Notes on Physics from Symmetry

Author: Juvid Aryaman
Last compiled: June 16, 2021

This document contains my personal notes on Jakob Schwichtenberg's Physics from Symmetry (Schwichtenberg, 2015), with a sprinkling of notes from my undergraduate physics course in quantum field theory (and, to a lesser extent, general relativity).

1. Special relativity

1.1. Definitions and postulates

In special relativity, **inertial frames of reference** are coordinate systems moving with constant velocity relative to each other. Special relativity has two basic postulates:

1. **The principal of relativity**: The laws of physics are the same in all inertial frames of reference.
2. **The invariance of the speed of light**: The velocity of light has the same value c in all inertial frames of reference.

Theorem 1.1 (Invariant of special relativity). *Consider two events A and B in an inertial observer O's frame of reference. Let the time interval measured by O between the two events be (Δt), and the three spatial intervals be (Δx), (Δy), (Δz). Then, the quantity*

$$(\Delta s)^2 := (\Delta ct)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta y)^2 \tag{1.1}$$

is **invariant** *between all frames of reference. I.e.*

$$(\Delta s') = (\Delta s) \tag{1.2}$$

for any inertial frame of reference O'.

Theorem 1.1 follows directly from the invariance of the speed of light (consider a pair of mirrors, for two observers with relative velocity).

Definition 1.1 (Proper time). *Proper time, τ, is the time measured by an observer in the special frame of reference where the object in question is at rest. In this frame of reference,*

$$(\Delta s)^2 = (c\Delta\tau)^2. \tag{1.3}$$

In the infinitesimal limit

$$(\mathrm{d}s)^2 = (c\,\mathrm{d}\tau)^2. \tag{1.4}$$

Physically, Defn. 1.1 means that all observers agree on the time interval between events for an observer who travels with the object in question. However, different observers **do not** in general agree on the time interval between events generally: $(\Delta t) \neq (\Delta t')$ – this is called **time dilation**.

1.2. c is an upper speed limit

All observers agree on the value of $(\mathrm{d}s)^2 = (c\,\mathrm{d}\tau)^2$. Furthermore, we commonly assume that there exists a minimal proper time of $\tau = 0$ for two events if $\Delta s^2 = 0$. We can therefore write that when $\tau = 0$

$$c^2 = \frac{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2}{(\mathrm{d}t)^2} \tag{1.5}$$

between two events with an infinitesimal distance. We can equate the right-hand side with a squared velocity, and hence

$$\tau = 0 \implies c^2 = v^2 \tag{1.6}$$

so

$$(\mathrm{d}s)^2 \geq 0 \implies c^2 \geq v^2 \tag{1.7}$$

for **any** pair of events (which are causally connected, although how this follows is not immediately clear to me right now).

1.3. Tensor notation and Minkowski spacetime

Definition 1.2 (Four-vector (contravariant)). *A position four-**vector** is defined as*

$$x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{1.8}$$

Definition 1.3 (Minkowski metric). *The Minkowski metric is defined as*

$$\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1). \tag{1.9}$$

η is used to compute distances and lengths in Minkowski space.

We define $\eta^{\mu\nu}$ through the relation

$$\eta^{\mu\nu}\eta_{\nu\sigma} = \delta^\mu{}_\sigma \tag{1.10}$$

where we have appled the **Einstein summation convention**, where a repeated Greek index implies a summation from 0 to 3 (where the zeroth index is time), and a repeated Roman index is summed from 1 to 3. Hence, for a matrix multiplication between two 3×3 matricies A and B, $(AB)_{ij} = A_{ik}B_{kj}$, and $(A^T)_{ij} = A_{ji}$.

Definition 1.4 (One-form (covariant vector)). *We define a one-form as*

$$x_\mu = \eta_{\mu\nu}x^\nu. \tag{1.11}$$

Thus,

$$\mathrm{d}s^2 = \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu. \tag{1.12}$$

Definition 1.5 (Scalar product). *A scalar product between four-vectors x and y is defined as*

$$x \cdot y := x^\mu y^\nu \eta_{\mu\nu} = x_\mu y_\nu \eta^{\mu\nu} = x^\mu y_\mu = x_\nu y^\nu \tag{1.13}$$

due to the symmetry of the metric: $\eta_{\mu\nu} = \eta_{\nu\mu}$.

Ordering (spacing) of indicies In order to be able to freely raise/lower indicies (without repeatedly writing the metric tensor), we can impose an ordering upon indicies of tensor fields – which we can represent typographically with spacing between tensor indicies. A metric g_{ij} (or g^{ij}) has the effect of lowering (or raising) a repeated index. For example,

$$g_{iq}T^{abcd}{}_{efgh}{}^{ijkl}{}_{mnop} = T^{abcd}{}_{efghq}{}^{jkl}{}_{mnop}. \tag{1.14}$$

(Proof of this, I imagine, requires background in differential geometry?)

1.4. Lorentz transformations

From the invariant of SR (Theorem 1.1), we have

$$\mathrm{d}s'^2 = \mathrm{d}x'_\mu \,\mathrm{d}x'_\nu \eta^{\mu\nu} = \mathrm{d}x_\mu \,\mathrm{d}x_\nu \eta^{\mu\nu} \tag{1.15}$$

for all reference frames. We denote Λ as a (1,1) tensor field, which transforms a four-vector from one reference frame to another:

$$\mathrm{d}x'^\mu = \Lambda^\mu{}_\nu \,\mathrm{d}x^\nu \tag{1.16}$$

which leaves the $\mathrm{d}s^2$ invariant, i.e. $\mathrm{d}s'^2 = ds^2$. It follows that

$$\begin{aligned}\eta_{\mu\nu} &= \Lambda^\sigma{}_\mu \Lambda^\delta{}_\nu \eta_{\sigma\delta} \\ \eta &= \Lambda^T \eta \Lambda.\end{aligned} \tag{1.17}$$

The physical meaning of Eq.(1.17) is that Lorentz transformations leave the scalar product of Minkowski spacetime invariant: i.e. changes between frames of reference that respect the two postualtes of special relativity (Section 1.1). Conservation of the scalar product is analogous to rigid rotation (O) in Euclidean space $(a \cdot b = a' \cdot b' = a^T O^T O b \implies O^T I O = I)$, which preserves orientation $(\det(\Lambda) = 1)$.

Note that $\Lambda^\mu{}_\nu \neq \Lambda_\nu{}^\mu$. Beginning with Eq.(1.17),

$$\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta_{\mu\nu} = \eta_{\rho\sigma}$$

we can raise one index, and lower one index, of $\Lambda^\nu{}_\sigma$

$$\begin{aligned}\Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma \eta_{\nu\mu} \eta^{\sigma\lambda} &= \eta_{\rho\sigma} \eta_{\nu\mu} \eta^{\sigma\lambda} \\ \Lambda^\mu{}_\rho \Lambda_\mu{}^\lambda \cancel{\eta_{\mu\nu}} &= \cancel{\eta_{\mu\nu}} \delta_\rho{}^\lambda \\ \Lambda^\mu{}_\rho \Lambda_\mu{}^\lambda &= \delta_\rho{}^\lambda\end{aligned} \tag{1.18}$$

so we see that $\Lambda_\nu{}^\mu$ is the inverse of $\Lambda^\mu{}_\nu$.

2. Lie group theory

2.1. Invariance, symmetry, and covariance

We call a quantity **invariant** if it does not change under particular transformations. E.g. if we transform $A, B, C, ... \to A', B', C', ...$ and we have

$$F(A', B', C', ...) = F(A, B, C, ...) \tag{2.1}$$

then we say F is invariant under this transformation. **Symmetry** is defined as invariance under a transformation (or class of transformations). An equation is covariant if it takes the same form when objects in it are transformed. *All physical laws must be covariant under Lorentz transformations.*

Group theory describes the properties of particular sets of transformations: the invariances under such groups allows us to mathematically describe symmetry. For example, the set of rotations about the origin of a square by $n\pi/2$ form a **discrete group**, and leave the set of points which constitute the square invariant under the transformation. The set of rotations about the origin of a circle form a **continuous group**. We can use group theory to work with *all* kinds of symmetries: symmetries which operate on vectors, equations, ...

2.2. Groups

Definition 2.1 (Group axioms). *A group $(G, \circ)$ is a set G, together with a binary operation $\circ$ defined on G, that satisfies the following axioms*

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

The set of all transformations that leave a given object invariant is called a **symmetry group**. For Minkowski spacetime, the object that is left invariant is the Minkowski metric, and the corresponding symmetry group is called the **Poincaré group**. Notice that the transformations which constitute a group are defined entirely independently from the object on which the transformations act.

2.2.1. Rotations in two dimensions and $SO(2)$

Consider the 2D rotation matrix

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{2.2}$$

and the two reflection matrices

$$P_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad P_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.3}$$

These matrices satisfy the group axioms. We can uncover this group from a symmetry perspective. The above transformations leave the length of a vector unchanged, i.e.

$$a.a = a'.a'. \tag{2.4}$$

Letting the transformation be represented by $a' = Oa$, it follows that all members of the group must satisfy

$$O^T O = I. \tag{2.5}$$

This condition defines the group $O(2)$, which is the group of all **orthogonal** 2×2 matrices. It follows that $\det(O) = \pm 1$ – i.e. the transformations are area-preserving. The subgroup with $\det(O) = 1$ is called $SO(2)$, which corresponds to rigid rotations preserving the orientation of the system – "S" denoting **special**.

2.2.2. Rotations with unit complex numbers and $U(1)$

A unit complex number is a complex number z which satisfies $|z|^2 = z^* z = 1$. The group $U(1)$ is the set of unit complex numbers, together with ordinary complex number multiplication. The U stands for '**unitary**', which generally stands for the condition

$$U^\dagger U = 1, \tag{2.6}$$

where $U^\dagger = (U^T)^*$ is the **Hermitian conjugate** of U. For scalars, the Hermitian conjugate is equivalent to the complex conjugate. Note that a unit complex number can also be denoted as

$$R_\theta = e^{i\theta} = \cos(\theta) + i\sin(\theta) \tag{2.7}$$

which makes the interpretation of $U(1)$ as rotations on the unit complex numbers evident.

We can connect this description of rotations ($U(1)$) to the previous ($SO(2)$) by defining

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.8}$$

For an arbitrary unit complex number $z = a + ib$, let

$$f(z) = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \tag{2.9}$$

Since $z = R_\theta = \cos(\theta) + i\sin(\theta)$, we can plug in the real and imaginary components of z into Eq.(2.9) to arrive at Eq.(2.2). We then have $z' = R_\theta z$, to perform rotations. There therefore exists an **isomorphism** between $SO(2)$ and $U(1)$:

Definition 2.2 (Group isomorphism). *Given two groups $(G, *)$, $(H, \odot)$, a group isomorphism is a bijective function $f : G \to H$ such that*

$$f(u * v) = f(u) \odot f(v) \ \forall \ u, v \in G. \tag{2.10}$$

which is written as

$$(G, *) \cong (H, \odot). \tag{2.11}$$

$f(z)$ in Eq.(2.9) is therefore a group isomorphism between $U(1)$ and $SO(2)$,

$$SO(2) \cong U(1). \tag{2.12}$$

This realization has an analogue in three dimensions, which will reveal something fundamental about nature.

2.2.3. Rotations in three dimensions and $SO(3)$

As with $SO(2)$, the defining conditions of $SO(3)$ are

$$R^T R = I \tag{2.13}$$

$$\det(R) = 1. \tag{2.14}$$

Matrices satisfying these two conditions represent rigid rotations of 3-dimensional vectors. These matrices can be described by the following three "basis rotations"

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad R_y = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$
$$R_z = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.15}$$

So, to rotate a vector v around the z-axis by θ, we would compute $R_z(\theta)v$. The set of (orientation-preserving) rotation matrices acting on 3-dimensional vectors is called $SO(3)$.

×	1	i	j	k
1	1	i	j	k
i	i	−1	k	−j
j	j	−k	−1	i
k	k	j	−i	−1

Figure 1. Quaternion multiplication table, read as row $\times$ column $=$ value. E.g. $\mathbf{ji} = -\mathbf{k}$. In general, the basic quaternions anti-commute.

2.2.4. Quaternions and $SU(2)$

To get a second description of rotations in three dimensions, we must generalize complex numbers in higher dimensions. Astonishingly, it turns out that there are no 3-dimensional complex numbers. Instead, we can find 4-dimensional complex numbers called quaternions, which will turn out to be able to describe rotations in 3-dimensions. The fact that quaternions are 4-dimensional will reveal something deep about the universe. We could have anticipated this result, because we will be using unit quaternions, which have 3 degrees of freedom.

To construct quaternions, we introduce three complex units satisfying the relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 - -1 \tag{2.16}$$

$$\mathbf{ijk} = -1 \tag{2.17}$$

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}. \tag{2.18}$$

All other relations can be computed from the above. For example, the relation $\mathbf{ij} = \mathbf{k}$ can be derived by multiplying both sides of Eq.(2.18) by $\mathbf{k}$. Notice that it follows that the **basic quaternions anticommute** with each other, see Fig. 1.

The set of unit quaternions satisfy

$$q^\dagger q = 1 \tag{2.19}$$

$$\implies a^2 + b^2 + c^2 + d^2 = 1. \tag{2.20}$$

As the unit complex numbers formed a group under complex number multiplication, the unit quaternions form a group under quaternion multiplication. There are several possible ways of representing the basic quaternions with 2D matrices, but one way is as follows:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \mathbf{i} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
$$\mathbf{j} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad , \quad \mathbf{k} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{2.21}$$

With these matrices, a generic quaternion $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ can be written in a matrix representation as

$$f(q) = \begin{pmatrix} a + di & b + ci \\ -b + ci & a - di \end{pmatrix}. \tag{2.22}$$

We also observe that $\det(f(q)) = 1$, and so we conclude that the unit quaternions are given by the set of matrices with the above form and unit determinant. The unit quaternions, written as $2{\times}2$ matrices U therefore fulfil the conditions

$$U^\dagger U = 1 \qquad \text{and } \det(U) = 1. \tag{2.23}$$

This defines the symmetry group $SU(2)$.

The map between $SU(2)$ and $SO(3)$ is not as simple as the one we saw between $U(1)$ and $SO(2)$. The mapping of a complex number onto a 2-dimensional vector is easy because a complex number has two degrees of freedom: $v = x + \mathbf{i}y$. But the mapping of a quaternion onto 3-dimensional vector is not so straightforward because a quaternion has four degrees of freedom. We will make the mapping of a 3-dimensional vector $(x, y, z)^T$ onto a quaternion v as

$$v \equiv x\mathbf{i} + y\mathbf{j} + z\mathbf{k}. \tag{2.24}$$

Using Eq.(2.21), we see that $\det(v) = x^2 + y^2 + z^2$. In order to perform transformations which preserve the length of the vector (x, y, z), we must use transformations which preserve determinants. Therefore, the restriction to *unit* quaternions means that we must restrict to matrices with *unit* determinants[1]. Naively, a first guess would be that simply multiplying a vector v by a unit quaternion u induces a rotation on v, but this is not the case because the product of u and v may not belong to $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$. It turns out that the following transformation can describe rotations in 3-dimensions

$$v' = qvq^{-1}. \tag{2.25}$$

Let t be a quaternion defining a rotation through ϕ, where

$$t = \cos(\frac{\phi}{2}) + \sin(\frac{\phi}{2})u \tag{2.26}$$

$$u = u_x\mathbf{i} + u_y\mathbf{j} + u_z\mathbf{k} \tag{2.27}$$

$$u^\dagger u = 1 \implies t^\dagger t = 1. \tag{2.28}$$

As an example, suppose we wish to rotate the vector $\vec{v} = (1, 0, 0)^T$ around the z-axis by ϕ. Then using Eq.(2.21)

$$\vec{v} = (1, 0, 0)^T \to v = 1\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.29}$$

From Eq.(2.26), defining $\theta = \phi/2$

$$R_z(\theta) = \cos(\theta)\mathbf{1} + \sin(\theta)\mathbf{k} = \begin{pmatrix} \cos(\theta) + i\sin(\theta) & 0 \\ 0 & \cos(\theta) - i\sin(\theta) \end{pmatrix}. \tag{2.30}$$

From Eq.(2.25), the rotated vector v' is

$$v' = R_z(\theta)vR_z(\theta)^{-1} = \begin{pmatrix} 0 & \cos(\phi) + i\sin(\phi) \\ -\cos(\phi) + i\sin(\phi) & 0 \end{pmatrix}. \tag{2.31}$$

Using the general quaternion matrix representation Eq.(2.22), we can equate

$$v'_x = \cos(\phi), \qquad v'_y = \sin(\phi), \qquad v'_z = 0 \tag{2.32}$$

as expected.

Inspection of Eq.(2.26) reveals that the mapping of unit quaternions onto 3-dimensional rotations is not one-to-one. For example, a rotation by $\phi = \pi$ is equivalent to a rotation by $\phi = 2\pi + \pi = 3\pi$. But,

$$t_{\phi=\pi} = \sin(\frac{\pi}{2})u = u \tag{2.33}$$

$$t_{\phi=3\pi} = \sin(\frac{3\pi}{2})u = -u. \tag{2.34}$$

[1]Since $\det(BA) = \det(B)\det(A)$

Hence, we call $SU(2)$ a **double-cover** of $SO(3)$, because every element of $SO(3)$ has two corresponding elements in $SU(2)$ [**TODO: I think**?]. It is therefore always possible to go unambiguously from $SU(2)$ to $SO(3)$, but not vice versa.

We will see later that groups which cover other groups are fundamental for quantum spin. Note that the above had one quaternion parameter too many, which may be interpreted as a hint towards relativity: a more natural identification may have been $v = t\mathbf{1} + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Rotations in 4-dimensions require 6 degrees of freedom[2]. However, there is no 7-dimensional generalisation of complex numbers. But two unit quaternions do have 6 degrees of freedom: we will see later how there is a close connection between two copies of $SU(2)$ and rotations in four dimensions.

3. Lie algebras

3.1. Intuition behind generators

Consider an element of a continuous group which is arbitrarily close to the identity

$$g(\epsilon) = I + \epsilon X \tag{3.1}$$

where ϵ is small, and I is the identity (think about 1D rotations for concreteness). We let X remain abstract for the moment. We wish to apply this transformation N times to achieve a finite transformation by a total amount θ, $h(\theta)$. As a result, we may write:

$$h(\theta) = (I + \frac{\theta}{N} X)^N. \tag{3.2}$$

We then take the limit of $N \to \infty$, which is just the exponential function

$$h(\theta) = \lim_{N\to\infty} (I + \frac{\theta}{N})^N = e^{\theta X}. \tag{3.3}$$

In this sense, the object X generates the finite transformation h, which is why we call X a **generator**.

We can then differentiate $h(\theta)$ to obtain the generator:

$$X = \left.\frac{\mathrm{d}h(\theta)}{\mathrm{d}\theta}\right|_{\theta=0}. \tag{3.4}$$

If we consider a continuous group of transformations that are given by matrices, we can also make a Taylor expansion of an element of the group about the identity

$$h(\theta) = \sum_n \frac{1}{n!} \left.\frac{\mathrm{d}^n h}{\mathrm{d}\theta^n}\right|_{\theta=0} \theta^n \tag{3.5}$$

which shows how generators generate transformations.

3.2. An intuitive definition for matrix Lie groups

For matrix Lie groups, the corresponding Lie algebra can be defined as the collection of objects that give an element of the group when exponentiated.

Definition 3.1 (Lie algebra for matrix Lie groups, non-rigorous). *For a matrix Lie group G (given by $n \times n$ matrices), the Lie algebra $\mathfrak{g}$ of G is given by those $n \times n$ matrices X such that*

$$e^{tX} \in G \tag{3.6}$$

[2] Using ordinary 4×4 matrices, with the constraints $O^T O = 1$ and $\det(O) = 1$ reduce the 16 components of an arbitrary 4×4 matrix to 6 independent components

for $t \in \mathbb{R}$, together with an operation called the Lie bracket, that defines how to combine elements of the Lie algebra

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}. \tag{3.7}$$

A Lie algebra is **closed** *under the Lie bracket.*

Note that, in general $X \circ Y \notin \mathfrak{g}$.

Theorem 3.1. *Consider a matrix Lie group G and corresponding Lie algebra $\mathfrak{g}$. If $X, Y \in \mathfrak{g}$ and $g, h \in G$, then*

$$g \circ h = e^X \circ e^Y = \underbrace{e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\dots}}_{\in G} \tag{3.8}$$

where the right hand side is the Baker-Campbell-Hausdorff formula[3]*. For* **matrix Lie groups***, the Lie bracket $[X, Y]$ is* **equivalent** *to the* **commutator** *of X and Y:*

$$[X, Y] = XY - YX. \tag{3.9}$$

Eq.(3.8) *connects combinations of elements of the group G to combinations of elements of the Lie algebra $\mathfrak{g}$.*

3.3. Generators and Lie algebra of $SO(3)$

Using the norm-preserving condition of $SO(3)$, Eq. (2.13), along with Eq.(3.6) $O = e^{\theta J}$, yields

$$O^T O = e^{\theta J^T} e^{\theta J} = 1 \implies J^T + J = 0, \tag{3.10}$$

The parity-preserving condition in Eq.(2.14) also yields[4]

$$\det(e^{\theta J}) = e^{\theta \operatorname{tr}(J)} = 1 \implies \operatorname{tr}(J) = 0. \tag{3.11}$$

A set of three **basis**[5] generators J_1, J_2, J_3 fulfilling both of these conditions can be written conveniently as

$$(J_i)_{jk} = -\epsilon_{ijk}, \tag{3.12}$$

where ϵ_{ijk} is the Levi-Civita symbol

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) = \{(1,2,3),(2,3,1),(3,1,2)\} \\ 0 & \text{if } i = j \text{ or } j = k \text{ or } k = i \\ -1 & \text{if } (i,j,k) = \{(1,3,2),(2,1,3),(3,2,1)\} \end{cases}. \tag{3.13}$$

The corresponding Lie bracket is

$$[J_i, J_j] = \epsilon_{ijk} J_k. \tag{3.14}$$

In physics, it is conventional to define the generators of $SO(3)$ with an extra i, so that we get Hermitian generators[6], to arrive at

$$O = e^{i\phi J} \tag{3.15}$$
$$\implies (J_i)_{jk} = -i\epsilon_{ijk} \tag{3.16}$$
$$[J_i, J_j] = i\epsilon_{ijk} J_k. \tag{3.17}$$

[3] This holds more generally, not only for matrix Lie groups. See also Defn. 3.2

[4] Because $\det(e^A) = e^{\operatorname{tr}(A)}$

[5] i.e. all other generators are linear combinations of the basis generators

[6] i.e. $J^\dagger = J$

We call the Lie bracket relation of the basis generators **the** Lie algebra of a given group, because everything that is important about a Lie algebra is encoded in the Lie bracket relation of the basis generators.

An alternative way to derive the basis generators is to begin with the finite transformation matrices and apply Eq.(3.4). However, the above method is more general, whose steps were:

1. Begin with the definition of the group;
2. Use those constraints to derive the basis generators;
3. Then derive the an explicit form for finite transformations.

3.4. Formal definition of a Lie algebra

Definition 3.2 (Lie algebra). *A Lie algebra is a vector space $\mathfrak{g}$ over some field*[7] F *together with a binary operation* $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *called the Lie bracket satisfying the following axioms:*

1. *Bilinearity:*

$$[aX + bY, Z] = a[X, Z] + b[Y, Z] \tag{3.18}$$

$$[Z, aX + bY] = a[Z, X] + b[Z, Y] \tag{3.19}$$

   for all scalars $a, b \in F$ *and all elements* $X, Y, Z \in \mathfrak{g}$.

2. *Anticommutativity:*

$$[X, Y] = -[Y, X] \tag{3.20}$$

   for all $X \in \mathfrak{g}$.

3. *The Jacobi identity*

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0 \tag{3.21}$$

   for all elements $X, Y \in \mathfrak{g}$.

Importantly, this definition makes no reference to any Lie group: the definition of a Lie algebra stands on its own.

3.5. Generators and Lie algebra of $SU(2)$

Using Equations (2.23) for the definition of $SU(2)$, and Eq.(3.15) for the definition of a generator, we arrive at

$$e^{-iJ^\dagger} e^{iJ_i} = e^{-iJ_i^\dagger + iJ_i + \frac{1}{2}[J_i^\dagger, J_i] + \cdots} = 1 \tag{3.22}$$

where the right-hand side uses the Baker-Campbell-Hausdorff theorem. This implies that

$$J = J^\dagger \tag{3.23}$$

$$\mathrm{tr}(J) = 0 \tag{3.24}$$

where the second equation uses $\det(U) = 1$. A basis for Hermitian traceless 2×2 matrices is given by the 3 **Pauli** matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.25}$$

[7]Informally, a field is a set, with operations of addition and multiplication defined on that set, along with inverses of those operations

We conventionally define

$$J_i \equiv \sigma_i/2 \tag{3.26}$$

to arrive at the Lie bracket (or, equivalently, commutator relationship)

$$[J_i, J_k] = i\epsilon_{ijk} J_k \tag{3.27}$$

which is **identical** to Eq.(3.17) for $SO(3)$! Hence we say that $SU(2)$ and $SO(3)$ have the same Lie algebra, since they share the same Lie bracket.

3.6. Lie groups and covering groups

An $SU(2)$ transformation doesn't necessarily have to be represented by 2×2 matrices. An abstract definition of a Lie group will enable us to see the connection between different descriptions of the same transformation.

A **Lie group** is a group that is also a differentiable manifold. A Lie group must have the group operations $\circ$ **induce** a **differentiable** map of the manifold onto itself. Concretely, this means that every group element, e.g. A, induces a differentiable map that takes any element of the group B to another element of the group $C = AB$.

Definition 3.3 (Simply connected Lie group). *A Lie group is said to be simply connected if every closed curve on the manifold can be shrunk smoothly to a point.*

Theorem 3.2 (Existence of a covering group). *There exists a unique simply connected Lie group corresponding to each Lie algebra. We call this Lie group the* **covering** *group. All other groups having the same Lie algebra are said to be "covered" by the simply connected Lie group.*[8]

Definition 3.4 (Unit n-sphere). *A unit n-sphere is denoted as the set of points satisfying*

$$S^n = \{x \in \mathbb{R}^{n+1} : ||x|| = 1\}. \tag{3.28}$$

S^n is therefore a sphere embedded in $n+1$ dimensions, but with a unit radius constraint, implying an n-dimensional object. For groups we have discussed so far, we may identify:

1. S^1, with group operation of multiplication, is the covering group of 1-dimensional rotations. $S^1 \cong U(1) \cong SO(2)$.
2. S^3, with group operation of multiplication, is the covering group of 3-dimensional rotations. $S^3 \cong SU(2) \underbrace{\rightarrow}_{\text{two-to-one}} SO(3)$.

Geometrically, we can think of $SO(3)$ as a 3-sphere but with antipodal points identified: this is because each point of $SO(3)$ is associated with two points of $SU(2)$ which is equivalent to S^3.

3.7. Representation theory

Representation theory deals with *representing* abstract algebraic structures (mainly groups, associative algebras, and Lie algebras) with linear transformations of vector spaces – i.e. matrices operating on columnar vectors. Considering just groups for now, a representation abides the following definition

Definition 3.5 (Representation for a group). *A representation is a map R (or* **homomorphism***) between any group element g of a group G onto a vector space V such that:*

- *$R(e) = I$, i.e. the identity element of the group is also the matrix identity*

[8] This is a result from differential geometry.

- $R(g^{-1}) = (R(g))^{-1}$, *i.e. inverse elements are mapped to inverse transforms*
- $R(g) \circ R(h) = R(gh)$.

For example, the usual rotation matrices in Eq.(2.15) are a representation of the group $SO(3)$ on the vector space $\mathbf{R}^3$. Representation theory allows us to examine the effect of the group on other vector spaces – of potentially different dimensionality than the definition of the group itself.

Theorem 3.3 (Similarity transformation). *If R is a representation of group G with elements g, then for any invertible matrix S,*

$$R'(g) = S^{-1}R(g)S \tag{3.29}$$

is also a representation.

Proof. Follows immediately from the definition of a representation. For example,

$$R'(g_1)R'(g_2) = S^{-1}R(g_1)SS^{-1}R(g_2)S = S^{-1}R(g_1)R(g_2)S = S^{-1}R(g_1g_2)S = R'(g_1g_2). \tag{3.30}$$

□

This freedom to perform similarity transformations corresponds to the freedom to choose a basis for the vector space the group acts on.

Definition 3.6 (Invariant subspace). *Given a representation R of a group G on vector space V, we call $V' \subseteq V$ an invariant subspace if for all $v \in V'$, we have $R(g) \in V'$.*

Definition 3.7 (Subrepresentation). *A subrepresentation of a representation (R, V) of a group G is a representation (R', V') such that $V' \subseteq V$ and*

$$R'(g)v = R(g)v \tag{3.31}$$

for all $v \in V'$.

Definition 3.8 (Irreducible representation). *An irreducible representation is a representation (R, V) of a group G that has no invariant subspace besides the zero space $\{0\}$ and V itself.*

Theorem 3.4 (Irreducible representations and block diagonal form). *An irreducible representation is a representation which cannot be rewritten, using a similarity transformation, into block diagonal form. In contrast, reducible representations can be rewritten in block diagonal form through similarity transformations.*

There are many representations for any particular group. How do we know which ones to choose to describe nature?

Definition 3.9 (General linear group of a vector subspace). *$GL(V)$ is the set of $n \times n$ invertible matrices, together with ordinary matrix multiplication on a vector space V. It is the group of automorphisms of V, i.e. the set of bijective linear transformations $V \to V$.*

Lemma 3.1 (Schur's Lemma). *If we have an irreducible representation $R : \mathfrak{g} \to GL(V)$, then any linear operator $T : V \to V$ that commutes with all operators $R(X)$ for $X \in \mathfrak{g}$, must be a scalar multiple of the identity operator.*

Definition 3.10 (Casimir element). *A Casimir element C is an object with the property*

$$[C, X] = 0 \tag{3.32}$$

for every $X \in \mathfrak{g}$ where $\mathfrak{g}$ is a Lie algebra. Note that if C commutes with each basis of $\mathfrak{g}$ then it follows that C obeys Eq.(3.32)*.*

Consequently, by Schur's Lemma, the Casimir element for any given representation of a Lie algebra must be a scalar multiple of the identity. This scalar provides us with a number to label representations of Lie algebras. **Physical mass and spin turn out to be examples of these constants of proportionality**.

3.8. $SU(2)$

On our way to describing the Poincaré group, we will first discuss irreducible representations of $SU(2)$, which will then allow us to start talking about special relativity again.

3.8.1. Finite dimensional irreducible representations

Definition 3.11 (Quadratic Casimir element). *Quadratic Casimir elements are the simplest Casimir elements and are constructed as quadratic polynomials of generators of a Lie algebra.*

For the Lie algebra $\mathfrak{su}(2)$, there exists a **unique** Casimir element

$$J^2 = J_1^2 + J_2^2 + J_3^2 \tag{3.33}$$

where J_i is the basis for the Lie algebra Eq.(3.26). Different dimensional representations will give different representations of the Casimir operator: in 2-D $J_{\mathrm{2d}}^2 = 3/4 \cdot I$ whereas $J_{\mathrm{3d}}^2 = 2 \cdot I$. Letting $J^2 v = bv$, for some vector v, then we will use b to label the representation.

Definition 3.12 (Cartan element). *Cartan elements of a representation of a Lie algebra are those basis generators which are simultaneously diagonalizable.*[9]

Only one basis generator of $\mathfrak{su}(2)$ can be diagonalized at a time, therefore the Cartan element of $\mathfrak{su}(2)$ is unique. It is customary to use J_3 for this purpose. We will show below that the eigenvalues of J_3 (m) label the elements of the vector space that the Lie algebra acts upon.

Definition 3.13 (Bra-ket notation). *If ϕ is a vector in a complex vector space, then we use $|\phi\rangle$ to denote this vector (pronounced "ket-phi"). Any convenient label can be used inside the ket $|\ \rangle$, which denotes that the object is simply a vector in a vector space. A bra $\langle\phi|$ is an element of the dual vector space, i.e. a bra is a linear functional which is a linear map from the vector space to the complex numbers.*

We will therefore label every basis vector, for every representation of a Lie algebra, in bra-ket notation as follows

$$\begin{aligned} J^2 |b,m\rangle &= b\,|b,m\rangle \\ J_3 |b,m\rangle &= m\,|b,m\rangle \end{aligned} \tag{3.34}$$

where m **labels the representation** and b **labels the vector within the representation**. We can obtain basis vectors for the representation by enumerating the eigenvalues of J^2 and J_3.

Definition 3.14 (Ladder operators for $\mathfrak{su}(2)$). *Define the following complex linear combination of the basis generators of $SU(2)$*

$$J_+ = J_1 + iJ_2 \tag{3.35}$$
$$J_- = J_1 - iJ_2. \tag{3.36}$$

where J_+ is a **raising** *operator and J_- is a* **lowering** *operator.*

In $\mathfrak{su}(2)$ we only consider real linear combinations of generators, so we thus expand our consideration to the **complexification** of $\mathfrak{su}(2)$, called $\mathfrak{sl}(2,\mathbb{C})$ where $SL(2,\mathbb{R})$ is the special linear group.

Applying Eq.(3.27) gives rise to two new commutation relations

$$[J_3, J_\pm] = \pm J_\pm \tag{3.37}$$
$$[J_+, J_-] = 2J_3. \tag{3.38}$$

[9] A set of matrices are simultaneously diagonalizable if there exists a matrix P such that $P^{-1}AP$ is a diagonal matrix for every A in the set.

As a result, we find that if $|b,m\rangle$ is an eigenvector of J_3 with eigenvalue m, then $J_\pm |b,m\rangle$ is also an eigenvector of J_3 with eigenvalue $m \pm 1$, i.e.

$$J_3 J_\pm |b,m\rangle = (m \pm 1) J_\pm |b,m\rangle \tag{3.39}$$
$$\implies J_\pm |b,m\rangle \propto |b, m \pm 1\rangle . \tag{3.40}$$

There can only be a finite number of eigenvectors of J_3 because we are dealing with a finite-dimensional representation, because the vector space is finite dimensional and therefore there can only be a finite number of linearly independent vectors. Therefore, there must exist an eigenvector with a maximum eigenvalue – let's call it j. It must have the property

$$J_+ |b,j\rangle = 0. \tag{3.41}$$

Using

$$J_- J_+ = J^2 - J_3^2 - J_3 \tag{3.42}$$

then it follows from the action of $J_- J_+ |b,j\rangle$ that

$$b = j(j+1). \tag{3.43}$$

If we let k be the minimum eigenvalue of J_3, then an analogous argument allows us to conclude that $k = -j$ and therefore

$$-j \leq m \leq j. \tag{3.44}$$

Hence we have derived a relationship between the whole representation and the number of states in the vector space.

Now, if we begin at $|b,j\rangle$ and by applying the lowering operator J_- we reduce the eigenvalue of J_3 by 1. We perform this operation some finite number of times, and hence $j - k$ must be an integer. But $k = -j$, so

$$2j = \text{integer} \implies j = \frac{\text{integer}}{2}. \tag{3.45}$$

Simply plugging in $j = 0, 1/2, 1, 3/2, ...$ allows us to enumerate all representations of $\mathfrak{su}(2)$.

Lastly, by assuming that our basis vectors $|b,m\rangle$ are normalized

$$|b,m\rangle^\dagger |b,m\rangle = 1 \tag{3.46}$$

we can derive the constant of proportionality in Eq.(3.40). By observing that $(J_+ |b,m\rangle)^\dagger J_+ |b,m\rangle = |C|^2$ where C is the constant of proportionality, and by using an analogous argument for J_-, we arrive at

$$J_+ |j(j+1), m\rangle = \sqrt{j(j+1) - m^2 - m}\, |j(j+1), m+1\rangle \tag{3.47}$$
$$J_- |j(j+1), m\rangle = \sqrt{j(j+1) - m^2 + m}\, |j(j+1), m-1\rangle . \tag{3.48}$$

Finally, note that the representation label $j(j+1)$ takes up a lot of space and is somewhat redundant, so we tend to label states simply as $|j,m\rangle$ rather than $|j(j+1), m\rangle$. The tools above will allow us to derive explicit representations of $SU(2)$ in different dimensions.

3.8.2. Recipe for explicit $\mathfrak{sl}(2,\mathbb{C})$ representations of $SU(2)$

First, we choose a particular value of j – the maximum value of the eigenvalue of the Casimir element of the group. In 2-dimensions, this is $j = 1/2$. We then take all the corresponding eigenvalues of the Casimir element, in this case $m = 1/2, -1/2$. We then construct a diagonal matrix for representation of the Cartan element:

$$J_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.49}$$

We therefore have the basis vectors of the Lie algebra as

$$|1/2, 1/2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1/2, -1/2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.50}$$

We can find the explicit matrix form of the other two $SU(2)$ generators by rewriting Eq.(3.36) in terms of J_1 and J_2. We can apply J_1 and J_2 on the two basis vectors above, and concatenate the corresponding vectors into the matrix representations of J_1 and J_2. We then recover (3.26). This can be repeated for any dimension you like.

3.9. The Lorentz group $O(1, 3)$

We now have the tools to derive a finite-dimensional representation of the Lorentz group, of which Lorentz transforms are elements.

The Lorentz group is the set of all transformations which preserve the inner product in Minkowski space

$$x^\mu x_\mu = x^\mu \eta_{\mu\nu} x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{3.51}$$

where $\eta_{\mu\nu}$ is defined in Eq.(1.9). The reason we call the group $O(1, 3)$ is due to the signature of the Minkowski metric: the group $O(4)$ preserves $(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2$.

Recall that Eq.(1.17)

$$\eta_{\mu\nu} = \Lambda^\sigma{}_\mu \Lambda^\delta{}_\nu \eta_{\sigma\delta}$$

provides the definition of a Lorentz transformation: it is the set of transformations which leaves the Minkowski inner product invariant, resulting in $\eta = \Lambda^T \eta \Lambda$. Taking the determinant of both sides, we get

$$\det \Lambda = \pm 1. \tag{3.52}$$

If we look at the time component, $\mu = \nu = 0$, then Eq.(1.17) gives

$$\Lambda^0{}_0 = \pm\sqrt{1 + \Lambda^i{}_0 \Lambda^i{}_0}. \tag{3.53}$$

As a consequence, we have four different sub-categories of Lorentz transform:

$$\begin{aligned}
L_+^\uparrow &: \det(\Lambda) = +1; \qquad \Lambda^0{}_0 \geq +1 \\
L_-^\uparrow &: \det(\Lambda) = -1; \qquad \Lambda^0{}_0 \geq +1 \\
L_+^\downarrow &: \det(\Lambda) = +1; \qquad \Lambda^0{}_0 \leq -1 \\
L_-^\downarrow &: \det(\Lambda) = -1; \qquad \Lambda^0{}_0 \leq -1.
\end{aligned}$$

Definition 3.15 (Proper orthochronous Lorentz group). *The proper orthochronous Lorentz group ($L_+^\uparrow$) is the set of Lorentz transformations (Eq.(1.17)) which satisfy $\det(\lambda) = +1$ (proper) and $\Lambda^0{}_0 \geq +1$ (orthochronous). Proper implies that transformations are parity-preserving. Orthochronous implies transformations preserve the direction of time. This group is denoted as $SO(1, 3)_+^\uparrow$, also known as the* **restricted** *Lorentz group.*

Lorentz transformations not from $L_+^\uparrow$ can be written as combinations of transformations from $L_+^\uparrow$ and the parity ($Lambda_P$) and time-reversal ($Lambda_T$) transformations[10]

$$\Lambda_P = \mathrm{diag}(1, -1, -1, -1) \tag{3.54}$$
$$\Lambda_T = \mathrm{diag}(-1, 1, 1, 1). \tag{3.55}$$

[10]Note, however, that Λ_P and Λ_T will look quite different in different representations.

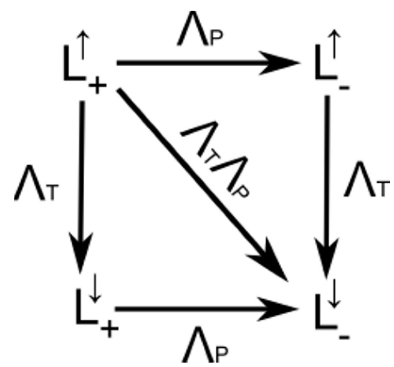

Figure 2. Components of the Lorentz group. Elements of the Lorentz group outside of the proper orthochronous Lorentz group are connected via the discrete "jumps" Λ_P and Λ_T.

Hence, we may write the complete Lorentz group as

$$O(1,3) = \{L_+^\uparrow, \Lambda_P L_+^\uparrow, \Lambda_T L_+^\uparrow, \Lambda_P \Lambda_T L_+^\uparrow\}. \tag{3.56}$$

Notice that only $L_+^\uparrow$ can be built up by infinitesimal transformations from the Lorentz group, because only these transformations are continuously connected to the identity element of the group. There exists a "gap" between the transformations in $L_+^\uparrow$ and those not in $L_+^\uparrow$, mediated by Λ_P and Λ_T (see Fig. 2). Hence, in what follows, we will concentrate on the Lie group $L_+^\uparrow$, since the covering group is a single topologically connected piece – whereas the full Lorentz group consists of four topologically separated pieces.

References

Schwichtenberg, J., 2015 *Physics from symmetry*. Springer.