Lectures Notes for MSc in Theoretical Physics
University of Bologna, 2023/24

Relativistic Quantum Mechanis — Module 2

Spacetime Symmetries and their Representations

Dr Ling Lin[1]
University of Bologna
Department of Physics and Astronomy

Some references:

- S. Weinberg, “The Quantum Theory of Fields, Vol. 1”,
- H. Georgi, “Lie Algebras in Particle Physics”,
- H. Osborn, “Symmetries and Particles”, lecture notes of Part III Math, Cambridge University.

[1] Please send comments and corrections to ling.lin@unibo.it.

Contents

1 Vectors and dual vectors in Relativity

In Newtonian mechanics, space and time are separate entities. Physical laws in flat Euclidean space are invariant under Galilean transformations that map one inertial frame to another,

$$(t, \vec{x}) \mapsto (t', \vec{x}') = (t + \tau, R\vec{x} + \vec{v}t + \vec{d}), \tag{1.1}$$

where τ and $\vec{d}$ are constant translations in time and space, respectively, $\vec{v}$ is a constant velocity vector for the relative motions of the two inertial systems, and $\vec{x}$ is a rotation matrix. The key characterization of this linear transformation is that it leaves invariant the duration of events in time, $t'_2 - t'_1 = t_2 - t_1$, as well as the Euclidean distance in space, $\|\vec{y} - \vec{x}\|^2 \equiv \|\vec{s}\|^2 = \langle \vec{s}, \vec{s} \rangle$; more precisely, it preserves the scalar product[2]

$$\begin{aligned} \langle \vec{s}, \vec{r} \rangle \equiv \vec{s} \cdot \vec{r} &= \sum_{i,j} s_i r_i \delta_{ij} \\ &\stackrel{!}{=} \vec{s}' \cdot \vec{r}' = \sum_{i,j,k,l} (R_{ki} s_i)(R_{lj} r_j) \delta_{kl} \\ \Leftrightarrow \qquad \delta_{ij} &= \sum_{k,l} R_{ik} R_{lj} \delta_{kl} = \sum_{k,l} (R^T)_{ik} \delta_{kl} R_{lj} \,. \end{aligned} \tag{1.2}$$

This restricts the rotation matrices to be orthogonal matrices,

$$R \in O(3) = \{M \in \mathbb{R}^{3\times 3} | M^T M = \mathbb{1}\} \,. \tag{1.3}$$

More mathematically speaking, space in Newtonian mechanics has a non-degenerate symmetric bilinear form $\langle \cdot, \cdot \rangle$, for which one can choose an orthonormal basis $\vec{e}_i$, i.e., one with $\langle \vec{e}_i, \vec{e}_j \rangle = \delta_{ij}$. Orthogonal transformations, which map one orthonormal basis to another, leave the equations of motions invariant.

In relativity, space and time are now combined into a single structure: spacetime, also called *Minkowski space*, with coordinates $\vec{x}^\mu \equiv (ct, \vec{x})$.[3] By convention, the index μ starts at 0 for the time component, so that $x^\mu = (\vec{x})_\mu$ for $\mu = 1, 2, 3$. The reason why we place the index as a superscript instead of a subscript will be

[2] Note that only the displacement vector $\vec{s} = \vec{y} - \vec{x}$ between two points has a well-defined norm. This is because flat space itself is not a vector space, but an affine space (a "vector space without origin"). Choosing a reference frame is equivalent of choosing the origin, making the set of points to a vector space, on top of which one can define the scalar product.

[3] Again, these are points in an affine space, and only their differences form a vector space.

explained shortly. For now, we can think of the superscript μ as identyfing x as a point in spacetime, just as the overset arrow indicates points in space. Because of length contraction and time dilation, spatial distances and temporal intervals are no longer invariants across different inertial reference frames. Instead, it is a specific combination of them, encoded in the *Minkowskian* scalar product. For this scalar product, one always find a basis e_μ in each reference frame such that

$$\langle e_\mu, e_\nu \rangle = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{\mu\nu} . \tag{1.4}$$

$\eta_{\mu\nu}$ is also called the Minkowski metric,[4] and δ_{ij} would be the Euclidean metric. For a vector $x_2^\mu - x_1^\mu \equiv \Delta x^\mu$, the invariance of

$$\langle \Delta x^\mu, \Delta x^\mu \rangle = -(\Delta x^0)^2 + (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 = -(c\Delta t)^2 + \|\Delta\vec{x}\|^2 \tag{1.5}$$

means that time dilates by the same amount as spatial length contracts.

Linear transformations that leave this scalar product invariant are, parallel to the Euclidean case, a combination of generalized rotation Λ and a constant shift vector d^μ,

$$x'^\mu = (\Lambda x)^\mu + d^\mu . \tag{1.6}$$

In complete analogy to (1.2), invariance of the scalar product with respect to η now requires that the generalized rotations satisfy

$$\Lambda^T \eta \Lambda = \eta . \tag{1.7}$$

Such transformations are called *Lorentz transformations*, and are denoted

$$O(1,3) = \{\Lambda \in \mathbb{R}^{4\times 4} \mid \Lambda^T \eta \Lambda = \eta\} . \tag{1.8}$$

They are the obvious generalization of orthogonal matrices to a scalar product of signature $(1,3)$.

[4] In many references, particularly by authors with particle physics backgrounds, you may find $\eta_{\mu\nu} = \mathrm{diag}[1,-1,-1,-1]_{\mu\nu}$. The convention used here is more frequent in General Relativity.

Dual vectors

Given a vector space V (over any field K), the set V^* of *dual vectors* are defined as

$$V^* = \{f : V \to K \mid f \text{ is linear}\} . \tag{1.9}$$

In the finite dimensional case, given any basis $\{e_i\}$ of V, dual vectors are completely determined by their actions on $\{e_i\}$:

$$f(v) = f(\sum_i v_i e_i) = \sum_i f(e_i) v_i = (f(e_1), f(e_2), ...) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} . \tag{1.10}$$

We can therefore think of dual vectors as row vectors after choosing a basis. The natural basis, $(1, 0, ...)$, $(0, 1, ...)$, etc, form the *dual basis* $\{e_i^*\}$, satisfying $e_i^*(e_j) = \delta_{ij}$. Since there are equally many dual basis vectors, this establishes an isomorphism $V^* \cong V$.

However, this isomorphism is not canonical, i.e., it is only defined with respect to a choice of basis. Having made such a choice, we can easily verify that, for a basis transformation $e_i = \sum_k S_{ki} d_k$ with S_{ki} an invertible matrix, the new dual basis must be $d_i^* = \sum_l S_{il} e_l^*$. This means that vectors and dual vectors transform with the "opposite" matrix:

$$\begin{aligned} v &= \sum_i v_i e_i = \sum_{i,k} v_i S_{ki} d_k = \sum_{i,k} (S \cdot v)_k d_k , \\ v^* &= \sum_i v * i e_i^* = \sum_{i,k} v * i (S^{-1})_{ik} d_k^* = (v^* \cdot S^{-1})_k d_k^* , \end{aligned} \tag{1.11}$$

where in the last term in the second row, $v^* \cdot S^{-1}$ should be thought of as a row vector v^* multiplying a matrix from the left.

Tensor product

Given two vector spaces, V and W over the same field K, the *tensor product* $V \otimes W$ can be thought of as the vector space spanned by $e_i \otimes u_j$, where $\{e_i\}$ and $\{u_j\}$ are a basis of V and W, respectively. For $x = \sum_i x_i e_i \in V$ and $y = \sum_j y_j u_j \in W$, we then have

$$x \otimes y = \sum_{i,j} x_i y_j \, (e_i \otimes u_j) \in V \otimes W . \tag{1.12}$$

The so-called Universal Property of the tensor product makes this seemingly basis-dependent definition compatible with basis changes on V and W.

In particular, we can now consider tensor products of V and its dual V^*. For example, take $M \in V \otimes V^*$. Assuming that we have picked dual bases $\{e_i\}$ and $\{e_i^*\}$, then we can expand M in the basis $e_i \otimes e_j^*$:

$$M = \sum_{i,j} M_{ij}\, e_i \otimes e_j^* \,. \tag{1.13}$$

Because of the dual vector factor, we can regard M also as a linear map $V \longrightarrow V$, with

$$M(v) := \sum_{i,j} M_{ij}\, e_i \times e_j^* \left(\sum_k v_k e_k \right) = \sum_{i,j} M_{ij} v_j\, e_i \,, \tag{1.14}$$

from which we see that the coefficients M_{ij} are nothing but the matrix entries of the linear map in the basis $\{e_i\}$. At the same time, because $V \cong (V^*)^*$, we can also regard M as a linear map $V^* \longrightarrow V^*$, by

$$M(v^*) := \sum_{i,j} M_{ij}\, e_i \left(\sum_k v_k^* e_k^* \right) \times e_j^* = \sum_{i,j} v_i^* M_{ij}\, e_j^* \,, \tag{1.15}$$

where now the final sum can be read as a row vector v^* multiplying the matrix M_{ij} from the left.

Under basis transformation, $e_i = \sum_k S_{ki} d_k$, and correspondingly $e_j^* = \sum_l (S^{-1})_{jl} d_l^*$, we have

$$M = \sum_{i,j} M_{ij}\, e_i \otimes e_j^* = \sum_{i,j,k,l} S_{ki} M_{ij} (S^{-1})_{jl}\, d_k \otimes d_l^* \equiv \sum_{k,l} M'_{kl}\, d_k \otimes d_l^* \,. \tag{1.16}$$

So, as expected, the first index of M_{ij} transforms as a vector index, while the second index transforms as a dual-vector index. It is not hard to see that this generalizes to a tensor product with arbitrary powers of V and V^*: the indices in the basis expansion corresponding to a factor of V (V^*) will transform as a (dual) vector index.

This motivates the following notation that will be prevalent in relativity. A tensor of type (n, m) is an element

$$T \in \underbrace{V \otimes \cdots \otimes V}_{n \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{m \text{ times}} \,, \tag{1.17}$$

whose coefficients have *n upstairs* and *m downstairs* indices. A *basis vector* in turn has a downstairs index, and a *basis dual vector* has an upstairs index. The modified Einstein sum convention is then stating that any index that appears once upstairs and once downstairs is summed over. Given a basis $\{e_i\}$ for V and $(e^*)^j$ for V^*, the above tensor will be written as

$$T = T^{i_1 \dots i_n}{}_{j_1 \dots j_m} e_{i_1} \otimes \cdots e_{i_n} \otimes (e^*)^{j_1} \otimes \cdots (e^*)^{j_m} \,. \tag{1.18}$$

For most purposes, the basis vectors will be omitted, because the index structure of the coefficients determine the type of the the tensor, and the above mentioned Universal Property guarantees the intuitive expectation, namely that the coefficients transform under basis transformation, $e_i = \sum_k S^k{}_i d_k$ and $(e^*)^j = \sum_k (S^{-1})^j{}_k (d^*)^k$, in the natural way:

$$(T')^{i_1 \dots i_n}{}_{j_1 \dots j_m} = S^{i_1}{}_{k_1} \cdots S^{i_n}{}_{k_n} \, (S^{-1})^{j_1}{}_{l_1} \cdots (S^{-1})^{j_m}{}_{l_m} \, T^{i_1 \dots i_n}{}_{j_1 \dots j_m} \,. \tag{1.19}$$

Vectors will then simply be denoted by their components v^i with an upstairs index, and dual vectors with v_i, having a downstairs index. In the context of relativity, vectors and upstairs indices are also often referred to as contravariant vectors/indices, as opposed to dual vectors / downstairs indices which are referred to as covariant.

Scalar product and raising/lowering indices

On a real vector space V, a scalar product is a symmetric bilinear form $B : V \times V \longrightarrow \mathbb{R}$ which is non-degenerate, i.e., if $B(v, u) = 0$ for all $u \in V$, then $v = 0$.[5] In a given basis $\{e_i\}$, the coefficient matrix $B(e_i, e_j) \equiv \mathcal{B}_{ij}$ are sometimes called the metric tensor.

Clearly, we have $B \in V^* \otimes V^*$. Under a basis transformation $e_i = \sum_k S^k{}_i d_k$, we therefore have

$$\begin{aligned} &\mathcal{B}_{ij} = S^k{}_i \, S^l{}_j \, B(d_k, d_l) = (S^T)_i{}^k \mathcal{B}_{kl} \, S^l{}_j = (S^T \, \mathcal{B} \, S)_{kl} \,, \\ \Longrightarrow \quad &\mathcal{B}'_{kl} := B(d_k, d_l) = (S^{-1})^i{}_k \, (S^{-1})^j{}_l \, \mathcal{B}_{ij} = ((S^{-1})^T \mathcal{B} \, S^{-1})_{kl} \,. \end{aligned} \tag{1.20}$$

The non-degeneracy of the bilinear form translates into the matrix $\mathcal{B}$ to have non-vanishing determinant.

[5] On a complex vector space, bilinearity is replaced with *sesquilinearity*, i.e., $B(v_1 + \lambda v_2, u) = B(v_1, u) + \overline{\lambda} B(v_2, u)$, and symmetric is replaced with Hermitian, i.e., $B(v, u) = \overline{B(u, v)}$; here, $\overline{x}$ denotes complex conjugation.

Such a scalar product now provides a canonical isomorphism between vectors and dual vectors. That is, the map

$$\iota : V \longrightarrow V^* , \quad v \mapsto B(v, \cdot) , \tag{1.21}$$

whose invertibility is a consequence of the non-degeneracy of B, is obviously independent of any choice of basis. Given a basis $\{e_i\}$, it is easy to verify that this map allows us to determine the dual basis as

$$(e^*)^i = \sum_j (\mathcal{B}^{-1})^{ij} B(e_j, \cdot) \quad \Leftrightarrow \quad \iota(e_j) = B(e_j, \cdot) = \mathcal{B}_{ji} (e^*)^i , \tag{1.22}$$

where the inverse metric tensor $\mathcal{B}^{-1}$ now carries upstairs index.

For a vector $v = v^i e_i$, we then have

$$\iota(v) = v^i B(e_i, \cdot) = v^i \mathcal{B}_{ij} \, (e^*)^j = \mathcal{B}_{ji} v^i \, (e^*)^j \equiv v_i \, (e^*)^j . \tag{1.23}$$

This means that, given a scalar product with metric tensor $\mathcal{B}_{ij}$ in some basis, any vector v^i is naturally associated with one specific dual vector v_i, whose components are obtained by "lowering the index with the metric tensor", $v_i = \mathcal{B}_{ij} v^j$. Given that the inverse map ι^{-1} must satisfy $\iota^{-1} \, \iota = \mathrm{id}$, it is evident that

$$\iota^{-1}(v_i) \equiv v^i = (\mathcal{B}^{-1})^{ij} v_j , \tag{1.24}$$

i.e., we "raise the index", using the inverse metric tensor.

These processes generalize in an obvious fashion to general tensors, e.g., $T_{ij} = T^k{}_j \mathcal{B}_{ik}$. Lowering an index changes a type (n, m) tensor to a type $(n-1, m+1)$ tensor, and raising an index changes type (n, m) to $(n+1, m-1)$.

In the Euclidean case, after choosing an orthonormal basis, the metric tensor is simply δ_{ij}, so v_i and v^i agree component-wise. For the Minkowski case, the minus sign in the temporal component for $\eta_{\mu\nu}$ and $(\eta^{-1})^{\mu\nu} \equiv \eta^{\mu\nu}$ leads to a significant difference. A peculiar consequence of this notation is the identity

$$\eta^\mu{}_\nu = \eta^{\mu\alpha} \eta_{\alpha\nu} = \delta^\mu{}_\nu . \tag{1.25}$$

Maxwell's equations

One important way to construct new tensors in physics is to take derivatives of other tensors. To verify the nature of the derivative operator $\frac{\partial}{\partial x^\mu}$ in regards to its index, let us check its transformation behavior under a basis transformation. Because the coordinates x^μ behave as vectors under this transformation,

$$x^\mu \mapsto y^\mu = S^\mu{}_\nu x^\nu \quad \Leftrightarrow \quad x^\nu = (S^{-1})^\nu{}_\mu y^\mu \,, \tag{1.26}$$

we can apply the chain rule, and find

$$\frac{\partial}{\partial y^\mu} = \sum_\nu \frac{\partial x^\nu}{\partial y^\mu}\frac{\partial}{\partial x^\nu} = \sum_\nu (S^{-1})^\nu{}_\mu \frac{\partial}{\partial x^\nu} \,. \tag{1.27}$$

This means that the derivative operator transforms as a dual vector, so it makes sense to use the short-hand notation $\partial_\mu = \frac{\partial}{\partial x^\mu}$.

Let us demonstrate the usefulness of the new notation in the relativistic formulation of Maxwell's equations. In their original form, we have

$$\begin{aligned} \text{inhomogeneous:} \quad & \vec{\nabla}\cdot\vec{E} = \rho \,, \quad \vec{\nabla}\times\vec{B} - \frac{1}{c}\frac{\partial \vec{E}}{\partial t} = \frac{1}{c}\vec{j} \,, \\ \text{homogeneous:} \quad & \vec{\nabla}\cdot\vec{B} = 0 \,, \quad \vec{\nabla}\times\vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0 \,. \end{aligned} \tag{1.28}$$

Now, since $\vec{E}$ and $\vec{B}$ field strength are transformed into each other under Lorentz transformations, they are combined into a single quantity in relativity, which is called the field strength tensor,

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}^{\mu\nu} = -F^{\nu\mu} \,. \tag{1.29}$$

Using this tensor, the equations can be recast into

$$\begin{aligned} \text{inhomogeneous:} \quad & \partial_\mu F^{\mu\nu} = -\frac{1}{c} J^\nu \quad \text{with } J^\nu = (c\rho, \vec{j})^\nu \,, \\ \text{homogeneous:} \quad & \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \,. \end{aligned} \tag{1.30}$$

Furthermore, we can also introduce a relativistic vector potential,

$$A_\mu = (A_0, \vec{A}) = (-\phi, \vec{A}) \,, \tag{1.31}$$

such that we have

$$\vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{t} \,, \quad \vec{B} = \vec{\nabla}\times\vec{A} \quad \Leftrightarrow \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{1.32}$$

2 Basics of Lie groups

In physics, traditional symmetries are transformations of the system which leave the physical laws invariant; they can be thought of as a change of mathematical description, $\mathcal{S} \xrightarrow{T} \mathcal{S}'$. Intuitively, these transformations can be composed, i.e., $T_2 \circ T_1$ should also be a transformation, and invertible, which are precisely the axioms for a group. Moreover, many important symmetry transformations in physics are continuous, and have a notion of "infinitesimal" transformations. Mathematically, these are so-called Lie groups, and the main examples we will look at will be the rotation group in 3d Euclidean space, and the Lorentz group of $(3+1)$d Minkowski space.

2.1 Groups

A group G is a set with a "composition" map,

$$\circ : G \times G \to G\,, \quad (g, h) \mapsto g \circ h \in G\,, \tag{2.1}$$

which satisfies the following axioms:

$$\begin{aligned} \text{associativity:} &\quad (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)\,, \\ \text{unique unit element:} &\quad \exists_1 e \in G\, \forall g \in G : \ e \circ g = g \circ e = g\,, \\ \text{invertibility:} &\quad \forall g \in G\, \exists g^{-1} : \ g \circ g^{-1} = g^{-1} \circ g = e\,. \end{aligned} \tag{2.2}$$

In addition, if the composition is commutative, i.e., $g \circ h = h \circ g$ for all $g, h \in G$, then G is also called a commutative, or *abelian* group.

Familiar examples of groups are rational, real, or complex numbers with $\circ = +$, and $e = 0$, or non-zero rational/real/complex numbers with $\circ = \cdot$, and $e = 1$. In fact, any field consists essentially of two groups (addition and multiplication, the later must exclude 0) which are "distributively compatible".

Though not the main subject in what follows, the theory of finite groups is rich and interesting. In principle, finite groups G, whose number of elements is called the order $|G|$, can be entirely characterized by a "multiplication table". That is, by specifying, for any ordered pair (g, h) the result of $g \circ h$, G is completely specified as a group. A finite abelian group would have a multiplication table that is symmetric along the diagonal. The smallest non-abelian group is of order 6, and is the permutation group S_3 of three ordered objects. Its multiplication table takes the following form:

$\circ$	e	A	B	C	D	E
e	e	A	B	C	D	E
A	A	B	e	D	E	C
B	B	e	A	E	C	D
C	C	E	D	e	B	A
D	D	C	E	A	e	B
E	E	D	C	B	A	e

(2.3)

Notice that the consistency with the group axioms require that in each column and row, each group element appears exactly once.

Given a group G, a *subgroup* $H \subset G$ is a subset which satisfies the group axioms with the composition given by the restriction of the composition of G to H. The trivial group, $\{e\}$, is always a subgroup of any group. For finite groups, a fundamental property of subgroups is that the order of the subgroup must divide the order of the parent group, $|G|/|H| \in \mathbb{N}$. This means that S_3 can have non-trivial subgroups of order 2 and 3. From the multiplication table, we can recognize subgroups as sub-columns and -rows which contain exactly the restricted set of elements. For example, there are three order 2 subgroups, $\{e, C\}$, $\{e, D\}$ and $\{e, E\}$, and one order 3 group, $\{e, A, B\}$. These are all abelian, and isomorphic to $\mathbb{Z}_2$ and $\mathbb{Z}_3$, respectively.[6] More generally, *any* finite group appears as a subgroup of some permutation group S_m for large enough m.

2.2 Matrix groups

Another class of groups which are generically non-abelian are matrix groups, with group composition given by matrix multiplication, and the unit element being the unit matrix. Since not all matrices are invertible, the most general subset of matrices that form a group is the *general linear group*,

$$GL(n; K) = \{M \in K^{n\times n} \mid \det(M) \neq 0\} \tag{2.4}$$

for any field K.

[6] The cyclic group of order n is the quotient $\mathbb{Z}_n \equiv \mathbb{Z}/n\mathbb{Z} \cong \{0, 1, ..., n-1\}$, with group composition given by $a \circ b = b \circ a := a + b \mod n$, and unit element $e = 0$.

Further specialization leads to various subgroups of $GL(n;K)$. For example, given any fixed, invertible $n \times n$ matrix $\mathcal{B}$, the subgroup of $GL(n;K)$ which preserves this particular matrix, $M^T \mathcal{B} M = \mathcal{B}$, form a group. We have already seen examples in the orthogonal group $O(n)$ with $K = \mathbb{R}$ and $\mathcal{B}_{ij} = \delta_{ij}$, or the Lorentz group $O(1, n-1)$ and $\mathcal{B}_{ij} = \mathrm{diag}(-1, 1, ..., 1)_{ij}$. Another example is when $n = 2m$ is even, and we have

$$\mathcal{B} = \begin{pmatrix} \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & 0 & \cdots & 0 \\ 0 & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \end{pmatrix}, \tag{2.5}$$

in which case the group is the sympletic group $Sp(m;K)$ (sometimes it is also labelled $Sp(2m;K)$. Over the complex numbers, we have the *unitary* groups,

$$U(n) = \{U \in \mathbb{C}^{n \times n} \mid U^\dagger U = \mathbb{1}\}, \tag{2.6}$$

where $\dagger$ denotes Hermitian conjugation, i.e., complex conjugation and transposition. For all these specialized groups G, their definition immediately implies that $|\det(M)| = 1$ for $M \in G$. In the case of the symplectic group, it is possible to use the concept of the *Pfaffian* to show that $\det(M) = 1$ for $M \in Sp(m)$.

In the other cases, we can impose $\det(M) = 1$ as a further restriction to give us so-called "special" groups. These are

$$\begin{aligned} \text{special linear group:} \quad & SL(n;K) = \{M \in GL(n;K) \mid \det(M) = 1\}, \\ \text{special orthogonal group:} \quad & SO(n) = \{R \in O(n) \mid \det R = 1\}, \\ \text{special Lorentz transf.:} \quad & SO(1,n) = \{\Lambda \in O(1,n) \mid \det \Lambda = 1\}, \\ \text{special unitary group:} \quad & SU(n) = \{U \in U(n) \mid \det(U) = 1\}. \end{aligned} \tag{2.7}$$

2.3 Infinitesimal transformations and Lie algebras

Matrix groups exemplify the class of groups known as Lie groups. A Lie group is a group G endowed with a manifold structure, such that group actions — composition and taking the inverse — are smooth maps, i.e.,

$$\mu : G \times G \to G, (x, y) \mapsto x^{-1} \circ y \tag{2.8}$$

is smooth. Intuitively, this means that it makes sense to talk about "close by" elements, and the existence of a sort of "power series" to approximate finitely separated elements.

More precisely, we focus on group elements that are infinitesimally away from the unit element. Thinking of the group as (symmetry) transformations, e.g., as matrices acting on vectors, $v \mapsto Mv$, the manifold structure of the group allows us to expand the transformation in a Taylor-like power series, $Mv = v + Xv + X^2 v + \ldots$ Just as for Taylor-expansions, there is a sense in which the first order term parametrizes an infinitesimal transformation $Mv = v + \delta v$, with $\delta v = Xv$. On a smooth manifold, the set $\mathfrak{g} \ni X$ of infinitesimal transformations is the *tangent space* at the origin $e \in G$, and naturally carries an $\mathbb{R}$-vector space structure (the tangent vectors).

However, because of the group structure on G, that tangent space also has a multiplication (which in general is neither commutative nor invertible). For matrix groups, these tangent vectors are themselves matrices, with vector addition given by component-wise addition, and the multiplication is induced by matrix multiplication. Note that the matrices X describing tangent vectors are not necessarily of the same kind as those describing the matrices in G, and they are in general not invertible. All in all, this gives $\mathfrak{g}$ the structure of an *algebra* (a vector space with generally non-invertible multiplication between vectors), and is therefore called the *Lie algebra* associated to G.

A particularly important property of Lie groups is the existence of a map, the so-called exponential,

$$\exp : \mathfrak{g} \to G\,. \tag{2.9}$$

This map is the aforementioned power series that constructs a group element g that is finitely far away as $g = \exp(X)$ for some $X \in \mathfrak{g}$. Geometrically, this map is tied to the existence of a unique *path* $\gamma : \mathbb{R} \to G$ in G with $\gamma(0) = e$ and $\gamma(1) = g$ which is a *one-parameter subgroup* $G_\gamma = \{\gamma(s) \mid s \in \mathbb{R}\} \subset G$, and whose tangent vector at e is X. Behind this more abstract mathematical definition is the simple intuition for matrix groups, where this map is literally the matrix exponential; we will see this in detail for the rotation and the Lorentz group later.

The bottom line is that this map allows us to study most of G from the properties of the Lie algebra. However, from the path description, it is also clear that we can only obtain group elements that lie in the *connected component* of e, also called the *identity component* of G. That is, only the subset of G that can be connected to e

with an uninterrupted path can be expressed as the exponential of some Lie algebra element.

For the identity component, we have the famous Baker–Campbell–Hausdorff (BCH) formula,

$$\exp(X) \circ \exp(Y) = \exp\left(X + Y + \tfrac{1}{2}[X, Y] + \tfrac{1}{12}[X, [X, Y]] - \tfrac{1}{12}[Y, [X, Y]] + \cdots\right), \tag{2.10}$$

which connects group composition with the Lie algebra multiplication. This multiplication operation is denoted by square brackets, and is called the *Lie bracket*, or *the commutator* in physics context,

$$[\cdot, \cdot] : \ \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}, \ (X, Y) \mapsto [X, Y] \, . \tag{2.11}$$

The commutator is an anti-symmetric multiplication, i.e., it satisfies

$$[\alpha X + \beta Y, Z] = -[Z, \alpha X + \beta Z] = \alpha[X, Z] + \beta[Y, Z] \, , \quad \alpha, \beta \in \mathbb{R} \, , \tag{2.12}$$

which furthermore must satisfy the *Jacobi identity*,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \, . \tag{2.13}$$

For a matrix group G (which we will see in examples below), $\mathfrak{g}$ consists of (in general non-invertible) matrices X such that $G \ni g = \exp(X)$ is the exponential of matrices. Then, the Lie bracket on $\mathfrak{g}$ takes the familiar form of the commutator,

$$[X, Y] = XY - YX \, , \tag{2.14}$$

with XY being the matrices multiplication between X and Y.

Given a vector space basis $\{T_i\}$ of a Lie algebra $\mathfrak{g}$, the Lie bracket structure can be encoded in so-called *structure constants* ${f_{ij}}^k$,

$$[T_i, T_j] = \sum_k f_{ijk} T_k \, . \tag{2.15}$$

Anti-symmetry of the commutator then becomes the condition $f_{ijk} = -f_{jik}$, and a more lengthy one for the Jacobi identity. Any consistent set of structure constants then define a Lie algebra, and through the exponential map with the BCH formula, one can express elements of the Lie group in terms of exponentials of linear combinations of the basis elements $\{T_i\}$. These elements will also be called the *generators* of the Lie algebra.

2.4 Symmetries in Quantum Mechanics

The fact that symmetry transformations are generated by infinitesimal actions should be familiar from the properties of position / momentum and Hamilton operators in quantum mechanics. For example, the momentum operator $\hat{P}$ is said to generate infinitesimal spatial translation:

$$\langle x|\hat{P}|\psi\rangle = -i\hbar\partial_x\psi(x)\,. \tag{2.16}$$

Finite translations are then generated by the exponential of the operator,

$$\langle x|\exp(-\tfrac{i}{\hbar}y\,\hat{P})|\psi\rangle = \psi(x+y)\,. \tag{2.17}$$

An even more prominent example is the Schrödinger equation,

$$\hat{H}|\psi(t)\rangle = i\hbar\frac{d}{dt}|\psi(t)\rangle, \tag{2.18}$$

which leads to finite time translations through the time evolution operator

$$|\psi(t+\tau)\rangle = \exp(-\tfrac{i}{\hbar}\hat{H}\tau)|\psi(t)\rangle\,. \tag{2.19}$$

Also in quantum mechanics, the key relations that determine the dynamics are the commutation relations between the operators.

3 Rotation and Lorentz groups as Lie groups

3.1 The Lie group $SO(3)$

We first study the rotation group in three dimensions. As we remarked earlier, from the defining property $R^T R = \mathbb{1}$ of $O(3)$, the matrices have $\det(R) = \pm 1$. Since these are discrete values, and the determinant is a continuous function in the matrix entries (since it can be expressed as a polynomial), there cannot be a continuous path in $O(3)$ that connects an element R_1 with $\det(R_1) = +1$ with an element R_2 with $\det(R_2) = -1$. Since $\det(\mathbb{1}) = 1$, this means that the subset of matrices with negative determinant is not connected to the identity element. The identity component of $O(3)$ is then just $SO(3)$, as defined above. Once we know $SO(3)$, we can also construct a negative determinant matrix in $O(3)$ by composing an $SO(3)$ element with a reflection along a plane.

To determine the Lie algebra $\mathfrak{so}(3)$, note that any $SO(3)$ rotation R can be parametrized by three Euler angles, or, equivalently, by specifying a rotation axis (labelled by a unit vector $\vec{n}$, which only has two independent parameters since $\vec{n} \cdot \vec{n} = 1$) and an angle θ around it. Notice that $R(\theta = 0, \vec{n}) = \mathbb{1}$. For infinitesimal angles $\delta\theta$, the action on an arbitrary vector $\vec{y} \in \mathbb{R}^3$ is an infinitesimal rotation given by

$$R(\delta\theta, \vec{n})\vec{y} = \vec{y} + \delta\theta \vec{n} \times \vec{y} = \vec{y} + \delta\theta \begin{pmatrix} n_2 y_3 - n_3 y_2 \\ n_3 y_1 - n_1 y_3 \\ n_1 y_2 - n_2 y_1 \end{pmatrix} . \tag{3.1}$$

If we now pick $\vec{n}$ to be the Euclidean basis vectors, we obtain

$$\begin{aligned}
R(\delta\theta, \vec{n} = (1,0,0)^T) &= \vec{y} + \delta\theta \begin{pmatrix} 0 \\ -y_3 \\ y_2 \end{pmatrix} = \vec{y} + \delta\theta \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}_{L_1} \vec{y} , \\
R(\delta\theta, \vec{n} = (0,1,0)^T) &= \vec{y} + \delta\theta \begin{pmatrix} y_3 \\ 0 \\ -y_1 \end{pmatrix} = \vec{y} + \delta\theta \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}}_{L_2} \vec{y} , \\
R(\delta\theta, \vec{n} = (0,0,1)^T) &= \vec{y} + \delta\theta \begin{pmatrix} -y_2 \\ y_1 \\ 0 \end{pmatrix} = \vec{y} + \delta\theta \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{L_3} \vec{y} .
\end{aligned} \tag{3.2}$$

The matrices L_i have components $(L_i)_{jk} = -\epsilon_{ijk}$, and correspond to a basis of the Lie algebra $\mathfrak{so}(3)$. Any finite rotation can then be written as

$$R(\theta, \vec{n}) = \exp\left(\theta \sum_i n_i L_i\right) \equiv \exp(\theta \vec{n} \cdot \vec{L}) , \tag{3.3}$$

where $\vec{L}$ is a vector whose components are themselves 3×3 matrices.

The definition of the group $SO(3)$ is also reflected in these matrices: Orthogonality requires $\mathbb{1} = R(\theta, \vec{n})^T R(\theta, \vec{n}) = \exp(\theta \sum_i n_i (L_i^T + L_i) + \ldots)$, so

$L_i^T + L_i = 0$ implies that L_i is anti-symmetric, and, by the formula $\det(\exp(M)) = \exp(\mathrm{tr}(M))$, the determinant condition $\det(R) = 1$ requires $\mathrm{tr}(L_i) = 0$. So $\mathfrak{so}(3) = \{$anti-symmetric traceless 3×3 matrices$\}$. This is a three-dimensional real vector space with basis $\{L_i\}$.

From the explicit forms of the generators L_i, we can compute the commutators explicitly, and find

$$[L_i, L_j] = \sum_k \epsilon_{ijk} L_k \, . \tag{3.4}$$

As an exercise, it is also instructive to derive this relation from applying the BCH formula to $R = R_2^{-1} R_1^{-1} R_2 R_1$, with $R_i \equiv R(\delta\theta_i, \vec{n}_i)$, and compare the lowest non-trivial order terms with the "direct computation" using (3.1).

As a final comment, note that in the physics literature, it is often common to use the generators $J_k = i L_k$, so that $R(\theta, \vec{n}) = \exp(-i\theta \vec{n} \cdot \vec{J})$. Because L_k is anti-symmetric, J_k is a Hermitian matrix, which satisfy the commutation relation

$$[J_i, J_j] = i \sum_k \epsilon_{ijk} J_k \, . \tag{3.5}$$

Later, when we come to representations of groups and algebras, we will explain why it makes sense to multiply vectors of a real vector space with imaginary numbers.

3.2 Proper orthochronous Lorentz group and its algebra

From the definition of the Lorentz transformations,

$$(\Lambda^T \eta \Lambda)_{\mu\nu} = \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \, \eta_{\alpha\beta} = \eta_{\mu\nu} \, , \tag{3.6}$$

we also have $\det(\Lambda) = \pm 1$. Analogously to the orthogonal matrices, the subset of matrices with negative determinant are not connected to the identity. One may naively assume that the positive determinant part, $SO(1,3)$, is then the identity component, but for Lorentz transformations there is an additional subtlety. From the $\mu = \nu = 0$ component of the above tensor equation, we see that

$$-1 = \eta_{00} = \Lambda^\alpha{}_0 \Lambda^\beta{}_0 \, \eta_{\alpha\beta} = -(\Lambda^0{}_0)^2 + \sum_{i=1}^{3} (\Lambda^i{}_i)^2 \implies (\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\Lambda^i{}_i)^2 \, . \tag{3.7}$$

This implies $(\Lambda^0{}_0)^2 \geq 1$, so either $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$. Again, it is obvious that there is no continuous path that can interpolate between values larger than 1 and smaller than -1 without crossing the region in between, which however would not be part of the Lorentz group. Since the unit matrix has $\Lambda^0{}_0 = 1$, the identity component is

$$SO(1,3)^+ = \{\Lambda \in SO(1,3) \mid \Lambda^0{}_0 \geq 1\}\,, \tag{3.8}$$

which is called the *proper orthochronous* Lorentz group. The full Lorentz group is obtained by composing proper orthochronous transformations with spatial and time reflections.

Let us directly inspect the Lie algebra. Again, we consider an infinitesimal Lorentz transformation, $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu$, where $\omega^\mu{}_\nu$ are some "small" parameters. Inserting this into (3.6), we find

$$\eta_{\mu\nu} = (\delta^\alpha{}_\mu \delta^\beta{}_\nu + \omega^\alpha{}_\mu \delta^\beta{}_\nu + \delta^\alpha{}_\mu \omega^\beta{}_\nu)\eta_{\alpha\beta} + O(\omega^2) = \eta_{\mu\nu} + \omega_{\nu\mu} + \omega_{\mu\nu} + O(\omega^2) \tag{3.9}$$

which implies $\omega_{\mu\nu} = -\omega_{\nu\mu}$ is anti-symmetric. So there are six independent ways to perform infinitesimal Lorentz transformations, i.e., the Lie algebra $\mathfrak{so}(1,3)$ should have six generators. We could use labels $A \in \{1, 2, \ldots, 6\}$, so that the generators are $\tilde{M}_A$, and a general element takes the form $\sum_A \omega_A \tilde{M}_A$. However, we can equivalently use the Lorentz indices α, β as labels, but impose $M_{\alpha\beta} = -M_{\beta\alpha}$ to reduce the number of independent combinations to six, and then write a general Lie algebra element as $\frac{1}{2}\omega^{\alpha\beta} M_{\alpha\beta}$, now using the convention establish in the first chapter to denote sums over indices (the factor of $\frac{1}{2}$ accounts for summing over the same terms twice due to anti-symmetry). Notice that $\tilde{M}_A$ and $M_{\alpha\beta}$ are related to each other in a non-trivial way, because we have to use the metric tensor to raise the indices on ω. However, this is just a simple redefinition of the basis we choose for the Lie algebra. Having established this basis, we can now write the Lorentz transformation $\Lambda^\mu{}_\nu x^\nu$ for any four-vector x^μ as

$$\Lambda^\mu{}_\nu x^\nu = \exp(-\tfrac{i}{2}\omega^{\alpha\beta} M_{\alpha\beta})^\mu{}_\nu\, x^\nu = (\delta^\mu{}_\nu + \omega^\mu{}_\nu) x^\nu + O(\omega^2)\,, \tag{3.10}$$

where $(M_{\alpha\beta})^\mu{}_\nu = -(M_{\beta\alpha})^\mu{}_\nu$ are some 4×4 matrices.[7]

[7] The indices may seem irritating at first. In the end, we are using Lorentz indices to both label different matrices (in this case each label is a pair (α, β)), as well as the entries of each single matrix (where μ enumerate the rows, and ν the columns).

This allows us to compute the commutation relations $[M_{\alpha\beta}, M_{\sigma\rho}]$ without explicitly determining the form of the matrices $(M_{\alpha\beta})^\mu{}_\nu$. To do this, we will compare the result for a Lorentz transformation of the form $(\tilde{\Lambda}^{-1}\Lambda^{-1}\tilde{\Lambda}\Lambda)^\mu{}_\nu x^\nu$, computed first using the infinitesimal action on vectors directly, and then using the BCH formula for the matrix exponential.

First, we have

$$\begin{aligned}
&(\tilde{\Lambda}^{-1}\Lambda^{-1}\tilde{\Lambda}\Lambda)^\mu{}_\nu = (\tilde{\Lambda}^{-1})^\mu{}_\alpha(\Lambda^{-1})^\alpha{}_\beta\tilde{\Lambda}^\beta{}_\gamma\Lambda^\gamma{}_\nu \\
&= (\delta^\mu{}_\alpha - \tilde{\omega}^\mu{}_\alpha + O(\tilde{\omega}^2))(\delta^\alpha{}_\beta - \omega^\alpha{}_\beta + O(\omega^2))\times \\
&\quad(\delta^\beta{}_\gamma + \tilde{\omega}^\beta{}_\gamma + O(\tilde{\omega}^2))(\delta^\gamma{}_\nu + \omega^\gamma{}_\nu + O(\omega^2)) \\
&= \delta^\mu{}_\nu + \tilde{\omega}^\mu{}_\alpha\,\omega^\alpha{}_\nu - \tilde{\omega}^\mu{}_\alpha\omega^\alpha{}_\nu - \omega^\mu{}_\beta\tilde{\omega}^\beta{}_\nu + \tilde{\omega}^\mu{}_\gamma\omega^\gamma{}_\nu + O(\omega^2) + O(\tilde{\omega}^2) \\
&= \delta^\mu{}_\nu - \omega^\mu{}_\beta\tilde{\omega}^\beta{}_\nu + \tilde{\omega}^\mu{}_\gamma\omega^\gamma{}_\nu + O(\omega^2) + O(\tilde{\omega}^2)\,,
\end{aligned} \tag{3.11}$$

where, from the infinitesimal expansion, we only keep track of terms proportional to $\omega\tilde{\omega}$ and $\tilde{\omega}\omega$. From (3.10), this means that we must have

$$(\tilde{\Lambda}^{-1}\Lambda^{-1}\tilde{\Lambda}\Lambda) = \exp(-\tfrac{i}{2}(\tilde{\omega}^\mu{}_\gamma\omega^\gamma{}_\nu - \omega^\mu{}_\beta\tilde{\omega}^\beta{}_\nu)\eta^{\nu\sigma}M_{\mu\sigma})\,. \tag{3.12}$$

Alternatively, we can express $\Lambda = \exp(-\frac{i}{2}\omega^{\alpha\beta}M_{\alpha\beta})$ and use the BCH formula. From this we obtain

$$\begin{aligned}
&\tilde{\Lambda}^{-1}\Lambda^{-1}\tilde{\Lambda}\Lambda \\
&= \exp(\tfrac{i}{2}\tilde{\omega}^{\alpha\beta}M_{\alpha\beta})\exp(\tfrac{i}{2}\omega^{\sigma\rho}M_{\sigma\rho})\exp(-\tfrac{i}{2}\tilde{\omega}^{\alpha\beta}M_{\alpha\beta})\exp(-\tfrac{i}{2}\omega^{\sigma\rho}M_{\sigma\rho}) \\
&= \exp([\tfrac{i}{2}\tilde{\omega}^{\alpha\beta}M_{\alpha\beta}, \tfrac{i}{2}\omega^{\sigma\rho}M_{\sigma\rho}] + O(\omega^2) + O(\tilde{\omega}^2))\,.
\end{aligned} \tag{3.13}$$

These two results should in the end be interpreted as a power series in the "variables" ω and $\tilde{\omega}$, after expanding the exponential. For them to agree for arbitrary ω and $\tilde{\omega}$, the coefficients in front of each "monomial" must agree, in particular those in front of $\tilde{\omega}\omega$ and $\omega\tilde{\omega}$. In our computations above, the only sources of these monomials are just the lowest order terms in the argument of the exponentials. So this implies the relation

$$\begin{aligned}
[\tfrac{i}{2}\tilde{\omega}^{\alpha\beta}M_{\alpha\beta}, \tfrac{i}{2}\omega^{\sigma\rho}M_{\sigma\rho}] &= -\tfrac{1}{4}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}[M_{\alpha\beta}, M_{\sigma\rho}] \\
&\overset{!}{=} -\tfrac{i}{2}(\tilde{\omega}^\mu{}_\gamma\omega^\gamma{}_\nu - \omega^\mu{}_\beta\tilde{\omega}^\beta{}_\nu)\eta^{\nu\sigma}M_{\mu\sigma} \\
&= -\tfrac{i}{2}\eta_{\gamma\lambda}(\tilde{\omega}^{\mu\gamma}\omega^{\lambda\nu} - \omega^{\mu\gamma}\tilde{\omega}^{\lambda\nu})M_{\mu\nu}\,.
\end{aligned} \tag{3.14}$$

To extract from this condition the commutation relations for the generators, we make the general ansatz

$$\begin{aligned}&[M_{\alpha\beta}, M_{\sigma\rho}]\\ &= T^{(1)}_{\alpha\beta} M_{\sigma\rho} + T^{(2)}_{\alpha\sigma} M_{\beta\rho} + T^{(3)}_{\alpha\rho} M_{\beta\sigma} + T^{(4)}_{\beta\sigma} M_{\alpha\rho} + T^{(5)}_{\beta\rho} M_{\alpha\sigma} + T^{(6)}_{\sigma\rho} M_{\alpha\beta} \,.\end{aligned} \tag{3.15}$$

Inserting into (3.14), we immediately see that $T^{(1)} = T^{(6)} = 0$, because there are no matching terms on the right-hand side once we contract this with $\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}$. Moreover, we can exploit the anti-symmetry $M_{\mu\nu} = -M_{\nu\mu}$:

$$\begin{aligned}0 &= [M_{\alpha\beta}, M_{\sigma\rho}] + [M_{\beta\alpha}, M_{\sigma\rho}]\\ &= [M_{\alpha\beta}, M_{\sigma\rho}] + T^{(2)}_{\beta\sigma} M_{\alpha\rho} + T^{(3)}_{\beta\rho} M_{\alpha\sigma} + T^{(4)}_{\alpha\sigma} M_{\beta\rho} + T^{(5)}_{\alpha\rho} M_{\beta\sigma} \,.\end{aligned} \tag{3.16}$$

Since this is to hold for all values of indices, this implies, together with (3.15), $T^{(2)}_{\alpha\sigma} + T^{(4)}_{\alpha\sigma} = 0$ and $T^{(3)}_{\alpha\rho} + T^{(5)}_{\alpha\rho} = 0$. Likewise, by using anti-symmetry for the second argument of the commutator, we find $T^{(2)}_{\alpha\sigma} + T^{(3)}_{\alpha\sigma} = 0$ and $T^{(4)}_{\beta\sigma} + T^{(5)}_{\beta\sigma} = 0$. So all in all we have

$$T^{(2)} = -T^{(3)} = -T^{(4)} = T^{(5)} \equiv T \tag{3.17}$$

for some tensor T.

Inserting this into (3.14), we find

$$\begin{aligned}&-\tfrac{1}{4}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}[M_{\alpha\beta}, M_{\sigma\rho}]\\ &= -\tfrac{1}{4}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}(T_{\alpha\sigma}M_{\beta\rho} - T_{\alpha\rho}M_{\beta\sigma} - T_{\beta\sigma}M_{\alpha\rho} + T_{\beta\rho}M_{\alpha\sigma})\\ &= -\tfrac{1}{4}\big(T_{\alpha\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}M_{\beta\rho} - T_{\alpha\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\rho\sigma}M_{\beta\rho}\\ &\qquad - T_{\beta\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}M_{\alpha\rho} + T_{\beta\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\rho\sigma}M_{\alpha\rho}\big)\,,\end{aligned} \tag{3.18}$$

where we have simply expanded the product, and exchanged (or relabelled) the indices $\sigma \leftrightarrow \rho$ in the second and fourth term; this is possible because all of the indices here are summed over, so they are just dummy variables that can be renamed arbitrarily.

In the next step, we exploit the anti-symmetry $\omega^{\sigma\rho} = -\omega^{\rho\sigma}$ (this is not a relabelling!) in the second and fourth term, which allows us to then group the terms as follows:

$$\begin{aligned}
& -\tfrac{1}{4}\big(T_{\alpha\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}M_{\beta\rho} - T_{\alpha\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\rho\sigma}M_{\beta\rho} \\
&\qquad\qquad - T_{\beta\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}M_{\alpha\rho} + T_{\beta\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\rho\sigma}M_{\alpha\rho}\big)\,, \\
&= -\tfrac{1}{2}\big(T_{\alpha\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}M_{\beta\rho} - T_{\beta\sigma}\tilde{\omega}^{\alpha\beta}\omega^{\sigma\rho}M_{\alpha\rho}\big) \\
&= -\tfrac{1}{2}\big(T_{\gamma\lambda}\tilde{\omega}^{\gamma\nu}\omega^{\lambda\mu}M_{\nu\mu} - T_{\gamma\lambda}\tilde{\omega}^{\mu\gamma}\omega^{\lambda\nu}M_{\mu\nu}\big) \\
&= -\tfrac{1}{2}T_{\gamma\lambda}\big(\tilde{\omega}^{\gamma\nu}\omega^{\mu\lambda}M_{\mu\nu} - \tilde{\omega}^{\mu\gamma}\omega^{\lambda\nu}M_{\mu\nu}\big) = \tfrac{1}{2}T_{\gamma\lambda}\big(\tilde{\omega}^{\mu\gamma}\omega^{\lambda\nu} - \omega^{\mu\lambda}\tilde{\omega}^{\gamma\nu}\big)M_{\mu\nu}\,.
\end{aligned} \tag{3.19}$$

Here, moving from the second to third line, we relabelled $(\alpha \to \gamma, \beta \to \nu, \sigma \to \lambda, \rho \to \mu)$ in the first term, and $(\alpha \to \mu, \beta \to \gamma, \sigma \to \lambda, \rho \to \nu)$ in the second; finally, moving to the last line, we have used in the first term the anti-symmetry of ω and M in their indices simultaneously, and then rearranged slightly. Comparing now to (3.14), we therefore find $T_{\gamma\lambda} = -i\eta_{\gamma\lambda}$, because that equation must hold for arbitrary ω. Going back to the ansatz (3.15) for the commutator, we therefore have

$$[M_{\alpha\beta}, M_{\sigma\rho}] = -i\big(\eta_{\alpha\sigma}M_{\beta\rho} - \eta_{\alpha\rho}M_{\beta\sigma} - \eta_{\beta\sigma}M_{\alpha\rho} + \eta_{\beta\rho}M_{\alpha\sigma}\big)\,. \tag{3.20}$$

Notice how we derived this relation without explicitly finding representations for $M_{\alpha\beta}$ in terms of explicit 4×4 matrices. Instead, as the argumentation above shows, this relation is a direct consequence of the defining property (3.6) of the group, together with the assumption that we can use the exponential map to express any finite group element. Similar to the discussion of the 3d rotation group, we could also have found an explicit form of $M_{\alpha\beta}$ from (3.10). We leave it as an exercise to derive that

$$(M_{\alpha\beta})^{\mu}{}_{\nu} = i(\delta^{\mu}_{\alpha}\eta_{\beta\nu} - \delta^{\mu}_{\beta}\eta_{\alpha\nu})\,. \tag{3.21}$$

Notice that we have put the matrix indices in a way that it is convenient for computation matrix multiplications, e.g.,

$$[M_{\alpha\beta}, M_{\sigma\rho}]^{\mu}{}_{\nu} = (M_{\alpha\beta})^{\mu}{}_{\gamma}(M_{\sigma\rho})^{\gamma}{}_{\nu} - (M_{\sigma\rho})^{\mu}{}_{\gamma}(M_{\alpha\beta})^{\gamma}{}_{\nu}\,. \tag{3.22}$$

3.3 Structure of the Lorentz algebra

To gain some intuition about the structure of this algebra, we can first restrict the indices $M_{\alpha\beta}$ to spatial indices. Conventionally, this is indicated by using Roman $(i, j, k... \in \{1, 2, 3\})$ instead of Greek letters. Because $\eta_{ij} = \delta_{ij}$, we then have

$$[M_{ij}, M_{kl}] = -i(\delta_{ik}M_{jl} - \delta_{il}M_{jk} - \delta_{jk}M_{il} + \delta_{jl}M_{ij})\,. \qquad (3.23)$$

This gives three independent generators $J_m = -\frac{1}{2}\epsilon_{mij}M_{ij}$, or equivalently, $M_{ij} = -\epsilon_{ijk}J_k$, for which one can verify that (now using standard Einstein summation convention)

$$[J_m, J_n] = i\epsilon_{mnk}J_k\,, \qquad (3.24)$$

which is precisely the commutation relation (3.5) of $\mathfrak{so}(3)$. This is to be expected, since for Lorentz transformations of the form $\Lambda = \left(\begin{smallmatrix}1 & 0\\ 0 & R\end{smallmatrix}\right) \in SO(1,3)^+$, with R a 3×3 matrix, the defining property (3.6) reduces to the defining property $R^T R = \mathbb{1}$ of $SO(3)$. That is, the proper orthochronous Lorentz group contains the 3d rotations as a subgroup, and consequently the Lie algebras must reflect this. From the explicit representation of the generators (3.21), we can see that also M_{ij} are block diagonal matrices of the form $M \sim \left(\begin{smallmatrix}1 & 0\\ 0 & N\end{smallmatrix}\right)$ with N a 3×3 matrix, so the exponentiation retains this form and produces the expected rotation matrices inside $SO(1,3)^+$.

There are three further generators $K_i = M_{0i}$ associated to the components with one temporal index. For these, we find

$$\begin{aligned}
[K_i, K_j] &= -i(\eta_{00}M_{ij} + \delta_{ij}M_{00}) = -iM_{ij} = -i\epsilon_{ijk}J_k\,,\\
[J_i, K_j] &= -\tfrac{1}{2}\epsilon_{imn}[M_{mn}, M_{0j}] = -\tfrac{i}{2}\epsilon_{imn}(-\delta_{mj}M_{n0} + \delta_{nj}M_{m0}) = i\epsilon_{ijn}K_n\,.
\end{aligned} \qquad (3.25)$$

The K_i generate so-called boosts in the i-th direction, i.e., we change into a reference frame that is moving along the i-th axis with a constant velocity. More precisely, it can be shown that, if the velocity vector is $\vec{v} = c\tanh(r)\vec{u}$ for some unit vector $\vec{u}$ and finite $r \in \mathbb{R}$ (r is also called *rapidity*), then the boost transformation is

$$\Lambda = \begin{pmatrix}\cosh(r) & \sinh(r)\,\vec{u}^T\\ \sinh(r)\,\vec{u} & \mathbb{1}_3 + (\cosh(r)-1)\vec{u}\,\vec{u}^T\end{pmatrix}\,. \qquad (3.26)$$

A general proper orthochronous Lorentz transformation can then be characterized by a 3d rotation $(\theta, \vec{n})$ and a boost $(r, \vec{u}) \cong \vec{v}$, and can be explicitly constructed using the exponential map,

$$\Lambda(\theta, \vec{n}; r, \vec{u}) = \exp\left(-i\theta\vec{n} \cdot \vec{J} - i\vec{v} \cdot \vec{K}\right) = \exp\left(-i\theta\vec{n} \cdot \vec{J} - ic\tanh(r)\vec{u} \cdot \vec{K}\right) . \quad (3.27)$$

4 Representations of groups and algebras

Even though groups are mathematically defined as objects in their own right, their appearance in physics is through their action on other objects. This is especially familiar for matrix groups, where $n \times n$ matrices naturally act on n-component vectors. However, there can be many other ways a group can act on objects; for example, a matrix group that is defined via $n \times n$ matrices can also act on vectors of other sizes. As we will see, this is closely related to the phenomenon of half-integer spin, which can not be described by rotations on 3-component vectors. These different ways of acting on objects is known as *representations* of a group.

More precisely, given a group G, a *linear representation*, or representation in short, is a group homomorphism $\rho : G \to \text{Aut}(V)$ between G and the group of automorphisms (the set of invertible linear maps) on a vector space V, i.e.,

$$\forall g, h \in G : \ \rho(e) = \text{id}_V , \quad \rho(g \circ h) = \rho(g)\rho(h) , \quad \rho(g^{-1}) = \rho(g)^{-1} . \quad (4.1)$$

For finite dimensional vector spaces, $\dim(V) = n < \infty$, $\text{Aut}(V) \cong GL(n; K)$ after picking a basis for V, so a representation literally represents elements of a group in terms of certain $n \times n$ square matrices, with the group law represented by the matrix group law. For all relevant cases, the vector space is over $\mathbb{R}$ or $\mathbb{C}$. The *dimension* of a the representation (ρ, V) is the dimension of V. Unless otherwise stated, we will only discuss finite dimensional representations

Given a representation (ρ, V), the group acts on the vectors of V as linear transformations,

$$V \ni v \overset{g \in G}{\mapsto} \rho(g)v . \quad (4.2)$$

A representation is called *reducible* if there is a non-zero subspace $\{0\} \neq U \subsetneq V$ such that

$$\forall u \in U : \ \ \rho(g)u \in U . \quad (4.3)$$

If no such subspace exist, ρ is called an *irreducible representation*, often abbreviated as "irrep". Given a reducible representation ρ, we can always find a basis for V such that $\rho(g) = \begin{pmatrix} \hat\rho(g) & \beta(g) \\ 0 & \rho'(g) \end{pmatrix}$ is a block triangle matrix, where the invariant subspace is $U = \{\binom{u}{0} \in V\}$. This gives rise to a representation $\hat\rho$ of smaller dimension. A reducible representation is called *completely* reducible if $\beta(g) = 0$. In this case, ρ decomposes into the direct sum of two representations, $\rho \cong \hat\rho \oplus \rho'$.

Two representation of the same dimension n, ρ_1 and ρ_2, are called *equivalent*, if there exists an invertible $(n \times n)$ matrix S, such that

$$\forall g \in G : \quad \rho_2(g) = S^{-1}\rho_1(g)S \,, \tag{4.4}$$

i.e., if there exists a basis change S on V that relate the representations. A representation ρ is called *faithful*, if

$$g_1 \neq g_2 \quad \Rightarrow \rho(g_1) \neq \rho(g_2) \,. \tag{4.5}$$

For a non-faithful representation ρ, there exists a subset $H \subset G$ for which $\rho(h) = \mathbb{1}$ for $h \in H$. It is easy to see that H is a subgroup.

An important class of representations for physics are *unitary* representations. These are complex representations, $\rho : G \to GL(n;\mathbb{C})$, such that $\rho(g)$ is a unitary matrix, i.e.,

$$\rho(g^{-1}) = \rho(g)^{-1} \stackrel{!}{=} \rho(g)^\dagger \,. \tag{4.6}$$

The intuition for their relevance is that, traditionally, symmetries should act on the Hilbert space while they preserve the norm,

$$\|\rho(g)|\psi\rangle\|^2 = \langle\psi|\rho(g)^\dagger\rho(g)|\psi\rangle \stackrel{!}{=} \langle\psi|\psi\rangle \quad\Rightarrow\quad \rho(g)^\dagger\rho(g) = \mathrm{id} \,. \tag{4.7}$$

Examples

For any group, there always exists the trivial, or *singlet* representation, $\rho(g) = 1$, which is clearly one-dimensional. For any matrix group G, there is another non-trivial one-dimensional representation, the *determinant* representation, $\rho(g) = \det(g)$.

As another example, consider $G = \mathbb{Z}_n$. This group always has $n-1$ inequivalent complex one-dimensional representations labelled by $k \in \{0, ..., n-1\}$,

$$\rho_k : \quad \mathbb{Z}_n \to GL(1,\mathbb{C}) \cong \mathbb{C}^\times = \mathbb{C} \setminus \{0\} \,, \quad 1 \mapsto \exp(\tfrac{2\pi i k}{n}) \,. \tag{4.8}$$

However, only those for which n is not divisble by k are faithful (why?). There are also higher-dimensional complex representations; however, these are always completely reducible. In fact, the irreps over the complex numbers of abelian group are all one-dimensional. As an example, consider the three dimensional representation for $\mathbb{Z}_3$,

$$\rho^{(3)}(0) = \begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix}, \quad \rho^{(3)}(1) = \begin{pmatrix} 0&0&1\\1&0&0\\0&1&0 \end{pmatrix}, \quad \rho^{(3)}(2) = \begin{pmatrix} 0&1&0\\0&0&1\\1&0&0 \end{pmatrix}. \tag{4.9}$$

We can verify that there are three one-dimensional invariant subspaces,

$$V_0 = \left\{\lambda \begin{pmatrix} 1\\1\\1 \end{pmatrix}\right\}, \quad V_1 = \left\{\lambda \begin{pmatrix} \xi^2\\ \xi\\ 1 \end{pmatrix}\right\}, \quad V_2 = \left\{\lambda \begin{pmatrix} \xi\\ \xi^2\\ 1 \end{pmatrix}\right\}, \tag{4.10}$$

where $\xi = \exp(\frac{2\pi i}{3})$. As an exercise, show that $\rho^{(3)}\big|_{V_k} \cong \rho_k$, and, hence, the three-dimensional representation is completely reducible, $\rho^{(3)} = \rho_0 \oplus \rho_1 \oplus \rho_2$.

For the permutation group S_3, it turns out that there is one irreducible representation $\rho^{(2)}$ over the complex numbers of dimension two, the so-called standard representation:

g	e	A	B	C	D	E
$\rho^{(2)}(g)$	$\begin{pmatrix} 1&0\\0&1 \end{pmatrix}$	$\begin{pmatrix} \xi&0\\0&\xi^2 \end{pmatrix}$	$\begin{pmatrix} \xi^2&0\\0&\xi \end{pmatrix}$	$\begin{pmatrix} 0&1\\1&0 \end{pmatrix}$	$\begin{pmatrix} 0&\xi\\ \xi^2&0 \end{pmatrix}$	$\begin{pmatrix} 0&\xi^2\\ \xi&0 \end{pmatrix}$

(4.11)

This is a faithful unitary representation. Moreover, there are two one-dimension irreps, the trivial irrep ρ_0, and the so-called alternating irrep ρ_A, which are unitary but not faithful. The alternating irrep is the sign of $g \in S_3$ as a permutation, which can also be conveniently expressed as the determinant representation associated to $\rho^{(2)}$, $\rho_A(g) = \det(\rho^{(2)}(g)) = \pm 1$. Notice how the $\mathbb{Z}_3$ subgroup $\{e, A, B\}$ has positive determinant, and the $\mathbb{Z}_2$ generators have negative determinant.

Another identity showing the depth of representation theory for finite groups is the relationship between the dimension of complex irreps and the order of the group,

$$|G| = \sum_{\rho \text{ inequiv. irrep}} \dim(\rho)^2, \tag{4.12}$$

which is obviously satisfied in the above examples.

4.1 Representation of Lie groups and Lie algebras

Let us now turn to matrix groups as examples of Lie groups. These have a natural, *defining representation*, given by $\rho(g) = g$. The defining representation of a matrix group of $n \times n$ matrices has dimension n. We will see later in examples of representations with different dimensions.

Given a Lie algebra — a vector space $\mathfrak{g}$ with an anti-symmetric product $[\cdot, \cdot]$ — a representation of $\mathfrak{g}$ is a vector space V with an algebra-homomorphism

$$\rho_{\mathfrak{g}} : \quad \mathfrak{g} \to \mathrm{End}(V) \tag{4.13}$$

into the set of endomorphisms on V (linear maps from V to V, not necessarily invertible). $\mathrm{End}(V)$ carries a natural vector space structure in terms of adding linear maps, and a product (not necessarily invertible) in terms of composition. Given a vector space basis of V, $\mathrm{End}(V)$ just becomes (not necessarily invertible) matrices, with multiplication given by matrix multiplication. Compatibility with the Lie algebra structure requires that

$$\rho_{\mathfrak{g}}([X, Y]) = \rho_{\mathfrak{g}}(X)\rho_{\mathfrak{g}}(Y) - \rho_{\mathfrak{g}}(Y)\rho_{\mathfrak{g}}(X)\,, \tag{4.14}$$

with the product on the right-hand side being matrix multiplication. Given a set of generators $\{T_i\}$ with structure constants f_{ijk}, we have $\rho_{\mathfrak{g}}([T_i, T_j]) = f_{ijk}\rho_{\mathfrak{g}}(T_k)$.

Any representation (ρ, V) of a Lie group G also induces a representation of its Lie algebra $\mathfrak{g}$. If $G \ni g = \exp(tX)$ for $t \in \mathbb{R}$ and $X \in \mathfrak{g}$, then $\rho(\exp(tX))$ defines a "path" of transformations on the representation space V. We can then define a representation of $\mathfrak{g}$ on the same space V, via

$$\forall v \in V : \quad \rho_{\mathfrak{g}}(X)(v) := [\tfrac{d}{dt}\rho(\exp(tX))(v)]\big|_{t=0}\,. \tag{4.15}$$

Hence, $\rho_{\mathfrak{g}}(X)$ is a matrix of the same size as $\rho(g)$, so they also act on V. One can show explicitly, using the BCH formulat, that $\rho_{\mathfrak{g}}$ respects the bracket structure. Thus, $(\rho_{\mathfrak{g}}(X), V)$ is a representation of the Lie algebra.

Given a representation (ρ, V) of a Lie group G over $\mathbb{C}$, the Lie algebra representations are a priori also some complex matrices. For a unitary representation, the Lie algebra representation must be given by "skew-Hermitian" matrices, i.e.,

$[\rho_{\mathfrak{g}}(X)]^\dagger = -\rho_{\mathfrak{g}}(X)$. Heuristically, this can be seen from (4.15) as follows. Because of unitarity, we have $\rho(\exp(tX))^\dagger\rho(\exp(tX)) = 1$, so

$$\begin{aligned} 0 &= \left(\tfrac{d}{dt}\left[\rho(\exp(tX))^\dagger\rho(\exp(tX))\right]\right)\Big|_{t=0} \\ &= \left[(\rho_{\mathfrak{g}}(X)^\dagger + \rho_{\mathfrak{g}}(X))\rho(\exp(tX))^\dagger\rho(\exp(tX))\right]\Big|_{t=0} = \rho_{\mathfrak{g}}(X)^\dagger + \rho_{\mathfrak{g}}(X)\,. \end{aligned} \tag{4.16}$$

In physics, it is then common to instead consider the related representation

$$\tilde{\rho}_{\mathfrak{g}}(X) = i\rho_{\mathfrak{g}}(X)\,, \tag{4.17}$$

which are now Hermitian matrices. Given a basis $\{T_a\}$ of $\mathfrak{g}$, the commutator is modified accordingly:

$$[\tilde{\rho}_{\mathfrak{g}}(T_a), \tilde{\rho}_{\mathfrak{g}}(T_b)] = -[\rho_{\mathfrak{g}}(T_a), \rho_{\mathfrak{g}}(T_b)] = -\sum_c f_{abc}\rho_{\mathfrak{g}}(T_c) = i\sum_c f_{abc}\tilde{\rho}_{\mathfrak{g}}(T_c)\,. \tag{4.18}$$

This is precisely the reason why physicists tend to use the basis $J_k = iL_k$ for the Lie algebra $\mathfrak{so}(3)$: they already have in mind a unitary representation of the group (for physical reasons we will explain momentarily), and it is (marginally) more convenient to work with Hermitian rather than skew-Hermitian matrices.

While any representation of a Lie group descends to a representation of its algebra, the converse is not true in general: given a Lie group G with algebra $\mathfrak{g}$, not all representations $\rho_{\mathfrak{g}}$ of the algebra necessarily extend to representations of G. The obstruction here is related to the *global topology* of G as a manifold, and is absent, i.e., all Lie algebra representations are also group representations, if G is *simply-connected* ("all closed paths in G are contractable"). If G is not simply-connected, then there is always another Lie group $\tilde{G}$ with *the same* Lie algebra $\mathfrak{g}$ which *is* simply-connected. $\tilde{G}$ is called the *universal cover* of G; "cover" means that there exists a projection map $\pi : \tilde{G} \to G$ which is a surjective group homomorphism, and "universal" refers to the fact that its simply-connected. Therefore, given a Lie group G with Lie algebra $\mathfrak{g}$, the representations of $\mathfrak{g}$ are in 1-to-1 correspondence to representations of the universal cover $\tilde{G}$. Another key property between $\tilde{G}$ and G is that their Lie algebras agree, $\tilde{\mathfrak{g}} = \mathfrak{g}$. This can be understood better geometrically: For any point $g \in \tilde{G}$, there is always an open neighborhood $g \in U \subset G$ such that $\pi|U : U \to \pi(U)$ is a diffeomorphism (smooth and invertible map). Because of this, the tangent spaces at $g \in \tilde{G}$ is isomorphic to $\pi(g) \in G$, which in particular applies $g = e$.

It turns out that neither the rotation group $SO(3)$ nor the Lorentz group $SO(1,3)^+$ are simply-connected; their universal covers are $SU(2)$ and $SL(2;\mathbb{C})$, respectively. This means that representations of their Lie algebras do not all lift to representations of the group. However, they are still relevant physically thanks to quantum mechanics.

4.2 Representations in Quantum Mechanics

To highlight the "specialness" of quantum mechanics, let us briefly recap the classical setting. There, we expect symmetries to act on the system given by points in a phase space (and functions on it, or "observables") in certain linear representations. Since the rotation group $SO(3)$ is non-simply-connected, there are some representations (to be discussed soon) of the Lie algebra $\mathfrak{so}(3)$ which are not valid representations of the group. These are precisely those representations which "do not close after a 2π rotation", i.e., half-integer spin representations. Indeed, we do not encounter such objects (which are the corresponding representation vectors) in the classical regime of nature.

However, things change drastically in the quantum world, due to a simple fact that is baked into the axioms of quantum mechanics. Namely, the state of the system is now characterized by a vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ *up to a non-zero scalar multiple*! That is, we consider $|\psi\rangle$ and $\lambda|\psi\rangle$ for some $\lambda \in \mathbb{C}^\times$ to be physically equivalent. This means that when a symmetry group $G \ni g$ now acts on $\mathcal{H}$ as some unitary operator $\rho(g) : \mathcal{H} \to \mathcal{H}$, it only needs to respect the group structure *up to a phase* (because the product must still be unitary, it cannot be an arbitrary scalar multiple),

$$\rho(g)\rho(h) = e^{i\varphi(g,h)}\rho(g \circ h)\,, \quad \varphi(g,h) \in \mathbb{R}\,. \tag{4.19}$$

A representation ρ that satisfies (4.19) is called a *projective* representation, as opposed to *linear* representations that we have discussed above. Thus, for a symmetry group G of a quantum system, the relevant representations are unitary projective representations of G.

Now we can make use of a powerful mathematical statement that *all* (finite dimensional) unitary projective representations of G arises from a unitary *linear* representation of the universal covering group $\tilde{G}$. These, in turn, come from representations of the Lie algebra. In the $SO(3)$ example, this means that there are half-integer spin representations allowed in quantum mechanics, and these can be thought of as linear representations of $SU(2)$.

In the following, we shall see how some of the above abstract statements work more explicitly in the concrete cases of $G = SO(3)$ and $G = SO(1,3)^+$.

5 Representations of $SO(3)$ and $SU(2)$

The construction of $SO(3)$ representations or, rather, of $\mathfrak{so}(3) = \mathfrak{su}(2)$ representations should be familiar from the quantum mechanical treatment of the hydrogen atom. In the following, we will review their construction, and then discuss their lift to representations of the group and its universal cover $SU(2)$. Borrowing the quantum mechanics jargon, we will sometimes call the representation of a group or algebra element an "operator".

5.1 Representations of the Lie algebra

With the physics motivation, we are interested in unitary representations of the groups, so we want representations of the Lie algebra

$$[L_i, L_j] = \epsilon_{ijk} L_k \tag{5.1}$$

in terms of skew-Hermitian matrices, or, equivalently, representations $\rho_{\mathfrak{g}}(L_i) \equiv J_i$ as Hermitian matrices that satisfy

$$[J_i, J_j] = i\epsilon_{ijk} J_k\,, \tag{5.2}$$

These act a vectors V which is a Hilbert space, i.e., it carries a (complex) scalar product $B(\cdot,\cdot)$ (see footnote 5). Adopting the physics notation, we will denote vectors of V as $|v\rangle$, where v is some label that characterizes the vector in a useful way. The scalar product is then written as $B(|u\rangle, |v\rangle) \equiv \langle u|v\rangle$.

Now, to study the possible finite dimensional representation spaces V, the first step is to define the "ladder" operators

$$J_\pm := J_1 \pm iJ_2 \quad \text{with} \quad [J_3, J_\pm] = \pm J_\pm\,, \quad [J_+, J_-] = 2J_3\,. \tag{5.3}$$

Because J_i were assumed to be Hermitian, we have $J_\pm^\dagger = J_\mp$. Furthermore, we define the operator $J^2 := J_1^2 + J_2^2 + J_3^2$, which is Hermitian. One can easily show that

$[J^2, J_i] = [J^2, J_\pm] = 0$. Mathematically, this operator an example of a *quadratic Casimir invariant* of the Lie algebra, which commutes with Lie algebra generators; in physics, it appears as the "total angular momentum operator".

Because J_3 and J^2 are Hermitian matrices on a Hilbert space V which commute, the spectral theorem tells us that there is a orthonormal basis $\{|\zeta, m\rangle\}$ of V, i.e., $\langle \zeta, m|\zeta', m'\rangle = \delta_{mm'}\delta_{\zeta\zeta'}$, which are simultaneous eigenvectors of J_3 and J^2. The labels m and ζ are chosen to be the eigenvalues of that vector: $J_3|\zeta, m\rangle = m|\zeta, m\rangle$, and $J^2|\zeta, m\rangle = \zeta|\zeta, m\rangle$. A priori, there could be degeneracies for these labels, i.e., more than one linearly independent basis elements with the same eigenvalues (ζ, m). However, as it turns out, the action of $J_\pm$ only depends on these values, so different basis elements with the same labels would correspond to invariant one-dimensional subspaces, i.e., we could restrict to one of these and obtain a valid representation. In the following, we will therefore assume without loss of generality that the eigenspace for each pair (ζ, m) is one-dimensional.

Since J^2 commutes with $J_\pm$, we have $J^2 J_\pm|\zeta, m\rangle = \zeta J_\pm|\zeta, m\rangle$. In addition, we also have

$$J_3 J_\pm|\zeta, m\rangle = ([J_3, J_\pm] + J_\pm J_3)|\zeta, m\rangle = (\pm J_\pm + m J_\pm)|\zeta, m\rangle = (m \pm 1) J_\pm|\zeta, m\rangle\,. \tag{5.4}$$

This means

$$J_\pm|\zeta, m\rangle = \lambda^\pm_{\zeta,m}|\zeta, m \pm 1\rangle\,. \tag{5.5}$$

Now we will proceed to determine the possible values of (ζ, m, λ). For that, we use $J^2 - J_3^2 = J_1^2 + J_2^2 = \frac{1}{2}(J_+J_- + J_-J_+)$ to show

$$\begin{aligned}\langle\psi|J^2 - J_3^2|\psi\rangle &= \tfrac{1}{2}\langle\psi|J_+J_- + J_-J_+|\psi\rangle = \tfrac{1}{2}\langle\psi|J_-^\dagger J_- + J_+^\dagger J_+|\psi\rangle \\ &= \tfrac{1}{2}(\|J_-|\psi\rangle\|^2 + \|J_+|\psi\rangle\|^2) \geq 0\,,\end{aligned} \tag{5.6}$$

for any $|\psi\rangle \in V$. So this implies

$$\langle\zeta, m|J^2 - J_3^2|\zeta, m\rangle = (\zeta - m^2)\langle\zeta, m|\zeta, m\rangle \overset{!}{\geq} 0\,. \tag{5.7}$$

Since J_+ increases the value of m for fixed ζ, there must be a maximal value, $m_{\max}(\zeta) \equiv m_{\max}$ such that $|\zeta, m_{\max}\rangle \neq 0$, but $J_+|\zeta, m_{\max}\rangle = 0$. In particular, this means that

$$0 = J_-J_+|\zeta, m_{\max}\rangle = (J^2 - J_3^2 - J_3)|\zeta, m_{\max}\rangle = (\zeta - m_{\max}^2 - m_{\max})|\zeta, m_{\max}\rangle\,. \tag{5.8}$$

Likewise, because J_- decreases the value of m, and thus also increasing m^2 once m becomes negative, there is also a $|\zeta, m_{\rm min}\rangle \neq 0$ such that $J_-|\zeta, m_{\rm min}\rangle = 0$. Analogously, we then get

$$0 = J_+J_-|\zeta, m_{\rm min}\rangle = (J^2 - J_3^2 + J_3)|\zeta, m_{\rm min}\rangle = (\zeta - m_{\rm min}^2 + m_{\rm min})|\zeta, m_{\rm min}\rangle \,. \tag{5.9}$$

All together, this means

$$\zeta = m_{\rm max}(m_{\rm max} + 1) = m_{\rm min}(m_{\rm min} - 1)\,, \tag{5.10}$$

which has solutions $m_{\rm max} = m_{\rm min} - 1 < m_{\rm min}$ and $m_{\rm max} = -m_{\rm min}$. The first is clearly not compatible with the assumption that $m_{\rm max} \geq m_{\rm min}$, so the second must hold. Moreover, since we can increase the value of m in steps of 1 by acting with J_+, $m_{\rm max} - m_{\rm min} = 2m_{\rm max}$ must be an integer. The common notation in the literature is to define $j = m_{\rm max}$, which is in general a non-negative integer or half-integer that is otherwise not constrained.

In summary, any representation space V of $\mathfrak{so}(3)$ has an orthonormal basis $|j, m\rangle \equiv |\zeta, m\rangle$ (we now replace the label $\zeta = j(j+1)$ by j, with $j \in \frac{\mathbb{N}_0}{2}$ and $m \in \{-j, -j+1, ..., j-1, j\}$) characterized by

$$J^2|j, m\rangle = j(j+1)|j, m\rangle\,, \quad J_3|j, m\rangle = m|j, m\rangle\,. \tag{5.11}$$

If $j \in \mathbb{N}_0$, then $m \in \mathbb{Z}$, and if $j \in \mathbb{N} + \frac{1}{2}$, then $m \in \mathbb{Z} + \frac{1}{2}$. Finally, there were the coefficients $\lambda^\pm_{\zeta,m} \equiv \lambda^\pm_{j,m}$ in (5.5) that remains to be determined. We leave it as an exercise to derive the result which is

$$J_\pm|j, m\rangle = \sqrt{(j \mp m)(j \pm m + 1)}|j, m \pm 1\rangle\,. \tag{5.12}$$

Notice that the value of j is not changed by the action of the Lie algebra generators. It means that different values of j correspond to different representations, which have dimensions $2j + 1$.[8] Given the explicit action of the Lie algebra generators on the basis elements, one can explicitly verify that these representations are irreducible. The common nomenclature is to call an irrep labelled by j a “spin-j” irrep of $\mathfrak{so}(3)$. For $j = 1$, we obtain the 3d representation space $V_{j=1}$, which is the Lie algebra representation induced by the defining representation of $SO(3)$. The integer spin

[8] This is an incarnation of a more fundamental fact for Lie algebras: the irreducible representations are characterized by eigenvalues of Casimir elements.

$j > 1$ representations can be shown to correspond to (irreps inside the) tensor products $V_1 \otimes \ldots \otimes V_1 = V_1^{\otimes j}$, which therefore are naturally $SO(3)$ representations. But the half-integer spin representations are not $SO(3)$ representations, as we now explain.

5.2 Why half-integer spin irreps are not $SO(3)$ irreps

Labelling the elements of the basis $\{|j, m\rangle\}$ by the value m, we can explicitly write down the representation matrices $M = \rho_j(X)$, whose rows and columns are also labelled by indices with the same range as m. Here and in the following, ρ_j denotes commonly the representation on the algebra and the (covering) group. Their entries of $M = \rho_j(X)$ for any $X \in \mathfrak{so}(3)$ is then

$$M_{ab} = \langle j, a|M|j, b\rangle \,, \tag{5.13}$$

and its action on the basis elements is

$$M|j, m\rangle = \sum_k |j, k\rangle M_{km} \,. \tag{5.14}$$

By exponentiating this matrix (with a factor of i because we are using a Hermitian representation of the generators), we get a representation of an element that in general is in the universal covering group.

Now recall that the element $R = \exp(-i\theta\vec{n} \cdot \vec{J})$ represents a rotation around $\vec{n}$ by an angle θ. Specifically, we can consider rotations around the y-axis, $\vec{n} = (0, 1, 0)^T$, whose matrix elements are

$$\rho_j[R(\theta)]_{ab} = \langle j, a| \exp(-i\theta J_2^{(j)})|j, b\rangle = \langle j, a| \exp(-\tfrac{1}{2}\theta(J_+^{(j)} - J_-^{(j)}))|j, b\rangle \,, \tag{5.15}$$

where $J_k^{(j)} \equiv \rho_j(L_k)$ now explicitly shows the representation we are in for the Lie algebra generators L_k. This exponential can be computed explicitly by expanding the power series, because large enough powers of J_- and J_+ annihilate $|j, b\rangle$, thus truncating the series to a finite sum. We simply quote the final result for $\theta = \pi$ (recall that in the above parametrization, $\theta \in [0, \pi]$),

$$\rho_j[R(\pi)]_{ab} = (-1)^{j-b}\delta_{a,-b} \,. \tag{5.16}$$

A full 2π rotation is then given by

$$\begin{aligned}
&\rho_j[R(\pi)R(\pi)]_{ab} = (\rho_j[R(\pi)]\rho_j[R(\pi)])_{ab} \\
&= \sum_c \rho_j[R(\pi)]_{ac}\rho_j[R(\pi)]_{cb} = \sum_c (-1)^{j-c}\delta_{a,-c}(-1)^{j-b}\delta_{c,-b} \\
&= \sum_c (-1)^{2j-c-b}\delta_{a,-c}\delta_{c,-b} = (-1)^{2j}\delta_{ab}\,.
\end{aligned} \tag{5.17}$$

We therefore see that for half-integer spin representations, a rotation by 2π is not quite the unit matrix. Therefore, it does not respect the $SO(3)$ group law, $R(\pi)R(\pi) = \mathrm{id}$, at least not linearly. Instead, since the difference to the unit matrix is precisely a phase (norm 1 prefactor), this is a projective representation of $SO(3)$.

5.3 Relationship between $SO(3)$ and $SU(2)$

Let us take a closer look at the spin $\frac{1}{2}$ representation. This is a two-dimensional representation with basis vectors $|e_1\rangle = |\frac{1}{2},\frac{1}{2}\rangle$ and $|e_2\rangle = |\frac{1}{2},-\frac{1}{2}\rangle$. From $J_3^{(1/2)}|\frac{1}{2},\pm\frac{1}{2}\rangle = \pm\frac{1}{2}|\frac{1}{2},\pm\frac{1}{2}\rangle$, we have

$$\begin{aligned}
&\langle e_1|J_3^{(1/2)}|e_1\rangle = -\langle e_2|J_3^{(1/2)}|e_2\rangle = -\tfrac{1}{2}\,, \quad \langle e_1|J_3^{(1/2)}|e_2\rangle = \langle e_2|J_3^{(1/2)}|e_1\rangle = 0\,, \\
&\implies \qquad J_3^{(1/2)} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}.
\end{aligned} \tag{5.18}$$

The ladder operators act on the basis as

$$\begin{aligned}
&J_-^{(1/2)}|e_1\rangle = |e_2\rangle\,, \quad J_-^{(1/2)}|e_2\rangle = 0 = J_+^{(1/2)}|e_1\rangle\,, \quad J_+^{(1/2)}|e_2\rangle = |e_1\rangle\,, \\
&\implies \quad J_-^{(1/2)} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad J_+^{(1/2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.
\end{aligned} \tag{5.19}$$

Combining these results with (5.3), we have

$$J_1^{(1/2)} = \tfrac{1}{2}(J_+^{(1/2)} + J_-^{(1/2)}) = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \quad J_2^{(1/2)} = \tfrac{1}{2i}(J_+^{(1/2)} - J_-^{(1/2)}) = \begin{pmatrix} 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix}. \tag{5.20}$$

Notice that $J_i^{(1/2)} = \frac{1}{2}\sigma_i$, with σ_i the i-th Pauli-matrix,

$$\sigma_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \quad \sigma_2 = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right), \quad \sigma_3 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right). \tag{5.21}$$

These satisfy $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$, which confirms that the $J_i^{(1/2)}$ matrices satisfies the $\mathfrak{so}(3)$ Lie algebra relation (3.5).

These three matrices are linearly independent over $\mathbb{R}$, and generate the full set of 2×2 Hermitian traceless matrices. To see this, simply notice that Hermiticity of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ requires $b = \overline{c}$, and a and d to be real; tracelessness then sets $a = -d$. In total, this leaves one free real parameter on the diagonal, and two free real parameters in b. Thus, the associated group elements, $U(\theta, \vec{n}) = \exp(-i\theta\vec{n} \cdot \vec{J}^{(1/2)}) = \exp(-i\frac{\theta}{2}\vec{n} \cdot \vec{\sigma})$, with $\vec{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3)$, are unitary and have determinant 1. These are precisely the matrices of $SU(2) = \{U \in \mathbb{C}^{2\times 2} \mid U^\dagger U = \mathbb{1}\,,\ \det(U) = 1\}$. For completeness, the explicit expression for these matrices is

$$U(\theta, \vec{n}) = \begin{pmatrix} \cos\frac{\theta}{2} - in_3 \sin\frac{\theta}{2} & -i(n_1 - in_2)\sin\frac{\theta}{2} \\ -i(n_1 + in_2)\sin\frac{\theta}{2} & \cos\frac{\theta}{2} + in_3 \sin\frac{\theta}{2} \end{pmatrix} . \tag{5.22}$$

Alternatively, we can also parametrize an $SU(2)$ matrix as

$$\begin{aligned} &U(a,b) = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} , \\ &\text{with } a, b \in \mathbb{C} \cong \mathbb{R}^2 \text{ such that } |a|^2 + |b|^2 = [\det U(a,b)] = 1 . \end{aligned} \tag{5.23}$$

This parametrization shows that $SU(2) \cong S^3 = \{\vec{v} \in \mathbb{R}^4 \mid ||\vec{v}||^2 = 1\}$ as a manifold. The action on the two-dimensional vector space spanned by $|\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$ is the defining, or fundamental representation, of $SU(2)$. As we have explained above, it is a projective representation of $SO(3)$. In the literature, it is typically referred to as the spinor representation of $SO(3)$.

The Pauli matrices allow us to further understand the relationship between $SU(2)$ and $SO(3)$. First, we can identify $\mathfrak{so}(3) = \mathfrak{su}(2)$ with $\mathbb{R}^3$, by

$$X = \sum_{i=1}^{3} x_i \sigma_i \mapsto \vec{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3 , \tag{5.24}$$

with the Euclidean scalar product given by

$$\vec{x} \cdot \vec{y} = \tfrac{1}{2}\mathrm{tr}(XY) . \tag{5.25}$$

Given an element $X \in \mathfrak{su}(2)$, which is represented by a traceless Hermitian matrix, we have for any $U \in SU(2)$ that

$$\mathrm{tr}(U\,X\,U^{-1}) = \mathrm{tr}(U^{-1}\,U\,X) = 0 , \quad (U\,X\,U^{-1})^\dagger = (U^{-1})^\dagger\, X^\dagger\, U^\dagger = U\,X\,U^{-1} , \tag{5.26}$$

so $X \mapsto U\,X\,U^{-1}$ defines an action $\mathfrak{su}(2) \cong \mathbb{R}^3 \to \mathbb{R}^3$. This action is orthogonal, because

$$\vec{x} \cdot \vec{y} = \tfrac{1}{2}\mathrm{tr}(XY) \mapsto \tfrac{1}{2}\mathrm{tr}(UXU^{-1}\,UYU^{-1}) = \tfrac{1}{2}\mathrm{tr}(XY) = \vec{x} \cdot \vec{y}\,. \tag{5.27}$$

Furthermore, because $SU(2)$ is connected, and $\mathrm{id} \in SU(2)$ is clearly mapped to $\mathrm{id} \in O(3)$, the orthogonal action induced by the $SU(2)$ elements must be connected to the identity, i.e., it is an $SO(3)$ action. This defines a three-dimensional representation $(\rho_{\mathrm{ad}}, V = \mathfrak{su}(2) \cong \mathbb{R}^3)$,

$$\rho_{\mathrm{ad}} : SU(2) \to SO(3) \subset \mathrm{Aut}(V)\,, \quad V \ni X \overset{U}{\mapsto} \rho_{\mathrm{ad}}(U)(X) := U\,X\,U^{-1}\,. \tag{5.28}$$

One can straightforwardly generalize this kind of action to any other matrix group G (and also to general Lie groups) by "conjugation" (U acting on X in the form UXU^{-1}) on the Lie algebra $\mathfrak{g}$, which defines the *adjoint representation* of G. The hallmark of this representation is that the representation vector space is the Lie algebra (as an abstract vector space) itself. In the case of $G = SU(2)$, the adjoint representation coincides with the spin 1 representation, which in turn is the defining representation of $SO(3)$.[9]

Lastly, we can explicitly compute the action in the parametrization (5.23) on the basis vectors,

$$\begin{aligned}
U(a,b)\sigma_1 U(a,b)^{-1} &= \mathrm{Re}(a^2 - b^2)\,\sigma_1 - \mathrm{Im}(a^2 - b^2)\,\sigma_2 + 2\mathrm{Re}(a\bar{b})\,\sigma_3\,,\\
U(a,b)\sigma_2 U(a,b)^{-1} &= \mathrm{Im}(a^2 + b^2)\,\sigma_1 + \mathrm{Re}(a^2 + b^2)\,\sigma_2 + 2\mathrm{Im}(a\bar{b})\,\sigma_3\,,\\
U(a,b)\sigma_3 U(a,b)^{-1} &= -\mathrm{Re}(ab)\,\sigma_1 + 2\mathrm{Im}(ab)\,\sigma_2 + (|a|^2 - |b|^2)\,\sigma_3\,.
\end{aligned} \tag{5.29}$$

Hence,

$$\rho_{\mathrm{ad}}(U(a,b)) = \begin{pmatrix} \mathrm{Re}(a^2 - b^2) & \mathrm{Im}(a^2 + b^2) & -\mathrm{Re}(ab) \\ -\mathrm{Im}(a^2 - b^2) & \mathrm{Re}(a^2 + b^2) & 2\mathrm{Im}(ab) \\ 2\mathrm{Re}(a\bar{b}) & 2\mathrm{Im}(a\bar{b}) & |a|^2 - |b|^2 \end{pmatrix}. \tag{5.30}$$

[9] The adjoint representation of $SO(3)$ also coincides with the spin 1 representation. In fact, the adjoint representation is the same for all Lie groups that have the same gauge algebra.

From this, we can obtain the $SO(3)$ rotation matrices around each axis,

$$
\begin{aligned}
R(\theta,(1,0,0)) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = \rho_{\mathrm{ad}}(U(\cos\tfrac{\theta}{2}, -i\sin\tfrac{\theta}{2}))\,, \\
R(\theta,(0,1,0)) &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} = \rho_{\mathrm{ad}}(U(\cos\tfrac{\theta}{2}, -\sin\tfrac{\theta}{2}))\,, \\
R(\theta,(0,0,1)) &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \rho_{\mathrm{ad}}(U(e^{-i\theta/2}, 0))\,.
\end{aligned}
\tag{5.31}
$$

Therefore, ρ_{ad} defines a surjective group homomorphism $SU(2) \to SO(3)$. To determine the kernel of this map, supposed $U(a,b)$ is mapped onto the identity in $SO(3)$. This requires $\mathrm{Re}(a\overline{b}) = 0$, which implies either $a = 0$ or $b = 0$. Since we also need $|a|^2 - |b|^2 = 1$, we must have $a \neq 0$, hence $b = 0$. Then, from $\mathrm{Re}(a^2 - b^2) = \mathrm{Re}(a^2) = |a|^2 = 1$, we must have $a = \pm 1$. Indeed, we have $\rho_{\mathrm{ad}}(\mathbb{1}_2) = \rho_{\mathrm{ad}}(-\mathbb{1}_2) = \mathbb{1}_3$. So we have $SO(3) \cong SU(2)/\ker(\rho_{\mathrm{ad}}) = SU(2)/\mathbb{Z}_2$. This establishes the universal cover $SU(2)$ as a *two-fold cover* of $SO(3)$.

This pattern extends to all rotation groups $SO(n \geq 3)$ in higher dimensions.[10] That is, all $SO(n)$ are non-simply-connected, and have a two-fold universal cover, which is called the *spin group* $Spin(n)$. For $n > 6$, $Spin(n)$ is not a conventional matrix group. For low n, we have the "coincidental" isomorphisms,

$$
\begin{aligned}
&Spin(3) \cong SU(2)\,, \quad Spin(4) \cong SU(2) \times SU(2)\,, \\
&Spin(5) \cong Sp(2)\,, \quad Spin(6) \cong SU(6)\,.
\end{aligned}
\tag{5.32}
$$

Each spin group $Spin(n)$ has an irrep analgous to the spin-$\frac{1}{2}$ representation of $SU(2)$, which is then the spinor representation of $SO(n)$. For $n = 2m$, the spinor has dimension 2^{m-1}, and for $n = 2m + 1$, the spinor has dimension 2^m.

6 Representation theory of the Lorentz group

We now move to the (proper orthochronous) Lorentz group $SO(1,3)^+$. There are some salient features that are distinct from the representation theory of $SO(3)$ in the

[10] We also have the abelian cases $SO(1) = \{1\}$, $SO(2) \cong U(1)$.

context of physics, which will be discussed in more detail in the next module of the course. The main difference comes from the fact that $SO(1,3)^+$ is not a *compact* manifold, unlike $SO(3)$. Without giving the explicit definition, the intuition here should be that, unlike for $SO(3)$, where the parameters θ and $\vec{n}$ have a finite range, Lorentz transformations have a parameter, the boost rapidity r in (3.27), with an infinite range.

For non-compact non-abelian Lie groups, any non-trivial unitary representation must be infinite dimensional. Thus the Hilbert space of any Lorentz invariant quantum system must also be infinite dimensional. The story of unitary representations for the Lorentz group, or more appropriately, the Poincaré group, will be explained in next terms Group Theory course.

Instead, we focus on the finite dimension representations. Since these are necessarily not unitary, a question is what their physical relevance is. In essence, in relativistic quantum mechanics, or its inevitable generalization to quantum field theory, space and time coordinates both become parameters on which the operators (the “fields” in quantum field theory) of the theory depend, so the operators will also transform in some Lorentz group representations. Since there is no natural scalar product for quantum operators which needs to be preserved, there is also no need for representations of quantum operators to be unitary.

We can see this in the defining representation of the Lie algebra explicitly,

$$(M_{\alpha\beta})^{\mu}{}_{\nu} = i(\delta^{\mu}_{\alpha}\eta_{\beta\nu} - \delta^{\mu}_{\beta}\eta_{\alpha\nu}) \,, \tag{6.1}$$

which we now interpreted as a representation over the complex numbers. The generators $J_m = -\frac{1}{2}\epsilon_{mij}M_{ij}$ and $K_i = M_{0i}$ have representations

$$\begin{aligned} &J_1 = -M_{23} = \begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&-i\\ 0&0&i&0 \end{pmatrix}, \quad J_2 = -M_{31} = \begin{pmatrix} 0&0&0&0\\ 0&0&0&i\\ 0&0&0&0\\ 0&-i&0&0 \end{pmatrix}, \quad J_3 = -M_{12} = \begin{pmatrix} 0&0&0&0\\ 0&0&-i&0\\ 0&i&0&0\\ 0&0&0&0 \end{pmatrix}, \\ &K_1 = \begin{pmatrix} 0&i&0&0\\ i&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}, \qquad K_2 = \begin{pmatrix} 0&0&i&0\\ 0&0&0&0\\ i&0&0&0\\ 0&0&0&0 \end{pmatrix}, \qquad K_3 = \begin{pmatrix} 0&0&0&i\\ 0&0&0&0\\ 0&0&0&0\\ i&0&0&0 \end{pmatrix}. \end{aligned} \tag{6.2}$$

While the 3d rotation generators J_i are Hermitian, the boost generators K_i are anti-Hermitian, $(K_i)^\dagger = -K_i$; this is related to the fact that the boost directions are the non-compact ones of the Lorentz group.

6.1 Irreps of the Lorentz algebra

Due to the limited scope of this module, we can only give a superficial explanation of how to construct the Lorentz algebra representations. The mathematical details will be discussed in the Group Theory course. Here we simply outline the main idea, and focus on the result.

To construct the Lorentz algebra representations, it is common to use a procedure called *complexification*, which we have already used before implicitly. We start by considering complex linear combinations of the generators of $\mathfrak{g}$, thereby passing to the complexified algebra $\mathfrak{g}_{\mathbb{C}} \equiv \mathfrak{g} \oplus i\mathfrak{g}$.

In the case of the Lorentz algebra, we define

$$A_m := \tfrac{1}{2}(J_m + iK_m)\,, \quad B_m := \tfrac{1}{2}(J_m - iK_m)\,. \tag{6.3}$$

These satisfy the familiar commutator relation

$$[A_m, A_n] = i\epsilon_{mnk}A_k\,, \quad [B_m, B_n] = i\epsilon_{mnk}B_k\,, \quad [A_m, B_n] = 0\,. \tag{6.4}$$

That is, we have $\mathfrak{so}(1,3)_{\mathbb{C}} = \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}}$ (note that the $\mathfrak{su}(2)$ algebras here are also complexified). Note that this implies

$$\begin{aligned} &\mathfrak{so}(1,3)_{\mathbb{C}} \cong \mathfrak{so}(1,3) \oplus i\mathfrak{so}(1,3) = \mathfrak{su}(2) \oplus i\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus i\mathfrak{su}(2) \\ &\implies \quad \mathfrak{so}(1,3) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)\,. \end{aligned} \tag{6.5}$$

The major advantage of this decomposition is that all the representations of the (complexified) $\mathfrak{su}(2)$ algebra are essentially the ones we have constructed in the previous section. The basis $|j, m\rangle$ we constructed there also serve as a complex basis for representation spaces of $\mathfrak{su}(2)_{\mathbb{C}}$; by restricting the general matrices $\exp(\vec{u} \cdot \vec{J}^{(j)})$ with complex parameters $\vec{u}$ to a real three-dimensional subspace, $\vec{u} = i\theta\vec{n}$, $\theta \in \mathbb{R}, \vec{n} \in S^2$, we obtain the unitary representations of $\mathfrak{su}(2)$.

While we cannot obtain any finite dimensional unitary representations for $\mathfrak{so}(1,3)$ this way (because the K_i and J_i, the generators of $\mathfrak{so}(1,3)$, cannot be made Hermitian simultaneously through choosing a real slice), it does give the finite dimensional irreps in terms of direct sums of $\mathfrak{su}(2)$ irreps. Namely, we can now label any representation of $\mathfrak{so}(1,3)$ by a pair (j_1, j_2), with $j_i = 0, \frac{1}{2}, 1, \ldots$.

More precisely, for a direct sum algebra $\mathfrak{g} \oplus \mathfrak{h}$, an irreducible representation (ρ, V) can be built from the tensor product of two irreducible representations $(\rho_{\mathfrak{g}}, V_{\mathfrak{g}})$ and $(\rho_{\mathfrak{h}}, V_{\mathfrak{h}})$:

$$\begin{aligned} &\rho : \mathfrak{g} \oplus \mathfrak{h} \to \mathrm{End}(V) \equiv \mathrm{End}(V_{\mathfrak{g}} \otimes V_{\mathfrak{h}}) \,, \\ &\rho(X+Y)(v \otimes w) := \rho_{\mathfrak{g}}(X)(v) \otimes w + v \otimes \rho_{\mathfrak{h}}(Y)(w) \,. \end{aligned} \tag{6.6}$$

For the Lorentz algebra, this means we will have irreducible representations

$$\begin{aligned} &\rho_{(j_1, j_2)} \left(\sum_m \lambda_m A_m + \sum_n \kappa_n B_n \right) \\ &= \sum_m \lambda_m \, \rho_{j_1}(A_m) \otimes \mathrm{id}_{V_{j_2}} + \sum_n \kappa_n \, \mathrm{id}_{V_{j_1}} \otimes \rho_{j_2}(B_n) \,, \end{aligned} \tag{6.7}$$

where we think of A_m and B_n as independent $\mathfrak{su}(2)$ generators, of which we take the spin j_1/j_2 representations. These have dimensions $\dim(V_{j_1} \otimes V_{j_2}) = \dim(V_{j_1}) \dim(V_{j_2}) = (2j_1 + 1)(2j_2 + 1)$.

We can now restrict the representation to the rotation algebra generated by $J_m = A_m + B_m$. Since this is a closed $\mathfrak{so}(3) = \mathfrak{su}(2)$ algebra inside of $\mathfrak{so}(1,3)$ (though not the same ones as generated by the A's and B's!), the (j_1, j_2) representation must restrict to a (possibly reducible) representation of $\mathfrak{su}(2)$. Using the technique of *Clebsch–Goran decomposition*, it is possible to show that this restriction decomposes into a sum of irreducible representations of spin j with

$$j \in \{j_1 + j_2 \,, j_1 + j_2 - 1 \,, \dots \,, |j_1 - j_2|\} \,, \tag{6.8}$$

where each value of j occuring exactly once. As a crosscheck, one can easily verify that the sum of the dimensions of these $\mathfrak{su}(2)$ representations add up to the total dimension $(2j_1 + 1)(2j_2 + 1)$ of $V_{(j_1, j_2)} = V_{j_1} \otimes V_{j_2}$.

In particular, if j_1 and j_2 are *both* integer or *both* half-integer, i.e., when $j_1 + j_2$ is integer, then the resulting rotation representations are all integer spin representation. Conversely, if one of the two j_i's is integer and the other half-integer, i.e., when $j_1 + j_2$ is half-integer, then the resulting rotation representations are all half-integer spins. This leads to an important distinction to the type of quantum fields each representation describes: bosonic fields must have integer $j_1 + j_2$, and fermionic fields must have half-integer $j_1 + j_2$.

As a last comment, we point out that the Lorentz group $SO(1,3)^+$, similarly to $SO(3)$, is also not simply-connected. We can then again ask for the universal covering space (thought now we do not really care about projective unitary representations). It turns out to be $SL(2,\mathbb{C})$, which is also a two-fold cover, i.e., $SO(1,3)^+ \cong SL(2,\mathbb{C})/\mathbb{Z}_2$. Similar to $SO(3)$ vs. $SU(2)$, $SL(2,\mathbb{C})$ would qualify as the spin group $Spin(1,3)$, which gives rise to spinor representations of the Lorentz group.

6.2 Spinor representations

Unlike for $SO(3)$, there are two spinor representations of the Lorentz group, characterized by $(j_1, j_2) = (\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$. They are both two-dimensional, and called the left- and right-handed *Weyl spinor* representations, respectively. To shorten the notation, we will abbreviate the representations a $\rho_{(1/2,0)} \equiv \rho_L$, and $\rho_{(0,1/2)} \equiv \rho_R$. The name "spinor" is justified because the restriction to the 3d rotation algebra $\mathfrak{su}(2)$ gives just one irreducible representation of spin $j = \frac{1}{2}$.

From the representation theory of $\mathfrak{su}(2)$, we know that the $j = \frac{1}{2}$ representation is generated by the Pauli matrices σ_m, and the $j = 0$ representation is just the trivial representation. So for the $(\frac{1}{2}, 0)$ spinor representation, we have the representation space $V_{(1/2,0)} = V_{1/2} \otimes V_0 = V_{1/2} \otimes \mathbb{C} \cong V_{1/2}$ being isomorphic to the spin $\frac{1}{2}$ representation space of $\mathfrak{su}(2)$, and the generators are

$$\rho_L(A_m) = \tfrac{1}{2}\sigma_m \otimes \mathbb{1}_1 \equiv \tfrac{1}{2}\sigma_m\,, \quad \rho_L(B_m) = \mathbb{1}_2 \otimes 0 \equiv 0\,. \tag{6.9}$$

For the $(0, \frac{1}{2})$ spinor we have the analogous structure, with $V_{(0,1/2)} \cong V_{1/2}$, and

$$\rho_R(A_m) = 0 \otimes \mathbb{1}_2 \equiv 0\,, \quad \rho_R(B_m) = \mathbb{1}_1 \otimes \tfrac{1}{2}\sigma_m \equiv \tfrac{1}{2}\sigma_m\,. \tag{6.10}$$

This allows us to reconstruct the original generators as

$$\begin{aligned}
(\tfrac{1}{2},0): &\quad \begin{cases} \rho_L(J_m) = \rho_L(A_m) + \rho_L(B_m) = \tfrac{1}{2}\sigma_m\,, \\ \rho_L(K_m) = -i(\rho_L(A_m) - \rho_L(B_m)) = -i\tfrac{1}{2}\sigma_m\,, \end{cases} \\
(0,\tfrac{1}{2}): &\quad \begin{cases} \rho_R(J_m) = \rho_R(A_m) + \rho_R(B_m) = \tfrac{1}{2}\sigma_m\,, \\ \rho_R(K_m) = -i(\rho_R(A_m) - \rho_R(B_m)) = +i\tfrac{1}{2}\sigma_m\,. \end{cases}
\end{aligned} \tag{6.11}$$

The associated (non-unitary) representation of the Lorentz group is then generated by *real* linear combinations of these matrices. E.g.,

$$\begin{aligned}
V_{(1/2,0)} \ni \phi \mapsto &\exp(-i\theta\vec{n} \cdot \rho_L(\vec{J}) - i\vec{v}\rho_L(\vec{K}))\phi \\
= &\exp(-\tfrac{1}{2}(i\theta\vec{n} + \vec{v}) \cdot \vec{\sigma})\phi\,.
\end{aligned} \tag{6.12}$$

Notice that these two representations are *complex conjugates* of each other. To see that, we make use of the identity

$$(i\sigma_2)\overline{\sigma_k}(-i\sigma_2) = (i\sigma_2)\overline{\sigma_k}(i\sigma_2)^{-1} = -\sigma_k\,. \tag{6.13}$$

Now take $\phi \in V_{(1/2,0)}$ as a representation vector of the left-handed spinors. Then, we consider the transformation properties of the new vector $\chi := i\sigma_2\overline{\phi}$, which is

$$\begin{aligned}
\chi \to & \, i\sigma_2\,\overline{\exp(-i\theta\vec{n}\cdot\rho_L(\vec{J}) - i\vec{v}\cdot\rho_L(\vec{K}))\phi} = i\sigma_2\,\overline{\exp(-\tfrac{1}{2}(i\theta\vec{n}+\vec{v})\cdot\vec{\sigma})\phi} \\
& = i\sigma_2\exp(-\tfrac{1}{2}(-i\theta\vec{n}+\vec{v})\cdot\overline{\vec{\sigma}})\overline{\phi} = i\sigma_2\,(i\sigma_2)^{-1}\exp(\tfrac{1}{2}(-i\theta\vec{n}+\vec{v})\cdot\vec{\sigma})\,(i\sigma_2)\overline{\phi} \\
& = \exp(-i\theta\vec{n}\cdot\rho_R(\vec{J}) - i\vec{v}\cdot\rho_R(\vec{K}))\,\chi\,.
\end{aligned} \tag{6.14}$$

So, up to the basis transformation by $i\sigma_2$, complex conjugation of ϕ gave rise to a vector χ which transforms as a vector in $V_{(0,1/2)}$.

Spinors are fundamental in QFT, because they describe fermion fields — the fields that describe all matter particles. Oftentimes, you will see them introduced slightly differently in QFT textbooks, namely, using the so-called *Dirac matrices* and *Dirac spinors*.

Dirac matrices and bi-spinors

Historically, Dirac introduced a set of (4×4) matrices γ_μ (the index μ here label different matrices, but also serve as a spacetime index, i.e., $\mu = 0,1,2,3$ which can be raised and lowered with the Minkowski metric tensor) which satisfies the *anti-commutation* relation

$$\{\gamma_\mu, \gamma_\nu\} := \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\eta_{\mu\nu}\,. \tag{6.15}$$

While the precise form of these matrices are not relevant here, this crucial property allows to straightforwardly construct a representation of the Lorentz algebra:

$$\begin{aligned}
& M_{\alpha\beta} = \tfrac{i}{4}[\gamma_\alpha, \gamma_\beta] = \tfrac{i}{4}(\gamma_\alpha\gamma_\beta - \gamma_\beta\gamma_\alpha) \\
\implies \quad & [M_{\alpha\beta}, M_{\sigma\rho}] = -i\big(\eta_{\alpha\sigma}M_{\beta\rho} - \eta_{\alpha\rho}M_{\beta\sigma} - \eta_{\beta\sigma}M_{\alpha\rho} + \eta_{\beta\rho}M_{\alpha\sigma}\big)\,.
\end{aligned} \tag{6.16}$$

As will be discussed in more detail in the next module, and also in QFT, this gives a four-dimensional representation which is not the defining representation of $SO(1,3)^+$. Dirac famously uses this representation in his “Dirac equation” to describe a massive fermion field.

In the framework of the (j_1, j_2) representation discussed above, this representation is actually the reducible representation $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$. In the representation theory jargon, this is also called a bi-spinor representation. By looking up the Dirac matrices (also called Gamma matrices) γ_μ from one's favorite source, one can in principle compare with the above Weyl spinor representations, and verify that the Dirac spinor representation space $V_{\text{Dirac}} \cong V_{(1/2,0)} \oplus V_{(0,1/2)}$, but expressed in a basis

$$\begin{aligned} \psi_1 &= \tfrac{1}{\sqrt{2}}(\varphi_1 + \chi_1)\,, \quad \psi_2 = \tfrac{1}{\sqrt{2}}(\varphi_2 + \chi_2)\,, \\ \psi_3 &= \tfrac{1}{\sqrt{2}}(\varphi_1 - \chi_1)\,, \quad \psi_4 = \tfrac{1}{\sqrt{2}}(\varphi_2 - \chi_2)\,, \end{aligned} \tag{6.17}$$

where (φ_1, φ_2) are the basis for $V_{(1/2,0)}$, and (χ_1, χ_2) that for $V_{(0,1/2)}$.

Despite being a reducible representation of $\mathfrak{so}(1,3)$, a massive fermion field cannot be thought of “just” as the sum of the two Weyl spinors. The reason is tied to the mass, which requires not only a treatment of the Lorentz algebra, but the full Poincaré algebra in order to properly explain.

6.3 Vector representation

Let us now consider the representation $(\frac{1}{2}, \frac{1}{2})$. The first thing to notice is that its restriction to the 3d rotation algebra gives to a spin $j = \frac{1}{2} + \frac{1}{2} = 1$ and a spin $j = \frac{1}{2} - \frac{1}{2} = 0$ representation, so it must describe bosonic fields.

It turns out that this representation is equivalent to the defining representation of $\mathfrak{so}(1,3)$. To see this explicitly, we must first pick a basis of the representation space $V_{(1/2,1/2)} = V_{1/2} \otimes V_{1/2}$. Since $V_{1/2}$ has basis $\{|\frac{1}{2}, \frac{1}{2}\rangle\,, |\frac{1}{2}, -\frac{1}{2}\rangle\} \equiv \{|+\rangle, |-\rangle\}$, a natural basis would be

$$|+\rangle \otimes |+\rangle \equiv |++\rangle\,, \quad |+\rangle \otimes |-\rangle \equiv |+-\rangle\,, \quad |-+\rangle\,, \quad |--\rangle\,. \tag{6.18}$$

However, we shall also consider another basis, which at this point may seem a bit ad hoc, and whose usefulness will be revealed momentarily:

$$\begin{aligned}
|e_1\rangle &= \tfrac{1}{\sqrt{2}}|+-\rangle - \tfrac{1}{\sqrt{2}}|-+\rangle\,, \\
|e_2\rangle &= \tfrac{1}{\sqrt{2}}|++\rangle - \tfrac{1}{\sqrt{2}}|--\rangle\,, \\
|e_3\rangle &= -\tfrac{i}{\sqrt{2}}|++\rangle - \tfrac{i}{\sqrt{2}}|--\rangle\,, \\
|e_4\rangle &= -\tfrac{1}{\sqrt{2}}|+-\rangle - \tfrac{1}{\sqrt{2}}|-+\rangle\,.
\end{aligned} \tag{6.19}$$

As for the spinor representations, the key point is to express the representation of the Lorentz generators J_m and K_m in terms of the $j = \frac{1}{2}$ representation of the $\mathfrak{su}(2)$ generators A_m and B_m. Using the short-hand notation $\rho_{(1/2,1/2)} \equiv \rho_v$, we first have

$$\begin{aligned}
\rho_v(A_m) &= \rho_{1/2}(A_m) \otimes \mathbb{1}_2 \equiv \mathcal{J}_m^{(1/2)} \otimes \mathbb{1}_2\,, \\
\rho_v(B_m) &= \mathbb{1}_2 \otimes \rho_{1/2}(B_m) \equiv \mathbb{1}_2 \otimes \mathcal{J}_m^{(1/2)}\,,
\end{aligned} \tag{6.20}$$

where $\mathcal{J}^{(1/2)}$ is the spin-$\frac{1}{2}$ representation for a single $\mathfrak{su}(2)$ generator. These act on the basis (6.18) in the obvious way, e.g.,

$$\begin{aligned}
\rho_v(A_m)|+-\rangle &= (\mathcal{J}_m^{(1/2)} \otimes \mathbb{1}_2)(|+\rangle \otimes |-\rangle) = (\mathcal{J}_m^{(1/2)}|+\rangle) \otimes |-\rangle\,, \\
\rho_v(B_m)|+-\rangle &= (\mathbb{1}_2 \otimes \mathcal{J}_m^{(1/2)})(|+\rangle \otimes |-\rangle) = |+\rangle \otimes (\mathcal{J}_m^{(1/2)}|-\rangle)\,.
\end{aligned} \tag{6.21}$$

From these, we can then write the representations of the Lorentz generators as

$$\begin{aligned}
\rho_v(J_m) &= \rho_v(A_m) + \rho_v(B_m) = \mathcal{J}_m^{(1/2)} \otimes \mathbb{1}_2 + \mathbb{1}_2 \otimes \mathcal{J}_m^{(1/2)}\,, \\
\rho_v(K_m) &= -i(\rho_v(A_m) - \rho_v(B_m)) = -i\mathcal{J}_m^{(1/2)} \otimes \mathbb{1}_2 + i\mathbb{1}_2 \otimes \mathcal{J}_m^{(1/2)}\,.
\end{aligned} \tag{6.22}$$

Given that we know all about the spin $\frac{1}{2}$ representation of $\mathfrak{su}(2)$, in particular,

$$\begin{aligned}
&\mathcal{J}_3^{(1/2)}|\pm\rangle \equiv \mathcal{J}_3^{(1/2)}|\tfrac{1}{2}, \pm\tfrac{1}{2}\rangle = \pm\tfrac{1}{2}|\pm\rangle\,, \\
&\mathcal{J}_1^{(1/2)}|+\rangle = \tfrac{1}{2}|-\rangle\,, \quad \mathcal{J}_1^{(1/2)}|-\rangle = \tfrac{1}{2}|+\rangle\,, \\
&\mathcal{J}_2^{(1/2)}|+\rangle = \tfrac{i}{2}|-\rangle\,, \quad \mathcal{J}_2^{(1/2)}|-\rangle = -\tfrac{i}{2}|+\rangle\,,
\end{aligned} \tag{6.23}$$

we can now also compute explicitly the action of these generators on the basis (6.18) of the representation space.

Let us look at the two simple examples of $\rho_v(J_3)$ and $\rho_v(K_3)$ for illustration. First, we have

$$\begin{aligned}
\rho_v(J_3)|++\rangle &= [(\mathcal{J}_3^{(1/2)}|+\rangle) \otimes |+\rangle] + [|+\rangle \otimes \mathcal{J}_3^{(1/2)}|+\rangle] \\
&= [\tfrac{1}{2}|+\rangle \otimes |+\rangle] + [|+\rangle \otimes \tfrac{1}{2}|+\rangle] = |++\rangle\,, \\
\rho_v(J_3)|--\rangle &= [(\mathcal{J}_3^{(1/2)}|-\rangle) \otimes |-\rangle] + [|-\rangle \otimes \mathcal{J}_3^{(1/2)}|-\rangle] \\
&= -\tfrac{1}{2}|--\rangle - \tfrac{1}{2}|--\rangle = -|--\rangle\,, \\
\rho_v(J_3)|+-\rangle &= \tfrac{1}{2}|+-\rangle - \tfrac{1}{2}|+-\rangle = 0 = \rho_v(J_3)|-+\rangle\,.
\end{aligned} \tag{6.24}$$

This may not be quite an illuminating result, but it can now help us to compute the action also in the "ad hoc" basis (6.19):

$$\begin{aligned}
\rho_v(J_3)|e_1\rangle &= \tfrac{1}{\sqrt{2}}\rho_v(J_3)|+-\rangle + \tfrac{i}{\sqrt{2}}\rho_v(J_3)|-+\rangle = 0\,, \\
\rho_v(J_3)|e_2\rangle &= \tfrac{1}{\sqrt{2}}\rho_v(J_3)|++\rangle - \tfrac{1}{\sqrt{2}}\rho_v(J_3)|--\rangle \\
&= \tfrac{1}{\sqrt{2}}|++\rangle + \tfrac{1}{\sqrt{2}}|--\rangle = i|e_3\rangle\,, \\
\rho_v(J_3)|e_3\rangle &= -\tfrac{i}{\sqrt{2}}|++\rangle + \tfrac{i}{\sqrt{2}}|--\rangle = -i|e_2\rangle\,, \\
\rho_v(J_3)|e_4\rangle &= 0\,.
\end{aligned} \tag{6.25}$$

So we see that, this basis, the representation matrix of J_3 is

$$\rho_v(J_3) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,, \tag{6.26}$$

which is precisely the form (6.2) we expect for the defining representation of $\mathfrak{so}(1,3)$. We can also look at $\rho(K_3) = -i\mathcal{J}_3^{(1/2)} \otimes \mathbb{1}_2 + i\mathbb{1}_2 \otimes \mathcal{J}_3^{(1/2)}$, which has

$$\begin{aligned}
\rho_v(K_3)|++\rangle &= -\tfrac{i}{2}|++\rangle + \tfrac{i}{2}|++\rangle = 0\,, \\
\rho_v(K_3)|+-\rangle &= -\tfrac{i}{2}|+-\rangle - \tfrac{i}{2}|+-\rangle = -i|+-\rangle\,, \\
\rho_v(K_3)|-+\rangle &= \tfrac{i}{2}|-+\rangle + \tfrac{i}{2}|-+\rangle = i|-+\rangle\,, \\
\rho_v(K_3)|--\rangle &= \tfrac{i}{2}|--\rangle - \tfrac{i}{2}|--\rangle = 0\,,
\end{aligned} \tag{6.27}$$

and hence

$$\begin{aligned}
\rho_v(K_3)|e_1\rangle &= -\tfrac{i}{\sqrt{2}}|+-\rangle - \tfrac{i}{\sqrt{2}}|-+\rangle = i|e_4\rangle\,, \\
\rho_v(K_3)|e_2\rangle &= \rho_v(K_3)|e_3\rangle = 0\,, \\
\rho_v(K_3)|e_4\rangle &= \tfrac{i}{\sqrt{2}}|+-\rangle - \tfrac{i}{\sqrt{2}}|-+\rangle = i|e_1\rangle\,, \\
\Longrightarrow\ \rho_v(K_3) &= \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}\,,
\end{aligned} \tag{6.28}$$

which also agrees with the expected form in (6.2).

Analogous computations will reveal that all the other generators will be represented as in (6.2) in the basis $|e_i\rangle$ given in (6.19). There, as claimed, the $(\frac{1}{2}, \frac{1}{2})$ representation of the Lorentz algebra is indeed the defining representation. Because it arises from considering vectors in Minkowski space, it is also called the vector representation. Quantum fields in this representation will carry one standard vector index, and are called vector fields. They describe non-gravitational "forces" in QFT, and give rise to, e.g., photon and gluon fields.

6.4 Higher spin representations

By increasing the spins j_1 and j_2, we obtain more representations; those that are most relevant in QFT are summarized in the following.

- $(j_1, j_2) = (0, 0)$: This is the trivial representation of the Lorentz group. Fields in this representation are called *scalar fields*, and are bosonic, i.e., have integer spin (0) under 3d rotations; you have by now seen them in the QFT course. This is probably the most-taught type of fields in QFT lectures around the world, but the only particle in the Standard Model in this representation is the Higgs boson.
- $(j_1, j_2) = (\frac{1}{2}, 0)$, $(0, \frac{1}{2})$, or $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$: These are the spinor representations which we have discussed in detail above. All elementary matter particles in the Standard Model transform in such representations.
- $(j_1, j_2) = (\frac{1}{2}, \frac{1}{2})$: As explained above, this is the defining, or vector representation, of $SO(1, 3)^+$. All "force bosons" in the Standard Model are described by fields in this representation.
- $(j_1, j_2) = (1, 0)$ or $(0, 1)$: These representations also restrict to spin 1 and 0 representations of 3d rotations, so one may be tempted to identify them with vectors as well. However, this turns out not to be the case. Instead, fields in the $(1, 0)$ representation are known as *self-dual 2-form fields*, and those in $(0, 1)$ representation are the *anti*-self-dual 2-form fields. The (anti-)self-dualness is related to transformation properties under time- and spatial reflections, which we have not discussed. In any event, there are no fundamental particles in the Standard Model in these representations. However, they are ubiquitous in extensions of the Standard Model, and more speculative theories such as supergravity or string theory.

- $(1,0) \oplus (0,1)$: This turns out to be the adjoint representation of $\mathfrak{so}(1,3)$. The fields in this representation are "parity-invariant" 2-forms; this has a realization in the Standard Model in terms of the field strengths of elementary forces (e.g., the electromagnetic field strength tensor $F_{\mu\nu}$).
- $(1, \frac{1}{2}) \oplus (\frac{1}{2}, 1)$: These describe parity-invariant fermionic fields with spin $\frac{3}{2}$ under 3d rotations, and are called *Rarita–Schwinger* fields. Again, no such elementary particle exists within the Standard Model, but they are plentiful in "Beyond the Standard Model" physics. The most famous example is the putative supersymmetric partner of the graviton, the gravitino.
- $(1,1)$: This is a bosonic representation containing spin 2 representations under 3d rotations. This representation describes *traceless symmetric tensor fields*. An example would be the *energy-momentum tensor*, which every QFT has, but which is not a fundamental particle field. Again, such particles are absent in the Standard Model, but if we include gravity, then the gravitons are in this representation. In fact, (under some mild assumptions) *any* massless particle with spin 2 must be a graviton.