FOUR VECTORS

SPECIAL AND GENERAL RELATIVITY

V.H. BELVADI*

THERE ARE TWO ironically *absolute* truths to special relativity. We call these the postulates of special relativity:

1. **The principle of relativity.** No experiment can measure the absolute velocity of an observer.

2. **The speed of light is universal.** The speed of light relative to any unaccelerated frame of reference is $3 \times 10^8\ \mathrm{ms}^{-1}$, regardless of its motion relative to an observer.

The first of these comes from Galileo while the second comes from Einstein. The second postulate can also be stated in a more interesting manner: *there is no frame of reference in which light is at rest*. We will soon see how this works.

1 FOUR VECTORS

The vectors you have come across so far look like $\mathbf{V} \equiv (V_x, V_y, V_z)$ and have three dimensions. Now we generalise it to $V^\mu \equiv (V^0, V^1, V^2, V^3)$ with four coordinates to get what is known as a **four-vector**. The usual three-dimensional vectors are called 'three-vectors', but this is not a term you will come across too often. In special relativity, working in Minkowski spacetime like we have been doing all this while, we will almost always use four-vectors. We will also refer to specific four-vectors with the prefix 'four-', e.g. four-momentum, four-velocity etc. There are three things to note before we begin:

1. Four-vectors have four components, all of which must carry the same dimensions or must all be dimensionless.

2. Four-vectors add just like three-vectors, so the triangle rule, the parallelogram rule etc. are all still valid here too.

3. You can always split a four-vector V^μ into two terms as $V^\mu \equiv (V^0, V^i)$, the 'temporal' component plus the 'spatial' three-vector, called so despite their dimensions being alike.

It might take you a while to quickly see how equations with indices translate into more familiar algebraic equations. It helps to keep in mind that the choice of indices is not special. You can choose to replace one index by another and, so long as you replace that index similarly everywhere in your equation, the meaning of the equation itself does not change. Practise by writing out expanded equations with numbered indices then converting them into general forms and vice versa.

For more visit `vhbelvadi.com/teaching`.

*vh@belvadi.com

Just like you have $\hat{i} \equiv (1,0,0)$, $\hat{j} \equiv (0,1,0)$ and $\hat{k} \equiv (0,0,1)$ as the unit vectors, or basis vectors or basis set or simply basis, of the standard three-dimensional Cartesian space, there are also four-dimensional basis vectors $e^0 \equiv (1,0,0,0)$ etc. You will notice that, taking into account a negative sign for ct, each row of the metric $\eta_{\mu\nu}$ simply corresponds to a basis four-vector. You can represent *the components* of a basis vector more simply using the Kronecker delta function. For e^β the components are given by

$$e^\beta_\alpha \equiv \delta^\beta_\alpha\,;\ \alpha = \{0,1,2,3\}$$

which means, for some $e^1_\alpha \equiv (e^1_0, e^1_1, e^1_2, e^1_3)$ the term $e^1_1 = 1$ while all others terms are zeroes, leaving us with $e^1 \equiv (0,1,0,0)$ as expected, corresponding to the row $\eta_{1\nu}$ of our metric tensor. Another way to look at the **Minkowski tensor** η is to think in terms of these basis vectors. That is, the nth row of η corresponds to e^n.

A reminder again: these indices are just notational superscripts, not powers.

There is another relationship between select three-vectors and four-vectors that will be of great use to us in the future. We can write down the proper time in a frame as follows:

*You probably know this result as the phenomenon called **time dilation**, where coordinate time lags behind proper time by γ.*

$$\begin{aligned} d\tau^2 &= dt^2 - \left(dx^2 + dy^2 + dz^2\right) \\ &= dt^2\left(1 - v^2\right) \\ \implies \frac{dt}{d\tau} &= \gamma \end{aligned} \tag{1}$$

Remember that γ is a function of the three-velocity and not the four-velocity, a term we will encounter later, but one that can nonetheless confuse those who are new to this field of study.

1.1 *Four-velocity*

The velocity four-vector, or simply the **four-velocity**, is the the proper-time derivative of the Minkowski coordinate x^μ. This works because x^μ transforms like a four-vector under Lorentz transformations. We define

$$U^\mu = \frac{dx^\mu}{d\tau} \tag{2}$$

where U^μ is the four-velocity described by the usual $\mu = 0\ldots3$ four-vector components. Subsequently the quantity $dU^\mu/d\tau$ is called the four-acceleration. Further, we can relate this to the three-velocity rather quickly using eq. (1) as follows:

$$U^\mu = \frac{dx^\mu}{d\tau} = \left(\frac{dx^\mu}{dt}\right)\frac{dt}{d\tau} = \left(\frac{dct}{dt}, \frac{dx^i}{dt}\right)\gamma = (c, \mathbf{v})\,\gamma \tag{3}$$

Besides the result also pay attention to the notation here: we will use the pair form (,) to refer to the zeroeth (or *timelike*) component separated from the

three-dimensional (or spatial) components, x, y, z. In this case $(c, \mathbf{v})$ suggests the *timelike* component is c while the *spatial* component is $\mathbf{v}$. Note, we dropped natural units here, for the time being, in favour of clarity.

1.2 *Four-acceleration*

What about acceleration as a four-vector then? The same technique works with that too—and we will employ our shorthand again now:

$$A^\mu = \frac{dU^\mu}{d\tau} = \left[\frac{dU^\mu}{dt}\right]\gamma = \left[\frac{d}{dt}\gamma(c, \mathbf{v})\right]\gamma = \left(c\frac{d\gamma}{dt}, \mathbf{v}\frac{d\gamma}{dt} + \mathbf{a}\gamma\right)\gamma$$

where $\mathbf{a} = d\mathbf{v}/dt$ is the three-acceleration and γ is the usual Lorentz factor. Keep in mind, γ is explicitly a function of the three-velocity alone i.e. $\gamma \equiv \gamma(v)$ and *not* $\gamma(U^\mu)$.

There is an interesting relationship between the four-velocity and the four-acceleration. For a particle at rest note that $U^\mu = (c, 0)$ and $A^\mu = (0, \mathbf{a})$. The former is obvious: $\mathbf{v} = 0$ for a particle at rest but information from it continues to travel at c; for the latter observe that photons at c cannot accelerate so there is no (or zero) acceleration for something already moving at c. Consequently we have

$$U^\mu A_\mu = 0 \tag{4}$$

as another invariant quantity. Geometrically this can be hard to visualise instinctively: this equation tells us that the four-velocity and four-acceleration of a particle are orthogonal to each other. In relation to a spacetime diagram, the velocity is tangential to the world line of a particle and the acceleration is perpendicular to that velocity.

1.3 *Four-momentum and the mass-energy-momentum equivalence*

Our next topic of discussion is the important idea of relativistic energy, or the mass-energy-momentum equivalence. En route we will also talk about four-momentum. The four-momentum p^μ is simply the product of mass and four-velocity much like the three-momentum you are already familiar with.

*On 30 June 1905 Albert Einstein published his famous paper on special relativity titled 'On the electrodynamics of moving bodies' in Annalen der Physik. This paper is simple enough for you to understand and I urge you to read it: see [**?**].*

$$p^\mu = m_0 U^\mu = \gamma m_0 (c, \mathbf{v}) \tag{5}$$

Focus on our use of m_0 for mass instead of just m; this refers to a quantity similar to proper and coordinate times but for masses. That mass of an object which is invariant across all frames of reference is called the **invariant mass** or **rest mass** and is usually denoted by m_0, the variant mass or **relativistic mass** is denoted by simply m. We will, henceforth, have to be careful about which mass we are referring to. A quick way to remember is that m_0 is what Newtonian mechanics measures as the absolute mass i.e. the mass of an object according to an observer who is moving along with that object, or the mass

Note the term 'lab frame' or 'laboratory frame of reference'. This is the reference frame of our 'lab' or room, which means we, and the 'lab', are still in this frame and other objects are moving relative to this. It is in the lab frame that measurements are usually made.

in the lab frame. By contrast the relativistic mass changes with velocity and the observer is external and not moving with the object in question. Invariant mass and relativistic mass are related as

$$m = \gamma m_0 \tag{6}$$

Keep in mind, all our measurements are made in the lab frame. Although we stated the four-momentum above we will now derive the same, along with a description of relativistic energy, from first principles. We will then end up with a slightly different form of eq. (5), one that is written in terms of relativistic energy. To do this we will have to return to the principle of least action discussed as part of analytical mechanics.

To help you recall, the principle of least action states that for every mechanical system there exists an integral S *called the action integral which has a minimum value for the actual path of motion. In our case now we do not concern ourselves with the minima specifically; we just seek an extremum value.*

By now we ought to be familiar with the action integral given by

$$\mathrm{S} = -\alpha \int_a^b \mathrm{d}s$$

where the integral is along the world line of an object and maxes out along a straight world line, i.e. for photons. Along a curved world line it has a smaller (but still positive) value and we use a negative sign to further push it towards a minimum. The term α is simply a parameter that characterises the system—we will determine what this is presently. First we will try to write this equation in a more familiar format, i.e. with an integral against $\mathrm{d}t$ rather than $\mathrm{d}s$. From eq. (1) we have

$$\mathrm{d}\tau = \frac{\mathrm{d}s}{c} = \frac{\mathrm{d}t}{\gamma}$$

which means our action integral can be written in terms of $\mathrm{d}t$ as

$$\mathrm{S} = \int_{t_1}^{t_2} -\frac{\alpha c}{\gamma}\,\mathrm{d}t$$

Compare this with the familiar action integral expression over time

$$\mathrm{S} = \int_{t_1}^{t_2} \mathscr{L}\,\mathrm{d}t$$

and we get our Lagrangian $\mathscr{L}$ as

$$\mathscr{L} = -\frac{\alpha c}{\gamma}$$

If we now expand γ^{-1} (approximately) as

$$\gamma^{-1} = \left(1 - \frac{v^2}{c^2}\right)^{1/2} = 1 - \frac{v^2}{2c^2}$$

and neglect higher order terms as they become too small, we get

$$\mathscr{L} = -\alpha c + \frac{\alpha v^2}{2c}$$

we realise that the second term looks like $\mathscr{L} = mv^2/2$, the classical energy expression, giving us $\alpha = mc$ so, substituting in $\mathscr{L} = -\alpha c/\gamma$,

$$\mathscr{L} = -\frac{mc^2}{\gamma} \tag{7}$$

From this point on the momentum is easy enough to arrive at. (Feel free to use to γ^{-1} approximation again.) Recall that $\mathbf{p} = \partial\mathscr{L}/\partial v$ giving us

It is left to you to solve the partial differential $\partial\mathscr{L}/\partial v$ as an exercise and arrive at eq. (8).

$$\mathbf{p} = \gamma m_0 \mathbf{v} \tag{8}$$

which is nothing but the three-momentum term in eq. (5). We get the p^0 term as expected by simply demanding that $v \to c$ in eq. (8).

Classically, the energy of a system is defined as

We will maintain the notation $\mathscr{E}$ to refer to the relativistic/total energy of a system and the usual E *for the purely non-relativistic energy term, although we will almost never use the latter.*

$$\mathscr{E} = \mathbf{p} \cdot \mathbf{v} - \mathscr{L}$$

which, upon substitution from eq. (7) and (8), gives us

$$\mathscr{E} = \gamma m_0 c^2 \tag{9}$$

This simple substitution too is left as an exercise. Try to square both sides of the previous equation for $\mathscr{E}$ after substitution and then expand γ to get the square of eq. (9).

This **relativistic energy** term is a particularly interesting one which shows us an important fact: the energy of a system does not go to zero even if the system is at rest. That is to say when $v = 0$ we have $\gamma = 1$ and we are left with

$$\mathscr{E} = m_0 c^2$$

We will return to this equation to make an important observation that tells us, at best, that this equation by itself is somewhat meaningless, i.e. that it only holds under certain extremely important conditions and is by no means general. Let us refer to this henceforth as the **rest energy** of our system.

For now we return our attention to eq. and expand $\gamma = \left(1 - v^2/c^2\right)^{-1/2} = 1 + v^2/2c^2$, along the same lines as our previous expansion, to get

$$\mathscr{E} = m_0 c^2 + \frac{1}{2} m_0 v^2$$

which tells us that all is well: the total energy of a system is its rest energy

plus its classical kinetic energy.

Another exercise: try to write eq. (8) in terms of v^2 then substitute this to get γ in terms of p^2, then substitute γ into eq. (9) to finally get the mass-energy-momentum relation.

Lastly, comparing eq. (8) and (9) will give us an important relation known as the **mass-energy-momentum equivalence**:

$$\mathscr{E}^2 = p^2c^2 + m_0^2c^4 \tag{10}$$

While reading other sources you will eventually come across the term ict used instead of ct for the temporal component. This appears in Minkowski's original solution but we have chosen not to use it here as it does not play well with General Relativity, which is what we will ultimately discuss in this course. The use of ct is thus more pragmatic.

This energy is also known as the Hamiltonian of the system and is sometimes written as $\mathscr{H} = c\sqrt{p^2 + m_0^2c^2}$ but we will not find the need to use this particular format in our course. Observe that in natural units this tells us that energy, mass and momentum have like dimensions: $\mathscr{E}^2 = p^2 + m_0^2$.

By comparing eq. (8) and (9) we can draw another relationship:

$$\mathbf{p} = \mathscr{E}\frac{\mathbf{v}}{c^2}$$

This does not interfere with our three-momentum, however, since we have already established that to be $\gamma m_0\mathbf{v}$, which means our *timelike* component p^0 is when $v = c$ and is therefore simply $\mathscr{E}/c$. We therefore end up with our contravariant four-momentum components as

$$p^\mu = \left(\frac{\mathscr{E}}{c}, \mathbf{p}\right) \tag{11}$$

1.4 *Four-force*

Our discussion on four-vectors culminates in the four-force term. This is sometimes also known as the **relativistic Lorentz force**. The motivation for defining this arises from the principle of relativity which, as we stated earlier, requires that the laws of physics be same in all reference frames.

Just as with three-momentum and Newton's classical second law we can define four-force as the proper-time derivative of four-momentum,

$$\mathrm{F}^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = \gamma\frac{\mathrm{d}p^\mu}{\mathrm{d}t} \tag{12}$$

Likewise we can also define four-force as $\mathrm{F}^\mu = m_0\mathrm{A}^\mu$ in terms of the four-acceleration. This gives us

$$\begin{aligned}
\mathrm{F}^\mu &= m_0\left(c\frac{\mathrm{d}\gamma}{\mathrm{d}t}, \mathbf{v}\frac{\mathrm{d}\gamma}{\mathrm{d}t} + \mathbf{a}\gamma\right)\gamma \\
&= \left(\frac{1}{c}\frac{\mathrm{d}(\gamma m_0c^2)}{\mathrm{d}t}, \frac{\mathrm{d}(\gamma m_0\mathbf{v})}{\mathrm{d}t} + \gamma m_0\mathbf{a}\right)\gamma \\
&= \left(\gamma\frac{\mathbf{F}\cdot\mathbf{v}}{c}, \gamma\mathbf{F}\right) \quad \text{or} \quad \left(\gamma\frac{\mathrm{d}\mathrm{E}}{\mathrm{d}t}, \gamma\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}\right)
\end{aligned} \tag{13}$$

However, keep in mind that for this relation to be valid in general m_0 would

have to be constant in proper time. The use of $\mathrm{d}p^\mu/\mathrm{d}\tau$, though, is valid everywhere.

This brings us to an important observation. Recall four-velocity from eq. (2). Whereas $\mathrm{U}^\mu \mathrm{A}_\mu = 0$ the magnitude of the four-velocity, given by $\mathrm{U}^\mu \mathrm{U}_\mu$, is one. This is quite easy to see:

To go from U^μ to U_μ use the Minkowski metric tensor shown before: $\mathrm{U}^\mu = \eta^\nu_\mu \mathrm{U}_\nu$.

$$\begin{aligned}
\mathrm{U}^\mu \mathrm{U}_\mu &= \sqrt{\sum_\mu (\gamma \mathrm{U}^\mu)^2} \\
&= \sqrt{\gamma^2 c^2 + \gamma^2 v_x^2 + \gamma y_y^2 + \gamma v_z^2} \\
&= \gamma \sqrt{c^2 + v^2} \\
&= \gamma c \sqrt{1 + \frac{v^2}{c^2}} \\
\implies \mathrm{U}^\mu \mathrm{U}_\mu &= c := 1
\end{aligned}$$

The definition of orthogonality must be clear by now. If we have vectors defined by (t^1, x^1, y^1, z^1) and (t^2, x^2, y^2, z^2) we say they are orthogonal iff they satisfy

$$t^1 t^2 + x^1 x^2 + y^1 y^2 + z^1 z^2 = 0$$

So if you have a purely *timelike* vector and a purely *spacelike* vector you can see that they must necessarily be orthogonal to each other.

Now that we have got that out of the way let us return to the four-force. We can show that four-force is orthogonal to the four-velocity. Follow these steps carefully as they also serve as an explainer of the orthogonality of four vectors:

$$\begin{aligned}
\mathrm{F}^\mu \mathrm{U}_\mu &= \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} \mathrm{U}_\mu \\
&= m \frac{\mathrm{d}\mathrm{U}^\mu}{\mathrm{d}\tau} \mathrm{U}_\mu \\
&= m \frac{\mathrm{d}\mathrm{U}^\mu \mathrm{U}_\mu}{\mathrm{d}\tau} \\
&= 0
\end{aligned}$$

1.5 *Photons*

Since light plays a particularly important part in relativity, including in how we describe spacetime diagrams, it is worth paying attention to what specifically happens to photons under the conditions we have discussed so far.

Four-velocity. The second postulate of relativity tells us that light is never at rest in any frame of reference. Further, we can never travel at c which

An MCRF is a Momentarily co-moving reference frame, which is just another name for an Instantaneously Co-moving Frame or ICF.

means there is no MCRF for a photon. Mathematically then $\mathrm{d}x \cdot \mathrm{d}x = 0$ which promptly puts it on the null line in a spacetime diagram as we expect.

Four-momentum. For a photon, like for any particle, the four-momentum is essentially a description of the energy and classical momentum of the particle in a given frame. If the zeroth component of four-momentum $p^0 = \mathscr{E}$ then, for a photon moving in, say, just the x-direction, $p^x = \mathscr{E}$ to ensure that it remains parallel to the null or lightlike world line. Consequently,

$$p^0 \cdot p^x = -\mathscr{E} \cdot \mathscr{E} = 0$$

the four-momentum of a photon is null. The main takeaway here of course is that photons have a spatial momentum equal to their energy, their timelike component of momentum having already been determined to be $\mathscr{E}$.

Rest mass. Any particle with zero four-momentum also has zero rest mass and the photon is no different (to say nothing of the fact that it is the only massless particle we know). Consider that its mass is given by

$$p^\mu \cdot p^\mu = m\mathrm{U}^\mu \cdot m\mathrm{U}^\mu = -m^2 \quad \text{so that} \quad m^2 = -p^\mu \cdot p^\mu = 0$$

which turns out to be just as we had expected.

2 VECTORS AND TENSORS

This section is best read with support from the handout on tensors.

When we speak of primed and unprimed reference frames we refer to two frames viewed from the perspective of a third observer in our hypothetical lab. According to this lab frame one of the frames is at rest—this is the unprimed frame—while another is in motion—this is the primed frame. We represent these without and with a prime (′) respectively.

The root of most of these results would have been covered in Paper 102 on mathematical methods.

Tensor formalism is an important part of relativity. Although there are ways to study relativity without tensors the use of tensors helps us keep our expressions compact which, in turn, lets us handle expressions across primed and unprimed reference frames with considerable ease. Remember that this review of tensors is just that, a review, and not a detailed look at why these results are true.

Before we begin, there are three things to keep in mind. These are important enough that we will number them here.

1. When a tensor T transforms we mean that the unprimed and primed frames transform into each other. When this occurs the new tensor will still be called T but with or without a prime depending on where we started.

2. Whenever a tensor transforms or, to broaden the rule, whenever any operation is performed on a tensor it is important to change its indices. Use any greek letter of your choice (for four components, or latin for three as has been our rule till now). Whether the new index is the same as the old index is secondary—they may, in fact, be equal—but we cannot make that assumption outright.

3. The indices themselves do not matter as much as their positions. So keep an eye on primed and unprimed notations and the positions of

the indices and do not obsess over the actual greek letter used. For all purposes, unless otherwise mentioned, writing T^{α}_{β} is no different from writing T^{σ}_{ρ} but it *is* different from writing, for example, T'^{α}_{β} or $T^{\alpha}\beta$.

3 SCALAR PRODUCT AND COVECTORS

As a last stop we will discuss the nature of tensors and why we call some 'covariant' and others 'contravariant' first. We already discussed the idea of a basis. Think of the basis $\mathbf{e}_\alpha$, with α taking values $0, 1, 2, 3$ as usual. Let us define two vectors $\mathbf{A}$ and $\mathbf{B}$ in this bases as

$$\mathbf{A} = A^\alpha \mathbf{e}_\alpha \qquad\qquad \mathbf{B} = B^\beta \mathbf{e}_\beta$$

This is of course a general way of writing an example that you are probably more familiar with. Think of $\mathbf{A}$ *as being a vector in cartesian space. Its components would then be* A^x, A^y *and* A^z, *and the bases of Cartesian space we know are* $e_1 = \hat{i}$, $e_2 = \hat{j}$ *and* $e_3 = \hat{k}$. *Consequently we see that our definition of* $\mathbf{A}$ *above essentially means* $\mathbf{A} = A^q \mathbf{e}_q$ *or* $\sum_q A^q \mathbf{e}_q = A^1\mathbf{e_1} + A^2\mathbf{e_2} + A^3\mathbf{e_3}$ *which is simply* $A^x\hat{i} + A^y\hat{j} + A^z\hat{k}$, *an expression you are already quite familiar with.*

In these definitions A^α is a component of $\mathbf{A}$ and $\mathbf{e}_\alpha$ is an appropriate basis vector. The fact that, on the right-hand side, one index is above and one below means we must sum over the repeating index α. The **scalar product**, sometimes called the dot product, is defined as

$$\begin{aligned}\mathbf{A}\cdot\mathbf{B} &= A^\alpha B^\beta \mathbf{e}_\alpha \cdot \mathbf{e}_\beta \\ &= A^\alpha B^\beta \eta_{\alpha\beta} \\ &= -A^0B^0 + A^1B^1 + A^2B^2 + A^3B^3 \end{aligned} \tag{14}$$

from our definition of η the components of the Minkowski metric tensor and from its bases e described on p.2. The scalar product serves another function: *it helps us determine the components of a tensor if we simply multiply its bases*. For example, consider the metric tensor η with bases $\mathbf{e}_\alpha$ and $\mathbf{e}_\beta$ and consider their dot product: $\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = \eta_{\alpha\beta}$ which gives back the components of g.

THE NEXT TOPIC of discussion is that of **covectors**, also known as **one forms** or **covariant tensors**. Our primary intention here is to see why they are called 'covariant', but we will also try to understand their general nature. We proceed with the understanding that a tensor is like a function that takes n vectors and produces one scalar real number in return. First of all recall that a basis $\mathbf{e}_{\alpha'}$ transforms as follows:

$$\mathbf{e}_{\alpha'} = \Lambda^{\beta}_{\alpha'} \mathbf{e}_\beta \tag{15}$$

and that a one-form or covector p when supplied with a vector will output a real number. As described above if we must determine the components of this tensor p we need only use its bases as arguments. So for the base $\mathbf{e}_\alpha$ we have $p(\mathbf{e}_\alpha) = p_\alpha$ (say) so that when we use some vector $\mathbf{A}$, and use the definition of $\mathbf{A}$ stated at the beginning of this section, as the argument we get

Remember that any covector p, along with another covector q and scalar alpha, must satisfy the following rules: a linear combination like p + q must yield another covector, and a scalar multiplication like αp too must yield a covector.

$$p(\mathbf{A}) = p(A^\alpha \mathbf{e}_\alpha) = A^\alpha p_\alpha = A^0 p_0 + A^1 p_1 + A^2 p_2 + A^3 p_3$$

Notice that, unlike the scalar product of two vectors, the output of a one-form

is a scalar number without the metric tensor components; that is to say, the signs are not $(-,+,+,+)$ as in eq. (14) but rather just $(+,+,+,+)$ giving a real number. Now what if we start with the $\mathbf{e}_{\alpha'}$ bases and wish to transform to the $\mathbf{e}_\beta$ bases? We might do something like this:

$$\begin{aligned} p_{\alpha'} &= p(\mathbf{e}_{\alpha'}) \\ &= p(\Lambda^{\beta}_{\alpha'}\mathbf{e}_\beta) \\ \implies p_{\alpha'} &= \Lambda^{\beta}_{\alpha'} p_\beta \end{aligned}$$

Compare this result with equation (15) and you notice the similarities immediately. The fact that a one-form p transforms exactly like its bases is why we say it is a **covariant tensor**, i.e. one that varies 'co' or 'similarly to' its bases. By contrast contravariant vector components transform inversely to a covector. To see this consider the transformation of a one-form or covector p_α and some contravariant vector A^α in the following manner: $A^{\alpha'} p_{\alpha'} \to A^\beta p_\beta$.

$$\begin{aligned} A^{\alpha'} p_{\alpha'} &= \Lambda^{\alpha'}_{\beta} A^\beta \Lambda^{\gamma}_{\alpha'} p_\gamma \\ &= \delta^{\gamma}_{\beta} A^\beta p_\gamma \\ \implies A^{\alpha'} p_{\alpha'} &= A^\beta p_\beta \end{aligned}$$

Observe that we have made use of the δ function (Kronecker) first as per its definition and then used the fact that a delta function δ^μ_ν vanishes everywhere except when $\mu = \nu$ and therefore we have imposed that condition on our equation too. More importantly, in the first step, observe that our contravariant tensor A^μ does not transform like the basis $\mathbf{e}_\alpha$ would, or like the covector p_β; instead it transforms with the position of its indices inverted. Lastly, as is evident from the end result of our transformation above, the consequence of all this is that the product of a covariant and a contravariant tensor, like $A^{\alpha'} p_{\alpha'}$, always remains frame-independent.

Generally, covariant and contravariant vectors are known now as 'vectors' and 'dual vectors' as these terms are less confusing.

FINAL THOUGHTS: We started with two frames of reference moving relative to each other with uniform velocity and established a transformation across these frames that ensures the validity of all laws of physics and, more important, that ensures that observers across such frames agree with each other as to the outcome of any event.

However we note that this theory is far from complete. For instance what if the frames are not moving with uniform relative velocity but are accelerating instead? Or what if the space in these frames is not flat but curved instead? Is all this only a specific case of a much wider, global theory of relativity? We will address these questions and discover more about non-inertial frames of reference, the curvature of spacetime and gravity as geometry in the next part of this course when we discuss the general theory of relativity.

References

[1] Chappell JM, Hartnett JG, Iannella N, Iqbal A, and Abbott D. Exploring the origin of Minkowski spacetime. arXiv:1501.04857

[2] Schutz B. A first course in general relativity. Cambridge University Press. 2nd edition, 2009.

[3] Landau L, Lifshitz E. The classical theory of fields. Pergamon Press.

[4] Weinberg S. Gravitation and cosmology: principles and applications of the general theory of relativity. Wiley.