$

and the Lie algebra[51] reads

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \tag{3.72}$$

We introduce the additional "i" in physics to get Hermitian generators, which means generators that satisfy[52] $J^\dagger = (J^\star)^T = J$. This is a nice property, because Hermitian matrices have *real* eigenvalues and this becomes important in quantum mechanics where the eigenvalues of the generators are the values we expect to measure in experiments[53]. Without the "i", the generators are anti-Hermitian $J^\dagger = (J^\star)^T = -J$ and the corresponding eigenvalues are imaginary. This would make it less intuitive to interpret the eigenvalues as something that we can observe in experiments.

We can derive the basis generators in another way, by starting with the well known rotation matrices and use Eq. 3.53: $X = \frac{dh}{d\theta}|_{\theta=0}$. For the first rotation matrix, as quoted in Eq. 3.22 and derived in Eq. 3.69, this yields

$$\begin{aligned} J_1 = \frac{dR_1}{d\theta}|_{\theta=0} &= \frac{d}{d\theta} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \Big|_{\theta=0} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin(\theta) & -\cos(\theta) \\ 0 & \cos(\theta) & -\sin(\theta) \end{pmatrix} \Big|_{\theta=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \end{aligned} \tag{3.73}$$

which is exactly the first generator in Eq. 3.61. Nevertheless, the first method is more general, because we will not always start with

[49] As explained above, the natural product of the Lie algebra is the Lie bracket. Here we compute how the basis generators behave, when put into the Lie bracket. All other generators can be constructed by linear combination of these basis generators. Therefore, if we know the result of the Lie bracket of the basis generators, we know automatically the result for all other generators. This behavior of the basis generators in the Lie bracket, will become incredibly important in the next section. Everything that is important about a Lie algebra, is encoded in the Lie bracket relation of the basis generators.

[50] For example, we have
$[J_1, J_2] = J_1 J_2 - J_2 J_1$
$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \underbrace{\epsilon_{12k}}_{=0 \text{ except for } k=3} J_k$
$= \epsilon_{123} J_3 = J_3$

[51] We will call the Lie bracket relation of the basis generators **the** Lie algebra, because everything important is encoded here.

[52] For example now we have $J_1^\star = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix}$ and therefore
$J_1^\dagger = (J_1^\star)^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} = J_1$

[53] This will be discussed in Section 8.3.

given finite transformation matrices. For the Lorentz group we will start with the definition of the group, derive the basis generators and only afterwards compute an explicit matrix form of the Lorentz transformations. If you already have explicit transformation matrices, you can always use Eq. 3.53 to derive the corresponding generators.

Before we discuss the Lorentz group in more detail, we take a small detour and have a look at the modern definition of a Lie algebra. This modern definition is essential to get a deep understanding of the symmetries of nature.

3.4.2 The Abstract Definition of a Lie Algebra

Up to this point we used a simplified definition: The Lie algebra consists of all elements X that result in an element of the corresponding group G, when put into the exponential function $e^X \in G$, and an operation, called Lie bracket $[,]$, that we use to combine the Lie algebra elements.

We already discussed that the last part of this definition is crucial. As for the group, we also need a rule to combine Lie algebra elements. By looking at the Baker-Campbell-Hausdorff-Formula, we learned how the rule for the combination of group elements, is connected to the rule for the combination of Lie algebra elements. An important observation was that the rule for the combination of Lie algebra elements is not simply matrix multiplication, but a more complicated rule called Lie bracket.

While this operation already appears in our simplified definition, we now introduce a more abstract definition where the Lie bracket is even more central[54]

[54] This definition will prove to be invaluable for the following sections and it will become clear in a moment in what sense the Lie bracket is "more central".

A Lie algebra is a vector space $\mathfrak{g}$ equipped with a binary operation $[,]$: $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. The binary operation satisfies the following axioms:

- Bilinearity: $[aX + bY, Z] = a[X, Z] + b[Y, Z]$ and $[Z, aX + bY] = a[Z, X] + b[Z, Y]$, for arbitrary number a, b and $\forall X, Y, Z \in \mathfrak{g}$
- Anticommutativity: $[X, Y] = -[Y, X] \; \forall \quad X, Y \in \mathfrak{g}$
- The Jacobi Identity: $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$ $\forall X, Y, Z \in \mathfrak{g}$

You can check that the commutator of two matrices fulfills all these conditions and of course this standard commutator was used to motivate these axioms. Nevertheless, there are quite different binary operations that fulfill these axioms, for example, the famous Poisson bracket of classical mechanics.

The important point is that this definition makes no reference to any Lie group. The definition of a Lie algebra stands on its own and we will see that this makes a lot of sense. In the next section we will have a look at the generators of $SU(2)$ and find that the basis generators, which is the set of generators we can use to construct all other generators by linear combinations, fulfill the same Lie bracket relation as the basis generators of $SO(3)$ (Eq. 3.72). This is interpreted as $SU(2)$ and $SO(3)$ having the **same** Lie algebra. This is an incredibly important result and it will tell us a lot about $SU(2)$ and $SO(3)$.

3.4.3 The Generators and Lie Algebra of $SU(2)$

We stumbled upon $SU(2)$ while trying to describe rotations in three dimensions and discovered that $SU(2)$ is the double cover[55] of $SO(3)$.

[55] Recall that this means that the map from $SU(2)$ to $SO(3)$ identifies two elements of $SU(2)$ with the same element of $SO(3)$.

Remember that $SU(2)$ is the group of unitary 2×2 matrices with unit determinant[56] :

[56] This is what the "S" stands for: Special = unit determinant.

$$U^\dagger U = UU^\dagger = 1 \tag{3.74}$$

$$\det(U) = 1\,. \tag{3.75}$$

The first thing we now want to take a look at is the Lie algebra of this group. Writing the defining conditions of the group in terms of the generators $J_1, J_2, \dots$ yields[57]

[57] As discussed above, we now work with an extra "i" in the exponent, in order to get Hermitian matrices, which guarantees that we get real numbers as predictions for experiments in quantum mechanics.

$$U^\dagger U = (e^{iJ_i})^\dagger e^{iJ_i} \stackrel{!}{=} 1 \tag{3.76}$$

$$\det(U) = \det(e^{iJ_i}) \stackrel{!}{=} 1 \tag{3.77}$$

The first condition tells us, using the Baker-Campbell-Hausdorff theorem (Eq. 3.55) and $[J_i, J_i] = 0$

$$\begin{aligned}(e^{iJ_i})^\dagger e^{iJ_i} &= e^{-iJ_i^\dagger} e^{iJ_i} \stackrel{!}{=} 1\\ &\rightarrow e^{-iJ_i^\dagger + iJ_i + \frac{1}{2}[J_i^\dagger, J_i] + \dots} \stackrel{!}{=} 1\\ &\underbrace{\rightarrow}_{e^0=1} J_i^\dagger \stackrel{!}{=} J_i.\end{aligned} \tag{3.78}$$

A matrix fulfilling the condition $J_i^\dagger = J_i$ is called Hermitian and we therefore learn here that the generators of $SU(2)$ must be Hermitian.

Using the identity $\det(e^A) = e^{\mathrm{tr}(A)}$, we see that the second condition yields

$$\det(e^{iJ_i}) = e^{i\,\mathrm{tr}(J_i)} = 1 \underbrace{\rightarrow}_{e^0=1} \mathrm{tr}(J_i) \stackrel{!}{=} 0. \tag{3.79}$$

We conclude that the generators of $SU(2)$ must be Hermitian *traceless* matrices. A basis for Hermitian traceless 2×2 matrices is given by the following 3 matrices[58]:

[58] A complex 2×2 matrix has 4 complex entries and therefore 8 degrees of freedom. Because of the two conditions the degrees of freedom are reduced to 3.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (3.80)$$

This means every Hermitian traceless 2×2 matrix can be written as a linear combination of these matrices that are called **Pauli matrices**.

We can put these explicit matrices for the basis generators into the Lie bracket and this yields

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \qquad (3.81)$$

where ϵ_{ijk} is again the Levi-Civita symbol. To get rid of the nasty 2 it is conventional to define the generators of $SU(2)$ as $J_i \equiv \frac{1}{2}\sigma_i$. The Lie algebra then reads

$$[J_i, J_j] = i\epsilon_{ijk}J_k\,. \qquad (3.82)$$

Take note that this is exactly the same Lie bracket relation we derived for $SO(3)$ (Eq. 3.72)! Therefore one says that $SU(2)$ and $SO(3)$ have the same Lie algebra, because we define Lie algebras by their Lie bracket. We will use the abstract definition of this Lie algebra, to get different descriptions for the transformations described by $SU(2)$. We will learn that an $SU(2)$ transformation doesn't need to be described by 2×2 matrices. To make sense of things like this, we need a more abstract definition of a Lie group. At this point $SU(2)$ is defined as a set of 2×2 matrices, and a description of $SU(2)$ by, for example, 3×3 matrices, makes little sense. The abstract definition of a Lie group will enable us to see the connection between different descriptions of the same transformation. We will identify with each Lie group a geometrical object (a manifold) and use this abstract object to define a group. This may seem like a strange thought, but will make a lot of sense after taking a second look at two examples we already encountered in earlier sections.

3.4.4 The Abstract Definition of a Lie Group

One of the first Lie groups we discussed was $U(1)$, the unit complex numbers. These are defined as complex numbers that satisfy $z^\star z = 1$. If we write $z = a + ib$ this condition reads

$$z^\star z = (a+ib)^\star(a+ib) = (a-ib)(a+ib) = a^2 + b^2 = 1. \qquad (3.83)$$

This is exactly the defining condition of the unit circle[59]. The set

[59] The unit circle S^1 is the set of all points in two dimensions with distance 1 from the origin. In mathematical terms this means all points (x_1, x_2) fulfilling $x_1^2 + x_2^2 = 1$.

of unit-complex numbers **is** the unit circle in the complex plane. Furthermore, we found that there is a one-to-one map[60] between elements of $U(1)$ and $SO(2)$. Therefore, for these groups it is easy to identify them with a geometric object: the unit circle. Instead of talking about different descriptions for $SO(2)$ or $U(1)$, which are defined by objects of given dimension, it can help to think about this group as the unit-circle. Rotations in two-dimensions are, as a Lie group, the unit-circle and we can represent these transformations by elements of $SO(2)$, i.e. 2×2 matrices or elements of $U(1)$, i.e. unit-complex numbers.

[60] To be precise: an isomorphism. To say two things are isomorphic is the mathematical way of saying that they are "the same thing". Two things are called isomorphic if there exists an isomorphism between them.

The next groups we discussed were $SO(3)$ and $SU(2)$. Remember that we found a one-to-one map between $SU(2)$ and the unit quaternions. The unit quaternions are defined as those quaternions $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ that satisfy the condition (Eq. 3.28)

$$a^2 + b^2 + c^2 + d^2 \stackrel{!}{=} 1, \tag{3.84}$$

which is the same condition that defines[61] the three sphere S^3! Therefore the quaternions provide us with a map between $SU(2)$ and the three sphere S^3. This map is an isomorphism (1-1 and onto) and therefore we can really think of $SU(2)$ as the three sphere S^3.

[61] Recall that the unit circle S^1 is defined as the set of points that satisfy the condition $x_1^2 + x_2^2 = 1$. Equally, the two-sphere S^2 is defined by the condition $x_1^2 + x_2^2 + x_3^2 = 1$ and analogously the three sphere S^3 by $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$. The number that follows the S denotes the dimension. In two dimensions, with one condition we get a one-dimensional object: S^1. Equally we get in four dimensions, with one condition $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$ a three dimensional object S^3. S^3 is the surface of the four-dimensional ball.

These observations motivate the modern definition of a Lie group[62]: *A Lie group is a group, which is also a differentiable manifold*[63]*. Furthermore, the group operation $\circ$ must induce a differentiable map of the manifold into itself. This is a compatibility requirement that ensures that the group property is compatible with the manifold property. Concretely this means that every group element, say A induces a map that takes any element of the group B to another element of the group $C = AB$ and this map must be differentiable. Using coordinates this means that the coordinates of AB must be differentiable functions of the coordinates of B.*

[62] The technical details that follow aren't important for what we want to do in this book. Especially, don't worry if the exact meaning of notions like "induces" or "differentiable" is not clear. The important message to take away is: Lie group = manifold.

[63] A manifold is a set of points, for example a sphere that looks locally like flat Euclidean space R^n. Another way of thinking about a n-dimensional manifold is that it's a set which can be given n independent coordinates in some neighborhood of any point. For some more information about manifolds, see the appendix in Section 3.11 at the end of this chapter.

We can now understand the remark at the end of Section 3.4:

"... there is precisely one **distinguished** Lie group for each Lie algebra."

a bit better[64]. From the geometric perspective, the distinguished group has the property of being **simply connected**. This means that, if we use the modern definition of a Lie group as a manifold, any closed curve on this manifold can be shrunk smoothly to a point[65].

[64] Recall that we already discovered that different groups can have the *same* Lie algebra. For example, using the abstract definition of a Lie algebra, we say that $SO(3)$ and $SU(2)$ have the same Lie algebra (Eq. 3.82).

[65] We will not discuss this any further, but you are encouraged to read about it, for example in the books recommended at the end of this chapter. For the purpose of this book it suffices to know that there is always **one** distinguished group.

To emphasize this important point:[66]

There is precisely one simply-connected Lie group corresponding to each Lie algebra.

[66] A proof can be found, for example, in Michael Spivak. *A Comprehensive Introduction to Differential Geometry, Vol. 1, 3rd Edition*. Publish or Perish, 3rd edition, 1 1999. ISBN 9780914098706

This simply-connected group can be thought of as the "mother" of all those groups having the same Lie algebra, because there are maps to all other groups with the same Lie algebra from the simply connected group, but not vice versa. We could call it the mother group of this particular Lie algebra, but mathematicians tend to be less dramatic and call it the covering group. All other groups having the same Lie algebra are said to be covered by the simply connected one. We already stumbled upon an example of this: $SU(2)$ is the double cover of $SO(3)$. This means there is a two-to-one map from $SU(2)$ to $SO(3)$.

Furthermore, $SU(2)$ is the three sphere, which is a simply connected manifold. Therefore, we have already found the "most important" group belonging to the Lie algebra in Eq. 3.82. We can get all other groups belonging to this Lie algebra through maps from $SU(2)$.

We can now understand what manifold $SO(3)$ is. The map from $SU(2)$ to $SO(3)$ identifies with **two** points of $SU(2)$, **one** point of $SO(3)$. Therefore, we can think of $SO(3)$ as the top half[67] of S^3.

We can see, from the point of view that Lie groups are manifolds that $SU(2)$ is a more complete object than $SO(3)$. $SO(3)$ is just "part" of the complete object.

In this book we take the view that to describe nature at the most fundamental level, we need to use the most fundamental groups. For rotations in three dimensions this group is $SU(2)$ and not $SO(3)$. We will discover something similar when considering the symmetry group of special relativity.

We will see that Nature agrees with such lines of thought! To describe elementary particles one uses the representations of the *covering group* of the Poincaré group, instead of just the usual representation one uses to transform four-vectors. To describe nature at the most fundamental level, we must use the covering group, instead of any of the other groups that one can map to from the covering group.

We are able to derive the representations[68] of the most fundamental group, belonging to a given Lie algebra, by deriving representations of the Lie algebra. We can then put the matrices representing the Lie algebra elements (the generators) into the exponential function to get matrices representing group elements.

Herein lies the strength of Lie theory. By using pure mathematics we are able to reveal something fundamental about nature. The standard symmetry group of special relativity hides something from

[67] This picture is a bit oversimplified. Strictly speaking $SO(3)$ as a manifold is still a sphere, but with antipodal points identified.

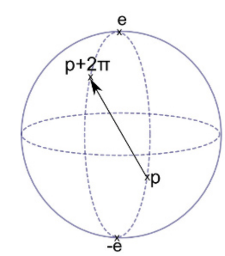


Fig. 3.7: Two-dimensional slice of the three Sphere S^3 (which is a three dimensional surface and therefore not drawable itself). We can see that the top half of the sphere is $SO(3)$, because to get from $SU(2)$ to $SO(3)$ we identify two points, for example, p and $p + 2\pi$, with each other.

[68] This notion will be made precise in the next section.

us, because it is not the most fundamental group belonging to this symmetry[69]. The covering group of the Poincaré group is the fundamental group and therefore we will use it to describe nature[70].

To summarize[71]

- $S^1 \hat{=} U(1) \underbrace{\leftrightarrow}_{\text{one-to-one}} SO(2)$
- $S^3 \hat{=} SU(2) \underbrace{\rightarrow}_{\text{two-to-one}} SO(3) \hat{=}$ "half" of S^3
  $\Rightarrow SU(2)$ is the distinguished group belonging to the Lie algebra $[J_i, J_j] = i\epsilon_{ijk}J_k$ (Eq. 3.82), because S^3 is simply connected.

Next, we will introduce another important branch of Lie theory, called representation theory. It is representation theory that enables us to derive from a given Lie group the tools that we need to describe nature at the most fundamental level.

3.5 Representation Theory

The important thing about group theory is that it is able to describe transformations without referring to any objects in the real world.

For theoretical considerations it is often useful to regard any group as an abstract group. This means defining the group by its manifold structure and the group operation. For example $SU(2)$ is the three sphere S^3, the elements of the group are points of the manifold and the rule associating a product point ab with any two points b and a satisfies the usual group axioms. In physical applications one is more interested in what the group actually does, i.e. the group action.

An important idea is that **one** group can act on **many** different kinds of objects[72]. This idea motivates the definition of a representation: A representation is a map[73] between any group element g of a group G and a linear transformation[74] $R(g)$ of some vector-space V

$$g \underbrace{\rightarrow}_{R} R(g) \tag{3.85}$$

in such a way that the group properties are preserved:

- $R(e) = I$ (The identity element of the group transforms nothing at all)
- $R(g^{-1}) = \big(R(g)\big)^{-1}$ (Every inverse element is mapped to the corresponding inverse transformation)

[69] For those who already know some quantum mechanics: the standard symmetry group hides spin from us!

[70] In the following, we will use one representation of the Poincaré group to derive the corresponding Lie algebra. Then we will use this Lie algebra to derive the representations of the one distinguished group that belongs to this Lie algebra, which doubly covers the Poincaré group.

[71] Maybe you wonder why S^2, the surface of the sphere in three dimensions, is missing. S^2 is not a Lie group and this is closely related to the fact that there are no three-dimensional complex numbers. Recall that we had to move from two-dimensional complex numbers with just **i** to the four-dimensional quaternions with **i,j,k**.

[72] This will make much more sense in a moment.

[73] The mathematical term for a map with these special properties is **homomorphism**. The definition of an isomorphism is then a homomorphism, which is in addition one-to-one.

[74] In the context of this book this will always mean that we map each group element to a matrix. Each group element is then given by a matrix that acts by usual matrix multiplication on the elements of some vector space.

- $R(g) \circ R(h) = R(gh)$ (The combination of transformations corresponding to g and h is the same as the transformation corresponding to the point gh)

A representation[75] identifies with each point (abstract group element) of the group manifold (the abstract group) a linear transformation of a vector space. Although we define a representation as a map, most of the time we will call a set of matrices a representation. For example, the usual rotation matrices are a representation of the group $SO(3)$ on the vector space[76] R^3. The rotation matrices are linear transformations on R^3. However, the important thing here is that we can examine the group action on other vector spaces, too!

[75] This concept can be formulated more generally if one accepts arbitrary (not necessarily linear) transformations of an arbitrary (not necessarily a vector) space. Such a map is called a realization. In physics one is concerned most of the time with linear transformations of objects living in some vector space (for example Hilbert space in quantum mechanics or Minkowski space for special relativity), therefore the concept of a representation is more relevant to physics than the general concept called realization.

[76] R^3 denotes three dimensional Euclidean space, where elements are ordinary 3 component vectors, as we use them for example in Appendix A.1.

Using representation theory, we are able to investigate systematically how a given group acts on very different vector spaces and that is where things start to get really interesting.

One of the most important examples in physics is $SU(2)$. For example, we can examine how $SU(2)$ acts on the complex vector space of dimension one $\mathbb{C}^1$, which is especially easy, as we will see later. We can also investigate how $SU(2)$ acts on $\mathbb{C}^2$. The objects living in $\mathbb{C}^2$ are complex vectors of dimension two and therefore $SU(2)$ acts on them as 2×2 matrices. The matrices (=linear transformations) acting on $\mathbb{C}^2$ are just the "usual" $SU(2)$ matrices that we already know. In addition, we can examine how $SU(2)$ acts on $\mathbb{C}^3$ or even higher dimensional vector spaces. There is a well defined framework for constructing such representations and as a result, $SU(2)$ acts, for example, on complex vectors of dimension three as 3×3 matrices. A basis for the $SU(2)$ generators when they act on $\mathbb{C}^3$ is given by[77]

[77] We will learn later in this chapter how to derive these. At this point just take notice that it is possible.

$$J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad J_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{3.86}$$

As usual, we can compute matrices that represent elements of the *group* $SU(2)$ by putting linear combinations of these generators into the exponential function.

One can go on and inspect how $SU(2)$ acts on higher dimensional vectors. This can be quite confusing and it would be better to call[78] this group S^3 instead of $SU(2)$, because usually $SU(2)$ is defined as the set of complex 2×2 (!) matrices satisfying (Eq. 3.32)

[78] In an early draft version of this book the group was consequently called S^3. Unfortunately, such a non-standard name makes it hard for beginners to dive deeper into the subject using the standard textbooks.

$$U^\dagger U = 1 \quad \text{and} \quad \det(U) = 1 \tag{3.87}$$

and now we write $SU(2)$ transformations as 3×3 matrices. Therefore one must always keep in mind that we mean the abstract group, instead of the 2×2 definition, when we talk about higher dimensional representation of $SU(2)$ or any other group.

Typically a group is defined in the first place with the help of an explicit representation. For example, we began our discussion of $SU(2)$ with explicit 2×2 matrices. This approach enables us to study the group properties concretely, as we did in the preceding sections. After this initial study it's often more helpful to regard the group as an abstract group[79], because it's possible to find other, useful representations of the group.

[79] For $SU(2)$ this means using S^3.

Before we move on to examples we need to define some abstract, but useful, notions. These notions will clarify the hierarchy of representations, because not every possible representation is equally fundamental.

The first notion we want to talk about is called **similarity transformation**. Given a matrix R and an invertible[80] matrix S then a transformation of the form

[80] A matrix S is called invertible, if there exists a matrix T, such that $ST = TS = 1$. The inverse matrix is usually denoted S^{-1}.

$$R \to R' = S^{-1} R S \tag{3.88}$$

is called a similarity transformation. The usefulness of this kind of transformation in this context lies in the fact that if we have a representation $R(G)$ of a group G, then $S^{-1}RS$ is also a representation. This follows directly from the definition of a representation: Suppose we have two group elements g_1, g_2 and a map R: $G \Rightarrow GL(V)$, i.e. $R(g_1)$ and $R(g_2)$. This is a representation if

$$R(g_1)R(g_2) = R(g_1 g_2) \tag{3.89}$$

If we now look at the similarity transformation of the representation

$$S^{-1}R(g_1)\underbrace{SS^{-1}}_{=1}R(g_2)S = S^{-1}R(g_1)R(g_2)S = S^{-1}R(g_1 g_2)S\,, \tag{3.90}$$

we see that this is a representation, too. Speaking colloquially, this means that if we have a representation, we can transform its elements wildly with literally any non-singular matrix S to get nicer matrices[81].

[81] The freedom to perform similarity transformations correspond to our freedom to choose a basis for the vector space our group acts on.

The next notion we want to introduce is called **invariant subspace**. When we have a representation R of a group G on a vector space V, we call $V' \subseteq V$ an invariant subspace if for[82] $v \in V'$ we have $R(g)v \in V'$ for all $g \in G$. This means, if we have a vector in the subspace V' and we act on it with arbitrary group elements, the

[82] Of course $v \in V$, too. The vector space V' must be part of the vector space V, which is mathematically denoted by $V' \subseteq V$. In other words this means that every element of V' is at the same time an element of V.

transformed vector will always be again part of the subspace V'. If we find such an invariant subspace we can define a representation R' of G on V', called a subrepresentation of R, by

$$R'(g)v = R(g)v \tag{3.91}$$

for all $v \in V'$. Therefore, one is led to the thought that the representation R, we talked about in the first place, is not fundamental, but a composite of smaller building blocks, called subrepresentations.

This leads us to the very important notion **irreducible representation**. An irreducible representation is a representation of a group G on a vector space V that has no invariant subspaces besides the zero space $\{0\}$ and V itself[83]. Such representations can be thought of as truly fundamental, because they are not made up by smaller representations. The irreducible representations of a group are the building blocks from which we can build up all other representations. There is another way to think about irreducible representation: An irreducible representation cannot be rewritten, using a similarity transformation, in block diagonal form[84]. In contrast, a reducible representation can be rewritten in block-diagonal form through similarity transformations. These notions are important because we use *irreducible* representations to describe elementary particles[85]. We will see later that the behavior of elementary particles under transformations is described by irreducible representations of the corresponding symmetry group.

[83] The subspace consisting solely of the identity element is always, trivially an invariant subspace.

[84] An example for a matrix in block-diagonal form is $\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix}$.

[85] What else?

There are many possible representations for each group[86], how do we know which one to choose to describe nature? There is an idea that is based on the Casimir elements. A Casimir element C is built from generators of the Lie algebra and its defining feature is that it commutes with every generator X of the group

[86] For example we already know two different representations for rotations in two-dimensions. One using complex numbers and one using 2×2 matrices. Both are representations of S^1 as a group.

$$[C, X] = 0. \tag{3.92}$$

What does this mean? A famous Lemma, called Schur's Lemma[87], tells us that if we have an irreducible representation $R : \mathfrak{g} \to GL(V)$, any linear operator $T : V \to V$ that commutes with all operators $R(X)$ must be a scalar multiple of the identity operator. Therefore, the Casimir elements give us linear operators with constant values for each representation. As we will see, these values provide us with a way of labelling representations naturally.[88] We can therefore start to investigate the irreducible representations, by starting with the representation with the lowest possible scalar value for the Casimir element.

[87] A basic result of group theory, which you can look up in any book about group theory. For example, Schur's lemma is proven at page 239 in Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148.

[88] This will become much clearer as soon as we look at an example.

Is there anything we can say about the vector space V mentioned in the definition of a representation above? An important observation that helps us to make sense of the vector space, is that for any Lie group, one or several of the corresponding generators can be diagonalized using similarity transformations. In physics we use these diagonal generators to get labels for the basis vectors that span our vector space. We use the eigenvectors of these diagonal generators as basis for our vector space and the corresponding eigenvalues as labels. This idea is incredibly important to actually understand the physical implications of a given group. If there is just one generator that can be diagonalized simultaneously, each basis vector is labelled by just one number: the corresponding eigenvalue. If there are several generators that can be diagonalized simultaneously, we get several numbers as labels for each basic vector. Each such number is simply the eigenvalue of a given diagonal generator that belongs to this basic vector (=eigenvector). In particular this is where the "charge labels" for elementary particles: electric charge, weak charge, color charge, come from. This idea will make a lot more sense, once we are ready to look at some explicit examples[89].

In the next sections, we will derive the irreducible representations of the Lie algebra of $SU(2)$, because, as we will see, the Lie algebra of the Lorentz group can be thought of as two copies of the algebra $\mathfrak{su}(2)$[90]. The Lorentz group is part of the Poincaré group and therefore we will talk about these groups in this order.

[89] A bit more background information: The set of diagonal generators is called Cartan subalgebra, and the corresponding generators Cartan generators. As already mentioned, these generators play a big role in elementary particle physics, because the eigenvalues of the Cartan generators are used to give charge labels to elementary particles. For example, to derive quantum chromodynamics, we use the group $SU(3)$, as we will see later, and there are two $SU(3)$ Cartan generators. Therefore, each particle that interacts via chromodynamics, carries two charge labels. Conventionally instead of writing two numbers, one uses the words red, blue, green, and calls the corresponding charge colour. Analogous, the theory of weak interactions uses the group $SU(2)$, which has only one Cartan generator. Therefore, each particle is labelled by the corresponding eigenvalue of this Cartan generator.

[90] Technically it's the *complexification* of the Lie algebra of the Lorentz group that can be understood as two copies of the *complexification* of the Lie algebra $\mathfrak{su}(2)$.

3.6 $SU(2)$

We used in Section 3.4.3 specific matrices (=a specific representation) to identify how the generators of $SU(2)$ behave, when put into the Lie bracket[91]. We can use this knowledge to find further representations. Unsurprisingly, we will derive again the representation that we started with, i.e. the set of unitary 2×2 matrices with unit determinant. However, we are then able to see that this is just one special case. Before we tackle this task, we want to take a moment to think about what representations we can expect.

[91] Recall that this is what we use to define the Lie algebra of a group in abstract terms. The final result was Eq. 3.82.

3.6.1 The Finite-dimensional Irreducible Representations of $SU(2)$

As noted earlier, there are special operators that we can build from a given set of generators that are useful to understand the representations of the Lie algebra in question. These operators are called Casimir operators and have the special property that they commute with all generators of the given Lie algebra. For the Lie algebra $\mathfrak{su}(2)$ there is one such operator[92]

[92] We restrict ourselves here to quadratic Casimir operators, which means operators that are quadratic in the generators. There are also higher order Casimir operators, but here we always mean quadratic Casimir operators.

$$J^2 = J_1^2 + J_2^2 + J_3^3 . \tag{3.93}$$

To illustrate this, let's consider again the 3-dimensional representation[93], as shown in Eq. 3.86. For, this representation the Casimir operator is

$$J^2_{3\text{-dim}} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} . \tag{3.94}$$

In other words, the Casimir operator is simply two times the 3×3 identity matrix. In contrast, for the 2-dimensional representation, as given by $J_i = \frac{1}{2}\sigma_i$, where σ_i are the Pauli matrices (Eq. 3.80), we get

$$J^2_{2\text{-dim}} = \begin{pmatrix} \frac{3}{4} & 0 \\ 0 & \frac{3}{4} \end{pmatrix} . \tag{3.95}$$

This illustrates that if the generators J_1, J_2, J_3 are given in some representation, we can calculate the corresponding Casimir operator explicitly. If we act with this Casimir operator on any element of the vector space that our generators act on[94], we get a number. This number is what we use to label different representations.

In addition to this label for different representations, we use additional special operators to label the different elements of the vector space that our groups acts on. The operators we use for this purpose are known as Cartan elements of the Lie algebra. In other words, while Casimir elements operators provide labels for different representations, the Cartan element provide labels within a given representation. The Cartan elements are all those generators that can be diagonalized simultaneously. For $\mathfrak{su}(2)$ there is only one such element and it is conventional to choose J_3 as diagonal generator[95]. In the 3-dimensional representation (Eq. 3.86) it is given by

$$J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{3.96}$$

and in the 2-dimensional representation it is given by

$$J_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} . \tag{3.97}$$

We use for each representation the eigenvectors of the diagonal generator J_3 as basis vectors for the vector space that our representation acts on. This means that every such basis vector is labeled by two numbers[96]

$$\begin{aligned} J^2 \left| b, m \right\rangle &= b \left| b, m \right\rangle \\ J_3 \left| b, m \right\rangle &= m \left| b, m \right\rangle . \end{aligned} \tag{3.98}$$

93 So far, we have not discussed where this representation comes from. We will understand this in a moment. Here we only use the final result to explain the basic idea behind the use of the Casimir operator J^2.

94 Recall that a representation is a map of the abstract group or Lie algebra elements to the linear operators that act on some vector space.

95 As noted above, for other algebras like $\mathfrak{su}(3)$, there is more than one Cartan element and therefore we get multiple labels for each vector.

96 We use here an abstract notation for the elements of the vector space that our generators act on. This notations emphasizes the labels that we use to distinguish different elements and is extremely popular in quantum mechanics, as will be discussed in Section 8.5.3.

The first label is the value of the quadratic Casimir operator J^2 for the representation in question and the second label corresponds to the value that we get when we act with the Cartan element J_3 on the vector. By looking at Eq. 3.95 and Eq. 3.97, we can see that the basis vectors for the 2-dimensional representation in this notation are[97] $|\frac{3}{4}, \frac{1}{2}\rangle, |\frac{3}{4}, -\frac{1}{2}\rangle$. Analogously, by looking at Eq. 3.94 and Eq. 3.96, we can see that the basis vectors for the 3-dimensional representation are[98] $|2,1\rangle, |2,0\rangle, |2,-1\rangle$.

[97] Take note that in the usual, non-abstract vector notation, we can use as basis vectors for the 2-dimensional representation $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$.

[98] Equally, in the usual, non-abstract vector notation we can use as our basis vectors for the 3-dimensional representation $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\0\\1\end{pmatrix}$.

After this preliminary discussion we are finally ready to understand the representations of $SU(2)$. To learn something about what finite-dimensional, irreducible representations of $SU(2)$ are possible, we define new operators from the ones we used in Section 3.4.3:[99]

$$J_+ = J_1 + iJ_2 \tag{3.99}$$
$$J_- = J_1 - iJ_2 \,. \tag{3.100}$$

[99] Take note that we use a complex linear combination here. This process of considering a complex linear combination, instead of the original generators is called a complexification, because usually we only allow real linear combinations of the generators. So from here on, we consider the complexification of the Lie algebra $\mathfrak{su}(2)$, which is also called $\mathfrak{sl}(2,\mathbb{C})$.

These new operators obey the following commutation relations, as you can check by using the commutator relations in Eq. 3.82[100]

$$[J_3, J_\pm] = \pm J_\pm \tag{3.101}$$
$$[J_+, J_-] = 2J_3. \tag{3.102}$$

[100] We can always diagonalize **one** of the generators. As mentioned above, we choose J_3 as diagonal and therefore yielding the basis vectors for our vector space.

If we now investigate how these operators act on an eigenvector $|b,m\rangle$ of J_3 with eigenvalue[101] m, we discover something remarkable:

[101] This means $J_3\,|b,m\rangle = m\,|b,m\rangle$ as explained in Appendix C.4.

$$\begin{aligned} J_3(J_\pm\,|b,m\rangle) &= J_3(J_\pm\,|b,m\rangle) + \underbrace{J_\pm J_3\,|b,m\rangle - J_\pm J_3\,|b,m\rangle}_{=0} \\ &= \underbrace{J_\pm J_3\,|b,m\rangle}_{=J_\pm m|b,m\rangle} + \underbrace{J_3 J_\pm\,|b,m\rangle - J_\pm J_3\,|b,m\rangle}_{=[J_3,J_\pm]|b,m\rangle} \\ &\underbrace{=}_{\text{Eq. 3.101}} (m\pm 1)J_\pm\,|b,m\rangle \,. \end{aligned} \tag{3.103}$$

We conclude that $J_\pm\,|b,m\rangle$ is again an eigenvector of J_3, but with eigenvalue $(m\pm 1)$ and thus we write[102] $J_\pm\,|b,m\rangle \equiv C\,|b,m\pm 1\rangle$:

[102] To be general, we include here a constant C. This constant will be discussed in a moment.

$$J_3\,|b,m\pm 1\rangle = (m\pm 1)\,|b,m\pm 1\rangle \,. \tag{3.104}$$

The operators J_- and J_+ are called raising and lowering operators or **ladder operators**. Starting from one J_3 eigenvector, we can construct more and more J_3 eigenvectors using the ladder operators $J_\pm$ repeatedly. This process must come to an end, because eigenvectors with different eigenvalues are linearly independent and we are dealing with finite-dimensional representations. This means that the corresponding vector space is finite-dimensional and therefore we can only find a finite number of linearly independent vectors.

We conclude that there must be an eigenvector with a maximum eigenvalue that we call j. For this maximum eigenvector, we must have

$$J_+ |b,j\rangle = 0\,, \tag{3.105}$$

because the only other possibility would be that, according to Eq. 3.103, we produce a new eigenvector with eigenvalue $j+1$ and thus j wouldn't be the maximum eigenvalue. We can calculate a relationship between this maximum J_3 eigenvalue j and the label b for the whole representation by using[103]

[103] This follows from $J^2 = J_1^2 + J_2^2 + J_3^3 \rightarrow J_1^2 + J_2^2 = J^2 - J_3^2$.

$$\begin{aligned} J_- J_+ &= (J_1 - iJ_2)(J_1 + iJ_2) \\ &= = \underbrace{J_1^2 + J_2^2}_{=J^2 - J_3^2} + i \underbrace{(J_1 J_2 - J_2 J_1)}_{=[J_1,J_2]=iJ_3 \text{ (Eq. 3.82)}} = J^2 - J_3^2 - J_3\,. \end{aligned} \tag{3.106}$$

Using this expression, we can write

$$0 \underbrace{=}_{\text{Eq 3.105}} J_- J_+ |b,j\rangle \underbrace{=}_{\text{Eq 3.106 and Eq 3.98}} (b - j^2 - j)\,|b,j\rangle\;. \tag{3.107}$$

Therefore, we can conclude $b - j^2 - j = 0$ and thus

$$b = j(j+1)\,. \tag{3.108}$$

Completely analogous to how we concluded that there is an eigenvector of J_3 with maximum eigenvalue j, we can conclude that there is an eigenvector with minimal eigenvalue. We call this minimal eigenvalue k and we must have $J_- |b,k\rangle = 0$, because otherwise k would not be the minimal eigenvalue. Analogous to the calculation in Eq. 3.107, we can calculate

$$0 = J_+ J_- |b,k\rangle = (b - k^2 + k)\,|b,k\rangle\;. \tag{3.109}$$

Therefore, we have $(b - k^2 + k) = 0$ and can conclude

$$b = -k(-k+1)\,. \tag{3.110}$$

By comparing Eq. 3.108 with Eq. 3.110, we can conclude

$$k = -j\,. \tag{3.111}$$

Now it's time to recall our discussion from the beginning of this section. We have one label b coming from the quadratic Casimir operator J^2, which we use to distinguish different representations. In addition, we have another label m coming from the diagonal generator J_3, which we use to label different elements of the vector space our representation acts on. We have derived above that for finite dimensional representations, there is a maximal value for m, which we called j. We introduced ladder operators $J_\pm$ that act on a given $|b,m\rangle$ and yield a new eigenvector of J_3 with eigenvalue $m \pm 1$. In practice, this means we can start with the eigenvector with maximum J_3 eigenvalue $|b,j\rangle$ for a given representation and then "move down the ladder" using J_-. The repeated application of J_- yields new eigenvectors of J_3. However, this process must come to an end at some eigenvector $|b,k\rangle$ with lowest J_3 eigenvalue k, because we are dealing with a finite dimensional representation and there can not be infinitely many eigenvectors. Acting with J_- on this eigenvector with minimal eigenvalue $|b,k\rangle$ yields simply zero. We calculated that this minimal eigenvalue is $k = -j$.

In summary, for each $SU(2)$ representation, labeled by $b = j(j+1)$ (Eq. 3.108), we get a set of J_3 eigenvectors with integer spaced eigenvalues in the range $-j \leq m \leq j$. As mentioned above, we can start climbing down the ladder at the vector $|b,j\rangle$ with maximum J_3 eigenvalue and get to the vector with minimum eigenvalue $|b,-j\rangle$ by repeated application of the operator J_-. After each application the eigenvalue is lowered by one. Therefore, the difference between the maximum J_3 eigenvalue j and the minimal value $-j$ is an integer: $j - (-j) = \text{integer}$. From this we can conclude immediately

$$2j = \text{integer} \rightarrow j = \frac{\text{integer}}{2}. \tag{3.112}$$

This is an incredibly important insight, because it allows us to understand the possible finite-dimensional representations of $\mathfrak{su}(2)$. If we simply start plugging in the allowed values of $j = 0, 1/2, 1, 3/2, \ldots$, we get all possible $\mathfrak{su}(2)$ representations. In addition, we can use what learned above to understand and construct these representations explicitly. This explicit construction will be the topic of the next sections. It is particular instructive to have a look at the possible J_3 eigenvalues m for the different allowed values of j and the value of the quadratic Casimir operator J^2 for these representations[104]

$$\begin{aligned}
&j = 0\,, && m = 0\,, && j(j+1) = 0 \\
&j = 1/2\,, && m = 1/2, -1/2\,, && j(j+1) = 3/4 \\
&j = 1\,, && m = 1, 0, 1\,, && j(j+1) = 2 \\
&j = 3/2\,, && m = 3/2, 1/2, -1/2, -3/2\,, && j(j+1) = 15/4 \\
&\ldots
\end{aligned} \tag{3.113}$$

We learn here that the $j = 0$ representation is one-dimensional, the $j = 1/2$ representation is two-dimensional, the $j = 1$ representation is three-dimensional and the $j = 3/2$ representation is four-dimensional[105]. Take note that the values for the quadratic Casimir

[104] In Eq. 3.98 we called the eigenvalue that belongs to J^2 simply b. However in Eq. 3.108, we calculated that we express b through the maximum J_3 eigenvalue that we called j: $b = j(j+1)$.

[105] We use the eigenvectors of the Cartan generator, here J_3, as basis vectors for the vector space that our representation acts on. For example, if there are two possible eigenvalues of J_3, i.e. m values, we have two basis vectors and thus the corresponding representation is two-dimensional.

operator J^2 for the two-dimensional and the three-dimensional representation are exactly those that we calculated at the beginning of this section.

There is one more thing we need to discuss, before we can move on to the actual explicit construction of these various representations. We calculated in Eq. 3.103 that if we act with J_+ on a given J_3 eigenvector $|b,m\rangle$, we get a new J_3 eigenvector with a higher eigenvalue: $J_3 J_+ |b,m\rangle = (m+1) J_+ |b,m\rangle$. However, we can not simply conclude that $J_+ |b,m\rangle = |b,m+1\rangle$. Instead, we only know that $J_+ |b,m\rangle$ is **proportional** to $|b,m+1\rangle$ and therefore we write

$$J_+ |b,m\rangle = C |b,m+1\rangle \ , \tag{3.114}$$

where C is some constant. We now want to calculate this constant, because only then we have fully understood the action of the ladder operators and can actually use them to construct the representations explicitly. The thing is that we want to work with normalized basis vectors and hence want that $|b,m\rangle$ and $|b,m+1\rangle$ are normalized:

$$\begin{aligned} |b,m\rangle^\dagger |b,m\rangle &= 1 \\ |b,m+1\rangle^\dagger |b,m+1\rangle &= 1 \,. \end{aligned} \tag{3.115}$$

If we act with J_+ or equally with J_- on a normalized basis vector like $|b,m\rangle$ the result will be, in general, not normalized. If we wouldn't care about normalization, we could write $J_+ |b,m\rangle = |b,m+1\rangle$. However, for us a basis vector like $|b,m+1\rangle$ means always a normalized basis vector and hence we need to include the additional constant C.

From the definition[106] of J_- and J_+ it follows that $J_-^\dagger = J_+$ and $J_+^\dagger = J_-$. Thus if we calculate the norm[107] of Eq. 3.114, we get

[106] Eq. 3.99 and Eq. 3.100: $J_+ = \frac{1}{\sqrt{2}}(J_1 + iJ_2)$ and $J_- = \frac{1}{\sqrt{2}}(J_1 - iJ_2)$

[107] It will become clear in a second why this is a clever thing to do.

$$\begin{aligned} (J_+ |b,m\rangle)^\dagger J_+ |b,m\rangle &= |C|^2 \underbrace{|b,m+1\rangle^\dagger |b,m+1\rangle}_{=1 \text{ (Eq. 3.115)}} \\ |b,m\rangle^\dagger J_+^\dagger J_+ |b,m\rangle &= |C|^2 \\ |b,m\rangle^\dagger \underbrace{J_- J_+}_{=J^2 - J_3^2 - J_3 \text{ (Eq. 3.106)}} |b,m\rangle &= |C|^2 \\ |b,m\rangle^\dagger \underbrace{(J^2 - J_3^2 - J_3) |b,m\rangle}_{=(b-m^2-m)|b,m\rangle \text{ (Eq. 3.98)}} &= |C|^2 \\ |b,m\rangle^\dagger (\underbrace{b}_{=j(j+1) \text{ (Eq. 3.108)}} - m^2 - m) |b,m\rangle &= |C|^2 \end{aligned}$$

$$\underbrace{|b,m\rangle^\dagger |b,m\rangle}_{=1 \text{ (Eq. 3.115)}}(j(j+1)-m^2-m) = |C|^2$$
$$(j(j+1)-m^2-m) = |C|^2 \,. \tag{3.116}$$

The final result of this calculation expresses the norm of the previously unknown constant C in terms of the labels j and m that we use to characterize our representation and the vectors within a representation. We can conclude[108] from Eq. 3.116:

[108] A possible complex phase of the constant C is irrelevant and we chose it to be real and positive.

$$\sqrt{(j(j+1)-m^2-m)} = C \,. \tag{3.117}$$

Thus, for a given representation, which means a given j and a given basis vector, which means given m, we can use Eq. 3.117 to calculate the full result of what happens if we act with J_+ on a given eigenvector:

$$J_+ |j(j+1),m\rangle = C\,|j(j+1),m+1\rangle$$
$$\underbrace{=}_{\text{Eq. 3.117}} \sqrt{(j(j+1)-m^2-m)}\,|j(j+1),m+1\rangle \,. \tag{3.118}$$

We already argued above that if we act with J_+ on the eigenvector with maximum J_3 eigenvalue $|j(j+1),j\rangle$, we must get zero. If we now plug in $m=j$ in Eq. 3.117 we can see that this indeed happens:

$$\sqrt{(j(j+1)-j^2-j)} = C$$
$$\sqrt{j^2+j-j^2-j} = C$$
$$0 = C \,. \tag{3.119}$$

In addition, following exactly the same steps as above, we can calculate the constant $\tilde{C}$ that we get if we act with J_- on a given eigenvector. The result is

$$J_- |j(j+1),m\rangle = \tilde{C}\,|j(j+1),m+1\rangle$$
$$= \sqrt{(j(j+1)-m^2+m)}\,|j(j+1),m+1\rangle \,. \tag{3.120}$$

We have now everything we need to calculate explicit matrix expressions for the various $SU(2)$ representations. In fact, it's possible to show that every irreducible representation of $SU(2)$ must be equivalent to one of these that we can construct by using the tools described above[109]. There is one small thing we need to discuss before we move on. The label $j(j+1)$ takes a lot of space and is somewhat redundant. It is therefore conventional to use simply j as a label instead. This means, we use $|j,m\rangle$ instead of $|j(j+1),m\rangle$.

[109] See, for example, page 190 in: Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

Now we look at specific examples for the representations. We start, of course, with the lowest dimensional representation.

3.6.2 The Representation of $SU(2)$ in one Dimension

The lowest possible value for j is zero. As already mentioned above, this representation acts on a one-dimensional vector space. We can see that this representation is trivial, because the only 1×1 "matrix" that fulfills the commutation relations of the $SU(2)$ Lie algebra $[J_l, J_m] = i\epsilon_{lmn}J_n$, are the number 0. If we exponentiate the generator 0, we always get the transformation $U = e^0 = 1$, which changes nothing at all.

3.6.3 The Representation of $SU(2)$ in two Dimensions

We now take a look at the next possible value $j = \frac{1}{2}$. This representation is $2\frac{1}{2} + 1 = 2$ dimensional[110]. The generator J_3 has eigenvalues $\frac{1}{2}$ and $\frac{1}{2} - 1 = -\frac{1}{2}$, as can be seen from Eq. 3.113 and is therefore given by

$$J_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.121}$$

because we choose J_3 to be the diagonal generator[111]. The eigenvectors corresponding to the eigenvalues $+\frac{1}{2}$ and $-\frac{1}{2}$ are[112]

$$|1/2, 1/2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1/2, -1/2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.122}$$

We can find the explicit matrix form of the other two $SU(2)$ generators J_1 and J_2 in this basis by rewriting them using the ladder operators

$$J_1 = \frac{1}{2}(J_- + J_+) \tag{3.123}$$

$$J_2 = \frac{i}{2}(J_- - J_+), \tag{3.124}$$

which we get directly from inverting the definitions of $J_\pm$ in Eq. 3.100 and Eq. 3.99. Recall that a basis for the vector space of this representation is given by the eigenvectors of J_3 and we therefore express the generators J_1 and J_2 in this basis. In other words, in this basis J_1 and J_2 are defined by their action on the eigenvectors of J_3. We compute

$$\begin{aligned} J_1 |1/2, 1/2\rangle &= \frac{1}{2}(J_- + J_+) |1/2, 1/2\rangle \\ &= \frac{1}{2}(J_- |1/2, 1/2\rangle + \underbrace{J_+ |1/2, 1/2\rangle}_{=0}) \\ &= \frac{1}{2} J_- |1/2, 1/2\rangle \\ &\underbrace{=}_{\text{Eq. 3.120 with } \tilde{C}=1} \frac{1}{2} |1/2, -1/2\rangle, \end{aligned} \tag{3.125}$$

[110] See Eq. 3.113.

[111] For $SU(2)$ only one generator is diagonal, because of the commutation relations. Furthermore, remember that we are able to transform the generators using similarity transformations and could therefore easily make another generator diagonal.

[112] As mentioned above, we use, for brevity, $|j, m\rangle$ instead of $|j(j+1), m\rangle$. Here $j = 1/2$.

where we used that $1/2$ is already the maximum J_3 eigenvalue and we cannot go higher. Similarly we get

$$J_1 |1/2, -1/2\rangle = \frac{1}{2}(J_- + J_+) |1/2, -1/2\rangle = \frac{1}{2} |1/2, 1/2\rangle \ . \qquad (3.126)$$

Using Eq. 3.125 and Eq. 3.126, we can write J_1 in matrix form:

$$J_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \qquad (3.127)$$

You can check that this matrix has the correct action on the basis vectors that we derived above[113]. In the same way, we find

$$J_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} . \qquad (3.128)$$

These are the same generators $J_i = \frac{1}{2}\sigma_i$, with the Pauli matrices σ_i, we found while investigating the Lie algebra of $SU(2)$ at the beginning of this chapter (Eq. 3.80). We can now see that the representation we used there was exactly this two dimensional representation. Nevertheless, there are many more, for example, in three-dimensions as we will see in the next section[114].

[113] We derived in Eq. 3.125: $J_1 |1/2, 1/2\rangle = \frac{1}{2} |1/2, -1/2\rangle$. Using the explicit matrix form of J_1 we get $J_1 |1/2, 1/2\rangle = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2} |1/2, -1/2\rangle \quad \checkmark$.

[114] Again, don't get confused by the name $SU(2)$, which we originally defined as the set of unitary 2×2 matrices with unit determinant. Here we mean the abstract group, defined by the corresponding manifold S^3 and we are going to talk about higher dimensional representations of this group, which result in, for example, a representation with 3×3 matrices. It would help if we could give this structure a different name (For example, using the name of the corresponding manifold S^3), but unfortunately $SU(2)$ is the conventional name.

3.6.4 The Representation of $SU(2)$ in three Dimensions

Following the same procedure[115] as in two-dimensions, we find:

$$J_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad (3.129)$$

This is the representation of the generators of $SU(2)$ in three dimensions. If you're interested, you can derive the corresponding representation for the group elements of $SU(2)$ in three dimensions, by putting these generators into the exponential function. We will not go any further and derive even higher dimensional representations, because at this point we already have everything we need to understand the most important representations of the Lorentz group.

[115] Again we start with the diagonal generator J_3, which we can write down immediately because we know its eigenvalues $(1, 0, -1)$ from Eq. 3.113. Afterwards, the other two generators J_1, J_2 can be derived by looking at how they act on the basis vectors, i.e. the eigenvectors of J_3. To calculate this, we use again that we can write J_1 and J_2 in terms of $J_\pm$.

3.7 The Lorentz Group $O(1,3)$

> "To arrive at abstraction, it is always necessary to begin with a concrete reality . . . You must always start with something. Afterward you can remove all traces of reality."

- Pablo Picasso[116]

[116] As quoted in Robert S. Root-Bernstein and Michele M. Root-Bernstein. *Sparks of Genius*. Mariner Books, 1st edition, 8 2001. ISBN 9780618127450

In this section we will derive **one** well-known representation of the Lorentz group. Then we will use this familiar representation to derive the Lie algebra of the Lorentz group. This is exactly the same route we followed for $SU(2)$. There we started with explicit 2×2 matrices to derive the corresponding Lie algebra. We will find that the complexified Lie algebra of the Lorentz group consists of two copies of the Lie algebra $\mathfrak{su}(2)$. This fact can be used to discover further representations of the Lorentz group, whereas the well-known vector representation, which is the representation of the Lorentz group by 4×4 matrices acting on four-vectors, will prove to be **one** of the representations. The new representations will provide us with tools to describe physical systems that cannot be described by the vector representation. This shows the power of Lie theory. Using Lie theory we are able to identify the hidden abstract structure of a symmetry and by using this knowledge, we are able to describe nature at the most fundamental level with the required tools.

We start with a characterisation of the Lorentz group and its subgroups. The Lorentz group is the set of all transformations that preserve the inner product of Minkowski space[117]

$$x^\mu x_\mu = x^\mu \eta_{\mu\nu} x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{3.130}$$

where $\eta_{\mu\nu}$ denotes the metric of Minkowski space

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.131}$$

This is the reason why we call the Lorentz group $O(1,3)$. The group $O(4)$ preserves $(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2$. Let's see what restriction this imposes. The conventional name for a Lorentz transformation is Λ ("lambda"). For the moment, Λ is just a name and we will derive now how these transformations look like explicitly. If we transform $x^\mu \to x'^\mu = \Lambda^\mu_\nu x^\nu$, we get the product

$$x^\mu \eta_{\mu\nu} x^\nu \to x'^\sigma \eta_{\sigma\rho} x'^\rho = (x^\mu \Lambda^\sigma_\mu) \eta_{\sigma\rho} (\Lambda^\rho_\nu x^\nu) \stackrel{!}{=} x^\mu \eta_{\mu\nu} x^\nu \tag{3.132}$$

and because this must hold for arbitrary x^μ we conclude

$$\Lambda^\sigma_\mu \eta_{\sigma\rho} \Lambda^\rho_\nu \stackrel{!}{=} \eta_{\mu\nu} \tag{3.133}$$

or written in matrix form[118]

$$\Lambda^T \eta \Lambda \stackrel{!}{=} \eta. \tag{3.134}$$

[117] This was derived in Chapter 2. Recall that this definition is analogous to our definition of rotations and spatial reflections in Euclidean space, which preserve the inner product of Euclidean space.

[118] Recall that in order to write the product of two vectors in matrix notation, the left vector is transposed. Therefore we get here Λ^T.

This is how the Lorentz transformations Λ are defined!

Starting from this definition, we will now derive a useful classification for all Lorentz transformations.

If we take the determinant of Eq. 3.133 and use $\det(AB) = \det(A)\det(B)$, we get the condition

$$\det(\Lambda)\underbrace{\det(\eta)}_{=-1}\det(\Lambda) \stackrel{!}{=} \underbrace{\det(\eta)}_{=-1} \rightarrow \det(\Lambda)^2 \stackrel{!}{=} 1 \tag{3.135}$$

$$\rightarrow \det(\Lambda) \stackrel{!}{=} \pm 1 \,. \tag{3.136}$$

Furthermore, if we look at the $\mu = \nu = 0$ component in Eq. 3.133[119], we get

$$\Lambda_0^\sigma \eta_{\sigma\rho} \Lambda_0^\rho \stackrel{!}{=} \underbrace{\eta_{00}}_{=1} \rightarrow \Lambda_0^\sigma \eta_{\sigma\rho} \Lambda_0^\rho = (\Lambda_0^0)^2 - \sum_i (\Lambda_0^i)^2 \stackrel{!}{=} 1 \tag{3.137}$$

and thus conclude

$$\Lambda_0^0 \stackrel{!}{=} \pm\sqrt{1 + \sum_i (\Lambda_0^i)^2}. \tag{3.138}$$

We now divide the Lorentz transformations into four "sub-categories", depending on the signs in Eq. 3.136 and Eq. 3.138:

$$\begin{aligned} L_+^\uparrow &: \det(\Lambda) = +1;\ \Lambda_0^0 \geq 1 \\ L_-^\uparrow &: \det(\Lambda) = -1;\ \Lambda_0^0 \geq 1 \\ L_+^\downarrow &: \det(\Lambda) = +1;\ \Lambda_0^0 \leq -1 \\ L_-^\downarrow &: \det(\Lambda) = -1;\ \Lambda_0^0 \leq -1 \,. \end{aligned} \tag{3.139}$$

This classification is useful, because, as we will see in a moment, only the transformations in one of these categories[120], $L_+^\uparrow$, can be generated through infinitesimal transformations. We, want to use the power of Lie theory and in particular learn as much as possible from the Lie algebra that belongs to a given group. Thus we focus on this sub-category $L_+^\uparrow$ and try to learn as much as possible using the corresponding generators. The transformations in the other categories are a combination of the transformations in this special category and one or both of two special transformations known as time-reversal Λ_T and space-inversion Λ_P, where[121]

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.140}$$

[119] We will see in a minute why this is useful.

[120] The technical term for this most important category $L_+^\uparrow$ is proper, orthochronous Lorentz group. "Proper" refers to the property det $(\Lambda) = +1$, which in physical terms means that transformations in this category do not change the spatial orientation. For example, a right-handed coordinate system stays right-handed and does not become a left-handed one. The word "orthochronous" refers to the property $\Lambda_0^0 \geq 1$ and means in physical terms that transformations with this property do not change the direction of time.

[121] At least for one representation, these operators look like this. We will see later that for different representations, these operators look quite different. The subscript P here denotes "parity" which is another name for space-inversion. In physical terms a parity transformation is a reflection at the spatial axes, whereas a time-reversal transformation is a reflection at the time axis.

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.141}$$

This is illustrated in figure 3.8. In this sense, we can understand the complete Lorentz group as the set

$$O(1,3) = \{L_+^\uparrow, \Lambda_P L_+^\uparrow, \Lambda_T L_+^\uparrow, \Lambda_P \Lambda_T L_+^\uparrow\}. \tag{3.142}$$

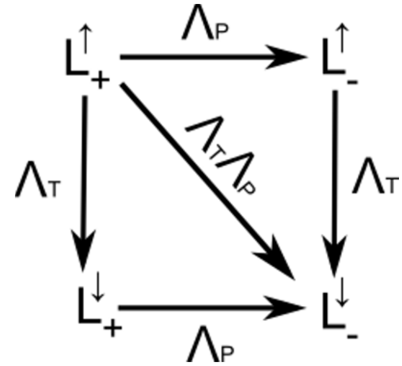


Fig. 3.8: The four components of the Lorentz group are connected through the two discrete transformations known as time-reversal Λ_T and space-inversion Λ_P.

So, why can only transformations in the $L_+^\uparrow$ category be built up by infinitesimal transformations? The transformations that can be built up by infinitesimal transformations are smoothly connected to the identity transformation, because the infinitesimal transformations are infinitesimally close to the identity element. Therefore, the identity element is in the same category as all the transformations that can be built up by repeating infinitesimal ones. We have

$$\det(\text{Id}) = +1; \ \text{Id}_0^0 = 1$$

and therefore the identity transformation and with it all transformations that can be generated through infinitesimal transformations, belong to the $L_+^\uparrow$ category.

An important related observation is that the four categories that we defined above: $L_+^\uparrow, L_-^\uparrow, L_+^\downarrow, L_-^\downarrow$, are not smoothly connected to each other. For example, there is a "gap" between the transformations with $\det(\Lambda) = +1$ and those with $\det(\Lambda) = -1$. There are no transformations in between. Equally, there is a "gap" between $\Lambda_0^0 \geq 1$ and $\Lambda_0^0 \leq -1$ and also no transformations in between. The jump across this gap can only be achieved by making use of the discrete transformations Λ_P and Λ_T. Therefore, the transformations in the other categories are not smoothly connected to the identity element, which is part of the $L_+^\uparrow$ category. For this reason, we can not get the transformations in these other categories, by solely using infinitesimal transformations. Instead, we always need the discrete "jumps" provided by Λ_P and Λ_T.

To summarize: We concentrate in the following on the transformations in the $L_+^\uparrow$ subcategory, because it contains all transformations that can be built up by repeating infinitesimal ones. Such transformations are especially nice, because we can understand them through the corresponding Lie algebra. In practice this means that in the next step, we investigate the generators[122] that generate all transformations in the $L_+^\uparrow$ category. The transformations in the other categories are simply a combination of the transformations that we derive this way, and the discrete operations Λ_P and Λ_T.

[122] Recall: Generators are the elements of the Lie algebra.

3.7.1 One Representation of the Lorentz Group

Let's see how we can use the defining condition of the Lorentz group (Eq. 3.133) to construct an explicit matrix representation of the allowed transformations. First let's think a moment about what we are trying to find. The Lorentz group, when acting on 4-vectors[123], is given by real 4×4 matrices. The matrices must be real, because we want to know how they act on elements of the real Minkowski space $R^{(1,3)}$. A generic, real 4×4 matrix has 16 parameters. The defining condition of the Lorentz group, which is in fact 10 conditions[124], restricts this to 6 parameters. In other words, to describe a most general Lorentz transformation, 6 parameters are needed. Therefore, if we find 6 linearly independent generators, we have found a complete basis for the Lie algebra of this group. This means every other generator can be written as a linear combination of these basis generators. In addition, we are then able to compute how these basis generators behave when put into the Lie bracket and therefore to derive the abstract definition of this Lie algebra.

[123] The usual vector space of special relativity is the real, four-dimensional Minkowski space $R^{(1,3)}$. We will look at the representation on this vector space first, because the Lorentz group is defined there in the first place, i.e. as the set of transformations that preserve the 4×4 metric. Equivalently $SU(2)$ was defined as complex 2×2 matrices in the first place and we tried to learn as much as possible about $SU(2)$ from these matrices, in order to derive other representations later .

[124] You can see this, by putting a generic 4×4 matrix Λ, in $\Lambda^T \eta \Lambda = \eta$.

First note that the rotation matrices of 3-dimensional Euclidean space, which involve only space and leave the time unchanged, fulfil the condition in Eq. 3.133[125]. This follows because the spatial part[126] of the Minkowski metric is proportional to the 3×3 identity matrix[127] and therefore for transformations involving only space, we have from Eq. 3.133 the condition

$$-R^T I_{3\times3} R = -R^T R \stackrel{!}{=} -I_{3\times3}$$

$$\rightarrow R^T I_{3\times3} R = R^T R \stackrel{!}{=} I_{3\times3}.$$

[125] $\Lambda^\sigma_\mu \eta_{\sigma\rho} \Lambda^\rho_\nu \stackrel{!}{=} \eta_{\mu\nu}$

[126] The spatial part are the components $\mu = 1,2,3$. Commonly this is denoted by η_{ij}, because Latin indices, like i, j always run from 1 to 3 and Greek indices, like μ and ν, run from 0 to 3.

[127] Recall $\eta_{11} = \eta_{22} = \eta_{33} = -1$ and $\eta_{ij} = 0$ for $i \neq j$.

This is exactly the defining condition of $O(3)$. Together with the condition (Eq. 3.139)

$$\det(\Lambda) \stackrel{!}{=} 1$$

these are the defining conditions of $SO(3)$. We conclude that the corresponding Lorentz transformations are given by

$$\Lambda_{\text{rot}} = \begin{pmatrix} 1 & \\ & R_{3\times3} \end{pmatrix}$$

with the rotation matrices $R_{3\times3}$ shown in Eq. 3.22 and derived in Section 3.4.1. The corresponding generators are therefore analogous to those we derived for three spatial dimensions in Section 3.4.1:

$$J_i = \begin{pmatrix} 0 & \\ & J_i^{3dim} \end{pmatrix} . \qquad (3.143)$$

For example, using Eq. 3.71 we now have

$$J_1 = \begin{pmatrix} 0 & \\ & J_1^{3dim} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}. \tag{3.144}$$

To investigate transformations involving time **and** space we will start, as always in Lie theory, with an infinitesimal transformation[128]

$$\Lambda^\mu_\rho \approx \delta^\mu_\rho + \epsilon K^\mu_\rho. \tag{3.145}$$

We put this into the defining condition (Eq. 3.133)

$$\Lambda^\mu_\rho \eta_{\mu\nu} \Lambda^\nu_\sigma \stackrel{!}{=} \eta_{\rho\sigma}$$

$$\rightarrow (\delta^\mu_\rho + \epsilon K^\mu_\rho)\eta_{\mu\nu}(\delta^\nu_\sigma + \epsilon K^\nu_\sigma) \stackrel{!}{=} \eta_{\rho\sigma}$$

$$\rightarrow \cancel{\eta_{\rho\sigma}} + \epsilon K^\mu_\rho \eta_{\mu\sigma} + \epsilon K^\nu_\sigma \eta_{\rho\nu} + \underbrace{\epsilon^2 K^\mu_\rho \eta_{\mu\nu} K^\nu_\sigma}_{\approx 0 \text{ because } \epsilon \text{ is infinitesimal } \rightarrow \epsilon^2 \approx 0} = \cancel{\eta_{\rho\sigma}}$$

$$\rightarrow K^\mu_\rho \eta_{\mu\sigma} + K^\nu_\sigma \eta_{\rho\nu} = 0 \tag{3.146}$$

which reads in matrix form[129]

$$K^T \eta = -\eta K. \tag{3.147}$$

Now we have the condition for the generators of transformations involving time and space. A transformation generated by these generators is called a **boost**. A boost means a change into a coordinate system that moves with a different constant velocity compared with the original coordinate system. We can boost the description that we have, for example in a frame of reference where the object in question is at rest, into a frame of reference where it moves relative to the observer. Let's go back to the example used in Chapter 2.1: A boost along the x-axis. Because we know that $y' = y$ and $z' = z$ the generator is of the form

$$K_1 = \begin{pmatrix} \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{\equiv k_1} & \\ & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \tag{3.148}$$

and we only need to solve a 2×2 matrix equation. Equation 3.147 reduces to

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

[128] With the Kronecker delta defined by $\delta^\mu_\rho = 1$ for $\mu = \rho$ and $\delta^\mu_\rho = 0$ for $\mu \neq \rho$. This means writing the Kronecker delta in matrix form is just the identity matrix.

[129] Recall that the first index denotes the row and the second the column. So far we have been a little sloppy with first and second index, by writing them above each other. In fact, we have $K^\mu_\rho \equiv K^\mu{}_\rho \rightarrow (K^T)^\mu{}_\rho = K_\rho{}^\mu$. Matrix multiplication always works by multiplying rows with columns. Therefore $K^\nu{}_\sigma \eta_{\rho\nu} = \eta_{\rho\nu} K^\nu{}_\sigma$, where the ρ-row of η is multiplied with the σ-column of K. This term then is in matrix notation ηK. Furthermore, $K^\mu_\rho \eta_{\mu\sigma} = K^\mu{}_\rho \eta_{\mu\sigma} = (K^T)_\rho{}^\mu \eta_{\mu\sigma}$. In order to write this index term in matrix notation we need to use the transpose of K, because only then we get a product of the form row times column. The ρ-row of K^T is multiplied with the σ-column of η. Therefore, this term is $K^T\eta$ in matrix notation. In index notation we are free to move objects around, because for example K^μ_ρ is just one element of K, i.e. a number.

which is solved by[130]

$$k_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

The complete generator for boosts along the x-axis is therefore

$$K_1 = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.149}$$

and equally we can find the generators for boosts along the y- and z-axis

$$K_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad K_3 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}. \tag{3.150}$$

Now, we already know from Lie theory how we get from the generators to finite transformations

$$\Lambda_1(\phi) = \mathrm{e}^{i\phi K_1}.$$

For brevity let's focus again on the exciting part of the generator K_1, i.e. the upper left 2×2 matrix k_1, which is defined in Eq. 3.148. We can then evaluate the exponential function using its series expansion and that[131] $(ik_1)^2 = 1$

$$\begin{aligned} \Lambda_1(\phi) = \mathrm{e}^{i\phi k_1} &= \sum_{n=0}^{\infty} \frac{i^n \phi^n k_1^n}{n!} = \sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \underbrace{(ik_1)^{2n}}_{=1} + \sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \underbrace{(ik_1^{2n+1})}_{=ik_1} \\ &= \left(\sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \right) I + i \left(\sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \right) k_1 = \cosh(\phi) I + i \sinh(\phi) k_1 \\ &= \begin{pmatrix} \cosh(\phi) & 0 \\ 0 & \cosh(\phi) \end{pmatrix} + \begin{pmatrix} 0 & -\sinh(\phi) \\ -\sinh(\phi) & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cosh(\phi) & -\sinh(\phi) \\ -\sinh(\phi) & \cosh(\phi) \end{pmatrix} \end{aligned} \tag{3.151}$$

This computation is analogous to the computation in Section 3.4.1, but observe that the sums here have no factor $(-1)^n$ and therefore these sums are not $\sin(\phi)$ and $\cos(\phi)$, but different functions called hyperbolic sine $\sinh(\phi)$ and hyperbolic cosine $\cosh(\phi)$. The complete 4×4 transformation matrix for a boost along the x-axis is therefore

$$\Lambda_1 = \begin{pmatrix} \cosh(\phi) & -\sinh(\phi) & 0 & 0 \\ -\sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.152}$$

[130] $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$ and $-\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = -\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$.

[131] As you can easily check: $(ik_1)^2 = i^2 \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, equally $k_1^4 = 1$ etc. for all even exponents and of course $(ik_1)^3 = ik_1$, $k_1^5 = ik_1$ etc. for all uneven exponents.

Analogously, we can derive the transformation matrices for boosts along the other axes:

$$\Lambda_2 = \begin{pmatrix} \cosh(\phi) & 0 & -\sinh(\phi) & 0 \\ 0 & 1 & 0 & 0 \\ -\sinh(\phi) & 0 & \cosh(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.153)$$

$$\Lambda_3 = \begin{pmatrix} \cosh(\phi) & 0 & 0 & -\sinh(\phi) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh(\phi) & 0 & 0 & \cosh(\phi) \end{pmatrix}. \quad (3.154)$$

An arbitrary boost can be composed by multiplication of these 3 transformation matrices.

3.7.2 Generators of the Other Components of the Lorentz Group

To understand how the generators for the transformations of the other components[132] of the Lorentz Group look like, we simply have to act with the parity operation Λ_P and the time reversal operator Λ_T on the matrices J_i, K_i we just derived. In index notation we have[133]

[132] Recall that the Lorentz group is in fact $O(1,3) = \{L_+^\uparrow, \Lambda_P L_+^\uparrow, \Lambda_T L_+^\uparrow, \Lambda_P \Lambda_T L_+^\uparrow\}$ and we derived in the last section the generators of $L_+^\uparrow$.

[133] We need two matrices Λ_P, one for each index. This is just the ordinary transformation behavior of operators under changes of the coordinate system.

$$(\Lambda_P)^\alpha_{\alpha'} (\Lambda_P)^\beta_{\beta'} (J_i)^{\alpha'\beta'} \underbrace{\widehat{=}}_{\text{switching to matrix notation}} \Lambda_P J_i (\Lambda_P)^T = J_i \widehat{=} (J_i)^{\alpha\beta} \quad (3.155)$$

$$(\Lambda_P)^\alpha_{\alpha'} (\Lambda_P)^\beta_{\beta'} (K_i)^{\alpha'\beta'} \underbrace{\widehat{=}}_{\text{switching to matrix notation}} \Lambda_P K_i (\Lambda_P)^T = -K_i \widehat{=} -(K_i)^{\alpha\beta}, \quad (3.156)$$

as we can check by a brute force computation, using the explicit matrices derived in the last section. For example,

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \rightarrow J_1' = \Lambda_P J_1 (\Lambda_P)^T = J_1, \quad (3.157)$$

because

$$\Lambda_P J_1 (\Lambda_P)^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^T$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \quad (3.158)$$

In contrast,

$$K_1 = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to K_1' = \Lambda_P K_1 (\Lambda_P)^T = -K_1, \tag{3.159}$$

because

$$\Lambda_P K_1 (\Lambda_P)^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^T$$

$$= - \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.160}$$

In conclusion, we have under parity transformations

$$J_i \underbrace{\to}_{\text{P}} J_i \qquad K_i \underbrace{\to}_{\text{P}} -K_i\,. \tag{3.161}$$

This will become useful later, because for different representations the parity transformations will not be as obvious as in the vector representation. Equally we can investigate the time-reversed generators and the result will be the same, because time-reversal involves only the first component, which only changes something for the boost generators K_i

$$(\Lambda_T)^{\alpha}_{\alpha'} (\Lambda_T)^{\beta}_{\beta'} (J_i)^{\alpha'\beta'} \underbrace{\widehat{=}}_{\text{switching to matrix notation}} \Lambda_T J_i (\Lambda_T)^T = J_i \widehat{=} (J_i)^{\alpha\beta} \tag{3.162}$$

$$(\Lambda_T)^{\alpha}_{\alpha'} (\Lambda_T)^{\beta}_{\beta'} (K_i)^{\alpha'\beta'} \underbrace{\widehat{=}}_{\text{switching to matrix notation}} \Lambda_T K_i (\Lambda_T)^T = -(K_i)^{\alpha\beta}, \tag{3.163}$$

Or shorter:

$$J_i \underbrace{\to}_{\text{T}} J_i \qquad K_i \underbrace{\to}_{\text{T}} -K_i. \tag{3.164}$$

3.7.3 The Lie Algebra of the Proper Orthochronous Lorentz Group

Now using the explicit matrix form of the generators[134] for $L_+^\uparrow$ we can derive the corresponding Lie algebra by brute force computation[135]

[134] See Eq. 3.149 for the boost generators and Eq. 3.61 for the rotation generators

[135] The Levi-Civita symbol ϵ_{ijk}, is defined in Appendix B.5.5.

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.165}$$

$$[J_i, K_j] = i\epsilon_{ijk}K_k \tag{3.166}$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k\,, \tag{3.167}$$

where again J_i denotes the generators of rotations and K_i are the generators of boosts. A general Lorentz transformation is of the form

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta}+i\vec{K}\cdot\vec{\Phi}} \tag{3.168}$$

Equation 3.166 tells us that the two types of generator (J_i and K_i) do not commute with each other. While the rotation generators are closed under commutation[136], the boost generators are not[137]. We can now define two new operator types from the old ones that are closed under commutation and commute with each other[138]

$$N_i^{\pm} = \frac{1}{2}(J_i \pm iK_i). \tag{3.169}$$

Working out the commutation relations yields

$$[N_i^+, N_j^+] = i\epsilon_{ijk}N_k^+ \tag{3.170}$$

$$[N_i^-, N_j^-] = i\epsilon_{ijk}N_k^- \tag{3.171}$$

$$[N_i^+, N_j^-] = 0. \tag{3.172}$$

These are precisely the commutation relations for the Lie algebra of $SU(2)$ and we have therefore discovered that the complexified Lie algebra of $L_+^{\uparrow}$ consists of two copies of the Lie algebra $\mathfrak{su}(2)$[139].

This is great news, because we already know how to construct all irreducible representations of the Lie algebra of $\mathfrak{su}(2)$. However we must be careful. The Lorentz group is, like $SO(3)$, not simply-connected[140] and Lie theory tells us that, for groups that aren't simply connected, there is no one-to-one correspondence between the irreducible representations of the Lie algebra and representations of the corresponding group[141]. Instead, **by deriving the irreducible representations of the complexified Lie algebra of the Lorentz group, we find the irreducible representations of the covering group of the Lorentz group**, if we put the corresponding generators into the exponential function. Some of these representations will be representations of the Lorentz group, but we will find more than that. This is good, because we need these additional representations to describe certain elementary particles.

136 Closed under commutation means that the commutator $[J_i, J_j] = J_iJ_j - J_jJ_i$, is again a rotation generator. From Eq. 3.165 we can see that this is the case.

137 Eq. 3.167 tells us that the commutator of two boost generators K_i and K_j isn't another boost generator, but a generator of rotations.

138 Again, take note that we use a complex linear combination here and recall that this process of considering a complex linear combination, instead of the original generators is called a complexification. Usually we only allow real linear combinations of the generators. So from here on, we consider the complexification of the Lie algebra of the Lorentz group.

139 Recall that when we discussed the representations of the Lie algebra $\mathfrak{su}(2)$, we also used the corresponding complexification. The complexification of $\mathfrak{su}(2)$ is $\mathfrak{sl}(2,\mathbb{C})$ and therefore technically we have $\mathfrak{so}(1,3)_{\mathbb{C}} \cong \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$, where $\mathfrak{so}(1,3)_{\mathbb{C}}$ denotes the complexification of the Lorentz group Lie algebra.

140 We will use this simply as a fact here, because a proof would lead us too far apart.

141 This can be quite confusing, but remember that there **is** always one distinguished group that belongs to a Lie algebra. This group is distinguished because it is simply connected. If we derive the irreducible representation of a Lie algebra, we get, by putting those Lie algebra elements (= the generators) in the exponential function, representations of the simply connected (= covering) group. Only for the simply connected group there is a one-to-one correspondence.

For brevity, we will continue to call the representations we will derive, representations of the Lorentz group instead of representations of the Lie algebra of the Lorentz group or representations of the double cover of the Lorentz group.

Each irreducible representation of the Lie algebra $\mathfrak{su}(2)$ can be labeled by the scalar value j of the $\mathfrak{su}(2)$ Casimir operator[142]. Therefore, we now know that we can label the irreducible representations of the covering group[143] of the Lorentz group by two integer or half integer numbers: j_1 and j_2. This means we will look at the (j_1, j_2) representations and use the $j_1, j_2 = 0, \frac{1}{2}, 1\ldots$ representations for the two $\mathfrak{su}(2)$ copies, which we derived earlier.

[142] To be precise: the scalar value that we get from the *quadratic* Casimir operator is $j(j+1)$. However, for brevity, we simply use j instead.

[143] The covering group of the proper orthochronous Lorentz group $L_+^\uparrow$ is $SL(2,C)$, which is defined as the set of 2×2 matrices with unit determinant and complex entries. The relationship $SL(2,C) \to L_+^\uparrow \equiv SO(1,3)_+^\uparrow$ is similar to the relationship $SU(2) \to SO(3)$ we discovered earlier in this text.

It is conventional to write the Lorentz algebra in a more compact way by introducing a new symbol $M_{\mu\nu}$, which is defined through the equations

$$J_i = \frac{1}{2}\epsilon_{ijk}M_{jk}. \tag{3.173}$$

$$K_i = M_{0i}\,. \tag{3.174}$$

With this new definition the Lorentz algebra reads

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}). \tag{3.175}$$

Next, we investigate different representations of the complexified Lie algebra of the Lorentz group and hence of the double cover of the Lorentz group in detail.

3.7.4 The $(0,0)$ Representation

The lowest dimensional representation is, as it is for $SU(2)$, trivial, because the vector space is one-dimensional for both copies of $\mathfrak{su}(2)$. Our generators must therefore be 1×1 matrices and the only 1×1 "matrix" that fulfills the commutation relations is the number 0:

$$N_i^+ = N_i^- = 0 \to e^{iN_i^+} = e^{iN_i^-} = e^0 = 1 \tag{3.176}$$

Therefore we conclude that the $(0,0)$ representation of the Lorentz group acts on objects that do **not** change under Lorentz transformations. This representation is called the **scalar representation**.

3.7.5 The $(\frac{1}{2},0)$ Representation

In this representation we use the[144] 2 dimensional representation for one copy of the $SU(2)$ Lie algebra N_i^+, i.e. $N_i^+ = \frac{\sigma_i}{2}$ and the 1

[144] Recall that the dimension of our vector space is given by $2j+1$. Therefore we have here $2\frac{1}{2}+1 = 2$ dimensions.

dimensional representation for the other N_i^-, i.e. $N_i^- = 0$. From the definition of N^- in Eq. 3.169 we conclude

$$N_i^- = \frac{1}{2}(J_i - iK_i) = 0 \tag{3.177}$$

$$\rightarrow J_i = iK_i. \tag{3.178}$$

Furthermore, we can use that we already derived in Section 3.6.3 the two dimensional representation of $SU(2)$:

$$N_i^+ = \frac{\sigma_i}{2}, \tag{3.179}$$

where σ_i denotes once more the Pauli matrices, which were defined in Eq. 3.80. On the other hand, we have

$$N_i^+ \underbrace{=}_{\text{Eq. 3.169}} \frac{1}{2}(J_i + iK_i) \underbrace{=}_{\text{Eq. 3.178}} \frac{1}{2}(iK_i + iK_i) = iK_i \tag{3.180}$$

Comparing Eq. 3.179 with Eq. 3.180 tells us that

$$iK_i = \frac{\sigma_i}{2} \rightarrow K_i = \frac{\sigma_i}{2i} = \frac{i\sigma_i}{2i^2} = \frac{-i}{2}\sigma_i \tag{3.181}$$

$$\text{Eq. 3.178} \rightarrow J_i = iK_i = \frac{-i^2}{2}\sigma_i = \frac{1}{2}\sigma_i. \tag{3.182}$$

We conclude that a Lorentz rotation in this representation is given by

$$R_\theta = \mathrm{e}^{i\vec{\theta}\cdot\vec{J}} = \mathrm{e}^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \tag{3.183}$$

and a Lorentz boost by

$$B_\phi = \mathrm{e}^{i\vec{\phi}\cdot\vec{K}} = \mathrm{e}^{\vec{\phi}\cdot\frac{\vec{\sigma}}{2}}. \tag{3.184}$$

By writing out the exponential function as series expansion we can easily get the representation of the double cover of the Lorentz group from the representation of the generators. For example, rotations about the x-axis e.g. are given by

$$R_x(\theta) = \mathrm{e}^{i\theta J_1} = \mathrm{e}^{i\theta\frac{1}{2}\sigma_1} = 1 + \frac{i}{2}\theta\sigma_1 + \frac{1}{2}\left(\frac{i}{2}\theta\sigma_1\right)^2 + \dots \tag{3.185}$$

And if we use the explicit matrix form of σ_1 (Eq. 3.80), together with the fact that $\sigma_1^2 = 1$, we get[145]

$$\begin{aligned} R_x(\theta) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{i}{2}\theta\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{1}{2}\left(\frac{\theta}{2}\right)^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \dots \\ &= \begin{pmatrix} \cos(\frac{\theta}{2}) & i\sin(\frac{\theta}{2}) \\ i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix}. \end{aligned} \tag{3.186}$$

[145] The steps are completely analogous to what we did in Section 3.4.1

Analogously, we can compute the transformation matrix for rotations around other axes or boosts. One important thing to notice is that we have here complex 2×2 matrices, representing the Lorentz transformations. These transformations certainly do not act on the four-vectors of Minkowski space, because these have 4 components. The **two-component**[146] objects this representation acts on are called **left-chiral spinors**[147]:

$$\chi_L = \begin{pmatrix} (\chi_L)_1 \\ (\chi_L)_2 \, . \end{pmatrix} \tag{3.187}$$

Spinors in this context are two component objects. A possible definition for left-chiral spinors is that they are objects that transform under Lorentz transformations according to the $(\frac{1}{2}, 0)$ representation of the Lorentz group. Take note that this is not just another way to describe the same thing, because spinors have properties that usual vectors do not have. For instance, the factor $\frac{1}{2}$ in the exponent. This factor shows us that a spinor is after a rotation by 2π not the same, but gets a minus sign[148]. This is a pretty crazy property, because we usually expect that objects are exactly the same after a rotation by $360^\circ = 2\pi$.

In the last section, we saw that the lowest-dimensional representation is trivial. The next representation in the "hierarchy" of representations is the spinor representation that we discovered in this section. In this sense, we can say that

> "... a spinor is the most basic sort of mathematical object that can be Lorentz-transformed."
>
> \- **A. M. Steane**[149]

[146] We will learn later that these two components correspond to **spin-up** and **spin-down** states.

[147] This name will make more sense after the definition of right-chiral spinors. Then we can see that parity transformations transform a left-chiral spinor transformation into a right-chiral spinor transformation and vice versa. These spinors are often called left-handed and right-handed, but this can be confusing, because these terms correspond originally to a concept called helicity, which is not the same as chirality. Recall what the parity operator does: changing a left-handed coordinate system into a right-handed coordinate system and vice versa. Hence the name.

[148] There is much more one can say about spinors. See, for example, chapter 3.2 in J. J. Sakurai. *Modern Quantum Mechanics*. Addison Wesley, 1st edition, 9 1993. ISBN 9780201539295

[149] Andrew M. Steane. An introduction to spinors. *ArXiv e-prints*, December 2013

3.7.6 The $(0, \frac{1}{2})$ Representation

This representation can be constructed analogous to the $(\frac{1}{2}, 0)$ representation but this time we use the 1 dimensional representation for N_i^+, i.e. $N_i^+ = 0$ and the two dimensional representation for N_i^-, i.e. $N_i^- = \frac{1}{2}\sigma_i$. A first guess could be that this representation looks exactly like the $(\frac{1}{2}, 0)$ representation, but this is not the case! This time we get from the definition of N^+ in Eq. 3.169

$$N_i^+ = \frac{1}{2}(J_i + iK_i) = 0 \tag{3.188}$$

$$\rightarrow J_i = -iK_i. \tag{3.189}$$

Take notice of the minus sign. Using the two-dimensional representation of $\mathfrak{su}(2)$ for N^+, which was derived in Section 3.6.3, yields

$$N_i^- = \frac{1}{2}\sigma_i = \frac{1}{2}(J_i - iK_i) \underbrace{=}_{\text{Eq. 3.189}} \frac{1}{2}(-iK_i - iK_i) = -iK_i\,. \tag{3.190}$$

Using this we can deduce the $(0, \frac{1}{2})$ representation of the boost generators

$$-iK_i = \frac{1}{2}\sigma_i \to K_i = \frac{-1}{2i}\sigma_i = \frac{-i}{2i^2}\sigma_i = \frac{i}{2}\sigma_i. \tag{3.191}$$

In addition, from Eq. 3.189 we get

$$J_i = -iK_i = \frac{1}{2}\sigma_i\,. \tag{3.192}$$

We conclude that in this representation a Lorentz rotation is given by

$$R_\theta = \mathrm{e}^{i\vec{\theta}\cdot\vec{J}} = \mathrm{e}^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \tag{3.193}$$

and a Lorentz boost by

$$B_\phi = \mathrm{e}^{i\vec{\phi}\cdot\vec{K}} = \mathrm{e}^{-\vec{\phi}\cdot\frac{\vec{\sigma}}{2}}. \tag{3.194}$$

Therefore, rotations are the same as in the $(\frac{1}{2}, 0)$ representation, but boosts differ by a minus sign in the exponent. We conclude both representations act on objects that are **similar but not the same**. We call the objects the $(0, \frac{1}{2})$ representation of the Lorentz group acts on **right-chiral spinors**:

$$\chi_R = \begin{pmatrix} (\chi_R)^1 \\ (\chi_R)^2 \end{pmatrix} \tag{3.195}$$

The generic name for left- and right-chiral spinors is **Weyl spinors**.

3.7.7 Van der Waerden Notation

Now we introduce a notation that makes working with spinors very convenient. We know that we have two kinds of objects that transform differently and therefore must be distinguished. In the last section we learned that they are different, but not too different. In a moment we will learn that there is a connection between the objects transforming according to the $(\frac{1}{2}, 0)$ representation (left-chiral spinors) and the objects transforming according to the $(0, \frac{1}{2})$ representation (right-chiral spinors). To be able to describe these different objects using one notation we introduce the notions of dotted and undotted indices, sometimes called Van der Waerden notation, after their inventor. This will help us to keep track of which object transforms in what way. This will become much clearer in a minute, as soon as we have set up the full formalism.

Let's define that a left-chiral spinor χ_L has a lower, undotted index

$$\chi_L = \chi_a \tag{3.196}$$

and a right-chiral spinor χ_R has an upper, dotted index

$$\chi_R = \chi^{\dot{a}}. \tag{3.197}$$

Next, we introduce the "spinor metric". The spinor metric enables us to transform a right-chiral spinor into a left-chiral and vice versa, but not alone as we will see. We define the spinor metric[150] as

[150] Take note that this is the Levi-Civita symbol in two dimensions as defined in Appendix B.5.5.

$$\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.198}$$

and show that it has the desired properties. Furthermore, we define[151]

[151] Maybe a short comment on the strange notation χ_L^C is in order. The superscript C denotes charge conjugation, as will be explained in Section 3.7.10 in more detail. Here we see that this operation flips one label, i.e. a left-chiral spinor becomes right-chiral. Later we will see this operation flips all labels, including, for example, the electric charge.

$$\chi_L^C \equiv \epsilon \chi_L^\star \tag{3.199}$$

where the $\star$ denotes complex conjugation. We will now inspect how χ_L^C transforms under Lorentz transformations and see that it transforms precisely as a right-chiral spinor. The defining feature of a right-chiral spinor is its transformation behavior and therefore we will conclude that χ_L^C *is* a right-chiral spinor. Let us have a look at how χ_L^C transforms under boosts, where we use

$$(-\epsilon)(\epsilon) = 1 \tag{3.200}$$

and

$$(\epsilon)\sigma_i^\star(-\epsilon) = -\sigma_i \tag{3.201}$$

for each Pauli matrix σ_i, as you can check by using Eq. 3.198. Transforming χ_L^C yields[152]

[152] We use the notation $\vec{\phi}\vec{\sigma} = \sum_i \phi_i\sigma_i \underbrace{=}_{\text{summation convention}} \phi_i\sigma_i$. The "vector" $\vec{\sigma}$ shouldn't be taken too seriously, because it's just a shorthand, conventional notation.

$$\begin{aligned}
\chi_L^C \to \chi_L'^C &= \epsilon(\chi')_L^\star \\
&= \epsilon\left((\mathrm{e}^{\frac{\vec{\phi}}{2}\vec{\sigma}}\chi_L\right)^\star \\
&= \epsilon\left(\mathrm{e}^{\frac{\vec{\phi}}{2}\vec{\sigma}}\underbrace{(-\epsilon)(\epsilon)}_{=1\ (\text{Eq. }3.200)}\chi_L\right)^\star \\
&= \underbrace{\epsilon\mathrm{e}^{\frac{\vec{\phi}}{2}\vec{\sigma}^\star}(-\epsilon)}_{=\mathrm{e}^{-\frac{\vec{\phi}}{2}\vec{\sigma}}\ (\text{Eq. }3.201)}(\epsilon)\chi_L^\star \\
&= \mathrm{e}^{-\frac{\vec{\phi}}{2}\vec{\sigma}}\underbrace{\epsilon\chi_L^\star}_{=\chi_L^C} \\
&= \mathrm{e}^{-\frac{\vec{\phi}}{2}\vec{\sigma}}\chi_L^C\,,
\end{aligned} \tag{3.202}$$

which is exactly the transformation behavior of a *right-chiral* spinor[153]. To get to the fifth line, we use the series expansion of $e^{\frac{\vec{\phi}}{2}\vec{\sigma}}$ and Eq. 3.201 on every term. You can check in the same way that the behavior under rotations is not changed by complex conjugation and multiplication with ϵ, as it should be, because χ_L and χ_R transform in the same way under rotations:

[153] The transformation behavior of right-chiral spinors under boosts was derived in Eq. 3.194: $B_\theta = e^{i\vec{\phi}\vec{K}} = e^{-\vec{\phi}\frac{\vec{\sigma}}{2}}$. Compare this to how left-chiral spinors transform under boosts, as derived in Eq. 3.184: $B_\theta = e^{i\vec{\phi}\vec{K}} = e^{\vec{\phi}\frac{\vec{\sigma}}{2}}$

$$\chi_L^C \to \chi_L'^C = \epsilon(\chi')_L^\star = \epsilon(e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\chi)_L^\star = e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\epsilon(\chi_L)^\star = e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\chi_L^C. \qquad (3.203)$$

Furthermore, you can check that ϵ is invariant under all transformations and that if we want to go the other way round, i.e. transform a right-chiral spinor into a left-chiral spinor, we have to use (-ϵ).

Therefore, we define in analogy with the tensor notation of special relativity that our "metric" raises and lowers indices

$$\epsilon\chi_L \underbrace{=}_{\text{written in index notation}} \epsilon^{ac}\chi_c = \chi^a \,, \qquad (3.204)$$

where summation over identical indices is implicitly assumed (Einstein summation convention). Furthermore, we know that if we want to get χ_R from χ_L we need to use complex conjugation as well

$$\chi_R = \epsilon\chi_L{}^\star \,. \qquad (3.205)$$

This means that complex conjugation transforms an undotted index into a dotted index:

$$\chi_R = \epsilon\chi_L{}^\star = \chi^{\dot{a}}. \qquad (3.206)$$

Therefore, we can get a lower, dotted index by complex conjugating χ_L:

$$\chi_L{}^\star = \chi_a{}^\star = \chi_{\dot{a}} \qquad (3.207)$$

and an upper, undotted index, by complex conjugating χ_R

$$\chi_R{}^\star = (\chi^{\dot{a}})^\star = \chi^a \,. \qquad (3.208)$$

It is instructive to investigate how $\chi_{\dot{a}}$ and χ^a transform, because these objects are needed to construct terms from spinors, which do not change under Lorentz transformations. Terms like this are incredibly important, because we need them to derive physical laws that are the same in all frames of reference. This will be made explicit in a moment. From the transformation behavior of a left-chiral spinor

$$\chi_L = \chi_a \to \chi_a' = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\;b}\chi_b \,, \qquad (3.209)$$

we can derive how a spinor with lower, *dotted* index transforms:

$$\chi_L^\star = \chi_a^\star = \chi_{\dot{a}} \to \chi'_{\dot{a}} = (\chi'_a)^\star = \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^{\ b} \right)^\star \chi_b^\star$$
$$= \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2} + \vec{\phi}\frac{\vec{\sigma}^\star}{2}} \right)_{\dot{a}}^{\ \dot{b}} \chi_{\dot{b}} \qquad (3.210)$$

Analogously, we use that we know how a right-chiral spinor transforms:

$$\chi_R \to \chi'_R = \chi'^{\dot{a}} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^{\dot{a}}_{\ \dot{b}} \chi^{\dot{b}} \qquad (3.211)$$

to derive how a spinor with upper, *undotted* index transforms:

$$\chi_R^\star = (\chi^{\dot{a}})^\star = \chi^a \to \chi'^a = (\chi'^{\dot{a}})^\star = \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^{\dot{a}}_{\ \dot{b}} \right)^\star (\chi^{\dot{b}})^\star$$
$$= \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2} - \vec{\phi}\frac{\vec{\sigma}^\star}{2}} \right)^a_{\ b} \chi^b . \qquad (3.212)$$

To be able to write products of spinors that do not change under Lorentz transformations, we need one more ingredient. Recall how the scalar product of two vectors is defined: $\vec{a} \cdot \vec{b} = \vec{a}^T \vec{b}$. In the same spirit we should transpose one of the spinors in a spinor product. We can see this, because at the moment we have the complex conjugate of the Pauli matrices $\sigma_i^\star$ in the exponent, for example, $e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2}}$. Together with transposing this becomes the Hermitian conjugate: $\sigma_i^\dagger = (\sigma_i^\star)^T$, where the symbol † is called "dagger". The Hermitian conjugate of every Pauli matrix, is again the same Pauli matrix

$$\sigma_i^\dagger = (\sigma_i^\star)^T = \sigma_i, \qquad (3.213)$$

as you can easily check by looking at the explicit form of the Pauli matrices (Eq. 3.80).

By comparing Eq. 3.209 with Eq. 3.212 and using Eq. 3.213, we see that the transformation behavior of a transposed spinor with *lower*, undotted index is exactly the opposite of a spinor with *upper*, undotted index. This means a term of the form $(\chi^a)^T \chi_a$ is invariant (=does not change) under Lorentz transformations, because[154]

[154] As explained in Appendix B.5.5, the symbol δ_b^c is called Kronecker symbol and denotes the unit matrix in index notation. This means $\delta_b^c = 1$ for $b = c$ and $\delta_b^c = 0$ for $b \neq c$.

$$
\begin{aligned}
(\chi^a)^T \chi_a \to (\chi'^a)^T \chi'_a &= \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2} - \vec{\phi}\frac{\vec{\sigma}^\star}{2}} \right)^a_{\ b} \chi^b \right)^T \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^c_{\ a} \chi_c \\
&= (\chi^b)^T \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^{\star T}}{2} - \vec{\phi}\frac{\vec{\sigma}^{\star T}}{2}} \right)^a_{\ b} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^c_{\ a} \chi_c \\
&\underbrace{=}_{\text{Eq. 3.213}} (\chi^b)^T \underbrace{\left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^a_{\ b} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^c_{\ a}}_{=\delta^c_b} \chi_c \\
&= (\chi^c)^T \chi_c \,. \qquad (3.214)
\end{aligned}
$$

In the same way we can combine an *upper*, dotted index with a *lower*, dotted index as you can verify by comparing Eq. 3.210 with Eq. 3.211. In contrast, a term of the form $(\chi^{\dot{a}})^T \chi_a \hat{=} \chi_R^T \chi_L$ isn't invariant under Lorentz transformations, because

$$
\chi_R^T \chi_L = (\chi^{\dot{a}})^T \chi_a \to (\chi'^{\dot{a}})^T \chi'_a = \chi^{\dot{b}} \underbrace{\left(e^{i\vec{\theta}\frac{\vec{\sigma}^T}{2} - \vec{\phi}\frac{\vec{\sigma}^T}{2}} \right)^{\dot{a}}_{\ \dot{b}} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^c_{\ a}}_{\neq \delta^c_b} \chi_c \qquad (3.215)
$$

Therefore a term combining a left-chiral with a right-chiral spinor is **not** Lorentz invariant. We conclude, we must always combine an upper with a lower index of the same type[155] in order to get Lorentz invariant terms. Or formulated differently, we must combine the Hermitian conjugate of a right-chiral spinor with a left-chiral spinor $\chi_R^\dagger \chi_L = (\chi_R^\star)^T \chi_L \hat{=} (\chi^a)^T \chi_a$, or the Hermitian conjugate of a left-chiral spinor with a right-chiral spinor $\chi_L^\dagger \chi_R = (\chi_L^\star)^T \chi_R = (\chi_{\dot{a}})^T \chi^{\dot{a}}$ to get Lorentz invariant terms. We will use this later, when need invariant terms that we can use to formulate our laws of nature.

[155] In this context dotted $^{\dot{a}}$ $_{\dot{a}}$ or undotted ^{a} $_a$.

In addition, we have now another justification for calling ϵ^{ab} the spinor metric, because the invariant spinor product in Eq. 3.214, can be written as

$$
\chi_a^T \chi^a \underbrace{=}_{\text{Eq. 3.204}} \chi_a^T \epsilon^{ab} \chi_b . \qquad (3.216)
$$

Compare this to how we defined in Eq. 2.31 the invariant product of Minkowsi space by using the Minkowski metric $\eta^{\mu\nu}$:

$$
x_\mu y^\mu = x_\mu \eta^{\mu\nu} y_\nu . \qquad (3.217)
$$

The spinor metric is for spinors indeed what the Minkowski metric is for four-vectors[156].

[156] Don't get confused why we have no transposition for the four-vectors here. These equations can be read in two ways. On the one hand as vector equations and on the other hand as component equations. It's conventional and sometimes confusing to use the same symbol x_μ for a four-vector and its components. If we read the equation as a component equation we need no transposition. The same is of course true for our spinor products. Nevertheless, we have seen above that we mustn't forget to transpose and in order to avoid errors we included the explicit superscript T, although the spinor equation here can be read as component equation that do not need it. In contrast, for three component vectors there is a clear distinction using the little arrow: $\vec{a}$ has components a_i.

After setting up this notation, we can now write the spinor "metric" with lowered indices

$$\epsilon_{ab} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \tag{3.218}$$

because we need[157] $(-\epsilon)$ to get from χ_R to χ_L. In addition, we can now write the two transformation operators as one object Λ. For example, when it has dotted indices we know it multiplies with a right-chiral spinor and we know which transformation operator to choose:

$$\chi_R \to \chi'_R = \chi'^{\dot{a}} = \Lambda^{\dot{a}}{}_{\dot{b}}\chi^{\dot{b}} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{a}}{}_{\dot{b}} \chi^{\dot{b}} \tag{3.219}$$

and analogously for left-chiral spinors

$$\chi_L \to \chi'_L = \chi'_a = \Lambda_a{}^b \chi_b = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a{}^b \chi_b . \tag{3.220}$$

Therefore:

$$\Lambda_{(\frac{1}{2},0)} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}}\right) \hat{=} \Lambda_a{}^b \tag{3.221}$$

and

$$\Lambda_{(0,\frac{1}{2})} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}}\right) \hat{=} \Lambda^{\dot{a}}{}_{\dot{b}} \tag{3.222}$$

This notation is useful, because, as we have seen, the two different objects χ_L and χ_R aren't so different after all. In fact we can transform them into each other and a unified notation is the logical result.

Now we move on to the next irreducible representation, which will turn out to be an old acquaintance.

3.7.8 The $(\frac{1}{2}, \frac{1}{2})$ Representation

For this representation we use the 2-dimensional representation for both copies of the $SU(2)$ Lie algebra[158] N_i^+ and N_i^-. This time let's have a look at what kind of object our representation is going to act on first. The copies will not interfere with each other, because N_i^+ and N_i^- commute, i.e. $[N_i^+, N_j^-] = 0$ (Eq. 3.172). Therefore, our objects will transform separately under both copies. Let's name the object we want to examine v. This object will have 2 indices v_a^b, each transforming under a separate two-dimensional copy of $\mathfrak{su}(2)$. Here the notation we introduced in the last section comes in handy.

We know that our object v will have 4 components, because each representation is 2 dimensional and this means that both indices can take on two values ($\frac{1}{2}$ and $-\frac{1}{2}$). Therefore, the objects can be 2×2 matrices, but it's also possible to enforce a four component vector form, as we will see[159].

[157] You can check this yourself, but it's not very important for what follows.

[158] Mathematically we have $(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$.

[159] Remember that when we talked about rotations of the plane we were in the same situation. The rotation could be described by complex numbers acting on complex numbers. Doing the map to real matrices we had real matrices acting on real matrices, but the same action could be described by a real matrix acting on a column vector.

First, let's look at the complex matrix choice. A general 2×2 matrix has 4 complex entries and therefore 8 free parameters. As noted above, we only need 4. We can write every complex matrix M as a sum of a Hermitian ($H^\dagger = H$) and an anti-Hermitian ($A^\dagger = -A$) matrix: $M = H + A$. Both Hermitian and anti-Hermitian matrices have 4 free parameters. In addition, we will see in a moment that our transformations in this representation always transform a Hermitian 2×2 matrix into another Hermitian 2×2 matrix and equivalently an anti-Hermitian matrix into another anti-Hermitian matrix. This means Hermitian and anti-Hermitian matrices are invariant subsets. As explained in Section 3.5 this means that working with a general matrix here, corresponds to having a *reducible* representation. Putting these observations together, we conclude that we can assume that our irreducible representation acts on Hermitian 2×2 matrices. A basis[160] for Hermitian 2×2 matrices is given by the Pauli matrices together with the identity matrix.

[160] This means that an arbitrary Hermitian 2×2 matrix can be written as a linear combination of the form: $a_0 1 + a_i \sigma_i$

Instead of examining $v_a^{\dot{b}}$, we will have a look at $v_{a\dot{b}}$, because then we can use the Pauli matrices as defined in Eq. 3.80. Take note that $v_a^{\dot{b}}$ and $v_{a\dot{b}}$ can be transformed into each other by multiplication with $\epsilon^{\dot{b}\dot{c}}$ and therefore if you want to work with $v_a^{\dot{b}}$, you simply have to use the Pauli matrices that have been multiplied with ϵ.

If we define $\sigma^0 = I_{2\times 2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, we can write

$$v_{a\dot{b}} = v_\nu \sigma^\nu_{a\dot{b}} = v^0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + v^1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + v^2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + v^3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.223}$$

As explained above, we could use[161] $v_a^{\dot{b}} = v_\mu \sigma^\mu_{a\dot{c}} \epsilon^{\dot{b}\dot{c}}$ instead, which means we would use the basis $(\tilde{\sigma}_a^{\dot{b}})^\mu = \sigma^\mu_{a\dot{c}} \epsilon^{\dot{b}\dot{c}}$. We therefore write a general Hermitian matrix as

[161] This is really just a basis choice and here we choose the basis that gives us with our definition of the Pauli matrices, the transformation behavior we derived earlier for vectors.

$$v_{ab} = \begin{pmatrix} v_0 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & v_0 - v_3 \end{pmatrix}. \tag{3.224}$$

Remember that we have learned in the last section that different indices transform differently. For example, a lower dotted index transforms differently than a lower undotted index.

Now we have a look at how $v_{a\dot{b}}$ transforms and use the transformation operators that we derived in the last sections

$$v \to v' = v'_{a\dot{b}} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2}+\vec{\phi}\frac{\vec{\sigma}^\star}{2}}\right)^{\dot{d}}_{\ \dot{b}}\right)^T$$

$$= \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2}+\vec{\phi}\frac{\vec{\sigma}^\dagger}{2}}\right)^{\dot{d}}_{\ \dot{b}}$$

$$\underbrace{=}_{\sigma_i^\dagger=\sigma_i} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{d}}_{\ \dot{b}} . \qquad (3.225)$$

We can now see that a Hermitian matrix is after such a transformation still Hermitian, as promised above[162]

$$\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{d}}_{\ \dot{b}} \to \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{d}}_{\ \dot{b}}\right)^\dagger$$

$$\underbrace{=}_{(ABC)^\dagger=C^\dagger B^\dagger A^\dagger} \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{d}}_{\ \dot{b}}\right)^\dagger v^\dagger_{c\dot{d}} \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c}\right)^\dagger$$

$$= \left(e^{i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2}+\vec{\phi}\frac{\vec{\sigma}^\dagger}{2}}\right)^{\dot{d}}_{\ \dot{b}} v^\dagger_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2}+\vec{\phi}\frac{\vec{\sigma}^\dagger}{2}}\right)_a^{\ c}$$

$$\underbrace{=}_{\text{if } v^\dagger_{c\dot{d}}=v_{c\dot{d}}} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{d}}_{\ \dot{b}} \quad \checkmark$$

$$(3.226)$$

[162] Exactly the same computation shows that an anti-Hermitian matrix is still anti-Hermitian after such a transformation. To see this, use in the last step instead of $v^\dagger_{c\dot{d}} = v_{c\dot{d}}$ that $v^\dagger_{c\dot{d}} = -v_{c\dot{d}}$.

The explicit computation for an arbitrary transformation is long and tedious[163] so instead, we will look at one specific example. Let's boost v along the z-axis[164]

$$v_{a\dot{b}} \to v'_{a\dot{b}} = \left(e^{\phi\frac{\sigma_3}{2}}\right)_a^{\ c} v_{c\dot{d}} \left(e^{\phi\frac{\sigma_3}{2}}\right)^{\dot{d}}_{\ \dot{b}}$$

$$= \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}} \end{pmatrix} \begin{pmatrix} v_0+v_3 & v_1-iv_2 \\ v_1+iv_2 & v_0-v_3 \end{pmatrix} \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}} \end{pmatrix}$$

$$= \begin{pmatrix} e^{\phi}(v_0+v_3) & v_1-iv_2 \\ v_1+iv_2 & e^{-\phi}(v_0-v_3) \end{pmatrix} , \qquad (3.227)$$

[163] See, for example, page 128 in Matthew Robinson. *Symmetry and the Standard Model*. Springer, 1st edition, August 2011. ISBN 978-1-4419-8267-4

[164] This means $\vec{\phi} = (0,0,\phi)^T$. Such a boost is the easiest because σ_3 is diagonal. For boosts along other axes the exponential series must be evaluated in detail.

where we have used the fact that σ_3 is diagonal[165] and that $e^A = \begin{pmatrix} e^{A_{11}} & 0 \\ 0 & e^{A_{22}} \end{pmatrix}$ holds for every diagonal matrix. We now compare the transformed object that we calculated in Eq. 3.227 with a generic object v':

$$v'_{a\dot{b}} = \begin{pmatrix} v'_0+v'_3 & v'_1-iv'_2 \\ v'_1+iv'_2 & v'_0-v'_3 \end{pmatrix} = \begin{pmatrix} e^{\phi}(v_0+v_3) & v_1-iv_2 \\ v_1+iv_2 & e^{-\phi}(v_0-v_3) \end{pmatrix} .$$

This tells us how the components of the transformed object are related to the untransformed components[166]

[165] $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

[166] We rewrite the equations using the connection between the hyperbolic sine, the hyperbolic cosine function and the exponential function $e^{-\phi} = (\cosh(\phi) - \sinh(\phi))$ and $e^{\phi} = (\cosh(\phi) + \sinh(\phi))$, which is conventional in this context. If you are unfamiliar with these functions you can either take notice of their definitions: $\cosh(\phi) \equiv \frac{1}{2}\left(e^{\phi} + e^{-\phi}\right)$ and $\sinh(\phi) \equiv \frac{1}{2}\left(e^{\phi} - e^{-\phi}\right)$ or rewrite the few equations here in terms of e^{ϕ} and $e^{-\phi}$, which is equally good.

$$\rightarrow v_0' + v_3' = e^{\phi}(v_0 + v_3) = (\cosh(\phi) + \sinh(\phi))(v_0 + v_3)$$

$$\rightarrow v_0' - v_3' = e^{-\phi}(v_0 - v_3) = (\cosh(\phi) - \sinh(\phi))(v_0 - v_3).$$

The addition and subtraction of both equations yields

$$\begin{aligned}\rightarrow v_0' &= \cosh(\phi)v_0 + \sinh(\phi)v_3 \\ \rightarrow v_3' &= \sinh(\phi)v_0 + \cosh(\phi)v_3 \,. \end{aligned} \tag{3.228}$$

This is exactly what we get using the 4-vector formalism[167]

$$\begin{pmatrix} v_0' \\ v_1' \\ v_2' \\ v_3' \end{pmatrix} = \begin{pmatrix} \cosh(\phi) & 0 & 0 & \sinh(\phi) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(\phi) & 0 & 0 & \cosh(\phi) \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$= \begin{pmatrix} \cosh(\phi)v_0 + \sinh(\phi)v_3 \\ v_1 \\ v_2 \\ \sinh(\phi)v_0 + \cosh(\phi)v_3 \end{pmatrix} . \tag{3.229}$$

[167] See 3.154 for the explicit form of the matrix for a boost along the z-axis.

This is true for arbitrary Lorentz transformations, as you can check by computing the other possibilities. What we have shown here is that the $(\frac{1}{2}, \frac{1}{2})$ representation *is* the vector representation. We can simplify our transformation laws by using the enforced vector form, because multiplying a matrix with a vector is simpler than the multiplication of three matrices. Nevertheless, we have seen how the familiar 4-vector is related to the more fundamental spinors. A 4-vector **is** a rank-2 spinor, which means a spinor with 2 indices that transforms according to the $(\frac{1}{2}, \frac{1}{2})$ representation of the Lorentz group. Furthermore, **we can now see that 4-vectors aren't appropriate to describe every physical system on a fundamental level, because they aren't fundamental**. There are physical systems they cannot describe.

We can now understand why some people say that "spinors are the square root of vectors". This is meant in the same way as vectors are the square root of rank-2 tensors[168]. A rank-2 tensor has two vector indices and a vector has two spinor indices. Therefore, the most basic object that can be Lorentz transformed is indeed a spinor.

[168] A rank-2 tensor is simply a matrix $M_{\mu\nu}$.

When we started our studies of the Lorentz group, we noted that it consists of four components[169]. These components are connected by the parity and the time-reversal operator[170]. Therefore, to be able to describe all transformations that preserve the speed of light, we need to find the parity and time-reversal transformation for each representation. In this text we restrict ourselves to parity transformations,

[169] Eq. 3.142: $O(1,3) = \{L_+^\uparrow, \Lambda_P L_+^\uparrow, \Lambda_T L_+^\uparrow, \Lambda_P \Lambda_T L_+^\uparrow\}$.

[170] See Eq. 3.142

because the discussion for the time-reversal transformation is very similar. We will discuss in in a later section that nature isn't always symmetric under parity transformations[171].

[171] Nature isn't invariant under time-reversal transformations either, but a satisfying discussion of this curious fact lies beyond the scope of this book.

3.7.9 Spinors and Parity

Up to this point, there is no justification why we call the objects transforming according to the $(\frac{1}{2},0)$ representation left-chiral and the objects transforming according to the $(0,\frac{1}{2})$ representation right-chiral. After talking a bit about parity transformation, this will make sense.

Recall that we already know the behavior of the generators of the Lorentz group under parity transformations. The result was Eq. 3.161, which we recite here for convenience

$$J_i \underbrace{\rightarrow}_{\text{P}} J_i \qquad K_i \underbrace{\rightarrow}_{\text{P}} -K_i \,. \tag{3.230}$$

By looking at the definition of the generators $N^{\pm}$ in Eq. 3.169, which we also recite here

$$N_i^{\pm} = \frac{1}{2}(J_i \pm iK_i)\,. \tag{3.231}$$

we can see that under parity transformations $N^+ \leftrightarrow N^-$. Therefore, the $(0,\frac{1}{2})$ representation of a transformation, becomes the $(\frac{1}{2},0)$ representation of this transformation and vice versa under parity transformations. This is the reason for talking about left- and right-chiral spinors[172]. Just as a right-handed coordinate system changes into a left-handed coordinate system under parity transformations, these two representations change into each other.

[172] The conventional name is left- and right-handed spinors, but this can be quite confusing, because the notions left-handed and right-handed are directly related to a concept called helicity, which is different from chirality. Anyway the name should make some sense, because something left is changed into something right under parity transformations.

Rotations are the same for both representations, but boost transformations differ by a sign and it is easy to make the above statement explicit:

$$(\Lambda_{\vec{K}})_{(\frac{1}{2},0)} = \mathrm{e}^{\vec{\phi}\vec{K}} \underbrace{\rightarrow}_{\text{P}} \mathrm{e}^{-\vec{\phi}\vec{K}} = (\Lambda_{\vec{K}})_{(0,\frac{1}{2})} \tag{3.232}$$

$$(\Lambda_{\vec{K}})_{(0,\frac{1}{2})} = \mathrm{e}^{-\vec{\phi}\vec{K}} \underbrace{\rightarrow}_{\text{P}} \mathrm{e}^{\vec{\phi}\vec{K}} = (\Lambda_{\vec{K}})_{(\frac{1}{2},0)}. \tag{3.233}$$

We learn here that if we want to describe a physical system that is invariant under parity transformations, we will always need right-chiral **and** left-chiral spinors. The easiest thing to do is to write them below each other into a single object called **Dirac spinor**

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_a \\ \xi^{\dot{a}} \end{pmatrix} \,. \tag{3.234}$$

Recalling that the generic name for left- and right-chiral spinors is **Weyl spinors**, we can say that a Dirac spinor Ψ consists of two Weyl spinors χ_L and ξ_R. Note that we want to stay general here and don't assume any a priori connection between χ and ξ. A Dirac spinor of the form

$$\Psi_M = \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} \tag{3.235}$$

is a special case, called Majorana spinor. A Dirac is *not a four-vector*, because it transforms completely different[173]. A Dirac spinor transforms according to the $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation[174] of the Lorentz group, which means nothing more than writing the corresponding transformations in block-diagonal form into one big matrix:

[173] Equally, a Majorana spinor is not a vector.

[174] This a *reducible* representation, which is obvious because of the block-diagonal form of the transformation matrix. In contrast, four-vectors transform according to the $(\frac{1}{2},\frac{1}{2}) = (\frac{1}{2},0) \otimes (0,\frac{1}{2})$ representation.

$$\Psi \to \Psi' = \Lambda_{(\frac{1}{2},0)\oplus(0,\frac{1}{2})}\Psi = \begin{pmatrix} \Lambda_{(\frac{1}{2},0)} & 0 \\ 0 & \Lambda_{(0,\frac{1}{2})} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.236}$$

For example, a boost transformation is in this representation

$$\Psi \to \Psi' = \begin{pmatrix} e^{\frac{\vec{\phi}}{2}\vec{\sigma}} & 0 \\ 0 & e^{\frac{-\vec{\phi}}{2}\vec{\sigma}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.237}$$

It is instructive to investigate how Dirac spinors behave under parity transformations, because once we know how Dirac spinors transform under parity transformations, we can check if a given theory is invariant under such transformations. We can't expect that a Dirac spinor is after a parity transformation still a Dirac spinor (an object transforming according to the $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation), because we know that under parity transformations $N^+ \leftrightarrow N^-$ and therefore

$$\left(0,\frac{1}{2}\right) \underbrace{\leftrightarrow}_{P} \left(\frac{1}{2},0\right). \tag{3.238}$$

We conclude that if a Dirac spinor transforms according to the $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation, the parity transformed object transforms according to the $(0,\frac{1}{2}) \oplus (\frac{1}{2},0)$ representation.

$$\Psi^P \to (\Psi^P)' = \Lambda_{(0,\frac{1}{2})\oplus(\frac{1}{2},0)}\Psi^P = \begin{pmatrix} \Lambda_{(0,\frac{1}{2})} & 0 \\ 0 & \Lambda_{(\frac{1}{2},0)} \end{pmatrix} \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix}. \tag{3.239}$$

Therefore

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad \to \quad \Psi^P = \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix}. \tag{3.240}$$

A parity transformed Dirac spinor contains the same objects ξ_R, χ_L as the untransformed Dirac spinor, only written differently. A parity transformation does nothing like $\xi_L \to \xi_R$, which is a different kind of transformation we will talk about in the next section.

3.7.10 Spinors and Charge Conjugation

In Section 3.7.7 we stumbled upon a transformation, which yields $\chi_L \to \chi_R$ and $\xi_R \to \xi_L$. The transformation is achieved by $\chi_L \to \chi_L^C = \epsilon\chi_L^\star = \chi_R$ and analogously for a right-chiral spinor $\xi_R \to \xi_R^C = (-\epsilon)\xi_R^\star = \xi_L$. We are now able to understand it from a quite different perspective.

Up to this point, we used this transformation merely as a computational trick in order to raise and lower indices. Now, how does a Dirac spinor transform under such a transformation? Naively we get:

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \to \tilde{\Psi} = \begin{pmatrix} \chi_L^C \\ \xi_R^C \end{pmatrix} = \begin{pmatrix} \chi_R \\ \xi_L \end{pmatrix}. \tag{3.241}$$

Unfortunately, this object does not transform like a Dirac spinor[175], which transform under boosts as follows

$$\Psi \to \Psi' = \begin{pmatrix} e^{\frac{\vec{\theta}}{2}\vec{\sigma}} & 0 \\ 0 & e^{\frac{-\vec{\theta}}{2}\vec{\sigma}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.242}$$

In contrast, the object $\tilde{\Psi}$ we get from the naive operation, transforms as

$$\tilde{\Psi} \to \tilde{\Psi}' = \begin{pmatrix} e^{-\frac{\vec{\theta}}{2}\vec{\sigma}} & 0 \\ 0 & e^{\frac{\vec{\theta}}{2}\vec{\sigma}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.243}$$

This is a different kind of object, because it transforms according to a different representation of the Lorentz group. Therefore we write

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \to \Psi^C = \begin{pmatrix} \xi_R^C \\ \chi_L^C \end{pmatrix} = \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}, \tag{3.244}$$

which incorporates the transformation behavior we observed earlier and transforms like a Dirac spinor. This operation is commonly called **charge conjugation**, which can be a little misleading. We know that this transformation transforms a left-chiral spinor into a right-chiral, i.e. flips one label we use to describe our elementary particles[176]. Later we will learn that this operator flips not only one, but all labels we use to describe fundamental particles. One such label is electric charge, hence the name charge conjugation, but before we are able to show this, we need of course to understand first what electric

[175] Unlike for parity transformations, we have a choice here and we prefer to keep working with the same kind of object. The object $\tilde{\Psi}$ can then be seen as a Dirac spinor that has been parity transformed and charge conjugated.

[176] For the more advanced reader: Recall that each Weyl spinor we are talking about here, is in fact a two component object. Later we will define a physical measurable quantity, called spin, that is described by $\frac{1}{2}\sigma_3$. The matrix ϵ, flips an object with eigenvalue $+\frac{1}{2}$ for the spin operator $\frac{1}{2}\sigma_3$ into an object with eigenvalue $-\frac{1}{2}$. This is commonly interpreted as spin flip, which means an object with spin $\frac{1}{2}$, becomes an object with spin $-\frac{1}{2}$.

charge is. Nevertheless, it's always important to remember that all labels get flipped, not only electric charge.

We could now go on and derive higher-dimensional representations of the Lorentz group, but at this point we already have every finite-dimensional irreducible representation we need for the purpose of this text. Nevertheless, there is another representation, the infinite-dimensional representation, that is especially interesting, because we need it to transform fields, like, for example, the electromagnetic field.

3.7.11 Infinite-Dimensional Representations

In the last sections, we talked about finite-dimensional representations of the Lorentz group and learned how we can classify them. These finite-dimensional representations acted on constant one-, two- or four-component objects. In physics the objects we are dealing with are dynamically changing in space and time, so we need to understand how such objects transform. So far, we have dealt with transformations of the form

$$\Phi_a \to \Phi'_a = M_{ab}(\Lambda)\Phi_b \,, \tag{3.245}$$

where $M_{ab}(\Lambda)$ denotes the matrix of the particular finite-dimensional representation of the Lorentz transformation Λ. This means $M_{ab}(\Lambda)$ is a matrix that acts, for example, on a two-component object like a Weyl spinor. The result of the multiplication with this matrix is simply that the components of the object in question get mixed and are multiplied with constant factors. If our object Φ changes in space and time, it is a function of coordinates[177] $\Phi = \Phi(x)$ and these coordinates are affected by the Lorentz transformations, too. In general we have

$$x^\mu \to \Lambda^\mu_\nu x^\nu, \tag{3.246}$$

where Λ^μ_ν denotes the vector representation (= $(\frac{1}{2}, \frac{1}{2})$ representation) of the Lorentz transformation in question. We have in this case[178]

$$\Phi_a(x) \to M_{ab}(\Lambda)\Phi_b(\Lambda x). \tag{3.247}$$

Our transformation will therefore consist of two parts. One part, represented by a finite-dimensional representation, acting on Φ_a and a second part that transforms the spacetime coordinates x. This second part will act on an infinite-dimensional[179] vector space and we therefore need an infinite-dimensional representation. The infinite-dimensional representation of the Lorentz group is given by differential operators[180]

$$M^{\text{inf}}_{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu)\,. \tag{3.248}$$

[177] Here x is a shorthand notation for all spacetime coordinates t, x, y, z

[178] Most books use the Wigner convention for symmetry operators: $\Phi_a(x) \to M_{ab}(\Lambda)\Phi_b(\Lambda^{-1}x)$, but unfortunately there is at this point no way to motivate this convention.

[179] Each component of Φ is now a function of x. The corresponding operators act on $\Phi_a(x)$, i.e. functions of the coordinates and the space of functions is in this context infinite-dimensional. The reason that the space of functions is infinite-dimensional is that we need an infinite number of basis functions. The expansion of an arbitrary function in terms of such an infinite number of basis functions is the idea behind the Fourier transform as explained in Appendix D.1.

[180] The symbols ∂^ν are a shorthand notation for the partial derivative $\frac{\partial}{\partial_\nu}$.

You can check by a straightforward computation that $M^{\text{inf}}_{\mu\nu}$ satisfies the Lorentz algebra (Eq. 3.175) and transforms the coordinates as desired.

The transformation of the coordinates is now given by[181]

$$\Phi'(\Lambda x) = \mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M^{\text{inf}}_{\mu\nu}}\Phi(x), \tag{3.249}$$

where the exponential function is, as usual, understood in terms of its series expansion. The complete transformation is then a combination of a transformation generated by the finite-dimensional representation $M^{\text{fin}}_{\mu\nu}$ and a transformation generated by the infinite-dimensional representation $M^{\text{inf}}_{\mu\nu}$ of the generators:

$$\Phi_a(x) \to \left(\mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M^{\text{fin}}_{\mu\nu}}\right)_a^{\;b} \mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M^{\text{inf}}_{\mu\nu}}\Phi_b(x). \tag{3.250}$$

Because our matrices $M^{\text{fin}}_{\mu\nu}$ are finite-dimensional and constant we can put the two exponents together

$$\Phi_a(x) \to \left(\mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M_{\mu\nu}}\right)_a^{\;b}\Phi_b(x) \tag{3.251}$$

with $M_{\mu\nu} = M^{\text{fin}}_{\mu\nu} + M^{\text{inf}}_{\mu\nu}$. This representation of the generators of the Lorentz group is called **field representation**.

We can now talk about a different kind of transformation: **translations**, which means transformations to another location in spacetime. Translations do not mix the components and therefore, we need no finite-dimensional representation. However, it's quite easy to find the infinite-dimensional representation for translations. These are not part of the Lorentz group, but the laws of nature should be location independent. The Lorentz group (boosts and rotations) plus translations is called the **Poincaré group**, which is the topic of the next section. Nevertheless, we will introduce the infinite-dimensional representation of translations already here. For simplicity, we restrict ourselves to one dimension. In this case an **infinitesimal** translation of a function, along the x-axis is given by[182]

$$\Phi(x) \to \Phi(x+\epsilon) = \Phi(x) + \underbrace{\partial_x\Phi(x)}_{\text{"rate of change" along the x-axis}}\epsilon,$$

which is, of course, the first term of the Taylor series expansion. It is conventional in physics to add an extra $-i$ to the generator and we therefore define

$$P_i \equiv -i\partial_i. \tag{3.252}$$

With this definition an arbitrary, finite translation is[183]

[181] Recall the definition of $M^{\mu\nu}$ in Eq. 3.173. The components of $\omega_{\mu\nu}$ can then be directly related to the usual rotation angles $\theta_i = \frac{1}{2}\epsilon_{ijk}\omega_{jk}$ and the boost parameters $\phi_i = \omega_{0i}$.

[182] For a non-infinitesimal ϵ, we would get here infinitely many terms. But for an infinitesimal ϵ, we use $\epsilon^2 \approx 0$ and therefore neglect all higher order terms.

[183] Using $i^2 = -1$.

$$\Phi(x) \to \Phi(x+a) = \mathrm{e}^{-ia^i P_i}\Phi(x) = \mathrm{e}^{-a^i \partial_i}\Phi(x)\,,$$

where a^i denotes the amount we want to translate in each direction. If we write the exponential function as Taylor series[184], this equation can simply be seen as the Taylor expansion[185] for $\Phi(x+a)$. If we want to transform to another point in time we use $P_0 = i\partial_0$, for a different location we use $P_i = -i\partial_i$.

[184] This is derived in Appendix B.4.1.

[185] This is derived in Appendix B.3.

Now, let's move on to the full spacetime symmetry group of nature: the Poincaré group.

3.8 The Poincaré Group

The Lorentz group includes rotations and boosts. Further transformations that leave the speed of light invariant are translations in space and time, because measuring the speed of light at a different point in spacetime does not change its value. If we add these symmetries to the Lorentz group we get the Poincaré group[186]

[186] The Poincaré group is not the direct, but the semi-direct, sum of the Lorentz group and translations, but for the purpose of this text we can neglect this technical detail.

$$\begin{aligned}\text{Poincaré group} &= \text{Lorentz group } \textbf{plus} \text{ translations} \\ &= \text{Rotations } \textbf{plus} \text{ boosts } \textbf{plus} \text{ translations}\end{aligned} \tag{3.253}$$

The generators of the Poincaré group are the generators of the Lorentz group J_i, K_i plus the generators of translations P_μ.

In terms of J_i, K_i and P_μ the Lie algebra reads[187]

[187] This is not very enlightening, but included for completeness.

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.254}$$

$$[J_i, K_j] = i\epsilon_{ijk}K_k \tag{3.255}$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k \tag{3.256}$$

$$[J_i, P_j] = i\epsilon_{ijk}P_k \tag{3.257}$$

$$[J_i, P_0] = 0 \tag{3.258}$$

$$[K_i, P_j] = i\delta_{ij}P_0 \tag{3.259}$$

$$[K_i, P_0] = -iP_i\,. \tag{3.260}$$

Because this looks like a huge mess it is conventional to write this in terms of $M_{\mu\nu}$, which was defined by

$$J_i = \frac{1}{2}\epsilon_{ijk}M_{jk} \tag{3.261}$$

$$K_i = M_{0i}. \tag{3.262}$$

With $M_{\mu\nu}$ the **Poincaré algebra** reads

$$[P_\mu, P_\nu] = 0 \tag{3.263}$$

$$[M_{\mu\nu}, P_\rho] = i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu) \tag{3.264}$$

and of course still have

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}). \tag{3.265}$$

For this quite complicated group it is very useful to label the representations by using the fixed scalar values of the Casimir operators. The Poincaré group has two Casimir operators[188] . The first one is:

$$P_\mu P^\mu =: m^2. \tag{3.266}$$

We give the scalar value the suggestive name m^2, because we will learn later that it coincides with the mass of particles[189].

The second Casimir operator is $W_\mu W^\mu$, where[190]

$$W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}P_\nu M_{\rho\sigma} \tag{3.267}$$

which is called the **Pauli-Lubanski four-vector**. In a lengthy computation it can be justified, that in addition to m, we use the number $j \equiv j_1 + j_2$ as a label. This label is commonly called **spin**. For the moment this is just a name. Later we will understand why the name spin is appropriate. Exactly as for the Lorentz group, we have one j_i for each of the two[191] representations of the $SU(2)$ algebra.

For example, the $(j_1, j_2) = (0,0)$ representation is called spin 0 representation[192]. The $(j_1, j_2) = (\frac{1}{2}, 0)$ and $(j_1, j_2) = (0, \frac{1}{2})$ are both called spin $\frac{1}{2}$ representations[193] and analogously the $(j_1, j_2) = (\frac{1}{2}, \frac{1}{2})$ representation is called spin 1 representation[194].

The message to take away is that each representation is labelled by two scalar values: m and j. While m can take on arbitrary values, j is restricted to half-integer or integer values.

188 Recall that a Casimir operator is defined as an operator, constructed from the generators, that commutes with all other generators.

189 Don't worry, this will make much more sense later.

190 $\epsilon^{\mu\nu\rho\sigma}$ is the four-dimensional Levi-Civita symbol, which is defined in Appendix B.5.5.

191 Recall that the complexified Lie algebra of the Lorentz group could be seen to consist of two copies of the Lie algebra of $SU(2)$. The representations of $SU(2)$ could be labelled by a number j and consequently we used for representations of the double cover of the Lorentz group two numbers (j_1, j_2).

192 $j_1 + j_2 = 0 + 0 = 0$

193 $j_1 + j_2 = \frac{1}{2} + 0 = 0 + \frac{1}{2} = \frac{1}{2}$

194 $j_1 + j_2 = \frac{1}{2} + \frac{1}{2} = 1$

3.9 Elementary Particles

The labels for the irreducible representations of double cover of the Poincaré group, **mass** m and by their **spin** ($= j$ here), are how elementary particles are **labeled** in physics[195]. An elementary particle with given labels m and spin, say $j = \frac{1}{2}$, is described by an object, which transforms according to the m, spin $\frac{1}{2}$ representation of the Poincaré group.

[195] Some physicists go even farther and say that elementary particles **are** the irreducible representations of the Poincaré group.

More labels, called charges, will follow later from internal symmetries. These labels are used to define an elementary particle. For example, an electron is defined by the following labels

- mass: $9,109 \cdot 10^{-31}$ kg,
- spin: $\frac{1}{2}$,
- electric charge: $1,602 \cdot 10^{-19}$ C,
- weak charge, called weak isospin: $-\frac{1}{2}$,
- strong charge, called color charge: 0.

These labels determine how a given elementary particle behaves in experiments. The representations we derived in this chapter define how we can describe them mathematically. An elementary particle with[196]

[196] Remember that in the introductory remarks about what we can't derive, it was said there is no real reason to stop here after three representations. We could go on to higher dimensional representations, but there are no elementary particles, for example, with spin $\frac{3}{2}$. Nevertheless, such representations can be used to describe composite objects. In addition, there are many physicists that believe the fundamental particle mediating gravity, called graviton, has spin 2 and therefore a corresponding higher dimensional representation must be used to describe it.

- spin 0 is described by an object Φ, called **scalar**, that transforms according to the $(0, 0)$, called **spin** 0 **representation** or scalar representation. For example, the Higgs particle is described by a scalar field.
- spin $\frac{1}{2}$ is described by an object Ψ, called **spinor**, that transforms according to the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation, called **spin** $\frac{1}{2}$ **representation** or spinor representation. For example, electrons and quarks are described by spinors.
- spin 1 is described by an object A, called **vector**, that transforms according to the $(\frac{1}{2}, \frac{1}{2})$, called **spin** 1 **representation** or vector representation. For example photons are described by vectors.

This is an incredibly important, deep and beautiful insight, so again:

What we get from deriving the irreducible representations of the Poincaré group are the **mathematical tools we need to describe all elementary particles**. To describe spin 0 particles, like the Higgs Boson, we use mathematical objects, called **scalars**, that transform according to the spin 0 representation. To describe **spin $\frac{1}{2}$ particles** like electrons, neutrinos, quarks etc. we use mathematical objects, called **spinors**, that transform according to the spin $\frac{1}{2}$ representation. To describe photons or other particles with **spin** 1 we use objects, called **vectors**, transforming according to the spin 1 representation.

An explanation for the very suggestive name spin, will be given in Section 8.5.5, after we talked about how and what we measure in experiments. We first have to know how we are able to find out if something is spinning, before we can justify the name spin. At this point, spin is merely a label.

Further Reading Tips

- **John Stillwell - Naive Lie Theory**[197] is a very readable, math orientated introduction to Lie Theory.
- **Nadir Jeevanjee - An Introduction to Tensors and Group Theory for Physicists**[198] is a very good introduction, with strong focus on the usage of group theory in physics.
- **Howard Georgi - Lie Algebras In Particle Physics**[199] is a great book to learn more about the parts of Lie group theory that are relevant for particle physics.

[197] John Stillwell. *Naive Lie Theory*. Springer, 1st edition, August 2008. ISBN 978-0387782140

[198] Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

[199] Howard Georgi. *Lie Algebras In Particle Physics: from Isospin To Unified Theories (Frontiers in Physics)*. Westview Press, 2 edition, 10 1999. ISBN 9780738202334

3.10 Appendix: Rotations in a Complex Vector Space

The concept of transformations that preserve the inner product can also be used in complex vector spaces. We want the inner product of a vector with itself to be a real number, because by definition this should result in the squared length of the vector. A complex number would make little sense as the length of the vector. Therefore, the inner product of complex vector spaces is defined with additional complex conjugation[200]

[200] Because for $z = a + ib$ we have $z^\star = a - ib$ and therefore $z^\star z = (a+ib)(a-ib) = a^2 + b^2$, which is real.

$$a \cdot a = (a^T)^\star a = a^\dagger a. \tag{3.268}$$

The symbol †, called dagger, denotes Hermitian conjugation, which means complex conjugation and transposing. We see that a transformation that preserves this inner product must fulfil the condition $U^\dagger U = 1$:

$$(Ua) \cdot (Ua) = a^\dagger U^\dagger U a = a^\dagger a\,. \tag{3.269}$$

Transformations like these form groups that are called $U(n)$, where n denotes the dimensions of the complex vector space and "U" stands for unitary. Again the groups $SU(n)$ are called "special", because their elements fulfil the additional condition $\det(U) = 1$.

3.11 Appendix: Manifolds

A manifold M is a set of points with a continuous 1-1 map from each open neighborhood onto an open set of R^n. In easy words this means that a manifold M looks **locally** like the standard R^n. This map from each open neighborhood of M onto R^n associates with each point P of M an n-tupel $(x_1(P), ...x_n(P))$ where the numbers $x_1(P), ...x_n(P)$ are called the coordinates of the point P. Therefore another way of thinking about a n-dimensional manifold is that it's a set, which can be given n independent coordinates ins some neighborhood of any point.

An example for a manifold is the surface of a ball. The surface of the three-dimensional ball is called two-sphere S^2 and is defined as the set of points in R^3 for which $x^2 + y^2 + z^2 = r$ holds, where r is the radius of the sphere. Take note that the surface of the three-dimensional ball is two-dimensional, because the definition involves 3 coordinates and one condition, which eliminates one degree of freedom. That is why it's called mathematically two-sphere. To see that the sphere is a manifold we need a map onto R^2. This map is given by the usual spherical coordinates.

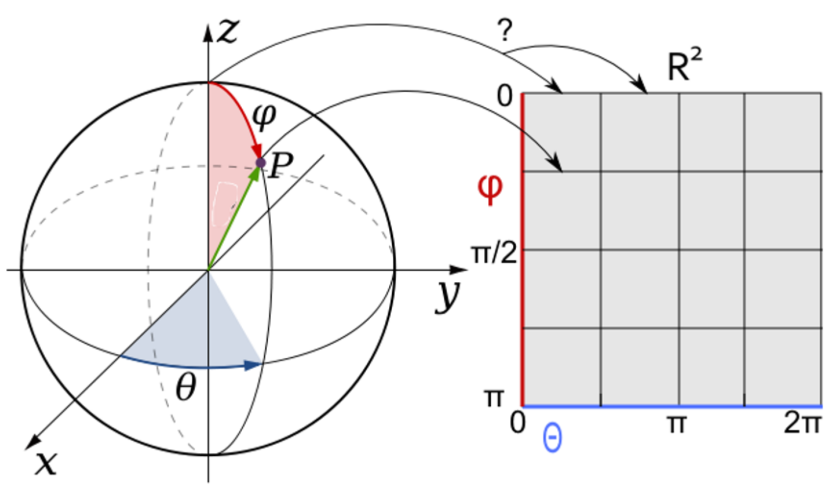


Fig. 3.9: Illustration of the map from one neighborhood of the sphere on to R^n.

Almost all points on the surface of the sphere can be identified unambiguously with a coordinate combination of the form (φ, θ). Almost all! Where is the pole $\varphi = 0$ mapped to? There is no one-to-one identification possible, because the pole is mapped to a whole line, as indicated in the image. Therefore this map does not work for the complete sphere and we need another map in the neighborhood of the pole to describe things there. A similar problem occurs for the map on the semicircle $\theta = 0$. Each point can be mapped in the R^2 to $\theta = 0$ and $\theta = 2\pi$, which is again not a one-to-one map. This illustrates the fact that for manifolds there is in general not one coordinate system for all points of the manifold, only local coordinates, which are valid in some neighborhood. This is no problem because the defining feature of a manifold is that it looks locally like R^n.

The spherical coordinate map is only valid in the open neighborhood $0 < \varphi < \pi, 0 < \theta < 2\pi$ and we need a second map to cover the whole sphere. We can use, for example, a second spherical coordinate system with different orientation, such that the problematic poles lie at different points for this map and no longer at $\varphi = 0$. With this second map every point of the sphere has a map onto R^2 and the two-sphere can be seen to be a manifold.

A trivial example for a manifold is of course R^n.