Jakob Schwichtenberg

Physics from Symmetry

Springer

3.6 $SU(2)$

We used in Sec. 3.4.3 specific matrices (=a specific representation) to identify how the generators of $SU(2)$ behave, when put into the Lie bracket. We can use this knowledge to find further representations. We will arrive again at the representation we started with, which means the set of unitary $2\times$ matrices with unit determinant and are then able to see that it is just one special case. Before we are going to tackle this task, we want to take a moment to think about what representations we can expect.

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

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

every Hermitian traceless 2×2 matrix can be written as a linear combination of these **Pauli matrices**.

We can put these explicit matrices for the basis generators into the Lie bracket, which yields

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \tag{3.82}$$

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

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.83}$$

ones we used in Sec. 3.4.3, by linear combination[83]

$$J_+ = \frac{1}{\sqrt{2}}(J_1 + iJ_2) \tag{3.94}$$

$$J_- = \frac{1}{\sqrt{2}}(J_1 - iJ_2) \tag{3.95}$$

These new operators obey the following commutation relations, as you can check by using the commutator relations in Eq. 3.83

$$[J_3, J_\pm] = \pm J_\pm \tag{3.96}$$

$$[J_+, J_-] = J_3. \tag{3.97}$$

If we now investigate how these operators act on an eigenvector v of J_3 with eigenvalue[84] b we discover something remarkable:

$$\begin{aligned} J_3(J_\pm v) &= J_3(J_\pm v) + \underbrace{J_\pm J_3 v - J_\pm J_3 v}_{=0} \\ &= \underbrace{J_\pm J_3 v}_{=J_\pm bv} + \underbrace{J_3 J_\pm v - J_\pm J_3 v}_{=[J_3,J_\pm]v} \\ &\underbrace{=}_{\text{Eq. 3.96}} (b \pm 1) J_\pm v \end{aligned} \tag{3.98}$$

We conclude that $J_\pm v$ is again an eigenvector, let's call him w, of J_3 with eigenvalue $(b \pm 1)$:

$$J_3 w = (b \pm 1) w \quad \text{with} \quad w = J_\pm v. \tag{3.99}$$

The operators J_- and J_+ are called raising and lowering or **ladder operators**. We can construct more and more eigenvectors of J_3 using the operators 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 there must be an eigenvector with a maximum eigenvalue v_{max}. After a finite number N of applications of J_+ we reach the maximum eigenvector v_{max}

$$v_{max} = J_+^N v \tag{3.100}$$

We have

$$J_+ v_{max} = 0, \tag{3.101}$$

[83] We can always diagonalize **one** of the generators. Following the convention we choose J_3 as diagonal and therefore yielding the basis vectors for our vector space. Furthermore, it is conventional to introduce the new operators $J_\pm$ in the way we do here.

[84] This means $J_3 v = bv$ as explained in appendix C.4.

because v_{max} is, by definition, the eigenvector with the highest eigenvalue. We call the maximum eigenvalue $j := b + N$. The same must be true for the other direction: There must be an eigenvector with minimum eigenvalue v_{min} for which the following relation holds

$$J_- v_{min} = 0 \tag{3.102}$$

Let us say we reach the minimum after operating M times with J_- on v_{max}

$$v_{min} = J_-^M v_{max}. \tag{3.103}$$

Therefore, v_{min} has eigenvalue j-M. To go further we need to know how exactly $J_\pm$ acts on eigenvectors. The computation above shows that $J_- v_k$ is, in general, a scalar multiplied by an eigenvector with eigenvalue $k-1$:

$$J_- v_k = \alpha_k v_{k-1}. \tag{3.104}$$

If we inspect in detail how J_- acts on v_{max} we get[85] the general rule for the scalar factor

$$\alpha_{j-k} = \frac{1}{\sqrt{2}}\sqrt{(2j-k)(k+1)} \tag{3.105}$$

Take note that this scalar factor becomes zero for $k = 2j$ and therefore, we have reached the end of the ladder after $2j$ steps if we start at the top. Therefore v_{min} has eigenvalue $j - 2j = -j$. We conclude that we have in general $2j+1$ eigenstates with eigenvalues

$$\{-j, -j+1, \ldots, j-1, j\} \tag{3.106}$$

This is only possible if j is an integer or an half-integer[86]. Now we know that our vector space V has $2j+1$ dimensions[87], because we have $2j+1$ linearly independent eigenvectors. Those eigenvectors of J_3 span the complete vector space V because J_1 and J_2 can be expressed in terms of J_+ and J_- and therefore take any linear combination $\sum_i a_i v_i$ into a possibly different linear combination $\sum_i b_i v_i$, with scalar factors a_i, b_i. Therefore, the span of the eigenvectors of J_3 is a non-zero invariant subspace of V and because we are looking for irreducible representations they span the complete vector space V.

We can use the construction above to define representations of $SU(2)$ on a vector space V_j with $2j+1$ dimensions and basis given by the eigenvectors v_k of J_3. Furthermore, it's possible to show that every irreducible representation of $SU(2)$ must be equivalent to one of these[88].

[85] See, for example, page 90 in Matthew Robinson. *Symmetry and the Standard Model*. Springer, 1st edition, August 2011. ISBN 978-1-4419-8267-4

[86] Try it with other fractions if you don't believe this!

[87] See, for example, page 189 in Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

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

3.6.2 The Casimir Operator of $SU(2)$

As described in Sec. 3.5, we can naturally label representations by using the Casimir operators[89] of the group. $SU(2)$ has exactly one Casimir operator:

$$J^2 := (J_1)^2 + (J_2)^2 + (J_3)^2 \tag{3.107}$$

that fulfils the defining condition:

$$[J^2, J_i] = 0. \tag{3.108}$$

We can re-express J^2 in terms of $J_\pm$ by using the definition of $J_\pm$ in Eq. 3.95 and Eq. 3.94:

$$\begin{aligned} J^2 &= J_+J_- + J_-J_+ + (J_3)^2 \\ &= \frac{1}{2}(J_1 + iJ_2)(J_1 - iJ_2) + \frac{1}{2}(J_1 - iJ_2)(J_1 + iJ_2) + (J_3)^2 \\ &= \frac{1}{2}\Big((J_1)^2 - iJ_1J_2 + iJ_2J_1 + (J_2)^2\Big) + \frac{1}{2}\Big((J_1)^2 + iJ_1J_2 - iJ_2J_1 + (J_2)^2\Big) \\ &\quad + (J_3)^2 \\ &= (J_1)^2 + (J_2)^2 + (J_3)^2 \quad \checkmark \end{aligned} \tag{3.109}$$

If we now use[90]

$$J_+v_k = \frac{1}{\sqrt{2}}\sqrt{(j+k+1)(j-k)}v_{k+1} \tag{3.110}$$

and

$$J_-v_k = \frac{1}{\sqrt{2}}\sqrt{(j+k)(j-k+1)}v_{k-1} \tag{3.111}$$

we can compute the fixed scalar value for each representation:

$$\begin{aligned} J^2v_k &= \left(\frac{1}{2}(J_+J_- + J_-J_+) + (J_3)^2\right)v_k \\ &= J_+J_-v_k + J_-J_+v_k + k^2v_k \\ &= J_+\frac{1}{\sqrt{2}}\sqrt{(j+k)(j-k+1)}v_{k-1} + J_-\frac{1}{\sqrt{2}}\sqrt{(j+k+1)(j-k)}v_{k+1} + k^2v_k \\ &= \frac{1}{\sqrt{2}}\sqrt{(j+k)(j-k+1)}J_+v_{k-1} + \frac{1}{\sqrt{2}}\sqrt{(j+k+1)(j-k)}J_-v_{k+1} + k^2v_k \\ &= \frac{1}{\sqrt{2}}\sqrt{(j+k)(j-k+1)}\frac{1}{\sqrt{2}}\sqrt{(j+(k-1)+1)(j-(k-1))}v_k \\ &\quad + \frac{1}{\sqrt{2}}\sqrt{(j+k+1)(j-k)}\frac{1}{\sqrt{2}}\sqrt{(j+(k+1))(j-(k+1)+1)}v_k + k^2v_k \\ &= \frac{1}{2}(j+k)(j-k+1) + \frac{1}{2}(j-k)(j+k+1)v_k + k^2v_k \\ &= (j^2+j)v_k = j(j+1)v_k \end{aligned} \tag{3.112}$$

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

[89] Recall that Casimir operators are defined as operators C, built from the generators of the group that commute with every generator X of the group: $[C, X] = 0$.

[90] These are just the normalization constants. If we act with $J_\pm$ onto a normalized state, the resulting state will in general not be normalized, too. Nevertheless, in physics we always prefer working with normalized states, for reasons that will become clear in the following chapters. The derivation is a bit tedious, but simply starts with $J_\pm v_k = cv_{k\pm1}$ where c is the normalization constant in question. The complete computation can be found in most books about quantum mechanics in the chapter about angular momentum and angular momentum ladder operators. If this is new to you, do not waste too much time here because the result of this section is not too important for everything that follows.

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

The lowest possible value for j is zero. In this case our representation acts on a $2j+1 = 2 \cdot 0 + 1 = 1$ dimensional vector space. We can see that this representation is trivial, because the only 1×1 matrices fulfilling the commutation relations of the $SU(2)$ Lie algebra
$[J_l, J_m] = i\epsilon_{lmn}J_n$, are trivially 0. If we exponentiate the generator 0 we always get the transformation $U = e^0 = 1$ which changes nothing at all.

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

We now take a look at the next lowest possible value $j = \frac{1}{2}$. This representation is $2\frac{1}{2} + 1 = 2$ dimensional. The generator J_3 has eigenvalues $\frac{1}{2}$ and $\frac{1}{2} - 1 = -\frac{1}{2}$, as can be seen from Eq. 3.106 and is therefore given by

$$J_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.113}$$

because we choose J_3 to be the diagonal generator[91]. The eigenvectors corresponding to the eigenvalues $+\frac{1}{2}$,$-\frac{1}{2}$ are:

$$v_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad v_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.114}$$

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

$$J_1 = \frac{1}{\sqrt{2}}(J_- + J_+) \tag{3.115}$$

$$J_2 = \frac{i}{\sqrt{2}}(J_- - J_+), \tag{3.116}$$

which we get directly from inverting the definitions of $J_\pm$ in Eq. 3.95 and Eq. 3.94. Recall that a basis four 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

$$J_1 v_{\frac{1}{2}} = \frac{1}{\sqrt{2}}(J_- + J_+)v_{\frac{1}{2}} = \frac{1}{\sqrt{2}}(J_- v_{\frac{1}{2}} + \underbrace{J_+ v_{\frac{1}{2}}}_{=0}) = \frac{1}{\sqrt{2}}J_- v_{\frac{1}{2}} = \frac{1}{2}v_{-\frac{1}{2}}, \tag{3.117}$$

where we used that $\frac{1}{2}$ is already the maximum value for $v_{\frac{1}{2}}$ and we cannot go higher. The factor $\frac{1}{2}$ is the scalar factor we get from Eq. 3.105. Similarly we get

$$J_1 v_{-\frac{1}{2}} = \frac{1}{\sqrt{2}}(J_- + J_+)v_{-\frac{1}{2}} = \frac{1}{2}v_{\frac{1}{2}} \tag{3.118}$$

[91] 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.

Written in matrix form, where our basis is given by $v_{\frac{1}{2}} = (1,0)^T$ and $v_{-\frac{1}{2}} = (0,1)^T$:

$$J_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.119}$$

You can check that this matrix has the action on the basis vectors we derived above[92]. In the same way, we find

$$J_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.120}$$

These are the same generators $J_i = \frac{1}{2}\sigma_i$, with the Pauli matrices σ_i, we found while investigating Lie algebra of $SU(2)$ at the beginning of this chapter (Eq. 3.81). 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[93].

[92] We derived in Eq. 3.117: $J_1 v_{\frac{1}{2}} = \frac{1}{2} v_{-\frac{1}{2}}$. Using the explicit matrix form of J_1 we get $J_1 v_{\frac{1}{2}} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2} v_{-\frac{1}{2}} \quad \checkmark$.

[93] 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.

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

Following the same procedure[94] as in two-dimensions, we find:

$$J_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.121}$$

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 deriving even higher dimensional representations, because at this point we already have everything we need to understand the most important representations of the Lorentz group.

[94] We start again with the diagonal generator J_3, which we can write down immediately because we know its eigenvalues $(1, 0, -1)$. Afterwards, the other two generators J_1, J_2 can be derived by their action, where we again use that we can write them in terms of $J_\pm$, on the eigenvectors of J_3.

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**[95]

[95] 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 use **one** known representation of the Lorentz group to derive the corresponding Lie algebra, which 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 this algebra can be seen to be constructed of two copies of the Lie algebra of $SU(2)$. 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[96]

$$x^\mu x_\mu = x^\mu \eta_{\mu\nu} x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{3.122}$$

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

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.123}$$

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

$$x^\mu \eta_{\mu\nu} x^\nu \to x'^\sigma \eta_{\sigma\rho} x'^\rho = (x^\mu \Lambda^\sigma_\mu) \eta_{\sigma\rho} (\Lambda^\rho_\nu x^\nu) \stackrel{!}{=} x^\mu \eta_{\mu\nu} x^\nu \tag{3.124}$$

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

$$\Lambda^\sigma_\mu \eta_{\sigma\rho} \Lambda^\rho_\nu \stackrel{!}{=} \eta_{\mu\nu} \tag{3.125}$$

or written in matrix form[97]

$$\Lambda^T \eta \Lambda \stackrel{!}{=} \eta. \tag{3.126}$$

This is how the Lorentz transformations Λ are defined! If we take the determinant of the equation and use
$\det(AB) = \det(A)\det(B)$ we get the defining condition

$$\det(\Lambda) \underbrace{\det(\eta)}_{=-1} \det(\Lambda) = \underbrace{\det(\eta)}_{=-1} \to \det(\Lambda)^2 \stackrel{!}{=} 1 \tag{3.127}$$

[96] This was derived in Chap. 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.

[97] 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.

$$\rightarrow \det(\Lambda) \stackrel{!}{=} \pm 1 \tag{3.128}$$

Furthermore, we get if we look at[98] the $\mu = \nu = 0$ component in Eq. 3.125

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

$$\Lambda_0^\sigma \eta_{\sigma\rho} \Lambda_0^\rho \stackrel{!}{=} \underbrace{\eta_{00}}_{=1} \rightarrow \Lambda_0^\sigma \eta_{\sigma\rho} \Lambda_0^\rho = (\Lambda_0^0)^2 - \sum_i (\Lambda_0^i)^2 \stackrel{!}{=} 1 \tag{3.129}$$

and we conclude

$$\Lambda_0^0 \stackrel{!}{=} \pm \sqrt{1 + \sum_i (\Lambda_0^i)^2}. \tag{3.130}$$

We divide the Lorentz group into four components, depending on the signs in the Eq. 3.128 and Eq. 3.130. The components that preserve the orientation[99] of the coordinate system are those two with $\det(\Lambda) = +1$. Furthermore, if we want to preserve the direction of time we need to restrict to $\Lambda_0^0 \geq 0$, because

[99] This means a right-handed coordinate system stays right-handed and a left-handed coordinate system stays left-handed. For the definition of left- and right-handed coordinate systems have a look at appendix A.5.

$$x^0 = t \rightarrow x'^0 = t' = \Lambda_\nu^0 x^\nu = \Lambda_0^0 t + \Lambda_1^0 x^1 + \Lambda_2^0 x^2 + \Lambda_3^0 x^3, \tag{3.131}$$

where we can see that, if $\Lambda_0^0 \geq 0$, then t' has the same sign as t. This component is called $SO(1,3)^\uparrow$ and we will talk about this subgroup most of the time. The fancy term for this subgroup is proper[100] orthochronous[101] Lorentz group. The four components of the Lorentz group are disconnected in the sense that it is not possible to get a Lorentz transformation of another component just by using the Lorentz transformations of one component. Other components can be obtained from $SO(1,3)^\uparrow$ by using[102]

[100] This term refers to the fact that the subgroup $SO(1,3)^\uparrow$ preserves orientation/parity.

[101] This means that this subgroup preserves the direction of time.

[102] At least for one representation, these operators look like this. We will see later that for different representations, these operators look quite different.

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.132}$$

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.133}$$

Λ_P is called the parity operator. A parity transformation is simply a reflection in a mirror. Λ_T is the time-reversal operator.

The complete Lorentz group $O(1,3)$ can then be seen as the set:

$$O(1,3) = \{SO(1,3)^\uparrow, \Lambda_P SO(1,3)^\uparrow, \Lambda_T SO(1,3)^\uparrow, \Lambda_P \Lambda_T SO(1,3)^\uparrow\} \tag{3.134}$$

Therefore, we can restrict our search for representations of the Lorentz group, to representations of $SO(1,3)^{\uparrow}$, because then we only need to find representations for Λ_P and Λ_T, to get representations of the other components.

3.7.1 One Representation of the Lorentz Group

Let's see how we can use the defining condition of the Lorentz group (Eq. 3.125) 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[103], 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[104], 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 the complete Lie algebra of this group. These generators form a basis for this Lie algebra, which 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.

[103] 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 .

[104] 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, involving only space and leaving time unchanged, fulfil the condition in Eq. 3.125. This follows because the spatial part[105] of the Minkowski metric is proportional to the 3×3 identity matrix[106] and therefore for transformations involving only space, we have from Eq. 3.125 the condition

[105] 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.

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

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

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

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

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

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

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

with the rotation matrices $R_{3\times 3}$ cited in Eq. 3.23 and derived in Sec. 3.4.1. The corresponding generators are therefore analogous to those we derived for three spatial dimension in Sec. 3.4.1:

$$J_i = \begin{pmatrix} 0 & \\ & J_i^{3dim} \end{pmatrix}. \tag{3.135}$$

For example, from Eq. 3.65 we now have

$$J_1 = \begin{pmatrix} 0 & \\ & J_1^{3dim} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3.136}$$

To investigate transformations involving time **and** space we will start, as always in Lie theory, with an infinitesimal transformation[107]

$$\Lambda^\mu_\rho \approx \delta^\mu_\rho + \epsilon K^\mu_\rho. \tag{3.137}$$

We put this into the defining condition (Eq. 3.125)

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

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

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

$$\rightarrow K^\mu_\rho \eta_{\mu\sigma} + K^\nu_\sigma \eta_{\rho\nu} = 0 \tag{3.138}$$

which reads in matrix form[108]

$$K^T \eta = -\eta K. \tag{3.139}$$

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 boost the description we have, for example in 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 Chap. 2.1: A boost along the x-axis. Because we know that $y' = y$ and $z' = z$ the generator is of the form

$$K_x = \begin{pmatrix} \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{\equiv k_x} & \\ & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \tag{3.140}$$

[107] 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.

[108] 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 $K^\nu{}_\sigma \eta_{\rho\nu} = \eta_{\rho\nu} K^\nu{}_\sigma$, were 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.

and we only need to solve a 2×2 matrix equation. Equation 3.139 reduces to

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

which is solved by[109]

$$k_x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

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

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

$$K_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.141}$$

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

$$K_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad K_z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{3.142}$$

Now, we already know from Lie theory how we get from the generators to finite transformations[110]

$$\Lambda_x(\phi) = e^{\phi K_x}$$

[110] Take not that the generators K_x,K_y and K_z are already Hermitian: $K_i^\dagger = K_i$. Therefore, we do not include an extra i here in the exponent, because then the generators would be anti-Hermitian.

For brevity let's focus again on the exciting part of the generator K_x, i.e. the upper left 2×2 matrix k_x, which is defined in Eq. 3.140. We can then evaluate the exponential function using its series expansion and that[111] $k_x^2 = 1$

$$\begin{aligned} \Lambda_x(\phi) = e^{\phi k_x} &= \sum_{n=0}^{\infty} \frac{\phi^n k_x^n}{n!} = \sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \underbrace{k_x^{2n}}_{=1} + \sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \underbrace{k_x^{2n+1}}_{=k_x} \\ &= \left(\sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \right) I + \left(\sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \right) k_x = \cosh(\phi) I + \sinh(\phi) k_x \\ &= \begin{pmatrix} \cosh(\phi) & 0 \\ 0 & \cosh(\phi) \end{pmatrix} + \begin{pmatrix} 0 & \sinh(\phi) \\ \sinh(\phi) & 0 \end{pmatrix} = \begin{pmatrix} \cosh(\phi) & \sinh(\phi) \\ \sinh(\phi) & \cosh(\phi) \end{pmatrix} \end{aligned} \tag{3.143}$$

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

This computation is analogous to the computation in Sec. 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

$$\Lambda_x = \begin{pmatrix} \cosh(\phi) & \sinh(\phi) & 0 & 0 \\ \sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.144}$$

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

$$\Lambda_y = \begin{pmatrix} \cosh(\phi) & 0 & \sinh(\phi) & 0 \\ 0 & 1 & 0 & 0 \\ \sinh(\phi) & 0 & \cosh(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.145}$$

$$\Lambda_z = \begin{pmatrix} \cosh(\phi) & 0 & 0 & \sinh(\phi) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(\phi) & 0 & 0 & \cosh(\phi) \end{pmatrix}. \tag{3.146}$$

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[112] 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[113]

$$(\Lambda_P)^\alpha_{\alpha'}(\Lambda_P)^\beta_{\beta'}(J_i)^{\alpha'\beta'} \underbrace{\hat{=}}_{\text{switching to matrix notation}} \Lambda_P J_i (\Lambda_P)^T = J_i \hat{=} (J_i)^{\alpha\beta} \tag{3.147}$$

$$(\Lambda_P)^\alpha_{\alpha'}(\Lambda_P)^\beta_{\beta'}(K_i)^{\alpha'\beta'} \underbrace{\hat{=}}_{\text{switching to matrix notation}} \Lambda_P K_i (\Lambda_P)^T = -K_i \hat{=} -(K_i)^{\alpha\beta}, \tag{3.148}$$

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

$$J_x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \to J'_x = \Lambda_P J_x (\Lambda_P)^T = J_x, \tag{3.149}$$

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

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

because

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

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{3.150}$$

In contrast,

$$K_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to K'_x = \Lambda_P K_x (\Lambda_P)^T = -K_x, \tag{3.151}$$

because

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

$$= - \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.152}$$

In conclusion, we have under parity transformations

$$J_i \underbrace{\to}_{\text{P}} J_i \qquad K_i \underbrace{\to}_{\text{P}} -K_i \tag{3.153}$$

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

$$(\Lambda_T)^{\alpha}_{\alpha'} (\Lambda_T)^{\beta}_{\beta'} (J_i)^{\alpha'\beta'} \underbrace{\hat{=}}_{\text{switching to matrix notation}} \Lambda_T J_i (\Lambda_T)^T = J_i \hat{=} (J_i)^{\alpha\beta} \tag{3.154}$$

$$(\Lambda_T)^{\alpha}_{\alpha'} (\Lambda_T)^{\beta}_{\beta'} (K_i)^{\alpha'\beta'} \underbrace{\hat{=}}_{\text{switching to matrix notation}} \Lambda_T K_i (\Lambda_T)^T = -(K_i)^{\alpha\beta}, \tag{3.155}$$

Or shorter:

$$J_i \underbrace{\rightarrow}_{\mathrm{T}} J_i \qquad K_i \underbrace{\rightarrow}_{\mathrm{T}} -K_i. \tag{3.156}$$

3.7.3 The Lie Algebra of the Proper Orthochronous Lorentz Group

Now using the explicit matrix form of the generators[114] for $SO(1,3)^{\uparrow}$ we can derive the corresponding Lie algebra by brute force computation[115]

[114] See Eq. 3.141 for the boost generators and Eq. 3.62 for the rotation generators

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

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.157}$$

$$[J_i, K_j] = i\epsilon_{ijk}K_k \tag{3.158}$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k \tag{3.159}$$

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

$$\Lambda = e^{iJ\theta + iK\Phi} \tag{3.160}$$

Equation 3.158 tells us that the two generator types (J_i and K_i) do not commute with each other. While the rotation generators are closed under commutation[116], the boost generators are not[117]. We can now define new operators from the old ones that are closed under commutation and commute with each other

[116] Closed under commutation means that the commutator $[J_i, J_j] = J_iJ_j - J_jJ_i$, is again a rotation generator. From Eq. 3.157 we can see that this is the case.

[117] Eq. 3.159 tells us that the commutator of two boost generators K_i and K_j isn't another boost generator, but a generator of rotations.

$$N_i^{\pm} = \frac{1}{2}(J_i \pm iK_i). \tag{3.161}$$

Working out the commutation relations yields

$$[N_i^+, N_j^+] = i\epsilon_{ijk}N_k^+ \tag{3.162}$$

$$[N_i^-, N_j^-] = i\epsilon_{ijk}N_k^- \tag{3.163}$$

$$[N_i^+, N_j^-] = 0. \tag{3.164}$$

These are precisely the commutation relations for the Lie algebra of $SU(2)$ and we have therefore discovered that the Lie algebra of $SO(1,3)_+^{\uparrow}$ consists of two copies of the Lie algebra of $SU(2)$.

This is great news, because we already know how to construct all irreducible representations of the Lie algebra of $SU(2)$. However the Lorentz group is, like $SO(3)$, not simply-connected[118] and Lie theory

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

tells us that there is, for groups that aren't simply connected, no one-to-one correspondence between the irreducible representations of the Lie algebra and representations of the corresponding group[119]. Instead, **by deriving the irreducible representations of the 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. It is a good thing that we find addition representations, because we need those representations to describe certain elementary particles.

[119] 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 the Lorentz group.

Each irreducible representation of the Lie algebra of $SU(2)$ can be labelled by the scalar value j of the Casimir operator of $SU(2)$. Therefore, we now know that we can label the irreducible representations of the covering group[120] 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 copies of the $SU(2)$, which we derived earlier.

[120] The covering group of the Lorentz group is $SL(2, C)$, the set of 2×2 matrices with unit determinant and complex entries. The relationship $SL(2, C) \rightarrow SO(1, 3)$ is similar to the relationship $SU(2) \rightarrow SO(3)$ we discovered earlier in this text.

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

$$J_i = \frac{1}{2}\epsilon_{ijk} M_{jk}. \tag{3.165}$$

$$K_i = M_{0i}. \tag{3.166}$$

With this definition the Lorentz algebra reads

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho} M_{\nu\sigma} - \eta_{\mu\sigma} M_{\nu\rho} - \eta_{\nu\rho} M_{\mu\sigma} + \eta_{\nu\sigma} M_{\mu\rho}). \tag{3.167}$$

Next, we want to take a look at what irreducible representations we can construct from the Lie algebra of the Lorentz group.

3.7.4 The $(0, 0)$ Representation

The lowest order representation is as for $SU(2)$ trivial, because the vector space is 1 dimensional for both copies of the Lie algebra of $SU(2)$. Our generators must therefore be 1×1 matrices and the only 1×1 matrices fulfilling the commutation relations are trivially 0:

$$N_i^+ = N_i^- = 0 \rightarrow \mathrm{e}^{N_i^+} = \mathrm{e}^{N_i^-} = \mathrm{e}^0 = 1 \tag{3.168}$$

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

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

In this representation we use the[121] 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 dimensional representation for the other N_i^-, i.e. $N_i^- = 0$ as explained in the last section. From the definition of N^- in Eq. 3.161 we conclude

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

$$N_i^- = \frac{1}{2}(J_i - iK_i) = 0 \tag{3.169}$$

$$\rightarrow J_i = iK_i. \tag{3.170}$$

Furthermore, we can use that we already derived in Sec. 3.6.4 the two dimensional representation of $SU(2)$:

$$N_i^+ = \frac{\sigma_i}{2} \tag{3.171}$$

where σ_i denotes once more the Pauli matrices, which were defined in Eq. 3.81. On the other hand, we have

$$N_i^+ \underbrace{=}_{\text{Eq. 3.161}} \frac{1}{2}(J_i + iK_i) \underbrace{=}_{\text{Eq. 3.170}} \frac{1}{2}(iK_i + iK_i) = iK_i \tag{3.172}$$

Comparing Eq. 3.171 with Eq. 3.172 tells us that

$$iK_i = \frac{\sigma_i}{2} \rightarrow K_i = \frac{\sigma_i}{2i} = \frac{i\sigma_i}{2i^2} = \frac{-i}{2}\sigma_i \tag{3.173}$$

$$\text{Eq. 3.170} \rightarrow J_i = iK_i = \frac{-i^2}{2}\sigma_i = \frac{1}{2}\sigma_i. \tag{3.174}$$

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

$$R_\theta = \mathrm{e}^{i\vec{\theta}\vec{J}} = \mathrm{e}^{i\vec{\theta}\frac{\vec{\sigma}}{2}} \tag{3.175}$$

and a Lorentz boost by

$$B_\theta = \mathrm{e}^{i\vec{\phi}\vec{K}} = \mathrm{e}^{\vec{\phi}\frac{\vec{\sigma}}{2}}. \tag{3.176}$$

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

$$R_x(\theta) = \mathrm{e}^{i\theta J_1} = \mathrm{e}^{i\theta\frac{1}{2}\sigma_1} = 1 + \frac{i}{2}\theta\sigma_1 + \frac{1}{2}\left(\frac{i}{2}\theta\sigma_1\right)^2 + \dots \tag{3.177}$$

And if we use the explicit matrix form of σ_1 as defined in Eq. 3.81, together with the fact that $\sigma_1^2 = 1$ we get[122]

$$R_x(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{i}{2}\theta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{1}{2}\left(\frac{\theta}{2}\right)^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \dots$$
$$= \begin{pmatrix} \cos(\frac{\theta}{2}) & i\sin(\frac{\theta}{2}) \\ i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix}. \tag{3.178}$$

[122] The steps are completely analogous to what we did in Sec. 3.4.1

Analogous we can compute the transformation matrix for rotations around other axes or boosts. One important thing to notice is 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**[123] objects this representation acts on are called **left-chiral spinors**[124]:

$$\chi_L = \begin{pmatrix} (\chi_L)_1 \\ (\chi_L)_2 \end{pmatrix} \tag{3.179}$$

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

[124] 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.

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[125] is after a rotation by 2π not the same, but gets a minus sign. This is a pretty crazy property, because all objects we deal with in everyday life are exactly the same after a rotation by $360° = 2\pi$.

> "One could say that a spinor is the most basic sort of mathematical object that can be Lorentz-transformed."
>
> \- **A. M. Steane**[126]

[125] 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

[126] 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. 3.161

$$N_i^+ = \frac{1}{2}(J_i + iK_i) = 0 \tag{3.180}$$

$$\rightarrow J_i = -iK_i. \tag{3.181}$$

Take notice of the minus sign. Using the two-dimensional representation of $SU(2)$ for N^+, which was derived in Sec. 3.6.4, yields

$$N_i^- = \frac{1}{2}\sigma_i = \frac{1}{2}(J_i - iK_i) \underbrace{=}_{\text{Eq. 3.181}} \frac{1}{2}(-iK_i - iK_i) = -iK_i \tag{3.182}$$

From this we get the $(0, \frac{1}{2})$ representation of the generators

$$-iK_i = \frac{1}{2}\sigma_i \rightarrow K_i = \frac{-1}{2i}\sigma_i = \frac{-i}{2i^2}\sigma_i = \frac{i}{2}\sigma_i. \tag{3.183}$$

And from Eq. 3.181 we get

$$J_i = -iK_i = \frac{1}{2}\sigma_i. \tag{3.184}$$

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

$$R_\theta = e^{i\vec{\theta}\vec{J}} = e^{i\vec{\theta}\frac{\vec{\sigma}}{2}} \tag{3.185}$$

and a Lorentz boost by

$$B_\theta = e^{i\vec{\phi}\vec{K}} = e^{-\vec{\phi}\frac{\vec{\sigma}}{2}}. \tag{3.186}$$

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**:

$$\chi_R = \begin{pmatrix} (\chi_R)^1 \\ (\chi_R)^2 \end{pmatrix} \tag{3.187}$$

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. We will learn in a moment that they are different, but not too different. In fact, 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

$$\chi_L = \chi_a \tag{3.188}$$

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

$$\chi_R = \chi^{\dot{a}}. \tag{3.189}$$

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[127] as

$$\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.190}$$

and show that it has the desired properties. Furthermore, we define[128]

$$\chi_L^C \equiv \epsilon \chi_L^\star \tag{3.191}$$

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 behaviour 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

$$(-\epsilon)(\epsilon) = 1 \tag{3.192}$$

and

$$(\epsilon)\sigma_i^\star(-\epsilon) = -\sigma_i \tag{3.193}$$

for each Pauli matrix σ_i, as you can check. Transforming yields[129]

$$\begin{aligned}
\chi_L^C \to \chi_L'^C &= \epsilon(\chi')_L^\star \\
&= \epsilon(\mathrm{e}^{\frac{\vec{\phi}}{2}\vec{\sigma}}\chi_L)^\star \\
&= \epsilon(\mathrm{e}^{\frac{\vec{\phi}}{2}\vec{\sigma}}\underbrace{(-\epsilon)(\epsilon)}_{=1 \text{ see Eq. 3.192}}\chi_L)^\star \\
&= \underbrace{\epsilon(\mathrm{e}^{\frac{\vec{\phi}}{2}\vec{\sigma}^\star}(-\epsilon)}_{\text{Eq. 3.193: } =\mathrm{e}^{-\frac{\vec{\phi}}{2}\vec{\sigma}}}(\epsilon)\chi_L^\star) \\
&= \mathrm{e}^{-\frac{\vec{\phi}}{2}\vec{\sigma}}\underbrace{\epsilon\chi_L^\star}_{=\chi_L^C} \\
&= \mathrm{e}^{-\frac{\vec{\phi}}{2}\vec{\sigma}}\chi_L^C \quad ,
\end{aligned} \tag{3.194}$$

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

[128] 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 Sec. 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.

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

which is exactly the transformation behaviour of a right-chiral spinor[130]. To get to the fifth line, we use the series expansion of $e^{\frac{\vec{\phi}}{2}\vec{\sigma}}$ and Eq. 3.193 on every term. You can check in the same way that the behaviour 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:

[130] The transformation behaviour of right-chiral spinors under boosts was derived in Eq. 3.186: $B_\theta = e^{i\vec{\phi}\vec{K}} = e^{-\vec{\phi}\frac{\vec{\sigma}}{2}}$. Compare this to how left-chiral spinors transform under boosts, as derived in Eq. 3.176: $B_\theta = e^{i\vec{\phi}\vec{K}} = e^{\vec{\phi}\frac{\vec{\sigma}}{2}}$

$$\chi_L^C \to \chi_L'^C = \epsilon(\chi')_L^\star = \epsilon(e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\chi)_L^\star = e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\epsilon(\chi_L)^\star. \tag{3.195}$$

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

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

$$\epsilon\chi_L \underbrace{=}_{\text{written in index notation}} \epsilon^{ac}\chi_c = \chi^a \tag{3.196}$$

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

$$\chi_R = \epsilon{\chi_L}^\star \tag{3.197}$$

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

$$\chi_R = \epsilon{\chi_L}^\star = \chi^{\dot{a}}. \tag{3.198}$$

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

$${\chi_L}^\star = {\chi_a}^\star = \chi_{\dot{a}} \tag{3.199}$$

and an upper, undotted index, by complex conjugating χ_R

$${\chi_R}^\star = (\chi^{\dot{a}})^\star = \chi^a \tag{3.200}$$

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 at all under Lorentz transformations[131]. From the transformation behaviour of a left-chiral spinor

[131] 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.

$$\chi_L = \chi_a \to \chi_a' = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ b}\chi_b \tag{3.201}$$

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

$$\begin{aligned}\chi_L^\star = \chi_a^\star = \chi_{\dot{a}} \to \chi_{\dot{a}}' = (\chi_a')^\star &= \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ b}\right)^\star \chi_b^\star \\ &= \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2}+\vec{\phi}\frac{\vec{\sigma}^\star}{2}}\right)_{\dot{a}}^{\ \dot{b}}\chi_{\dot{b}}\end{aligned} \tag{3.202}$$

Analogously, we use the transformation behaviour of a right-chiral spinor:

$$\chi_R \to \chi'_R = \chi'^{\dot{a}} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}-\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{a}}_{\dot{b}} \chi^{\dot{b}} \tag{3.203}$$

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

$$\begin{aligned}\chi_R^\star = (\chi^{\dot{a}})^\star = \chi^a \to \chi'^a = (\chi'^{\dot{a}})^\star &= \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}-\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^{\dot{a}}_{\dot{b}}\right)^\star (\chi^{\dot{b}})^\star \\ &= \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2}-\vec{\phi}\frac{\vec{\sigma}^\star}{2}}\right)^a_b \chi^b \end{aligned} \tag{3.204}$$

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 mustn't forget to 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

$$\sigma_i^\dagger = (\sigma_i^\star)^T = \sigma_i, \tag{3.205}$$

as you can easily check by looking at the explicit form of the Pauli matrices, as defined in Eq. 3.81.

By comparing Eq. 3.201 with Eq. 3.204 and using Eq. 3.205, we see that the transformation behaviour 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[132]

[132] As explained in appendix B.5.5, the symbol δ^c_b is called Kronecker symbol and denotes the unit matrix in index notation. This means $\delta^c_b = 1$ for $b = c$ and $\delta^c_b = 0$ for $b \neq c$.

$$\begin{aligned}(\chi^a)^T \chi_a \to (\chi'^a)^T \chi'_a &= \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2}-\vec{\phi}\frac{\vec{\sigma}^\star}{2}}\right)^a_b \chi^b\right)^T \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^c_a \chi_c \\ &= (\chi^b)^T \left(e^{-i\vec{\theta}\frac{(\vec{\sigma}^\star)^T}{2}-\vec{\phi}\frac{(\vec{\sigma}^\star)^T}{2}}\right)^a_b \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^c_a \chi_c \\ &\underbrace{=}_{\text{Eq. 3.205}} (\chi^b)^T \underbrace{\left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2}-\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^a_b \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right)^c_a}_{=\delta^c_b} \chi_c \\ &= (\chi^c)^T \chi_c \end{aligned} \tag{3.206}$$

In the same way we can combine an upper, dotted index with a lower, dotted index as you can verify by comparing Eq. 3.202 with

Eq. 3.203. 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

$$\chi_R^T \chi_L = (\chi^{\dot{a}})^T \chi_a \to (\chi'^{\dot{a}})^T \chi'_a = \chi^{\dot{b}} \underbrace{\left(e^{i\vec{\theta}\frac{\vec{\sigma}^T}{2} - \vec{\phi}\frac{\vec{\sigma}^T}{2}}\right)_{\dot{b}}^{\ \dot{a}} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a^{\ c}}_{\neq \delta_b^c} \chi_c \tag{3.207}$$

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[133] in order to get Lorentz invariant terms. Or formulated differently, we must combine the complex 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 complex 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 we look for invariant terms that we can use to formulate our laws of nature.

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

In addition, we have now another justification for calling ϵ^{ab} the spinor metric, because the invariant spinor product in Eq. 3.206, can be written as

$$\chi_a^T \chi^a \underbrace{=}_{\text{Eq. 3.196}} \chi_a^T \epsilon^{ab} \chi_b. \tag{3.208}$$

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

$$x_\mu y^\mu = x_\mu \eta^{\mu\nu} y_\nu. \tag{3.209}$$

The spinor metric is indeed what the Minkowski metric is for four-vectors[134].

[134] 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

$$\epsilon_{ab} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{3.210}$$

because we need[135] $(-\epsilon)$ 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:

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

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

and analogous for left-chiral spinors

$$\chi_L \to \chi'_L = \chi'_a = \Lambda_a{}^b \chi_b = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}}\right)_a{}^b \chi_b \quad . \tag{3.212}$$

Therefore:

$$\Lambda_{(\frac{1}{2},0)} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}+\vec{\phi}\frac{\vec{\sigma}}{2}}\right) \hat{=} \Lambda_a{}^b \tag{3.213}$$

and

$$\Lambda_{(0,\frac{1}{2})} = \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2}-\vec{\phi}\frac{\vec{\sigma}}{2}}\right) \hat{=} \Lambda^{\dot{a}}{}_{\dot{b}} \tag{3.214}$$

This notation will prove to be very 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[136] 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. $[N_i^+, N_j^-] = 0$. 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 $v_a^{\dot{b}}$, each transforming under a separate two-dimensional copy of $SU(2)$. Here the notation we introduced in the last section comes in handy.

We know from the fact that both indices can take on two values ($\frac{1}{2}$ and $-\frac{1}{2}$), because each representation is 2 dimensional, that our object v will have 4 components. Therefore, the objects can be 2×2 matrices, but it's also possible to enforce a four component vector form, as we will see[137].

But 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 and as explained in Sec. 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

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

[137] 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.

that our irreducible representation acts on Hermitian 2×2 matrices.

A basis[138] for Hermitian 2×2 matrices is given by the σ matrices[139] together with the identity matrix.

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

[139] Defined in Eq. 3.81

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. 3.81. 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

$$v_{a\dot{b}} = v_\nu \sigma^\nu_{a\dot{b}} = v^0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + v^1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + v^2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + v^3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.215}$$

As explained above, we could use[140] $v_a^{\dot{b}} = v_\mu \sigma^\mu_{a\dot{c}} \epsilon^{\dot{b}\dot{c}}$ instead, which means we would use the basis $(\tilde{\sigma}^{\dot{b}})^\mu = \sigma^\mu_{a\dot{c}} \epsilon^{\dot{b}\dot{c}}$. We therefore write a general Hermitian matrix as

[140] 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 behaviour we derived earlier for vectors.

$$v_{a\dot{b}} = \begin{pmatrix} v_0 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & v_0 - v_3 \end{pmatrix}. \tag{3.216}$$

Remember that we have learned in the last section that different indices transform differently. To be specific: 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 operator for an lower undotted index as derived in Eq. 3.202

$$\begin{aligned} v \to v' = v'_{a\dot{b}} &= \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^{\ c} v_{c\dot{d}} \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\star}{2} + \vec{\phi}\frac{\vec{\sigma}^\star}{2}} \right)_{\dot{b}}^{\ \dot{d}} \right)^T \\ &= \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2} + \vec{\phi}\frac{\vec{\sigma}^\dagger}{2}} \right)^{\dot{d}}_{\ \dot{b}} \\ &\underbrace{=}_{\sigma_i^\dagger = \sigma_i} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^{\ c} v_{c\dot{d}} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)^{\dot{d}}_{\ \dot{b}} \end{aligned} \tag{3.217}$$

We can now see that a Hermitian matrix is after such a transformation still Hermitian, as promised above[141]

[141] 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 $v^\dagger_{c\dot{d}} = -v_{c\dot{d}}$.

$$\left(e^{i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_a^{\ c} v_{c\dot d}\left(e^{-i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_{\dot b}^{\ \dot d} \to \left(\left(e^{i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_a^{\ c} v_{c\dot d}\left(e^{-i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_{\dot b}^{\ \dot d}\right)^\dagger$$
$$\underbrace{=}_{(ABC)^\dagger=C^\dagger B^\dagger A^\dagger}\left(\left(e^{-i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_{\dot b}^{\ \dot d}\right)^\dagger v^\dagger_{c\dot d}\left(\left(e^{i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_a^{\ c}\right)^\dagger$$
$$= \left(e^{i\vec\theta\frac{\vec\sigma^\dagger}{2}+\vec\phi\frac{\vec\sigma^\dagger}{2}}\right)_{\dot b}^{\ \dot d} v^\dagger_{c\dot d}\left(e^{-i\vec\theta\frac{\vec\sigma^\dagger}{2}+\vec\phi\frac{\vec\sigma^\dagger}{2}}\right)_a^{\ c}$$
$$\underbrace{=}_{\text{if } v^\dagger_{c\dot d}=v_{c\dot d}}\left(e^{i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_a^{\ c} v_{c\dot d}\left(e^{-i\vec\theta\frac{\vec\sigma}{2}+\vec\phi\frac{\vec\sigma}{2}}\right)_{\dot b}^{\ \dot d} \quad \checkmark \tag{3.218}$$

The explicit computation[142] for an arbitrary transformation is long and tedious so we will look at one specific example. Let's boost v along the z-axis[143]

$$v_{a\dot b} \to v'_{a\dot b} = \left(e^{\phi\frac{\sigma_3}{2}}\right)_a^{\ c} v_{c\dot d}\left(e^{\phi\frac{\sigma_3}{2}}\right)_{\dot b}^{\ \dot d}$$
$$= \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}}\end{pmatrix}\begin{pmatrix} v_0+v_3 & v_1-iv_2 \\ v_1+iv_2 & v_0-v_3\end{pmatrix}\begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}}\end{pmatrix}$$
$$= \begin{pmatrix} e^{\phi}(v_0+v_3) & v_1-iv_2 \\ v_1+iv_2 & e^{-\phi}(v_0-v_3)\end{pmatrix} \tag{3.219}$$

where we have used the fact that σ_3 is diagonal[144] and that $e^A = \begin{pmatrix} e^{A_{11}} & 0 \\ 0 & e^{A_{22}}\end{pmatrix}$ holds for every diagonal matrix. Comparing the transformed object we computed in Eq. 3.219 with a generic object v' yields

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

This tells us how the components of the transformed object are related to the untransformed components[145]

$$\to v'_0+v'_3 = e^{\phi}(v_0+v_3) = (\cosh(\phi)+\sinh(\phi))\,(v_0+v_3)$$
$$\to v'_0-v'_3 = e^{-\phi}(v_0-v_3) = (\cosh(\phi)-\sinh(\phi))\,(v_0-v_3).$$

The addition and subtraction of both equations yields

$$\to v'_0 = \cosh(\phi)v_0+\sinh(\phi)v_3$$

$$\to v'_3 = \sinh(\phi)v_0+\cosh(\phi)v_3 \tag{3.220}$$

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

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

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

[145] 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.

which is exactly what we get using the 4-vector formalism[146]

$$\begin{pmatrix} v'_0 \\ v'_1 \\ v'_2 \\ v'_3 \end{pmatrix} = \begin{pmatrix} \cosh(\phi) & 0 & 0 & \sinh(\phi) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(\phi) & 0 & 0 & \cosh(\phi) \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix}$$
$$= \begin{pmatrix} \cosh(\phi)v_0 + \sinh(\phi)v_3 \\ v_1 \\ v_2 \\ \sinh(\phi)v_0 + \cosh(\phi)v_3 \end{pmatrix} . \qquad (3.221)$$

[146] See 3.146 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[147]. 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.

[147] 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. These components are connected by the parity and the time-reversal operator[148]. 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 will restrict to parity transformations, because it turns out that nature isn't always symmetric under parity transformations, which we will discuss in later chapters. Similar to what we discuss in the next section it's possible to derive representations of the time-reversal operator.

[148] See Eq. 3.134

3.7.9 Spinors and Parity

Up to this point, there is no justification for why we called 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 behaviour of the generators of the Lorentz group under parity transformations. The result was Eq. 3.153, which we recite here for convenience

$$J_i \underbrace{\rightarrow}_{\text{P}} J_i \qquad K_i \underbrace{\rightarrow}_{\text{P}} -K_i \tag{3.222}$$

By looking at the definition of the generators $N^{\pm}$ in Eq. 3.161, which we recite here, too

$$N_i^{\pm} = \frac{1}{2}(J_i \pm iK_i) \tag{3.223}$$

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[149]. Just as a right-handed coordinate system changes into a left-handed coordinate system under parity transformations, these two representations change into each other.

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

$$(\Lambda_{\vec{K}})_{(\frac{1}{2},0)} = e^{\vec{\phi}\vec{K}} \underbrace{\rightarrow}_{\text{P}} e^{-\vec{\phi}\vec{K}} = (\Lambda_{\vec{K}})_{(0,\frac{1}{2})} \tag{3.224}$$

$$(\Lambda_{\vec{K}})_{(0,\frac{1}{2})} = e^{-\vec{\phi}\vec{K}} \underbrace{\rightarrow}_{\text{P}} e^{\vec{\phi}\vec{K}} = (\Lambda_{\vec{K}})_{(\frac{1}{2},0)}. \tag{3.225}$$

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**

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_a \\ \xi^{\dot{a}} \end{pmatrix} \tag{3.226}$$

Recalling 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

$$\Psi_M = \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} \tag{3.227}$$

[149] 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.

is a special case, called Majorana spinor. A Dirac or Majorana spinor is **not a four-vector**, because it transforms completely different. A Dirac spinor transforms according to the $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation[150] of the Lorentz group, which means nothing more than writing the corresponding transformations in block-diagonal form into one big matrix:

[150] 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.

$$\Psi \to \Psi' = \Lambda_{(\frac{1}{2},0)\oplus(0,\frac{1}{2})}\Psi = \begin{pmatrix} \Lambda_{(\frac{1}{2},0)} & 0 \\ 0 & \Lambda_{(0,\frac{1}{2})} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.228}$$

For example, a boost transformation is in this representation

$$\Psi \to \Psi' = \begin{pmatrix} e^{\frac{\vec{\phi}}{2}\vec{\sigma}} & 0 \\ 0 & e^{\frac{-\vec{\phi}}{2}\vec{\sigma}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.229}$$

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 $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation), because we know that under parity transformations $N^+ \leftrightarrow N^-$ and therefore

$$\left(0,\frac{1}{2}\right) \underbrace{\leftrightarrow}_{\text{P}} \left(\frac{1}{2},0\right). \tag{3.230}$$

We conclude, 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.

$$\Psi^P \to (\Psi^P)' = \Lambda_{(0,\frac{1}{2})\oplus(\frac{1}{2},0)}\Psi^P = \begin{pmatrix} \Lambda_{(0,\frac{1}{2})} & 0 \\ 0 & \Lambda_{(\frac{1}{2},0)} \end{pmatrix} \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix}. \tag{3.231}$$

Therefore

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad \to \quad \Psi^P = \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix}. \tag{3.232}$$

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 Sec. 3.7.7 stumbled upon a transformation, which yields $\chi_L \to \chi_R$ and $\xi_R \to \xi_L$. The transformation is $\chi_L \to \chi_L^C = \epsilon \chi_L^\star = \chi_R$ and analogously for a right-chiral spinor $\xi_R \to \xi_R^C = (-\epsilon)\xi_R^\star = \xi_L$. This transformation is not part of the Lorentz group and 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:

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \to \tilde{\Psi} = \begin{pmatrix} \chi_L^C \\ \xi_R^C \end{pmatrix} = \begin{pmatrix} \chi_R \\ \xi_L \end{pmatrix}. \tag{3.233}$$

Unfortunately, this object does not transform like a Dirac spinor[151], which transform under boosts

$$\Psi \to \Psi' = \begin{pmatrix} e^{\frac{\vec{\theta}}{2}\vec{\sigma}} & 0 \\ 0 & e^{\frac{-\vec{\theta}}{2}\vec{\sigma}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.234}$$

The object $\tilde{\Psi}$ we get from the naive operation, transform as

$$\tilde{\Psi} \to \tilde{\Psi}' = \begin{pmatrix} e^{-\frac{\vec{\theta}}{2}\vec{\sigma}} & 0 \\ 0 & e^{\frac{\vec{\theta}}{2}\vec{\sigma}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}. \tag{3.235}$$

This is a different kind of object, because it transforms according to a different representation of the Lorentz group. Therefore we write

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \to \Psi^C = \begin{pmatrix} \xi_R^C \\ \chi_L^C \end{pmatrix} = \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}, \tag{3.236}$$

which incorporates the transformation behaviour 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[152]. 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 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

[151] 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.

[152] 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 $-\frac{1}{2}$.

of this text. Nevertheless, there is another representation, the infinite-dimensional representation, that is especially interesting, because we need it to transform physical fields.

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 so far. 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

$$\Phi_a \to \Phi'_a = M_{ab}(\Lambda)\Phi_b \tag{3.237}$$

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[153] $\Phi = \Phi(x)$ and these coordinates are affected by the Lorentz transformations, too. In general we have

$$x^\mu \to \Lambda^\mu_\nu x^\nu, \tag{3.238}$$

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[154]

$$\Phi_a(x) \to M_{ab}(\Lambda)\Phi_b(\Lambda x). \tag{3.239}$$

Our transformation will therefore consist of two parts. One part, represented by a finite-dimensional representation, acting on Φ_a and a second part acting on the coordinates. This second part will act on an infinite-dimensional[155] vector space and we therefore need an infinite-dimensional representation. The infinite-dimensional representation of the Lorentz group is given by differential operators[156]

$$M^{\text{inf}}_{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu) \tag{3.240}$$

you can check by straightforward computation that $M^{\text{inf}}_{\mu\nu}$ satisfies the Lorentz algebra (Eq. 3.167) and transforms the coordinates as desired.

[153] Here x is a shorthand notation for all spacetime coordinates t, x, y, z

[154] 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.

[155] 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.

[156] The symbols ∂^ν are a shorthand notation for the partial derivative $\frac{\partial}{\partial_\nu}$.

The transformation of the coordinates is now given by[157]

$$\Phi(\Lambda x) = \mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M^{\mathrm{inf}}_{\mu\nu}}\Phi(x), \tag{3.241}$$

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 $M^{\mathrm{fin}}_{\mu\nu}$ and a transformation generated by the infinite-dimensional representation $M^{\mathrm{inf}}_{\mu\nu}$ of the generators:

$$\Phi_a(x) \to \left(\mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M^{\mathrm{fin}}_{\mu\nu}}\right)_a^{\;b} \mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M^{\mathrm{inf}}_{\mu\nu}}\Phi_b(x). \tag{3.242}$$

Because our matrices $M^{\mathrm{fin}}_{\mu\nu}$ are finite-dimensional and constant we can put the two exponents together

$$\Phi_a(x) \to \left(\mathrm{e}^{-i\frac{\omega^{\mu\nu}}{2}M_{\mu\nu}}\right)_a^{\;b}\Phi_b(x) \tag{3.243}$$

with $M_{\mu\nu} = M^{\mathrm{fin}}_{\mu\nu} + M^{\mathrm{inf}}_{\mu\nu}$. This representation of the generators of the Lorentz group is called field representation.

[157] Recall the definition of $M^{\mu\nu}$ in Eq. 3.165. 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 $\phi_i = \omega_{0i}$

We can now talk about a different kind of transformation: **translations**, which means transformations to another location in spacetime. Translations do not result in a mixing of components and therefore, we need no finite-dimensional representation, but 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 Poincare group, which is the topic of the next section. Nevertheless, we will introduce the infinite-dimensional representation for this kind of transformation 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

$$\Phi(x) \to \Phi(x+\epsilon) = \Phi(x) + \underbrace{\partial_x\Phi(x)}_{\text{"rate of change" along the x-axis}}\epsilon,$$

which is, of course, again 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

$$P_i \equiv -i\partial_i. \tag{3.244}$$

With this definition an arbitrary, finite translation is

$$\Phi(x) \to \Phi(x+a) = \mathrm{e}^{-ia^iP_i}\Phi(x) = \mathrm{e}^{a^i\partial_i}\Phi(x)$$

where a^i denotes the amount we want to translate in each direction. If we write the exponential function as Taylor series[158], this equation can simply be seen as the Taylor expansion[159] 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 $P_i = -i\partial_i$.

[158] This is derived in appendix B.4.1.

[159] As derived in appendix B.3.

3.8 The Poincare Group

Let's move on to the full spacetime symmetry group of nature: the Poincare 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. Or equivalently, the speed of light does not depend on the choice of where we put the origin of the coordinate system we use to describe some process. When we add these symmetries to the Lorentz group we get the Poincare group[160]

[160] The Poincare 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.

$$\begin{aligned} \text{Poincare group} &= \text{Lorentz group } \textbf{plus} \text{ translations} \\ &= \text{ Rotations } \textbf{plus} \text{ boosts } \textbf{plus} \text{ translations} \end{aligned} \tag{3.245}$$

The generators of the Poincare group are the generators of the Lorentz group J_i, K_i plus the generators of translations in Minkowski space P_μ.

In terms of J_i, K_i and P_μ the algebra reads[161]

[161] This is not very enlightening, but included for completeness.

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.246}$$

$$[J_i, K_j] = i\epsilon_{ijk}K_k \tag{3.247}$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k \tag{3.248}$$

$$[J_i, P_j] = i\epsilon_{ijk}P_k \tag{3.249}$$

$$[J_i, P_0] = 0 \tag{3.250}$$

$$[K_i, P_j] = i\delta_{ij}P_0 \tag{3.251}$$

$$[K_i, P_0] = -iP_i \tag{3.252}$$

Because this looks like a huge mess it is conventional to write this in terms of $M_{\mu\nu}$, which was defined by

$$J_i = \frac{1}{2}\epsilon_{ijk}M_{jk} \tag{3.253}$$

$$K_i = M_{0i}. \tag{3.254}$$

With $M_{\mu\nu}$ the Poincare algebra reads

$$[P_\mu, P_\nu] = 0 \tag{3.255}$$

$$[M_{\mu\nu}, P_\rho] = i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu) \tag{3.256}$$

and of course again

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}) \tag{3.257}$$

For this quite complicated group it is very useful to label the representations by using the fixed scalar values of the Casimir operators. The Poincare group has two Casimir operators[162] . The first one is:

$$P_\mu P^\mu =: m^2. \tag{3.258}$$

We give the scalar value the suggestive name m^2, because we will learn later that it coincides with the mass of particles[163].

The second Casimir operator is $W_\mu W^\mu$ with[164]

$$W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}P_\nu M_{\rho\sigma} \tag{3.259}$$

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 $j \equiv j_1 + j_2$, which 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[165] representations of the $SU(2)$ algebra.

For example, the $(j_1, j_2) = (0, 0)$ representation is called spin 0 representation[166]. 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[167] and analogously the $(j_1, j_2) = (\frac{1}{2}, \frac{1}{2})$ representation is called spin 1 representation[168].

The message to take away is that each representation is labelled by two scalar values: m and j. m can take on arbitrary values, but j is restricted to half-integer or integer values.

3.9 Elementary Particles

The labels for the irreducible representations of the Poincare group are how elementary particles are **labelled** in physics[169]: by their **mass** m and by their **spin** ($= j$ here). 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 Poincare group.

[162] Recall that a Casimir operator is defined as an operator, constructed from the generators, that commutes with all other generators.

[163] Don't worry, this will make much more sense later.

[164] $\epsilon^{\mu\nu\rho\sigma}$ is the four-dimensional Levi-Civita symbol, which is defined in appendix B.5.5.

[165] Recall that the 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).

[166] $j_1 + j_2 = 0 + 0 = 0$

[167] $j_1 + j_2 = \frac{1}{2} + 0 = 0 + \frac{1}{2} = \frac{1}{2}$

[168] $j_1 + j_2 = \frac{1}{2} + \frac{1}{2} = 1$

[169] Some prefer to say: Elementary particles **are** the irreducible representations of the Poincare 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

- mass: $9,109 \cdot 10^{-32}$ 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[170]

- 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 Poincare group are the **mathematical tools we need to describe all elementary particles**. To describe **scalar** 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 Sec. 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.

[170] 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.

Further Reading Tips

- **John Stillwell - Naive Lie Theory**[171] is a very readable, math orientated introduction to Lie Theory.
- **N. Jeevanjee - An Introduction to Tensors and Group Theory for Physicists**[172] is a very good introduction, with focus on the usage of Group Theory in physics.

[171] John Stillwell. *Naive Lie Theory*. Springer, 1st edition, 8 2008b. ISBN 9780387782140

[172] Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

3.10 Appendix: Rotations in a Complex Vector Space

The concept of transformations that preserve the inner product can be used with complex vector spaces, too. 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, where a complex number would make little sense. Therefore, the inner product of complex vector spaces is defined with additional complex conjugation[173]

$$a \cdot a = (a^T)^\star a = a^\dagger a. \tag{3.260}$$

[173] 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.

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 $U^\dagger U = 1$:

$$(Ua) \cdot (Ua) = a^\dagger U^\dagger U a = a^\dagger a \tag{3.261}$$

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 extra condition $\det(U) = 1$.

3.11 Appendix: Manifolds

A manifold M is a set of points if there exists 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 sphere. The surface of the three-dimensional sphere 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 sphere 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.

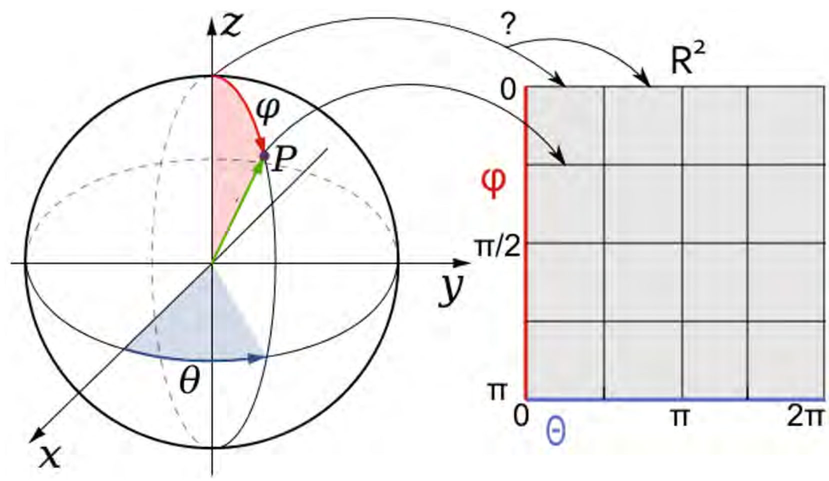

Fig. 3.8: Illustration of the map from one neighborhood of the sphere on to R^n.

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.