Special Relativity: Lecture 3

1 Four-Vector Formulation

A great advance was made in the presentation of classical physics with the introduction of vector notation. Consider equations such as $\mathbf{F} = m\mathbf{a}$ or $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$. In addition to being clear and concise the formulation is also independent of a particular choice of coordinate frame. Vector equations are covariant with respect to a rotation of the coordinate axes such as $\mathbf{a}' = \mathbf{R}.\mathbf{a}$ (or $a'_i = \sum_j R_{ij} a_j$), since

$$\begin{aligned}
\text{if } \mathbf{F} &= m\mathbf{a}, \text{ then} \\
\sum_j R_{ij} F_j &= m \sum_j R_{ij} a_j, \\
\text{or } \mathbf{F}' &= m\mathbf{a}',
\end{aligned}$$

so the covariance is manifest.

SUMMATION CONVENTION: Since one is repeatedly writing formulae involving summation over indices it is useful to have a convention where the Σ sign can be dispensed with. *Repeated indices are summed.* Usually Roman indices will be used for 3-vectors and Greek indices for 4-vectors.

Mathematically we can use the way an object transforms under a rotation to decide whether it is a 3-vector or scalar etc. Consider $\mathbf{a}.\mathbf{b}$ for example; if $\mathbf{a}$ and $\mathbf{b}$ are 3-vectors then $\mathbf{a}.\mathbf{b}$ is a scalar under a coordinate rotation.

Our aim is to extend these ideas to a four-dimensional world where time is taken to be the fourth component of a vector whose first 3 components are spatial. To keep things dimensionally correct we will use ct rather than t.

In Euclidean 3-space the length of a vector is an invariant: $r^2 = \mathbf{r}.\mathbf{r} = x^2 + y^2 + z^2$. Now we have seen that the interval between two space-time points is an invariant wrt the Lorentz transformation

$$S_{12}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2 (t_1 - t_2)^2 .$$

Apart from the minus sign the extension from 3- to 4-vectors would seem to be straightforward.

Strictly the presence of the minus sign means that we have to work with *non-Euclidean space*. Suppose we define a 4-vector by $a^\mu = (\mathbf{a}, a_4)$ — we call this a *contravariant vector*. To define the 4-vector scalar product we also need to define a *covariant vector* $b_\mu = (\mathbf{b}, -b_4)$ (NB: "covariant" here is another use of the word!) and a metric tensor $g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$.

The scalar product is then

$$\begin{aligned} a.b &= a^{\mu}b_{\mu} = \mathbf{a}.\mathbf{b} - a_4 b_4 \\ &= g_{\mu\nu}a^{\mu}b^{\nu}. \end{aligned}$$

This is the correct way to do things and if one is going to work a lot with relativistic EM it is sensible to use it. For general relativity it is essential.

However for special relativity — particularly elementary mechanics and EM — it is not essential and we shall not use the full contra-covariant formalism. Instead we shall cheat slightly and introduce $i = \sqrt{-1}$ into the definition of the fourth component: $x_\mu = (\mathbf{x},\, ict)$. In effect we have defined a pseudo-Euclidean 4-vector which can be manipulated exactly as a 3-vector. So

$$a.b = (\mathbf{a},\, ia_4)\,.\,(\mathbf{b},\, ib_4) = \mathbf{a}.\mathbf{b} - a_4 b_4.$$

2 Lorentz Transformation

Writing $x = (\mathbf{x},\, ict)$ and $x' = (\mathbf{x}',\, ict')$, eqs.(II.6) for the LT between IFs S and S' can be expressed as a matrix equation `(compare with a 3-vector rotation)`:

$$x'_\mu = L_{\mu\nu}x_\nu,$$

or

$$\begin{pmatrix} x' \\ y' \\ z' \\ ict' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ ict \end{pmatrix}. \qquad (III.1)$$

We can establish properties of the LT matrix $L_{\mu\nu}$. Since we defined the 4-vector in such a way that its length is invariant:

$$x'_\mu x'_\mu = L_{\mu\nu}x_\nu\, L_{\mu\sigma}x_\sigma = x_\sigma x_\sigma,$$

we see that

$$L_{\mu\nu}L_{\mu\sigma} = \delta_{\nu\sigma}, \qquad (III.2)$$

where

$$\delta_{\nu\sigma} = \begin{matrix} 1 \\ 0 \end{matrix} \quad \text{if} \quad \begin{matrix} \nu = \sigma \\ \text{otherwise} \end{matrix} \qquad \begin{pmatrix} \text{Kronecker} \\ \delta \text{ function} \end{pmatrix}.$$

From this relation the inverse of L can be defined:

$$\begin{aligned} L^{-1} &= L^{\mathrm{T}}, \\ x_\mu &= L^{\mathrm{T}}_{\mu\nu}x'_\nu. \quad \texttt{(check it!)} \end{aligned}$$

All this is of course exactly what is done for 3 vectors under rotations.

3 Covariance

A most important use of the 4-vector formalism is that if an equation can be written in terms of 4-vectors then we know that it will transform correctly under the LT — it is said to be *manifestly covariant.* `(This is also true for tensors - essential for formulating General Relativity.)`

Definitions

- **Scalar** A quantity that has the same numerical value in all IFs, e.g. $A.B$ or the length of a 4-vector.
- **4-vector:** Any quantity A_μ with 3 real components and 1 imaginary component transforming as $A'_\mu = L_{\mu\nu} A_\nu$ under the LT $L_{\mu\nu}$.
- **Tensor:** A 2nd rank tensor $F_{\mu\nu}$ for example is a 16-component object that transforms by two LFs: $F'_{\mu\nu} = L_{\mu\sigma} L_{\nu\lambda} F_{\sigma\lambda}$.

Theorem:

If A is a 4-vector and the quantity $A.B$ (where B is a 4 component object) is a scalar (wrt the LT $L_{\mu\nu}$) then B is also a 4-vector. `(can you prove this?)`

These ideas give us quick and powerful ways to get at physical results as we shall see repeatedly from now on.

Example:

Consider the phase ϕ of an EM wave $\phi = \mathbf{k}.\mathbf{x} - \omega t$. As we have already argued (for the Galilean transformation but the same will apply here) ϕ must be an invariant wrt Lorentz transformations — ϕ is proportional to a number of wave crests and will be the same in any frame. Since $(\mathbf{x},\, ict)$ is a 4-vector the combination $(\mathbf{k},\, i\omega/c)$ must also be a 4-vector. This allows us to write down immediately how it will transform:

$$k = \left(\mathbf{k},\, \frac{i\omega}{c}\right), \qquad k' = \left(\mathbf{k}',\, \frac{i\omega'}{c}\right)$$

and so

$$\begin{pmatrix} k'_x \\ k'_y \\ k'_z \\ i\omega'/c \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} k_x \\ k_y \\ k_z \\ i\omega/c \end{pmatrix}$$

or, in component form,

$$\begin{aligned} k'_x &= \gamma\left(k_x - \beta\frac{\omega}{c}\right) \\ k'_y &= k_y \\ k'_z &= k_z \\ \frac{\omega'}{c} &= \gamma\left((\frac{\omega}{c} - \beta k_x\right) \end{aligned} \qquad (III.3)$$

The last equation can be used to give the relativistic Doppler formula

$$\omega' = \gamma\omega(1 - \beta\cos\theta).$$

For EM radiation propagating at an angle θ wrt Ox in S, notice that ω and θ are measured in the *same* IF. Even if $\theta = \pi/2$ there is still an effect from the γ factor — the *transverse Doppler shift*. We shall return to this later.

The 4-vector k is an example of a *null vector* since $k^2 = k.k = \mathbf{k}^2 - \omega^2/c^2 = 0$. (However this does not necessarily imply that $k = 0$!)

Derivatives:

1. The derivative wrt a scalar parameter e.g. time $\dfrac{\partial x_\mu}{\partial t}$ can be defined in an IF S but as t is not a Lorentz scalar it will *not* transform as a 4-vector. Use, instead of t, the proper time τ which is a scalar. Then $\dfrac{\partial x_\mu}{\partial \tau}$ *is* a 4-vector.

2. $\dfrac{\partial\phi}{\partial x_\mu}$ where ϕ is a scalar function:

   By the chain rule $\dfrac{\partial}{\partial x'_\mu} = \dfrac{\partial x_\nu}{\partial x'_\mu}\dfrac{\partial}{\partial x_\nu}$ (NB: summation convention)

$$\begin{aligned} \text{now } x'_\mu &= L_{\mu\nu}x_\nu, \text{ but first invert this as} \\ L_{\mu\sigma}x'_\mu &= L_{\mu\sigma}L_{\mu\nu}x_\nu = \delta_{\sigma\nu}x_\nu = x_\sigma, \\ \text{so that } \frac{\partial x_\sigma}{\partial x'_\mu} &= L_{\mu\sigma}, \\ \text{and therefore } \frac{\partial}{\partial x'_\mu} &= L_{\mu\nu}\frac{\partial}{\partial x_\nu}\ , \end{aligned} \qquad (III.4)$$

   which shows that $\dfrac{\partial\phi}{\partial x_\mu}$ transforms as 4-vector.

3. $\dfrac{\partial V_\mu}{\partial x_\mu}$, the 4-divergence, transforms as a scalar. `(can you show this?)`

4. $\dfrac{\partial^2}{\partial x_\mu \partial x_\mu}$, the *d'Alembertian*, often denoted by $\Box$:

$$\Box \equiv \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \qquad (III.5)$$

   is a scalar under the Lorentz transformation.