Quantum Field Theory: PH6418/ EP4618 (Spring 2022)

Lecture 4 notes*

February 28, 2022

1 Special Relativity Reloaded

- **Minkowski spacetime and the Invariant (squared) spacetime interval**: According to special relativity (or equivalently in the absence of gravity), spacetime is the 4 dimensional manifold, $\mathbb{R}^{1,3}$ made up of points (events). Let's consider a pair of events occurring at spacetime location x and $x + dx$ according to some inertial frame (coordinate system). Then the spacetime interval between this pair of points is given by,

$$ds^2 = c^2 dt^2 - d\boldsymbol{x} \cdot d\boldsymbol{x}. \tag{1}$$

The same pair of events appear to occur at x' and $x' + dx'$ in another inertial frame. In special relativity the time interval of two nearby events is **not same** in all inertial frames, and neither is the spatial separation i.e. $dt' \neq dt$ and/or $|d\mathbf{x}'| \neq |d\mathbf{x}|$[1]. But instead the squared spacetime interval remains unchanged

$$ds^2 = c^2 dt^2 - d\boldsymbol{x} \cdot d\boldsymbol{x} = c^2 dt'^2 - d\boldsymbol{x}' \cdot d\boldsymbol{x}'.$$

In a sense ds^2 captures how far in both space and time two events occur and whether there is any causal connection between them i.e. if is it possible for them to be connected as cause and effect. If for two nearby events, $ds^2 > 0$, the two events are said to be *timelike separated* and in fact it is possible a signal to travel from one event to the other and thus they can be connected as cause and effect. If $ds^2 = 0$, then the two events are said to be *lightlike separated*, and they can be joined by a light signal traveling from either one event to the other. However, if $ds^2 < 0$, the two events are said to be *spacelike separated* and no signal can be sent from one event to the other, which means they cannot be connected by cause and effect.

*Typos and errors should be emailed to the Instructor: Shubho Roy (email: sroy@phy.iith.ac.in)

[1] In Galilean relativity, one *would* have $dt' = dt$ (Newton's notion of absolute time) and $|d\mathbf{x}'| \neq |d\mathbf{x}|$.

W can express the squared interval in matrix form which is coordinate/index free,

$$ds^2 = \begin{pmatrix} dx^0 & dx^1 & dx^2 & dx^3 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix} = dx^T \, \eta \, dx.$$

where we have introduced Minkowski metric as the matrix η,

$$\eta = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix},$$

which is a symmetric matrix, $\eta = \eta^T$ and in particular diagonal. Note that it is self-inverse, $\eta^2 = 1$.

- **Poincaré transformations** are those coordinate changes (diffeomorphisms) under which the Minkowski metric (1) remains invariant. In the language of GR, we would say that the Poincaré transformations are the isometries of Minkowski space, $\mathbb{R}^{1,3}$. Physically Poincaré transformations can be expressed in a column vector form,

$$x \to x' = \Lambda \, x + a$$

where x is the 4 row column vector, $x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$, Λ is a 4×4 matrix which represents a Lorentz transformation (boosts, rotations, etc.) and a is a 4-row column vector representing the shift of origin of spacetime coordinates (translations). In index notation, we

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu \, x^\nu + a^\mu. \tag{2}$$

a^μ's are all real and so are the matrix elements $\Lambda^\mu{}_\nu$. Sometimes it's better to work to coordinate differences/differentials so the shift vector, a drop out and we have[2] ,

$$\begin{aligned} dx'^\mu &= \Lambda^\mu{}_\nu \, dx^\nu, & (3) \\ \Rightarrow \frac{\partial x'^\mu}{\partial x^\nu} &= \Lambda^\mu{}_\nu. & (4) \end{aligned}$$

So the Lorentz transformation matrix is the Jacobian matrix of the transformation.

[2]Note that when we go over to GR, Lorentz transformations (LT) would generalize to general coordinate transformations (GCT), BUT this formula will still hold except the $\Lambda^\mu{}_\nu$ will not be constants but functions of spacetime,

$$\frac{\partial x'^\mu}{\partial x^\nu} = [\Lambda(x)]^\mu{}_\nu.$$

- **The Lorentz Group:** To determine what is the nature of the Lorentz transformation matrices, Λ's, we start by equating the expression for the invariant squared interval in the two frames,

$$\begin{aligned} ds^2 = dx^T\,\eta\,dx &= dx'^T\,\eta\,dx'. \\ &= (\Lambda\,dx)^T\;\eta\;(\Lambda\,x) \\ &= dx^T\;\left(\Lambda^T\eta\Lambda\right)\;dx. \end{aligned}$$

This implies,

$$\Lambda^T\,\eta\,\Lambda = \eta. \tag{5}$$

If in place of η one had the identity matrix, it would have meant Λ is an orthogonal matrix. But the $\eta_{11} = \eta_{22} = \eta_{33} = -1$ spoils this. Instead we call Λ an $O(1,3)$ matrix (O for orthogonal and the$(1,3)$ refers to the metric signature i.e. number of positive and negative eigenvalues respectively of η which is $+---$).
Taking determinants of both sides, we get

$$(\det\Lambda)^2 = 1 \implies \det\Lambda = \pm 1.$$

Since the Lorentz transformation is continuously connected to unity (i.e. no transformation), we will stick to $\det\Lambda = +1$. $O(1,3)$ matrices which are unit determinant are called "Special orthogonal" and we denote them by putting an extra "S" in front: $SO(1,3)$[3]. The fact that Lorentz transformation matrices are not orthogonal will have an important consequence as we will see momentarily, namely, one will have to introduce two different vector species in special relativity, one with upper indices (superscripts) and the other with lower indices (subscripts). It will also turn out that this is the precise reason we had to denote the Lorentz transformation with one up and one down index. Thus coordinate differentials, dx^μ and coordinate partials ∂_μ do not transform identically under a Lorentz transformation despite the fact that they both have one index, namely μ.
(This is unlike the case of Galilean transformation matrices, O_{ij}, as they being rotation matrices orthogonal matrices, $O = O^{-T}$ and hence coordinate differentials dx^i and partial derivatives ∂_i transform identically)

2 Geometry of the Lorentz Group

2.1 Orthochronous and Non-orthochronous Lorentz transformations

The defining equation for the Lorentz transformation matrices is,

$$\Lambda^T\eta\,\Lambda = \eta,$$

[3]One can generalize this to some different η which has p number of diagonal elements which are 1 and q number of diagonal elements which are -1. Then we would call the transformation matrix, Λ to be an $SO(p,q)$ type matrix. This is the relevant for example in the context of **Anti de sitter (AdS) space** (we need$SO(2,3)$) which is of great relevance in current day string theory research.

or in component notation,

$$\Lambda^\rho{}_\mu \, \Lambda^\sigma{}_\nu \, \eta_{\rho\sigma} = \eta_{\mu\nu}. \tag{6}$$

Since μ, ν are free indices, they can take any values. In particular consider the case when $\mu = \nu = 0$, i.e.,

$$\Lambda^\rho{}_0 \, \Lambda^\sigma{}_0 \, \eta_{\rho\sigma} = \eta_{00}.$$

Using $\eta_{00} = +1$, $\eta_{ij} = -\delta_{ij}$, we can simplify both sides and get,

$$\left(\Lambda^0{}_0\right)^2 - \left(\Lambda^i{}_0\right)^2 = 1 \Rightarrow \Lambda^0{}_0 = \pm\sqrt{1 + (\Lambda^i{}_0)^2}.$$

Thus, either $\Lambda^0{}_0 \geq +1$ or $\Lambda^0{}_0 \leq -1$. Since these two range of values for $\Lambda^0{}_0$ are non-overlapping, one cannot start from $\Lambda^0{}_0 \geq +1$ and smoothly reach a transformation, $\Lambda^0{}_0 \leq -1$ by continuously changing some real parameter. In some sense these two ranges give us two different/disjoint/unrelated Lorentz transformations. The Lorentz transformations for which $\Lambda^0{}_0 \geq +1$ are called ***orthochronous*** transformations, while those for which $\Lambda^0{}_0 \leq -1$ is called ***non-orthochronous***. Physically speaking, orthochronous transformations keep the direction of time same while non-orthochronous transformations reverse the direction of time. Examples of orthochronous Lorentz transformations are boosts, rotations, as well as parity and space inversion; while examples of non-orthochronous transformations are time-reversal or any product of a time-reversal operation with any number of boosts, rotations or parity/space inversion. Orthochronous Lorentz transformation matrices are denoted by $\Lambda^\uparrow$, while non-orthochronous Lorentz transformations are denoted by $\Lambda^\downarrow$.

2.2 Lorentz group has four disconnected components

We already noted before that Lorentz transformations could be either ***proper***, i.e. have unit determinant, $|\Lambda| = +1$, or be ***improper*** have determinant, $|\Lambda| = -1$. Since these two values of the determinants are also non-overlapping, it is not possible to start from one type, say a transformation of the kind $|\Lambda| = +1$, and reach a transformation of the other kind, $|\Lambda| = -1$ by continuously changing some real parameter. Thus topologically speaking these two subsets of the Lorentz group are disconnected in the space of all transformations. Following standard notation we shall refer to the set of proper Lorentz transformations by Λ_+ while the set of improper Lorentz transformations will be denoted by Λ_-. Thus in all Lorentz transformations can be divided into four following disconnected components (i.e. one cannot go from one component to the other by continuously changing some parameters),

- Proper Orthochronous $\left(\Lambda^\uparrow_+\right)$: e.g., boosts, rotations or products of any no. of boosts and rotations

- Improper Orthochronous $\left(\Lambda^\uparrow_-\right)$: e.g., Parity ($P$) or joint transformation of Parity and proper Orthochronous ($P \times \Lambda^\uparrow_+$)

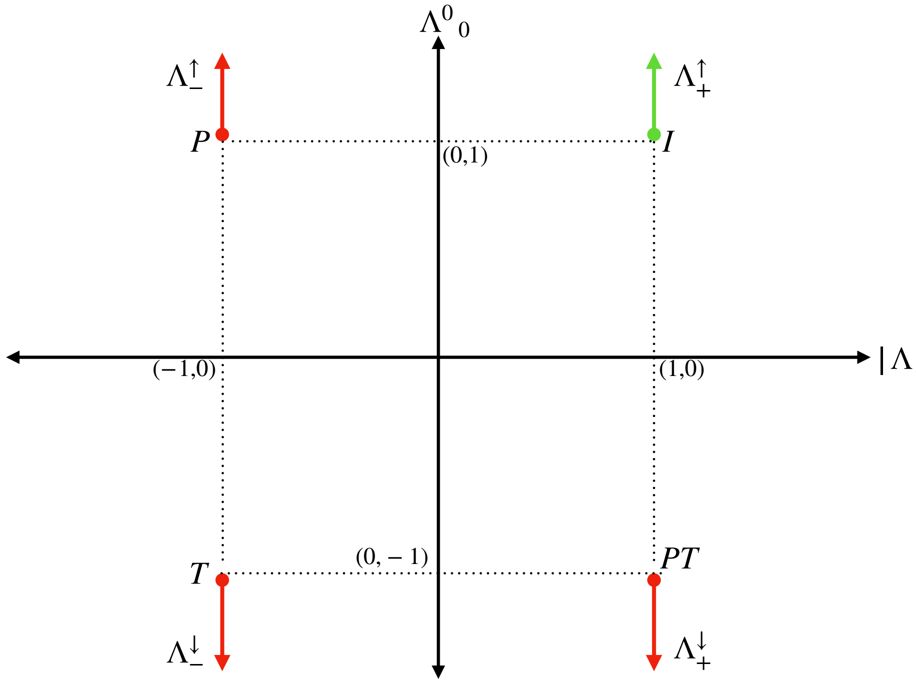

Figure 1: Lorentz group has four disconnected components in the space of matrices indicated by the red and green semi-infinite lines. The x-axis shows the determinant, $|\Lambda|$ while the vertical axis plots the element $\Lambda^0{}_0$. The component shaded in green is the restricted Lorentz group, $\Lambda_+^\uparrow$ or $SO^\uparrow(1,3)$. The components shaded in red are either improper, or Non-orthochronous or both.

- Improper Non-orthochronous $\left(\Lambda_-^\downarrow\right)$: e.g., Time-reversal ($T$) or joint transformation of time-reversal and a proper Orthochronous ($T \times \Lambda_+^\uparrow$)

- Proper Non-orthochronous $\left(\Lambda_+^\downarrow\right)$: e.g. PT or $PT \times \Lambda_+^\uparrow$.

These disconnected components of the Lorentz group are displayed in figure (1). The subset $\Lambda_+^\uparrow$ is continuously connected to the identity element and hence forms a group on its own right. This subgroup of the full Lorentz group is referred to as the ***Restricted Lorentz Group***. It is alternatively denoted by $SO^\uparrow(1,3)$ in some books.

Comment:

Note that if one allows complex Lorentz matrices, i.e. matrices with complex elements which satisfy (5) by extending the angles and boost parameters to the complex plane, then the two disconnected components $\Lambda_+^\uparrow$ and $\Lambda_+^\downarrow$. E.g. consider a boost along the x-direction by a rapidity parameter $\phi = i\pi$ (corresponding to a boost velocity, $\beta_1 = \cosh\phi = -1$), followed by a rotation around the x-axis by π radians, then the combined effect is same as the PT (Parity times Time reversal) transformation matrix.

$$PT = \begin{pmatrix} -1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

Thus allowing for complex valued proper Lorentz transformation connects the proper orthochronous with the proper nonorthochronous transformations.

3 Infinitesimal form of the restricted Lorentz transformations

Since the restricted Lorentz transformations are continuously connected to unity (i.e. they can be tuned to unity for some value of a continuous parameter, e.g, for the case of boosts by making the boost parameter/velocity zero), we can express them in the following *infinitesimal* form,

$$\Lambda^\mu{}_\nu \approx \delta^\mu_\nu + \omega^\mu{}_\nu. \tag{7}$$

Here $\omega^\mu{}_\nu$ are the infinitesimal or small boost velocity/ rotation angle. We have used an $\approx$ sign because we have dropped the $O(\omega^2)$ terms. Now using this infinitesimal form in the defining property of Lorentz transformation matrices (6) , we have

$$\left(\delta^\rho_\mu + \omega^\rho{}_\mu\right)\ \left(\delta^\sigma_\nu + \omega^\sigma{}_\nu\right)\ \eta_{\rho\sigma} = \eta_{\mu\nu}.$$

Expanding out the LHS,

$$\left(\delta^\rho_\mu + \omega^\rho{}_\mu\right)\ \left(\delta^\sigma_\nu + \omega^\sigma{}_\nu\right)\ \eta_{\rho\sigma} = \eta_{\mu\nu} + \omega_{\mu\nu} + \omega_{\nu\mu} + O(\omega^2).$$

Thus, equating linear order in ω on both sides we have,

$$\omega_{\mu\nu} + \omega_{\nu\mu} = 0. \tag{8}$$

So the matrix corresponding to covariant $\omega_{\mu\nu}$ is antisymmetric.

Further let's contract both sides of the anti-symmetry condition (8) by the inverse metric, namely, $\eta^{\mu\nu}$. This gives,

$$\eta^{\mu\nu}\omega_{\mu\nu} = 0,$$

or,

$$\omega^{\nu}{}_{\nu} = 0. \tag{9}$$

Thus, the mixed rank Lorentz parameter matrix, $\omega^{\mu}{}_{\nu}$ is traceless. This condition on ω can also be seen to arise from the condition of unit determinant,

$$\det\Lambda = +1.$$

For this we need the definition of the determinant,

$$\varepsilon_{\mu\nu\rho\sigma}\Lambda^{\mu}{}_{0}\Lambda^{\nu}{}_{1}\Lambda^{\rho}{}_{2}\Lambda^{\sigma}{}_{3} = \det\Lambda, \tag{10}$$

and we plug in the infinitesimal form (7) to get to first order in ω,

$$1 + \varepsilon_{\mu\nu\rho\sigma}\omega^{\mu}{}_{0}\delta^{\nu}_{1}\delta^{\rho}_{2}\delta^{\sigma}_{3} + \varepsilon_{\mu\nu\rho\sigma}\delta^{\mu}_{0}\omega^{\nu}{}_{1}\delta^{\rho}_{2}\delta^{\sigma}_{3} + \varepsilon_{\mu\nu\rho\sigma}\delta^{\mu}_{0}\delta^{\nu}_{1}\omega^{\rho}{}_{2}\delta^{\sigma}_{3} + \varepsilon_{\mu\nu\rho\sigma}\delta^{\mu}_{0}\delta^{\nu}_{1}\delta^{\rho}_{2}\omega^{\sigma}{}_{3} = 1$$

$$\implies \omega^{0}{}_{0} + \omega^{1}{}_{1} + \omega^{2}{}_{2} + \omega^{2}{}_{2} = 0. \tag{11}$$

i.e. $\omega^{\mu}{}_{\nu}$-matrix is traceless.

4 A Compendium of results on 4-vector component notation

- **Contravariant vectors:** We already have seen coordinate differentials transform homogeneously under Poincaré transformations,

  $$dx^{\mu} \to dx'^{\mu} = \Lambda^{\mu}{}_{\nu}\, dx^{\nu}.$$

  Contravariant vectors are four-component objects, $V^{\mu} = (V^0, V^1, V^2, V^3)$ which transform under Poincaré transformations exactly as coordinate differentials do,

  $$V^{\mu} \to V'^{\mu} = \Lambda^{\mu}{}_{\nu}\, V^{\nu}. \tag{12}$$

  We will often represent them as column vectors or row vectors and use a condensed notation for them, $V^{\mu} = (V^0, \boldsymbol{V})$ where V^0 is the time or temporal component and $\boldsymbol{V}$ is a 3-vector denoting the spatial part, $\boldsymbol{V} = (V^1, V^2, V^3)$.

- **Transformation of coordinate partials:** We already know how coordinate differentials, dx^{μ} transform under Lorentz transformation, namely Eq. (3). Here we are curious to know how do the coordinate partial derivative operators $\frac{\partial}{\partial x^{\mu}}$ transform after a Lorentz transformation? We will use a condensed notation denoting,

  $$\partial_{\mu} \equiv \frac{\partial}{\partial x^{\mu}}.$$

Let's say it is some transformation rule like this,

$$\frac{\partial}{\partial x'^{\rho}} = \Lambda_{\rho}{}^{\sigma}\frac{\partial}{\partial x^{\sigma}},$$

where we are yet to determine $\Lambda_{\mu}{}^{\nu}$. This can be easily determined by acting both sides on x'^{μ},

$$\begin{aligned}
\frac{\partial}{\partial x'^{\rho}}x'^{\mu} &= \Lambda_{\rho}{}^{\sigma}\frac{\partial}{\partial x^{\sigma}}x'^{\mu} \\
\delta^{\mu}_{\rho} &= \Lambda_{\rho}{}^{\sigma}\frac{\partial}{\partial x^{\sigma}}\left(\Lambda^{\mu}{}_{\nu}x^{\nu}\right) \\
&= \Lambda_{\rho}{}^{\sigma}\,\Lambda^{\mu}{}_{\nu}\frac{\partial x^{\nu}}{\partial x^{\sigma}} \\
&= \Lambda_{\rho}{}^{\sigma}\,\Lambda^{\mu}{}_{\nu}\delta^{\nu}_{\sigma} \\
\delta^{\mu}_{\rho} &= \Lambda_{\rho}{}^{\sigma}\,\Lambda^{\mu}{}_{\sigma}
\end{aligned}$$

This implies,

$$\Lambda_{\mu}{}^{\nu} = \left(\Lambda^{-1}\right)^{\nu}{}_{\mu} \tag{13}$$

- **Contravariant vectors:** Covariant vectors are 4-component objects denoted by a lowered index, $W_{\mu} = (W_0, W_1, W_2, W_3)$, which are defined by their transformation rule under a Poincaré transformation. These transform exactly like coordinate partials do,

$$W_{\mu} \to W'_{\mu} = \Lambda_{\mu}{}^{\nu}\,W_{\nu}. \tag{14}$$

- **Rank (p,q) tensors:** One can define a tensor of rank **(p,q)** as an object which has p-upper indices and q-lower indices. e.g. $T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_2}$. Such mixed upper and lower indexed tensor will transform with p-factors of the Lorentz transformation, Λ and q-factors of the inverse (transposed) transformation, Λ.

$$T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} \to T'^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} = \left(\underbrace{\Lambda^{\mu_1}{}_{\rho_1}\ldots\Lambda^{\mu_p}{}_{\rho_p}}_{p\text{-factors}}\right)\left(\underbrace{\Lambda_{\nu_1}{}^{\sigma_1}\ldots\Lambda_{\nu_q}{}^{\sigma_q}}_{q\text{-factors}}\right)T'^{\rho_1\ldots\rho_p}{}_{\sigma_1\ldots\sigma_q}$$

One such higher rank tensor which will prominent feature in our course is the Maxwell field strength tensor, $F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$, which is a $\mathbf{(0,2)}$ rank tensor. Evidently contravariant vectors are $(1,0)$ rank tensors and $(0,1)$ rank tensors while scalars are $(0,0)$ tensors.

- **Index contraction in products of vectors and tensors:** Consider the following combination which has no free indices,

$$W_\mu V^\mu.$$

How does this index-less object transform under a Lorentz transformation? Let's check,

$$\begin{aligned}
W_\mu V^\mu \to W'_\mu V'^\mu &= \left(\Lambda_\mu{}^\nu\, W_\nu\right)\left(\Lambda^\mu{}_\lambda\, V^\lambda\right)\\
&= \left(\Lambda_\mu{}^\nu\, \Lambda^\mu{}_\lambda\right)\, W_\nu V^\lambda\\
&= \left(\underbrace{\left(\Lambda^{-1}\right)^\nu{}_\mu\, \Lambda^\mu{}_\lambda}_{=\delta^\nu_\lambda}\right) W_\nu V^\lambda\\
&= W_\nu V^\nu.
\end{aligned}$$

Whenever we have such an expression with dummy indices, i.e. the same index appearing once upstairs and once downstairs, their Lorentz transformation cancel each other, we use the phrase that particular dummy index is **contracted**. This will be a generic feature, in general, when we will have a product of a certain number of tensors and/or vectors with some contracted (i.e. all repeated indices). The product will transform under a Lorentz transformation determined only by the free indices. E.g., the product,

$$A_{\mu\nu} B^\nu C^\mu{}_{\lambda\kappa}$$

will transform as if it only had the free indices λ and κ. The dummy indices μ and ν get contracted and will not matter at all! (HW: Check this)

- **The Metric as an INVARIANT (0,2) rank tensor**: Let's write this invariant squared spacetime interval in 4-vector component notation,

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = \eta_{\mu\nu} dx'^\mu dx'^\nu, \tag{15}$$

where we have introduced, the **Minkowski (flat) metric** $\eta_{\mu\nu}$. Comparing the RHS of (**??**) and (15), we see that we have to set

$$\begin{aligned}
\eta_{00} &= -1,\\
\eta_{11} = \eta_{22} = \eta_{33} &= +1,\\
\eta_{0i} = \eta_{i0} &= 0,\ i = 1, 2, 3\\
\eta_{ij} &= 0,\ i \neq j.
\end{aligned}$$

Right now there is no rigorous reason why we have chosen downstairs indices for η, because we have not checked whether it transforms like a $\mathbf{(0,2)}$-rank tensor yet! In fact, on the contrary we are attributing its components a fixed set of values, 0's and ±1's, in *all inertial frames!* It seems like the metric does not transform at all. Right now we have written two downstairs indices for η is because the interval ds^2 is a scalar, and for a scalar all indices contracted would look nice and dandy!

Now we show that the metric is indeed a tensor despite not changing form or values from frame to frame. Not all objects with indices can qualify to be a tensor so we need to check each time whether an object with indices attached to it is transforming like a tensor would do under Lorentz transformation. Let's look at the metric, $\eta_{\mu\nu}$. We know in all frames it is identical i.e. has the same diagonal matrix form. But it also has two indices downstairs, so can it be a (0,2) type tensor? To check this we momentarily assume that the metric is a$\mathbf{(0,2)}$-rank tensor and then look at what it transforms to under a LT. If it remains same, then we can indeed call the metric a (0,2) tensor. So let's see what the metric transforms into when after we apply a LT assuming it is a (0,2) tensor. Using (**??**),

$$\eta_{\mu\nu} \to \eta'_{\mu\nu} = \Lambda_\mu{}^\alpha\, \Lambda_\nu{}^\beta\, \eta_{\alpha\beta} \tag{16}$$

If we look a bit hard at the RHS we see that, we can rewrite it a bit using (13),

$$\begin{aligned}\text{RHS} &= \Lambda_\mu{}^\alpha\, \Lambda_\nu{}^\beta\, \eta_{\alpha\beta} \\ &= \left(\Lambda^{-1}\right)^\alpha{}_\mu \left(\Lambda^{-1}\right)^\beta{}_\nu\, \eta_{\alpha\beta} \\ &= \left(\Lambda^{-T}\right)_\mu{}^\alpha\, \eta_{\alpha\beta}\, \left(\Lambda^{-1}\right)^\beta{}_\nu \\ &= \left(\Lambda^{-T}\eta\, \Lambda^{-1}\right)_{\mu\nu} \\ &= \eta_{\mu\nu}\end{aligned}$$

where in going from the second last line to the last line we had replaced $\eta \to \Lambda^T\eta\,\Lambda$ using the defining equation of Lorentz matrices (5). Thus, we have

$$\eta_{\mu\nu} \to \eta'_{\mu\nu} = \eta_{\mu\nu}!$$

So indeed the metric is an invariant $\mathbf{(0,2)}$-rank symmetric tensor.

- **Invariant scalar product** of two contravariant vectors: One can define a scalar i.e. Lorentz invariant quantity which is bilinear in two contravariant vectors V and W as follows

$$V.W \equiv \eta_{\mu\nu} V^\mu W^\nu. \tag{17}$$

Since all the tensor indices are contracted it is obvious that this product is a Lorentz invariant (scalar), but one can readily check the invariance i.e. show that $V.W = V'.W'$ using (**??**) and (16).

- **"Lowering the index": Covariant vectors revisited**: Looking at (17) sort of prompts us to invent a covariant object (quantity with a downstairs index) from a contravariant vector by contracting it with the metric tensor,

$$W_\mu = \eta_{\mu\nu} W^\nu.$$

Again it is obvious from the index structure of the RHS of above equation that the LHS is a **(0, 1)**-type or covariant vector. This general rule i.e. contraction by $\eta_{\mu\nu}$ allows us to convert contravariant/upstairs indices into covariant/downstairs ones. E.g.

$$A_{\mu\nu} = \eta_{\mu\alpha}\eta_{\nu\beta}A^{\alpha\beta}$$

- **Inverse of the metric**: We had already noted before that the Minkowski metric is self-inverse, $\eta^{-1} = \eta$. Does this mean the that we can express η^{-1} by the same component notation as η i.e. $\eta_{\mu\nu}$. The answer is negative! Lets look at the index structure on both sides of the following equation.

  $$\eta^{-1}\eta = \mathcal{I}.$$

  The RHS is identity matrix and has component notation, Kronecker delta: δ^{μ}_{ρ} [4]. Therefor the LHS must have an unsummed/free upstairs index μ and a downstairs free index ρ. If ν be the repeated index needed for matrix multiplication, then we have,

  $$\left(\eta^{-1}\eta\right)^{\mu}_{\rho} = \left(\eta^{-1}\right)^{\mu\nu}\eta_{\nu\rho}$$

  So the inverse must have BOTH indices upstairs. Since η is self-inverse we will drop the $(\ )^{-1}$, and **denote the inverse by** $\boldsymbol{\eta^{\mu\nu}}$.

- **"Raising the index" with inverse metric:** Now that we have the metric inverse with purely upstairs indices, we can use that to convert covariant stuff to contravariant stuff.

  $$V^{\mu} = \eta^{\mu\nu}V_{\nu}.$$

  Another example would be to convert the Maxwell tensor with covariant indices, $F_{\mu\nu}$ to contravariant indices, $F^{\mu\nu}$,

  $$F^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}F_{\alpha\beta}.$$

- **Origin of the notation** $\boldsymbol{\Lambda_{\mu}{}^{\mu}}$**:** Here we justify why we were using the index structure/notation $\Lambda_{\mu}{}^{\nu}$ to represent the inverse transpose of the Lorentz matrices, Λ. First, multiplying both sides of (5) from the left by η^{-1}, we get,

  $$\left(\eta^{-1}\Lambda^{T}\eta\right)\Lambda = \mathcal{I}$$

[4] This follows from the identity transformation: If $x' = x$, then the transformation matrix should be identity. In component notation,

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} = \frac{\partial x^{\mu}}{\partial x^{\nu}} = \delta^{\mu}_{\nu}.$$

which means

$$\Lambda^{-1} = \eta^{-1}\Lambda^T\eta \tag{18}$$

Writing this is component form,

$$\left(\Lambda^{-1}\right)^{\mu}{}_{\nu} = \eta^{\mu\alpha}\,\Lambda^{\beta}{}_{\alpha}\,\eta_{\beta\nu}.$$

But since we have been using η and η^{-1} to lower and raise indices, we can do so for the RHS of the above equation and thus arrive at the equation,

$$\left(\Lambda^{-1}\right)^{\mu}{}_{\nu} = \Lambda_{\nu}{}^{\mu}. \tag{19}$$

Comparing this with the definition of the $\bar{\Lambda}$-matrices in (13) we arrive at the conclusion,

$$\bar{\Lambda}_{\nu}{}^{\mu} = \Lambda_{\nu}{}^{\mu}!$$

Henceforth we will get rid of the bar and use the symbol, $\Lambda_{\mu}{}^{\nu}$ for covariant transformation matrices.