A First Look at Indices and Tensors

Patrick Wang
patrick.wang@physics.org

June 10, 2021

Contents

1 Changing Bases

Any arbitrary none-null vector $\vec{v} \in V$, where V is a vector space over $\mathbb{R}^n$ and has finite dimension $n = \dim V$ can be expressed in component form as follows:

$$\vec{v} = v_i \vec{e}_i$$

Suppose we now change from our old basis $\{\vec{e}_i\}$ to a new basis $\{\vec{f}_i\}$, such that the new basis is related to our old basis via the following relation:

$$\vec{f}_j = \vec{e}_j S_{ij} \tag{1}$$

By definition, (1) codifies a **linear transformation**. We can define an inverse matrix T_{jk} to the matrix S_{ij} that carries out the inverse transformation such that the following is true:

$$\vec{e}_k = \vec{f}_j T_{jk} = \vec{e}_i S_{ij} T_{jk} \tag{2}$$

Note that we assume $\det(S) \neq 0$.
Treating the change of basis as a *passive transformation*, *i.e.* leaving the vectors themselves untouched, it should be obvious that the components of the vector will correctly transform via the matrix T.

$$x'_j = T_{ji} x_i \tag{3}$$

1.1 Gradient... a vector yet?

Take a scalar field $\varphi : V \to \mathbb{R}$, unsurprisingly, the components in our $\vec{e}$-basis contain derivatives after the coordinates x_i, *i.e.:* $(\nabla\varphi)_i = \partial_i \varphi$; consequently, in the new system, nabla should contain derivatives w.r.t. the new components: $(\nabla\varphi)'_j = \partial'_i \varphi$. The transformation rule we are after is therefore one that can transform x'-derivatives to x-derivatives. Applying the chain rule will get us started:

$$\frac{\partial \varphi}{\partial x'_j} = \frac{\partial x_i}{\partial x'_j}\frac{\partial \varphi}{\partial x_i} = S_{ij}\frac{\partial \varphi}{\partial x_i} \tag{4}$$

Astonishingly, nabla does not transform the way we expected a vector to transform! We can reconcile with this discovery by considering what is meant by the gradient operator, a directional derivative. For example: if we start at a point $\vec{r}_0$ in space and would like to know how $\varphi(\vec{r}_0)$ varies as we travel in different directions, call that $\vec{h}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(\vec{r}_0 + t\vec{h}) = \frac{\partial \varphi}{\partial x_i} h_i =: \nabla_{\vec{h}}\varphi \tag{5}$$

$\nabla\varphi$ is a map, in which we insert a direction vector $\vec{h}$, and receive a real number, *i.e.* the speed $\nabla_{\vec{h}}\varphi$ with which φ changes in that direction. This map is necessarily linear in $\vec{h}$

We can classify the n-component objects we encountered so far by how they transform, either:

1. like a vector, or
2. like nabla

We can build on this by defining some new objects.

1.2 Linear Forms and Dual Space

Definition
A *linear form* on a real vector space V is a linear map $\bar{a}: V \to \mathbb{R}$. The set of all linear forms on a vector space V, $\{\bar{a}_i\}$ is called the *dual space* V^*.

The sum of any two linear forms is again a linear form, a telltale sign that we are in a vector space! And indeed the dual space is a vector space.

Suppose we devise a linear form $\bar{\omega}_j$, such that it returns the j-th component of a vector:

$$\bar{\omega}_j(\vec{x}) = x_j \tag{6}$$

Then the following can be inferred:

$$\bar{\omega}_j(\vec{e}_i) = \delta_{ji} \tag{7}$$

$\bar{\omega}_j$ is said to be the dual basis to $\vec{e}_i$. And it is also important to note that $\dim V^* = \dim V$. By defining a new dual basis $\bar{\chi}_k$ via $\bar{\chi}_k(\vec{f}_l) = \delta_{kl}$, we can extract the new components of α in $\{\bar{\chi}_k\}$:

$$\alpha'_j = \bar{\alpha}(\vec{f}_j) = \bar{\alpha}(\vec{e}_i S_{ij}) = S_{ij}\alpha_i \tag{8}$$

This is exactly the way that gradient operator transforms, as it is a linear form.

Key Point
Objects that transform according to (1) is called covariant and denoted with a lower index.
Objects that transform according to (3) is called contravariant and is assigned an upper index.

Note that Einstein summation convention will always contract an upper and a lower index.

- Component-form of vectors: $\vec{v} = v^i\vec{e}_i$, $\bar{\alpha} = \alpha_j\bar{\omega}^j$
- Vector insertion with one-form: $\bar{\alpha}(\vec{v}) = \alpha^j\bar{\omega}_j v^i\vec{e}_j = \alpha v^i\delta^i_j = \alpha_i v^i$.
- Change of basis: $\vec{f}_j = S^i_j\vec{e}_i$.
- Inverse matrices: S^i_j and T^i_j are matrix inverses of each other: $S^i_j T^i_j = \delta^i_j$, and $T^a_b S^b_c = \delta^a_c$.

We can now introduce a more abstract, multi-slotted linear machines:

1.3 Tensors

Definition
A *tensor of rank* $\begin{bmatrix} p \\ q \end{bmatrix}$ is a real valued function:

$$\mathfrak{T}: (V^* \times \ldots \times V^* \times V \times \ldots \times V) \to \mathbb{R} \tag{9}$$

The tensor $\mathfrak{T}$ has p one-form slots and q vector slots, which is linear in every slot.

1.3.1 Scalar Products

Using our newly devised object, a *tensor*, we can define the scalar product (aka metric) as a two-slot map:

$$g: V \times V \to \mathbb{R} \tag{10}$$

The metric has the following properties:

- symmetric, *i.e.* $g(\vec{u},\vec{v}) = g(\vec{v},\vec{u})$ for $\vec{u},\vec{v} \in V$.
- bilinear, i.e., linear in both slots: for vectors $\vec{u},\vec{v},\vec{w} \in V$, and number $\lambda,\mu \in \mathbb{R}$. $g(\lambda\vec{u} + \mu\vec{v},\vec{w}) = \lambda g(\vec{u},\vec{w}) + \mu g(\vec{v},\vec{w})$
- positive-definite, for any vector $\vec{v}: g(\vec{v},\vec{v}) \geq 0$.

Equipped with the metric in our 'newly devised' formalism, we can rigorously define what is meant by an *orthonormal basis.*

Definition
For a vector space V and a scalar product $g: V \times V \to \mathbb{R}$. A basis with the property:

$$g(\vec{e}_i,\vec{e}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

is called an *orthonormal basis.*

In an ONB, the matrices g_{ij} and g_{ji} are simply the unit matrix and thus: $a_i = a^i$. The isomorphism between V and V^*, in this case, is canonical.

1.4 Rotations

A linear transformation on V which send an ONB to another ONB is a rotation. Rotations are characterised by orthogonal matrices. By means of extracting components from a vector: *i.e.*: $g(\vec{v}, \vec{e}_j)$, we can extract the components of the rotation matrix S^i_j: $S^i_j = g(\vec{e}_i, S^k_j \vec{e}_k) = \vec{e}_i \cdot \vec{f}_j$. From this, we can write down a general form of S:

$$S = (S^i_j) = \begin{bmatrix} \vec{e}_1 \cdot \vec{f}_1 & \dots & \vec{e}_1 \cdot \vec{f}_i \\ \vdots & \ddots & \vdots \\ \vec{e}_i \cdot \vec{f}_1 & \dots & \vec{e}_i \cdot \vec{f}_i \end{bmatrix} \quad (12)$$

Note that the column vectors are orthogonal to each other, and they are normalised. The columns contain the information f_i in the old system, and the rows contain information on e_i in the new system. This gives origin to the term **orthogonal matrix.**

Looking at the determinant of S, and the fact that $SS^T = \mathbb{1}$, we arrive at the fact that $\det S = \pm 1$. The positive determinant correspond to a *special orthogonal rotation,* in which there are no reflections. Consequently, the negative determinant corresponds to the reflective rotations. **The set of all orthogonal matrices is called SO(n).**

2 Special Relativity

Orthogonal transformations are important in SR: namely to see it as the transformation of a rank $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ tensor g that is our scalar product. To sum up, *orthogonal transformations are exactly those that leave the scalar product invariant.* This is intuitive as we think of rotations as active transformations of a rigid body, requiring the distances and angles between points to remain constant, *i.e.* non-deforming. By applying what we have obtained up to this point, we can carry our physics to the new arena 'spacetime' where velocity, momentum, energy, and fields become four-dimensional quantities: vectors and tensors in $\mathbb{R}^4$!

2.1 Indices in 4D

So far, our setup has been in a n-dimensional Euclidean geometry, where:

- the underlying space V is flat and rigid,
- distances, lengths, and angles are defined through a positive-definite scalar product.

Definition

Spacetime M is a four-dimensional real vector space.[a] Its elements are events that we specify by one value for the temporal coordinate $X^0 = t$ and three values for the spatial coordinates $X^i = (x, y, z)$. **Note the contravariant nature of the components.** The convention is to use Greek indices μ, ν, which run from 0 to 2.

[a] M in honour of H. Minkowski

Returning a previously meaningful physical quantity in this relativistic regime of ours, the interval. For two events $A, B \in M$ located at A^μ and B^μ respectively, they are connected by a vector $X^\mu = B^\mu - A^\mu$; the interval between these two events is given by: $I(A,B) = (X^0)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2 = (\Delta t)^2 - (\Delta x)^2$. This is usually written with the help of the Minkowski metric:

$$\eta_{\mu\nu} = \begin{bmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (13)$$

Similar to g, η is a symmetric rank-2 tensor, but it is not positive definite: there are vectors $X^\mu \in M$ such that $\eta(\vec{X}, \vec{X}) = 0$ or even less than 0! (Flash back to time-like, space-like separation, and the light cone!) The change to the 'definiteness' if precisely shown in Sylvester's Theorem of Inertia.

In this new regime, certain aspects of the Euclidean paradigm are no longer true: In Euclidean 4D geometry, which is defined through the signature[1] of g, which is $[1,1,1,1]$, whereas in Minkowski geometry, the signature, define through η, is $[1,-1,-1,-1]$. This signature is known as the *Lorentzian*.

Interpretation of angles and lengths, as we are familiar in plane-geometry is no longer possible. We shall denote length with a new term 'squared length', *i.e.* $\eta_{\mu\nu} X^\mu X^\nu$. This is used to characterise the relationship between two events connected by $\vec{X}$. For $\eta_{\mu\nu} X^\mu X^\nu > 0$, the events are timelike separated (*i.e.* X^μ is within either the future of past

[1] The signature is composed of the entries on the diagonal of a matrix.

light cone of the origin), and when $\eta_{\mu\nu}X^\mu X^\nu < 0$, the events are spacelike separated (where no exchange of causal influence is possible). Forcing the Euclidean idea that $\sqrt{\eta_{\mu\nu}X^\mu X^\nu}$ is some kind of length would require us to live with some vectors having imaginary length, and some vectors to have zero length despite them not being the zero vector (namely the lightlike ones).

Another difference we will have to account for is that despite the fact that we are in an ONB, our contravariant and covariant components are no longer equal. Recall the transformation of indices using g_{ij}:

$$x_i = g_{ij}x^j \tag{14}$$

$$g_{ij} = \mathbb{1} \text{ From ONB} \tag{15}$$

In Minkowski space, $X_0 = X^0$, but $X_i = -X^i$, this is again, due to the sign-change that we introduced in the Lorentzian signature.

2.2 Boosts

A change of inertial observer will relate to a change of coordinate system. And since the observers are in inertial FoR, they will need to agree if a worldline is straight or not, *i.e.* a linear change is necessary. Thus we can devise the following transformation:

$$X'^\delta = \Lambda^\delta_\gamma X^\gamma \tag{16}$$

where Λ is the analogue to T from Euclidean geometry, similarly, Λ transforms the contravariant components. By finding the inverse of Λ, we can find the object that transform the covariant components. Recall:

$$S^T S = S^T g S = g' = \mathbb{1} \tag{17}$$

which we can adapt to the Lorentzian signature, by adding a factor of -1 whenever there is a change in the position of a spatial index. Suppose we call the inverse V^ρ_σ, the Minkowski-analogue to S. The transform will look like the following:

$$\eta'_{\mu\nu} = V^\mu_\alpha V^\nu_\beta \eta_{\mu\nu} \tag{18}$$

$$= V_{\nu\alpha} V^\mu_\beta \tag{19}$$

$$= \eta_{\alpha\gamma} V^\gamma_\nu V^\nu_\beta \tag{20}$$

$$= \eta_{\alpha\gamma} \delta^\gamma_\beta \tag{21}$$

Given a Lorentz boost Λ^γ_ν, its inverse is Λ^ν_γm this is the equivalent of the Euclidean $T^{-1} = T$, which is true for orthogonal matrices. Two properties arise from this definition:

$$\Lambda^\gamma_\nu \Lambda^\nu_\beta = \delta^\gamma_\beta \tag{22}$$

$$\eta_{\alpha\beta} = \Lambda^\mu_\alpha \Lambda^\nu_\beta \eta_{\mu\nu} \tag{23}$$

Key Point
Lorentz boosts are transformations that leave the Minkowski metric invariant.

2.3 Lorentz Group

Similar to SO(3), the full set of 4×4 Lorentz transformations also form a group. Let R be an arbitrary 3×3 rotation matrix. Then:

$$\Theta^\alpha_\beta = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & R & \\ 0 & & & \end{array}\right]$$

is a 4 by 4 matrix which leave the time untouched, but rotating space. Such is the case then, the set of all those matrices can be parametrised by the three numbers (*e.g.* the Euler angles or Cayley-Klein parameters). Similarly, the set of all Lorentz boosts can also be parametrised by three numbers, (u_x, u_y, u_z). In general, the Lorentz boost may be written as follows:

$$\Lambda^\alpha_\beta = \left[\begin{array}{c|ccc} \gamma & \gamma u_x & \gamma u_y & \gamma u_z \\ \hline \gamma u_x & & & \\ \gamma u_y & & \mathbb{1} + \frac{\gamma-1}{u^2}\vec{u}\cdot\vec{u}^T & \\ \gamma u_z & & & \end{array}\right] \tag{24}$$

where $(\vec{u}\cdot\vec{u}^T) = u_i u_j$, a 3 by 3 matrix.

If we build all possible products of boosts Λ and rotations Θ, we can span the full six-dimensional *Lorentz group*: the group of all linear spacetime transformations that leave the Minkowski metric invariant.

Key Point

The Lorentz group is called SO(1,3) and it is one of the most important objects in mathematical physics. Special relativity is nothing more than the study of the symmetry group SO(1,3). From here, making the switch to general relativity becomes not impossible.
Suppose the Minkowski metric is replaced by a physical field $g_{\mu\nu}$, such that the components vary from place to time, or event to event. Such a scalar product could encode, amongst other things, the local curvature of spacetime. As long as the symmetric property is enforced everywhere, there will be an additional ten degrees of freedom. The 'only' thing left to do now is to specify the dynamics of these ten new degrees of freedom, which, if successfully completed, will yield general relativity.

3 Electrodynamics

Recall:

$$\nabla \cdot \vec{B} = 0 \tag{25}$$

$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \tag{26}$$

$$\nabla \cdot \vec{E} = 4\pi\rho \tag{27}$$

$$\nabla \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c}\vec{J} \tag{28}$$

where the first two of the four are the homogeneous Maxwell equations, and the later two are the inhomogeneous equations. Charge density ρ and current density $\vec{J}$ are connected through the requirement of the total charge $Q = \int \rho \, \mathrm{d}V$ remain constant. The charge can only change by a flux of charged flowing across the boundary $\partial\Omega$ of our volume V that we are integrating over. This leads to the charge conservation equation:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \nabla \cdot \vec{J} = 0 \tag{29}$$

Also recall, for a field $\vec{B}$that is solenoidal, *i.e.* $\nabla \cdot B = 0$ $\exists \vec{A}$ such that $\vec{B} = \nabla \times \vec{A}$. Similarly, for a conservative field, $\vec{E}$, *i.e.* $\nabla \times \vec{E} = 0, \exists \varphi$ such that $\vec{E} = \nabla\varphi$.

3.1 Gauge Transformation

By expressing the Maxwell equations using the scalar and vector potentials, we discover that these potentials are 'many-to-one'. *i.e.* there are many different potentials that lead to exactly the same fields. More precisely: two sets of potentials will yield the same $\vec{E}$ and $\vec{B}$ exactly if they are related by a **gauge transformation:**

$$\varphi_2 = \varphi_1 + \frac{1}{c}\frac{\partial \zeta}{\partial t} \tag{30}$$

$$\vec{A}_2 = \vec{A}_1 - \nabla\zeta \tag{31}$$

for $\zeta(t, \vec{x})$ is a scalar function. Inserting the potentials into the inhomogeneous equations, we recover the wave equations:

$$\Box\varphi = 4\pi\rho \tag{32}$$

$$\Box\vec{A} = 4\pi\vec{J} \tag{33}$$

To arrive at the equations above, we need to gauge-transform the potentials judiciously, such that they satisfy the *Lorenz gauge condition*:

$$\frac{1}{c}\frac{\partial \varphi}{\partial t} = \nabla \cdot \vec{A} = 0 \tag{34}$$