PH2102 - Mechanics II

Rohit Verma

1 Index Notation for Relativity

To understand index notation used in special relativity, and other advanced fields of physics like quantum field theory, general relativity, etc., we first first start by understand the standard index notation for matrix. And then we'll change that a bit to get index notation we use of special relativity.

1.1 Index Notation for Matrices

Suppose you have a matrix A of size $m \times n$ (m rows and n columns). Then we write that matrix as

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

Now, we denote the ij-th element (the element which is in i-th row and j-th column) of this matrix A, as $[A]_{ij}$. For the above matrix $[A]_{ij} = A_{ij}$.

e.g., Suppose you have a matrix

$$R = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then $[R]_{11} = \cos(\theta), [R]_{13} = 0, [R]_{21} = -\sin(\theta), [R]_{33} = 1$, and so on. Now, we can use this index notation to write the component of product matrices and much more.

Suppose we have two ($n \times n$) matrices A and B, then we can write the ij-th component of the product matrix as

$$[AB]_{ij} = \sum_{k=1}^{n} [A]_{ik}[B]_{kj} \tag{1}$$

If, we have product of three square matrices A, B and C, then

$$[ABC]_{ij} = \sum_{k=1}^{n} \sum_{l=1}^{n} [A]_{ik}[B]_{kl}[C]_{lj} \tag{2}$$

When you have product of more and more matrices, which is very common in advanced physics, there will be more and more summation symbols. This makes simple expression look extremely long, and doesn't add much information to the equation. That's why adopt the convention where we drop summation symbol, and assume the repeated indices to be summed over, unless explicitly mentioned. So we can write above expressions using this convention as

$$[AB]_{ij} = [A]_{ik}[B]_{kj}, \qquad \& \qquad [ABC]_{ij} = [A]_{ik}[B]_{kl}[C]_{lj}$$

But when we write the indices downstairs (as subscripts) only, then another problem arises. We cannot distinguish between vectors (contravariant vectors) and co-vectors (covariant vectors, or dual vectors). That's why we adopt convention of writing indices upstairs (as superscript) and downstairs (as subscript) to distinguish the two cases.

1.2 Index Notation in Special Relativity

As sir has taught in class, the objects which we call vectors and scalars, are vectors and scalar under certain kind of transformations. For example, the 3 component vectors (or the 3-vectors), which we studied in school days are vector under rotation transformation. And the 4 component vectors (called 4-vectors) which are discussed in the class are vectors under Lorentz transformations.

Now, let's first write Lorentz transformation in index notation. Note that Lorentz transformation is a linear transformation, relating the coordinates (spacetime coordinates) of two inertial frames, sharing the same spacetime origin. So, we can write the spacetime coordinates or 4-coordinates (or 4-position vector), as a 4-vector $X = (ct, x, y, z)$. We can write the same in the matrix notation as

$$[X] = \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

Note that from here onwards, when we write X, we mean the 4-position vector, but when we write $[X]$, it means the same thing but in matrix representation. Similarly when I write $[A]$, this represents the mathematical object A (be it a vector, a covector, or a transformation) as a matrix.

We are writing ct instead of t to make the physical dimension same for all components. And ct is as good as t alone. Now, since Lorentz transformations is a linear transformation, we can write it as a matrix (4×4 size). Thus the 4-coordinates of in the new frames, say X' are related to that of old frame via some matrix Λ:

$$[X'] = [\Lambda][X], \quad \text{or} \quad \begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = [\Lambda] \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} \tag{3}$$

Note that sir used L to denote Lorentz transformation matrix, but I'm not doing so. Let's first write the component of 4-vector X in index notation. Note that X is a 4-vector, in fact it's the prototype of all 4-vectors. And opposite to what you have studied earlier, we write the index of vector (aka contravariant vector) upstairs or as a subscript. This is the convention. So, we write

$$[X] = \begin{bmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{bmatrix}$$

Note that this is different from standard index notation in matrices. In maths when you study matrices, you might write it as

$$[X] = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

Note the difference between the two:

1. When you write component of matrices, the index is written downstairs instead of upstairs. But we are writing it upstairs to denote the X^0, X^1, X^2 and X^3 are component of a vector, which is different from covector (which we will come to later).

2. In normal matrices, the index runs from 1 to some n, 4 in this case. But in case of our 4 vectors, the indexing starts from 0. This is again just a convention, which proves to be useful. 0 index represents the temporal component, and non-zero indices represent the spatial components.

So, we can write $[X]_0 = X^0$, $[X]_1 = X^1$, $[X]_2 = X^2$, and $[X]_3 = X^3$. The indices which start from 0 and go upto 3 are denoted by lower case Greek letters instead of Latin letter. So we write $[X]_\mu = X^\mu$. This means that the μ-th component of X, denoted as $[X]_\mu$ in matrix notation, is denoted as X^μ in our new notation.

Now, let's write the component of X' in matrix notation, and then in our new notation, and compare the two. The μ-th component of X' is

$$[X']_\mu = \sum_{\nu=0}^{3} [\Lambda]_{\mu\nu} [X]_\nu \tag{4}$$

Note that expression (4) is same as (3), just written in index notation of matrices. Next, lets go to the summation convention and drop the summation sign:

$$X'^\mu = [\Lambda]_{\mu\nu} X^\nu$$

Since, on LHS μ is upstairs, and downstairs on RHS. It looks inconsistent. Because by looking at RHS, you cannot tell whether the μ index will be upstairs, or downstairs on LHS. So we want μ upstairs on RHS as well. So in our new notations, we write

$$[\Lambda]_{\mu\nu} = \Lambda^\mu{}_\nu \tag{$\star_1$}$$

Thus our expression becomes

$$X'^\nu = \Lambda^\mu{}_\nu X^\nu \tag{5}$$

Note that expression (5) is same as expression (4), just written in our new notation instead of matrix notation used in mathematics. So, we write the $[\Lambda]$ matrix as

$$[\Lambda] = \begin{bmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{bmatrix}$$

For example, let's write second component of X' in terms of X as an example

$$X'^2 = \Lambda^2{}_\nu X^\nu = \Lambda^2{}_0 X^0 + \Lambda^2{}_1 X^1 + \Lambda^2{}_2 X^2 + \Lambda^2{}_3 X^3$$

Next, we define a 4-vector as a 4 component object, which transforms exactly like 4-coordinates under Lorentz transformation. Suppose A is a 4-vector. Then we denote the μ-th component of A as A^μ. Now, in different inertial frame, the μ-th component of A' is

$$A'^\mu = \Lambda^\mu{}_\nu A^\nu = \Lambda^\mu{}_0 A^0 + \Lambda^\mu{}_1 A^1 + \Lambda^\mu{}_2 A^2 + \Lambda^\mu{}_3 A^3$$

Exercise 1: Suppose we have 4-vector in one frame is $A = (2, 3, -1, 6)$, and consider the special Lorentz transformation with $\beta = 4/5$, thus $\gamma = 5/3$. Find all the components of vector A in transformed frame.

1.3 Inner Product

Inner product is also sometime known as scalar products. Because inner product of two vectors produces a scalar. But the first question is, what is a scalar? A scalar is a quantity, a single component object, whose value doesn't change under given transformation. A scalar is an invariant under the transformation. So a scalar under Lorentz transformation (let's call it 4-scalar, analogous to 4-vector) is a quantity which remains invariant under Lorentz transformation.

Next, how do we construct a quantity out of 4-vector which remains invariant under Lorentz transformation?

Again, let's start with X vector. We already know a combination of components of X, which is invariant under Lorentz transformation. That is

$$c^2t^2 - x^2 - y^2 - z^2$$

This is what we call as 'spacetime' interval, which is equivalent of distance between two points in 3d space. Just how distance between two points in 3d space doesn't change under rotation, similarly, this combination of spacetime coordinates doesn't change under Lorentz transformation.

Let's call this quantity as inner product of X with itself. That is

$$\langle X, X\rangle = (X^0)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2$$

If we write this inner product in terms of Matrices, then have introduce a new matrix, which we will called the *Minkowski metric* (or simply the metric). The Minkowski metric is diagonal matrix

$$[\eta] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Then the inner product becomes

$$\langle X, X\rangle = [X]^T[\eta][X] \tag{6}$$

Next, we will find how we can write this matrix multiplication in the index notation of matrices as

$$\langle X, X\rangle = [X^T]_\mu[\eta]_{\mu\nu}[X]_\nu$$

Now the components of X and X^T are the same, so we write it as

$$\langle X, X\rangle = [\eta]_{\mu\nu}X^\mu X^\nu$$

Since in our notation, repeated indices are summed when one in upstairs and one is downstairs. And both μ, ν indices of X are already upstairs. So we have to keep both μ, ν indices in η downstairs. That is $[\eta]_{\mu\nu} = \eta_{\mu\nu}$ $(\star_2)$, and the inner product in physicist's index notation becomes

$$\langle X, X\rangle = \eta_{\mu\nu}X^\mu X^\nu \tag{7}$$

Since this inner product two vectors is invariant under Lorentz transformation, that means

$$\langle X', X'\rangle = \langle X, X\rangle$$

Now using this, we can find a relation between Λ and η matrix

$$\langle X', X'\rangle = \langle \Lambda X, \Lambda X\rangle = [\Lambda X]^T[\eta][\Lambda X] = [X]^T[\Lambda]^T[\eta][\Lambda][X] = [X]^T[\eta][X] = \langle X, X\rangle$$

This implies that

$$\boxed{[\Lambda]^T[\eta][\Lambda] = [\eta]} \tag{8}$$

I've boxed this equation, because this is most important property of Lorentz transformation. Or you can say that this is the defining property of Lorentz transformation. Now let's write this particular relation index notation.

$$[\eta]_{\mu\nu} = [\Lambda^T]_{\mu\rho}[\eta]_{\rho\sigma}[\Lambda]_{\sigma\nu} = [\eta]_{\rho\sigma}[\Lambda]_{\rho\mu}[\Lambda]_{\sigma\nu}$$

We already know how the components of Λ and η matrix are written in physicist's index notation

$$\boxed{\eta_{\mu\nu} = \eta_{\rho\sigma}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu} \tag{9}$$

Now, because of this relation between Λ and η matrices, we can define an inner product between two 4-vectors, which will produce a scalar (an invariant quantity). The inner product between two 4-vectors A, B can be defined as

$$\langle A, B\rangle = [A]^T[\eta][B] = \eta_{\mu\nu}A^\mu B^\nu \tag{10}$$

Exercise 2: Verify that $\langle A', B'\rangle = \langle A, B\rangle$ using the above property of Λ (expression (8)).

This invariance of inner product under Lorentz transformation is very crucial. It tells you that once you know the value of inner product in one inertial frame, it is going to be the same in all other inertial frames.

Exercise 3: Starting from definition of inner product in index notation, i.e.,

$$\langle A, B\rangle = \eta_{\mu\nu}A^\mu B^\nu$$

And using the invariance of inner product, i.e., find the relation (9). *Warning: Don't use the matrix notation at all.*

Note that, since η is a symmetric matrix, which in index notation means $\eta_{\mu\nu} = \eta_{\nu\mu}$, that's why $\langle A, B\rangle = \langle B, A\rangle$

1.4 Fields and Vector Spaces

Let's take a little detour and talk about fields and vector spaces. The field which we will talk about is mathematician's field, and it is different from what a physicist called as field. A field is a set $\mathbb{F}$, where you can do addition, subtraction, multiplication, and division between any two elements (except division by zero). Most common example of fields are set of all rational numbers $\mathbb{Q}$, set of all real numbers $\mathbb{R}$, and set of all complex numbers $\mathbb{C}$. Next, we define a vector space

Definition: A vector space is a set V over the field $\mathbb{F}$ with two binary operations, $+ : V \times V \to V$ (*vector addition*), and $\cdot : \mathbb{F} \times V \to V$ (*scaling* or *scalar multiplication*) defined on it, which satisfies the following axioms:

1 For Vector Addition:

- $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (*Commutation*)
- $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ (*Association*)

- For every $\boldsymbol{u} \in V$, $\exists\, \mathbf{0} \in F \mid \boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$ (*Existence of Additive Identity*)
- For every $\boldsymbol{u} \in V$, $\exists\, \boldsymbol{v} \in V$ such that $\boldsymbol{u} + \boldsymbol{v} = \mathbf{0}$. This $\boldsymbol{v}$ is represented as $-\boldsymbol{u}$ (*Existence of Additive Inverse* or *Negative*)

2 For Scaling: ($a, b \in \mathbb{F}$ and $\boldsymbol{u}, \boldsymbol{v} \in V$)

- $(a \cdot b) \cdot \boldsymbol{u} = a \cdot (b \cdot \boldsymbol{u})$ for all $a, b \in \mathbb{F}$ and $\boldsymbol{u} \in V$ (*Association*)
- $a \cdot (\boldsymbol{u} + \boldsymbol{v}) = a \cdot \boldsymbol{u} + a \cdot \boldsymbol{v}$ for all $a \in \mathbb{F}$ and $\boldsymbol{u}$, $\boldsymbol{v} \in V$ (*Distribution over vector addition*)
- $(a + b) \cdot \boldsymbol{u} = a \cdot \boldsymbol{u} + b \cdot \boldsymbol{u}$ for all $a, b \in \mathbb{F}$ and $\boldsymbol{u} \in V$ (*Distribution over field addition*)
- $\exists\, 1 \in \mathbb{F} \mid 1 \cdot \boldsymbol{u} = \boldsymbol{u}$ for all $\boldsymbol{u} \in V$ (*Existence of Scaling Identity*)

The elements of this vector space set, is what mathematicians call as *vectors*. Now, the set of all spacetime coordinate (X^0, X^1, X^2, X^3) form a vector space. That vector space is our familiar set $\mathbb{R}^4$, since you can write the spacetime coordinate of any event as 4 real numbers, one component denoting time, and remaining 3 denoting the spatial position. So, our 4-vector is one kind of vector for them. Several other objects, which physicist's don't generally call as vectors, are vectors for mathematician. *e.g.,* Matrices of any dimensions.

Similarly, space of all possible 4-velocity vectors forms a vector space, which is again $\mathbb{R}^4$, but this vector space of velocity and the vector space of spacetime coordinates are not identical. They have the same mathematical structure (that of $\mathbb{R}^4$), but physically they are different. Because given a vector space, you can add any two elements of that, and get a new vector within that space. So if vector space of 4-position and vector space of 4-velocity were the same $\mathbb{R}^4$, then that would mean you can add a position vector with velocity vector, which is physically absurd.

1.5 Covectors

In linear algebra, there exist a kind object which are called dual vector, which sort of helps in defining inner product. Now I'm gonna make a bunch of statements without proof, which are standard results from linear algebra. You have to take that for granted right now. If you're interested in proof, you can look into any standard linear algebra textbook.

First the definiton: A *dual vector* is a map, which takes vectors as input, and gives you an output in a *linear fashion*. Suppose we have one such map, let's call it Φ, then $\Phi(A) \in \mathbb{R}$ where A is a vector general, or 4-vector in our case because we are studying Lorentz transformations. Φ is a linear map, means that for $c_1, c_2 \in \mathbb{R}$ and for two 4-vectors A and B (of same vector space), we have

$$\Phi(c_1 A + c_2 B) = c_1 \Phi(A) + c_2 \Phi(B) \tag{11}$$

Now, we can prove (but won't) that the set of all possible dual vectors form a vector space in it's own right. Suppose the vector space, from which Φ is taking input is V. Then set of all possible dual vectors on V, i.e., $\{\Phi : V \to \mathbb{R}\}$ is denoted by V^*. We claim here, without proof, that V^* is also a vector space of same dimension as V. And for every vector $\boldsymbol{v} \in V$, you can find a corresponding dual vector $\boldsymbol{v}^* \in V^*$.

Suppose we take the vector space of all possible spacetime coordinates. Let's denote it by $\mathbb{R}^4_X$, then the dual of it $(\mathbb{R}^4_X)^*$ is also same as $\mathbb{R}^4$, because we need four real number to describe a dual position vector, because dimension of dual vector space V^* is same as dimension of the vector space V. Thus, if we denote dual of position vector as X^*, then $(\mathbb{R}^4_X)^* = \mathbb{R}^4_{X^*}$. This dual vector is also called the *covector* or covariant vector.

Now, we have a very natural way of mapping the vector to a dual vector using the inner product. Given a inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$, we can define a map from $V \to V^*$ as

$$\boldsymbol{v} \mapsto \boldsymbol{v}^*(\cdot) \equiv \langle \boldsymbol{v}, \cdot \rangle$$

The job of a dual vector is to give a number from the field, when given a vector as input. So let's feed vector $\boldsymbol{u}$ to the dual vector $\boldsymbol{v}^*$

$$\boldsymbol{v}^*(\boldsymbol{u}) = \langle \boldsymbol{v}, \boldsymbol{u} \rangle \in \mathbb{F}$$

Thus $\boldsymbol{v}^*(\cdot) \equiv \langle \boldsymbol{v}, \cdot \rangle$ is a valid way of defining dual vectors.

Now, let's return to our 4-vector world. Using this same argument, we can say that the dual of 4-position vector, aka the spacetime covector X^* can me written as

$$X^*(\cdot) = \langle X, \cdot \rangle \tag{12}$$

Since X^* is a linear map, it can be realized as a matrix. That is, if $[X^*]$ denotes the column matrix representation of X^* covector, then

$$X^*(X) = \langle X, X \rangle = [X^*]^T [X] \tag{13}$$

But we already know that $\langle X, X \rangle = [X]^T [\eta][X]$. That means we can identify

$$[X^*]^T = [X]^T [\eta] \implies [X^*] = [\eta][X] \tag{14}$$

Let's find the μ-th index of X^*

$$[X^*]_\mu = [\eta X]_\mu = [\eta]_{\mu\nu}[X]_\nu = \eta_{\mu\nu} X^\nu$$

Now, in our convention, the μ-th index in downstairs (in the rightmost expression), so for consistency, we'll denote the component of dual 4-position vector by writing the indices downstairs. That is, $[X^*]_\mu = X_\mu$. We don't need to add star to denote covector, since the placing of index downstairs is sufficient to tell that it's a covector, and not a vector. We also got the rule of finding the components of covector from the given vector.

$$X_\mu = \eta_{\mu\nu} X^\nu \tag{15}$$

Note that the η, which was the metric that defines inner product also provide a natural mapping between vectors and covectors. For example, if we have to find the 0-th component of 4-position covector, the we write

$$X_0 = \eta_{00} X^0 + \eta_{01} X^1 + \eta_{02} X^2 + \eta_{03} X^3$$

Since only $\eta_{00} = 1$ is the only non-zero term in the above expression. Thus we find

$$X_0 = X^0$$

Next, let's find the X_1 component of 4-position covector

$$X_1 = \eta_{10} X^0 + \eta_{11} X^1 + \eta_{12} X^2 + \eta_{13} X^3 = -X^1$$

Since $\eta_{11} = -1$ is the only non zero term in this expression. Similarly we can find that $X_2 = -X^2$ and $X_3 = -X^3$. Thus, the 4-position covector $X^* = (X_0, X_1, X_2, X_3) = (X^0, -X^1, -X^2, -X^3)$.

Similarly, you can define the covector A^* corresponding to any 4-vector A, and find the components of A^* by application of η. The μ-th component of dual vector A^*, denoted as A_μ can be found as

$$A_\mu = \eta_{\mu\nu} A^\nu \tag{16}$$

This operation is usually called lowering of indices (done via application of the metric). Because we went from $A^\nu \to A_\mu$. Also, we can say that the inner product of 4-vectors A and B is basically product of A^* covector matrix, with matrix of B vector, that is,

$$\langle A, B\rangle = [A^*][B] = [A^*]_\mu [B]_\mu = A_\mu B^\mu = A_0B^0 + A_1B^1 + A_2B^2 + A_3B^3 \tag{17}$$

You can find that for metric η, The components of A^* covectors, in terms of components of vector A are

$$A^* = (A^0, -A^1, -A^2, -A^3)$$

This means

$$\langle A, B\rangle = A_0B^0 + A_1B^1 + A_2B^2 + A_3B^3 = A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3 \tag{18}$$

which is exactly how we defined inner product earlier.

Exercise 4: Suppose instead of η, we take a different metric g, given by following matrix

$$[g] = \begin{bmatrix} 1 & 1/2 & 0 & 0 \\ 1/2 & 2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

And we define the lowering of indices via this matrix as $A_i = g_{ij}A^j$. Now, given a 4 component vector $A = (A^1, A^2, A^3, A^4)$, find the components of corresponding dual vector in terms of component of this vector. Basically you have to compute A_i for $i = 1, 2, 3, 4$.

Note: This example is an arbitrary mathematical construct to help you understand importance of metric in lowering indices. It does not correspond to any physics of special relativity, because Lorentz transformation doesn't preserve this metric g.

Lastly, we know that $[X^*] = [\eta][X]$. Thus we can write $[X] = [\eta^{-1}][X^*]$. Now in matrix notation

$$[X]_\mu = [\eta^{-1}]_{\mu\nu}[X^*]_\nu \implies X^\mu = [\eta^{-1}]_{\mu\nu}X_\nu \tag{19}$$

Now, the μ on LHS is upstairs. So, for consistency, μ on RHS should be upstairs. Since ν is summed over, and we have already fixed that ν for covector component should be downstairs. That's why ν in η^{-1} should be upstairs as well. Thus, we write

$$X^\mu = \eta^{\mu\nu}X_\nu \tag{20}$$

Note that $\eta^{\mu\nu} = [\eta^{-1}]_{\mu\nu}$ ($\star_3$). We don't need to write -1 in superscript explicitly because indices written upstairs is sufficient to denote that we are talking about inverse of η. In case of special relativity, η and η^{-1} turns out to be the same. But that is not the case always, especially when you are dealing with GR.

In summary, η lowers the index (takes you from vector to covector), and η^{-1} raises the index (takes you from covector to vector.

1.6 Transformation of Covectors

In this section, we'll see how does the components of a covector changes under a Lorentz transformation. We already know that $[X^*] = [\eta][X]$. So, after Lorentz transformation, we get

$$[X'^*] = [\eta][X'] = [\eta][\Lambda][X]$$

Now, we know that

$$[\Lambda]^T[\eta][\Lambda] = [\eta] \implies [\eta][\Lambda] = ([\Lambda]^T)^{-1}[\eta]$$

This gives

$$[X'^*] = ([\Lambda]^T)^{-1}[\eta][X] = ([\Lambda]^T)^{-1}[X^*] = [(\Lambda^T)^{-1}][X^*] \tag{21}$$

That means, the transformation matrix for covectors is $(\Lambda^T)^{-1}$. Now let's find the $\mu\nu$-th component of $(\Lambda^T)^{-1}$. From above we write

$$[\eta][\Lambda] = [(\Lambda^T)^{-1}][\eta] \implies [\eta][\Lambda][\eta^{-1}] = [(\Lambda^T)^{-1}] \tag{22}$$

Thus

$$[(\Lambda^T)^{-1}]_{\mu\nu} = [\eta]_{\mu\rho}[\Lambda]_{\rho\sigma}[\eta^{-1}]_{\sigma\nu} = \eta_{\mu\rho}\Lambda^\rho{}_\sigma\eta^{\sigma\nu} \tag{23}$$

Since η lowers the index and η^{-1} raises the index, we write

$$[(\Lambda^T)^{-1}]_{\mu\nu} = \eta_{\mu\rho}\Lambda^\rho{}_\sigma\eta^{\sigma\nu} = \Lambda_\mu{}^\nu \tag{$\star_4$}$$

Note the difference:

- $\mu\nu$ component of Λ is written as $\Lambda^\mu{}_\nu$ (first index upstairs and second one downstairs).
- $\mu\nu$ component of $(\Lambda^T)^{-1}$ is written as $\Lambda_\mu{}^\nu$ (first index downstairs and second one upstairs).

Using ($\star_4$), we can find that

$$(\Lambda^T)^{-1} = \begin{bmatrix} \Lambda_0{}^0 & \Lambda_0{}^1 & \Lambda_0{}^2 & \Lambda_0{}^3 \\ \Lambda_1{}^0 & \Lambda_1{}^1 & \Lambda_1{}^2 & \Lambda_1{}^3 \\ \Lambda_2{}^0 & \Lambda_2{}^1 & \Lambda_2{}^2 & \Lambda_2{}^3 \\ \Lambda_3{}^0 & \Lambda_3{}^1 & \Lambda_3{}^2 & \Lambda_3{}^3 \end{bmatrix} = \eta\Lambda\eta^{-1} = \begin{bmatrix} \Lambda^0{}_0 & -\Lambda^0{}_1 & -\Lambda^0{}_2 & -\Lambda^0{}_3 \\ -\Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ -\Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ -\Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{bmatrix}$$

Exercise 5: Suppose that a Lorentz transformation matrix is given by

$$\Lambda = \frac{1}{64}\begin{bmatrix} 125 & -48 & -60 & 75 \\ -75 & 80 & 36 & -45 \\ -60 & 0 & 80 & -36 \\ 48 & 0 & 0 & 80 \end{bmatrix}$$

What is the value of $\Lambda^1{}_0, \Lambda^2{}_1, \Lambda^3{}_2, \Lambda^0{}_2$? Find the corresponding $(\Lambda^T)^{-1}$ matrix. What is the value of $\Lambda_1{}^0, \Lambda_2{}^1, \Lambda_3{}^2, \Lambda_0{}^2$?

1.7 Summary

- Vector: $[A]_\mu = A^\mu$
- Vector transformation matrix: $[\Lambda]_{\mu\nu} = \Lambda^\mu{}_\nu$
- Covector: $[A^*]_\mu = A_\mu$
- Covector transformation matrix: $[(\Lambda^T)^{-1}]_{\mu\nu} = \Lambda_\mu{}^\nu$
- Metric: $[\eta]_{\mu\nu} = \eta_{\mu\nu}$
- Inverse of Metric: $[\eta^{-1}]_{\mu\nu} = \eta^{\mu\nu}$
- Identity Matrix: $[1]_{\mu\nu} = \delta^\mu{}_\nu$