1
Lorentz and Poincaré Invariance

1.1 Lorentz Invariance

To begin with we will very briefly review some aspects of Einstein's theory of relativity that are of particular importance here.

The theory of relativity states that physical laws are the same in two systems that move with respect to each other with uniform velocity. Furthermore, the speed of light is a constant. These two statements lead to the concept of invariance under Lorentz transformations. We now first investigate Lorentz transformations in some detail. Some of the rather basic mathematics involved is summarized in an appendix.

Lorentz transformations can be understood as rotations in four-dimensional space (three-dimensional space + time). A rotation can be specified by a matrix L with the property that $\widetilde{L}$ (the reflected of L) is its inverse:

$$\widetilde{L} = L^{-1} \qquad \text{or} \qquad L\widetilde{L} = 1\,.$$

Writing indices explicitly:

$$L_{\mu\nu}\widetilde{L}_{\nu\lambda} = \delta_{\mu\lambda} \quad \text{or} \quad L_{\mu\nu}L_{\lambda\nu} = \delta_{\mu\lambda}\,.$$

We have used here Einstein's summation convention: twice occurring indices (such as ν here) are summed over (in this case $\nu = 1, \ldots, 4$). At this point we must also settle some conventions. We take the fourth dimension as imaginary, $x_4 = ict$. This leads to the fact that the matrix-elements of a Lorentz transformation are imaginary if one (but not both) of the indices is four. With this convention a particle at rest has the four-momentum $(0, 0, 0, iM)$, where c has already been taken to be one.

Let us emphasize that there is no physics in the choice of metric. Some physicists prefer to work with real space/time but define their dot-product with a metric involving minus signs. It is really of no relevance where you hide your minus signs, at most it is a matter of convenience. Which is usually what you are used to. It is a matter though that you can debate hotly at lunch time (real time). See appendix on metric.

Examples of Lorentz transformations are:

- ordinary rotations in three dimensions such as a rotation over an angle ϕ around the third axis;

$$L = \begin{bmatrix} \cos\phi & \sin\phi & 0 & 0 \\ -\sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- transformation to a system moving with velocity v along the first axis:

$$L = \begin{bmatrix} \cos\theta & 0 & 0 & \sin\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sin\theta & 0 & 0 & \cos\theta \end{bmatrix} .$$

where θ is imaginary and such that $\sin\theta = iv/c\beta$, $\beta = \sqrt{1 - v^2/c^2}$. It follows that $\cos\theta = \sqrt{1 - \sin^2\theta} = 1/\beta$. This transformation is a rotation over an imaginary angle.

In addition to the Lorentz transformations that have determinant 1, such as the ordinary rotations and the velocity transformations there are also transformations with determinant -1. These are the space or time reflections. These are not transformations that you can actually do: nobody has ever managed to reflect himself, transforming himself from, say, a right handed person into a left handed person. In particle physics it has been discovered that the laws of nature are not invariant with respect to these reflections, although large parts of the interactions are. The reflections remain therefore important tools in classifying interactions and establishing selection rules.

A reflection is the combination of any ordinary Lorentz transformation and a space reflection P or time reversal T:

$$P = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

There are a number of fundamental differences about rotations over real versus imaginary angles. Rotating over a real angle of 2π gives the identity, because $\sin(2\pi) = 0$ and $\cos(2\pi) = 1$. Thus two successive rotations may lead back to the original (for example a rotation of 30° followed by a rotation of 330°). Since no such thing holds for imaginary angles this is not true for the rotations over imaginary angles. In fact:

$$\sin\theta = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right),$$

and if θ is imaginary, $\theta = i\alpha$ with real α, then

$$\sin\theta = \frac{1}{2i}\left(e^{-\alpha} - e^{\alpha}\right),$$

which grows exponentially for θ going to either $\pm i\infty$. Thus while the domain of the angle ϕ for spatial rotations is finite, $0 \le \phi < 2\pi$, the domain of θ is infinite. The rotations in three dimensions form a compact group (finite domain of the parameters) while the full set of Lorentz transformations is non-compact. To judge on compact versus non-compact on the basis of the domain of the parameters one must first specify how the parameters are to be chosen; the requirement is that they be chosen such that two successive applications of the same transformation described by the parameters $\alpha_1, \ldots, \alpha_n$ must be given by the transformation described by the parameters $2\alpha_1, \ldots, 2\alpha_n$. For example, two successive rotations over an angle θ equals a rotation over an angle 2θ, and that means that the parameter θ is appropriate for a judgment on compactness. The importance of compactness relates to group theory: compact groups have unitary representations. The Lorentz group is non-compact, and its representations are not necessarily unitary. At this point there is no need to understand this mathematically in any detail.

A general Lorentz transformation can be seen as a combination of a rotation in three dimensional space followed by a transformation to a system moving with some velocity in some direction. A rotation can be specified by three parameters, for example by a vector whose direction is the axis of rotation while its magnitude equals the magnitude of the rotation. Thus the vector $(0, 0, \pi/2)$ specifies a rotation over 90° around the third axis. The "velocity" transformation can be specified by giving the velocity, which is also a three component vector. It would be nice if we could talk in terms of an axis and an imaginary angle also in this case, but that is not important at this moment. The important point is that we observe that a Lorentz transformation is specified by six parameters. Three have a finite, three an infinite domain. The Lorentz-transformations form a six-parameter group.

Since any finite rotation can be seen as an infinite sequence of infinitesimal rotations it is sufficient for most purposes to understand infinitesimal Lorentz transformations. Let us first consider a rotation over an angle ϕ around the third axis. Its form has been given above, and we will denote it by $L(\phi)$. This rotation can be obtained also by applying n times a rotation over an angle ϕ/n :

$$L(\phi) = \left[L\left(\frac{\phi}{n}\right)\right]^n .$$

Let us now consider a rotation over an angle ϕ/n with very large n. We may then expand $\sin(\phi/n)$ and $\cos(\phi/n)$ to get:

$$L(\phi/n) = \begin{bmatrix} 1 & \phi/n & 0 & 0 \\ -\phi/n & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \mathcal{O}\left(\phi^2/n^2\right)$$
$$= I + \frac{\phi}{n} L_3 + \mathcal{O}\left(\phi^2/n^2\right)$$

with

$$L_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and I denoting the unit matrix. In the limit of large n:

$$L(\phi) = \lim_{n\to\infty} \left[L\left(\frac{\phi}{n}\right)\right]^n = e^{\phi L_3}.$$

Exercise 1.1 Read the appendix on matrices or else show that

$$\left[1 + \frac{\alpha}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right]^n = e^\alpha + \mathcal{O}\left(\frac{1}{n}\right).$$

We have now written this Lorentz transformation in exponential form. The great advantage is that the parameter ϕ is directly visible, and the property $L(\phi)L(\phi) = L(2\phi)$ is manifest:

$$e^{\phi L_3} e^{\phi L_3} = e^{2\phi L_3}.$$

Similarly other rotations may be treated. A general infinitesimal rotation in three dimensions differs infinitesimally from the unit matrix:

$$R = \begin{bmatrix} 1+g & a & b & 0 \\ d & 1+h & c & 0 \\ e & f & 1+k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with $a \ldots k$ infinitesimal. For a rotation the equation $R\tilde{R} = 1$ holds, and this leads to the result $h = g = k = 0$, $d = -a$, $e = -b$ and $f = -c$ (ignoring higher order terms in a, b, etc.).

Exercise 1.2 Prove this assertion.

We therefore can write:

$$R = I + cL_1 - bL_2 + aL_3$$

with

$$L_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad L_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$L_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The reason for the sign choices above will become clear shortly. Since a finite transformation can be obtained by exponentiation of an infinitesimal one we so find a representation in terms of three parameters for any rotation in three dimensions:

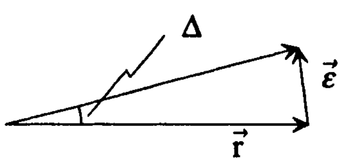
$$R = e^{\alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3} = e^{\alpha_i L_i}.$$

The three quantities α_i are precisely equal to the vector used to describe rotations introduced above. Thus the direction of α is the axis of rotation, the magnitude is the magnitude of the rotation in radians. The sense of the rotation is this: if $\vec{\alpha}$ points upwards, along the positive third axis, then a small rotation will turn a vector along the positive first axis into a vector having a small negative second component. From the above this connection is obvious for the special cases of rotations around first, second, or third axis. For the general case this becomes obvious by considering an infinitesimal rotation:

$$\begin{bmatrix} 1 & \alpha_3 & -\alpha_2 & 0 \\ -\alpha_3 & 1 & \alpha_1 & 0 \\ \alpha_2 & -\alpha_1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \alpha_i \text{ infinitesimal.}$$

This matrix describes an infinitesimal rotation over an angle $\Delta = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$ around an axis in the direction $\vec{\alpha}$. The choice of signs for the L_i was made such as to obtain this.

Exercise 1.3 Verify the above by showing that a vector in the direction of $\vec{\alpha}$ is invariant, while a vector perpendicular to $\vec{\alpha}$, for example the unit vector $\vec{r}$ with components $\lambda(\alpha_2,-\alpha_1,0,0)$ with $\lambda = 1/\sqrt{(\alpha_1^2 + \alpha_2^2)}$, is changed by an amount corresponding to a rotation over an angle Δ. Thus compute the effect of the infinitesimal rotation on $\vec{r}$, writing the result in the form $\vec{r}+\vec{\epsilon}$. Show that $\vec{\epsilon}$ is orthogonal to $\vec{r}$ and $\vec{\alpha}$, and has the magnitude Δ.



This treatment can be extended trivially to include the "velocity" transformations. A general "velocity" transformation will be of the form:

$$V = e^{\beta_1 M_1 + \beta_2 M_2 + \beta_3 M_3}$$

with imaginary β_1, β_2 and β_3 and real M_1, M_2 and M_3:

$$M_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$M_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

Of course we could equally well have used real $\vec{\beta}$ and imaginary

M, but we will do it this way. Writing $\vec{\beta} = i\vec{v}$, a particle of mass m at rest is transformed into a particle moving with momentum $\vec{p} = -m\vec{v}$. Note that the $\vec{v}$ here is the conventional velocity divided by the relativistic factor β.

The general Lorentz transformation is of the form

$$L = e^{\alpha_i L_i + \beta_i M_i}$$

but the interpretation of the α_i and β_i in terms of axis of rotation or velocity is no more as easy as above.

At this point it is necessary to introduce a new and sometimes slightly confusing notation. We write:

$$L = e^{\frac{1}{2}\alpha_{\mu\nu} K_{\mu\nu}}, \quad \mu, \nu = 1, \ldots 4.$$

The matrices K are defined by the prescription that $K_{\mu\nu}$ is a matrix with 1 in row μ, column ν, and -1 in row ν, column μ. Otherwise its elements are zero. Note that $K_{\mu\nu} = -K_{\nu\mu}$. The $\alpha_{\mu\nu}$ are chosen such as to give the correct result. Thus, given that $L_1 = K_{23}$, $L_2 = -K_{13}$ and $L_3 = K_{12}$ the correspondence is:

$$\begin{array}{ll} \alpha_1 \leftrightarrow \alpha_{23} & \beta_1 \leftrightarrow \alpha_{14} \\ \alpha_2 \leftrightarrow \alpha_{31} & \beta_2 \leftrightarrow \alpha_{24} \\ \alpha_3 \leftrightarrow \alpha_{12} & \beta_3 \leftrightarrow \alpha_{34} \end{array}$$

while the remaining α are defined by $\alpha_{\mu\nu} = -\alpha_{\nu\mu}$.

The confusion may arise by not being careful about indices. The K are 4×4 matrices, the α are numbers with $\alpha_{\mu\nu}$ real if $\mu, \nu = 1, 2, 3$, or $\mu = \nu = 4$, and imaginary if μ or $\nu = 4$. To be very explicit, the matrix-element i, j of the matrix K_{13} could be written as

$$(K_{13})_{ij} \, .$$

1.2 Structure of the Lorentz Group

We must now study the structure of the Lorentz group, by which we mean the following. Two successive Lorentz transformations equals another Lorentz transformation, and we must understand this connection in terms of the parameters $\alpha_{\mu\nu}$. Thus, let there be given two Lorentz transformations described by parameters $\alpha_{\mu\nu}$ and $\beta_{\mu\nu}$ respectively. The product of these two is another Lorentz transformation described by parameters $\gamma_{\mu\nu}$ and we would like to know the γ as function of the α and β. Unfortunately this relation is quite complicated. On the infinitesimal level it is relatively easy to compute γ, but the complete expression is more difficult.

The starting equation is:

$$L(\gamma) = L(\beta)\, L(\alpha).$$

In lowest order:

$$\gamma_{ij} = \beta_{ij} + \alpha_{ij} + \text{ terms of higher order.}$$

It is not that difficult to go to the next order. First consider the product

$$e^A\, e^B,$$

where A and B are matrices, to second order in A and B.

$$\left(1 + A + \tfrac{1}{2}A^2\right)\left(1 + B + \tfrac{1}{2}B^2\right) \simeq 1{+}A{+}B{+}AB{+}\tfrac{1}{2}B^2{+}\tfrac{1}{2}A^2{+}\ldots$$

Note that AB need not to be equal to BA. It is easy to see that up to higher order terms this is equal to

$$1 + C + \tfrac{1}{2}C^2$$

with $C = A + B + \frac{1}{2}AB - \frac{1}{2}BA = A + B + \frac{1}{2}[A, B]$. Thus, writing

$$e^C \;=\; e^A e^B$$

we have up to second order in A and B:

$$C = A + B + \tfrac{1}{2}[A, B] + \ldots$$

with $[A, B] = AB - BA$. In the same way subsequent terms of C may be worked out:

$$\begin{aligned} C = A + B + \tfrac{1}{2}[A, B] &+ \tfrac{1}{12}[A, [A, B]] \\ &+ \tfrac{1}{12}[[A, B], B] + \tfrac{1}{48}[A, [[A, B], B]] + \ldots \end{aligned}$$

involving multiple commutators such as

$$\begin{aligned} [[A, B], B]] &= [A, B]B - B[A, B] \\ &= ABB - 2BAB + BBA. \end{aligned}$$

It may be proven that all higher order terms can be written as multiple commutators. The equation is called the **Campbell–Baker–Hausdorff** formula, and in the appendix on matrices this assertion is proven. We do not need the explicit form, but only the fact that all terms are multiple commutators.

To work out the multiplication laws for the Lorentz transformations we must work out multiple commutators for the matrices K.

Here it turns out that the commutator of any two K matrices is a linear combination of K matrices. This is really a consequence of the fact that the transformations form a group, the product of two transformations is again a Lorentz transformation and can also be expressed as an exponential involving K matrices. Thus we need to work out only the commutators of any pair of K matrices. That is still some work, but not too much. Things are somewhat easier in terms of the original matrices L_i and M_i, for example

$$[L_1, L_2] = -L_3$$

and cyclically. Also the commutators between the M and the L and M are easy to do. Nevertheless, writing everything in terms of the K leads to a formidable expression:

$$[K_{ab}, K_{cd}] = c^{ij}_{abcd} K_{ij}$$

(summation over i and j from 1...4) with the coefficients c given by

$$c^{ij}_{abcd} = \delta_{bc}\delta_{ai}\delta_{dj} - \delta_{bd}\delta_{ai}\delta_{cj} - \delta_{ac}\delta_{bi}\delta_{dj} + \delta_{ad}\delta_{bi}\delta_{cj}$$

Exercise 1.4 Check this equation.

For example, for $a = 2$, $b = 3$, $c = 3$, $d = 1$ we have

$$c^{ij}_{2331} = \delta_{2i}\delta_{1j}$$

thus

$$[K_{23}, K_{31}] = K_{21} = -K_{12}$$

which is the same as $[L_1, L_2] = -L_3$.

Since all terms in the Campbell–Baker–Hausdorff formula are given in terms of multiple commutators we can express everything in terms of the coefficients c. For example, going back to our original equation:

$$\gamma^{ij} = \beta^{ij} + \alpha^{ij} + \tfrac{1}{2}\, c^{ij}_{abcd}\beta^{ab}\alpha^{cd} + \tfrac{1}{12} cc\beta\beta\alpha + \ldots$$

Summation over a, b, c, and d from 1 to 4 is understood. In the last term we did not write all indices, but just indicated which quantities are involved. A certain symmetry in notation has been achieved by putting the indices on the α, β and γ on top.

The constants c are called structure constants since knowledge of the c implies knowledge of the structure of the Lorentz group.

They are the central quantities of the group. Knowing the c one knows everything about products of Lorentz transformations. One can say that knowing a group amounts to knowing its structure constants.

1.3 Poincaré Invariance

Laws of nature seem not only invariant with respect to Lorentz invariance, they also appear to be invariant under translations. Thus Maxwell's laws appear equally valid in Europe or the U.S. (translation in space) or yesterday or today (translation in time). We observe thus invariance under the transformation

$$\vec{x} \rightarrow \vec{x} + \vec{b}$$
$$t \rightarrow t + b_0$$

or

$$x \rightarrow x + b$$

where now x and b denote 4-vectors. The combined invariance under Lorentz transformations and translations is called **Poincaré invariance**. It should be noted that Lorentz transformations and translations do not commute (first a translation $T(b)$ over a vector b followed by a Lorentz transformation L is in general not the same as L followed by $T(b)$).

1.4 Maxwell Equations

The Maxwell equations describe electricity and magnetism, and also light, and we must face the problem of quantizing these equations. We know that light is quantized in photons, thus since light is nothing but a combination of electric and magnetic fields we conclude that these fields are quantized. As it happens the Maxwell equations, while well understood classically, are quite difficult in the context of quantum field theory. The problem is with gauge invariance. The classical equations for the potentials are:

$$\partial_\mu F_{\mu\nu} = j_\nu$$

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu ,$$