# Notes on Physics from Symmetry

Author: Juvid Aryaman
Last compiled: May 21, 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 $(\Delta t)$, and the three spatial intervals be $(\Delta x)$, $(\Delta y)$, $(\Delta z)$. Then, the quantity*

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

*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, $\tau$, 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*

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

*$\eta$ 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*

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

*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 $\Lambda$ 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 $(a\cdot b = a'\cdot b' = a^TO^TOb \implies O^TIO = I)$, 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 $SO(2)$. This realization has an analogue in three dimensions, which will reveal something fundamental about nature.

### 2.2.3. Rotations in three dimensions

Rotations in three dimensions can be described by the following "basis rotations"

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad R_y = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$
$$R_z = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.12}$$

So, to rotate a vector $v$ around the $z$-axis by $\theta$, we would compute $R_z(\theta)v$.

### 2.2.4. Quaternions

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.

## References

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