

9.4: More on 4-vectors and 4-tensors

This is a good moment to introduce a formalism that will allow us, in particular, to solve the same proton collision problem in one more (and arguably, the most elegant) way. Much more importantly, this formalism will be virtually necessary for the description of the Lorentz transform of the electromagnetic field, and its interaction with relativistic particles – otherwise the formulas would be too cumbersome.

Let us call the 4-vectors we have used before,

$$\text{Contravariant 4-vectors} \quad A^\alpha \equiv \{A_0, \mathbf{A}\}, \quad (9.84)$$

contravariant, and denote them with top indices, and introduce also covariant vectors,

$$\text{Covariant 4-vectors} \quad A_\alpha \equiv \{A_0, -\mathbf{A}\}, \quad (9.85)$$

marked by lower indices. Now if we form a scalar product of these two vectors using the standard (3D-like) rule, just as a sum of the products of the corresponding components, we immediately get

$$A_\alpha A^\alpha \equiv A^\alpha A_\alpha \equiv A_0^2 - A^2. \quad (9.86)$$

Here and below the sign of the sum of four components of the product has been dropped.³⁷ The scalar product (86) is just the norm of the 4-vector in our former definition, and as we already know, is Lorentz-invariant. Moreover, the scalar product of two different vectors (also a Lorentz invariant), may be rewritten in any of two similar forms:³⁸

$$\text{Scalar product's forms} \quad A_0 B_0 - \mathbf{A} \cdot \mathbf{B} \equiv A_\alpha B^\alpha = A^\alpha B_\alpha; \quad (9.87)$$

again, the only caveat is to take one vector in the covariant, and the other one in the contravariant form.

Now let us return to our sample problem (Fig. 10). Since all components (\mathcal{E}/c and \mathbf{p}) of the total 4-momentum of our system are conserved at the collision, its norm is conserved as well:

$$(p_a + p_b)_\alpha (p_a + p_b)^\alpha = (4p)_\alpha (4p)^\alpha. \quad (9.88)$$

Since now the vector product is the usual math construct, we know that the parentheses on the left-hand side of this equation may be multiplied as usual. We may also swap the operands and move constant factors through products as convenient. As a result, we get

$$(p_a)_\alpha (p_a)^\alpha + (p_b)_\alpha (p_b)^\alpha + 2(p_a)_\alpha (p_b)^\alpha = 16p_\alpha p^\alpha. \quad (9.89)$$

Thanks to the Lorentz invariance of each of the terms, we may calculate it in the reference frame we like. For the first two terms on the left-hand side, as well as for the right hand side term, it is beneficial to use the frames in which that particular proton is at rest; as a result, according to Eq. (77b), each of the two left-hand-side terms equals $(mc)^2$, while the right-hand side equals $16(mc)^2$. On the contrary, the last term on the left-hand side is more easily evaluated in the lab frame, because in it, the three spatial components of the 4-momentum p_b vanish, and the scalar product is just the product of the scalars \mathcal{E}/c for protons a and b . For the latter proton, being at rest, this ratio is just mc so that we get a simple equation,

$$(mc)^2 + (mc)^2 + 2 \frac{\mathcal{E}_{\min}}{c} mc = 16(mc)^2, \quad (9.90)$$

immediately giving the final result $\mathcal{E}_{\min} = 7mc^2$, already obtained earlier in two more complex ways.

Let me hope that this example was a convincing demonstration of the convenience of representing 4-vectors in the contravariant (84) and covariant (85) forms,³⁹ with Lorentz invariant norms (86). To be useful for more complex tasks, this formalism should be developed a little bit further. In particular, it is crucial to know how the 4-vectors change under the Lorentz transform. For contravariant vectors, we already know the answer (54); let us rewrite it in our new notation:

$$A^\alpha = L^\alpha_\beta A'^\beta. \quad \text{Lorentz transform: contravariant vectors} \quad (9.91)$$

where L^α_β is the matrix (51), generally called the mixed Lorentz tensor:⁴⁰

$$L_{\beta}^{\alpha} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{Mixed Lorentz tensor} \quad (9.92)$$

Note that though the position of the indices α and β in the Lorentz tensor notation is not crucial, because this tensor is symmetric, it is convenient to place them using the general index balance rule: the difference of the numbers of the upper and lower indices should be the same in both parts of any 4-vector/tensor equality. (Please check yourself that all the formulas above do satisfy this rule.)

In order to rewrite Eq. (91) in a more general form that would not depend on the particular orientation of the coordinate axes (Fig. 1), let us use the contravariant and covariant forms of the 4-vector of the time-space interval (57),

$$dx^{\alpha} = \{cdt, d\mathbf{r}\}, \quad dx_{\alpha} = \{cdt, -d\mathbf{r}\}; \quad (9.93)$$

then its norm (58) may be represented as⁴¹

$$(ds)^2 \equiv (cdt)^2 - (dr)^2 = dx^{\alpha} dx_{\alpha} = dx_{\alpha} dx^{\alpha}. \quad (9.94)$$

Applying Eq. (91) to the first, contravariant form of the 4-vector (93), we get

$$dx^{\alpha} = L_{\beta}^{\alpha} dx'^{\beta}. \quad (9.95)$$

But with our new shorthand notation, we can also write the usual rule of differentiation of each component x^{α} , considering it as a function (in our case, linear) of four arguments x'^{β} , as follows:⁴²

$$dx^{\alpha} = \frac{\partial x^{\alpha}}{\partial x'^{\beta}} dx'^{\beta}. \quad (9.96)$$

Comparing Eqs. (95) and (96), we can rewrite the general Lorentz transform rule (92) in a new form,

$$\text{Lorentz transform: general form} \quad A^{\alpha} = \frac{\partial x^{\alpha}}{\partial x'^{\beta}} A'^{\beta}. \quad (9.97a)$$

which does not depend on the coordinate axes' orientation.

It is straightforward to verify that the reciprocal transform may be represented as

$$\text{Reciprocal Lorentz transform} \quad A'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\beta}} A^{\beta}. \quad (9.97b)$$

However, the reciprocal transform has to differ from the direct one only by the sign of the relative velocity of the frames, so that for the coordinate choice shown in Fig. 1, its matrix is

$$\frac{\partial x'^{\alpha}}{\partial x^{\beta}} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (9.98)$$

Since according to Eqs. (84)-(85), covariant 4-vectors differ from the contravariant ones by the sign of their spatial components, their direct transform is given by the matrix (98). Hence their direct and reciprocal transforms may be represented, respectively, as

$$A_{\alpha} = \frac{\partial x'^{\beta}}{\partial x^{\alpha}} A'_{\beta}, \quad A'_{\alpha} = \frac{\partial x^{\beta}}{\partial x'^{\alpha}} A_{\beta}, \quad \text{Lorentz transform: covariant vectors} \quad (9.99)$$

evidently satisfying the index balance rule. (Note that primed quantities are now multiplied, rather than divided as in the contravariant case.) As a sanity check, let us apply this formalism to the scalar product $A_{\alpha} A^{\alpha}$. As Eq. (96) shows, the implicit-sum notation allows us to multiply and divide any equality by the same partial differential of a coordinate, so that we can write:

$$A_{\alpha} A^{\alpha} = \frac{\partial x'^{\beta}}{\partial x^{\alpha}} \frac{\partial x^{\alpha}}{\partial x'^{\gamma}} A'_{\beta} A'^{\gamma} = \frac{\partial x'^{\beta}}{\partial x'^{\gamma}} A'_{\beta} A'^{\gamma} = \delta_{\beta\gamma} A'_{\beta} A'^{\gamma} = A'_{\gamma} A'^{\gamma}, \quad (9.100)$$

i.e. the scalar product $A_{\alpha} A^{\alpha}$ (as well as $A^{\alpha} A_{\alpha}$) is Lorentz-invariant, as it should be.

Now, let us consider the 4-vectors of derivatives. Here we should be very careful. Consider, for example, the following 4-vector operator

$$\frac{\partial}{\partial x^\alpha} \equiv \left\{ \frac{\partial}{\partial(ct)}, \nabla \right\}, \quad (9.101)$$

As was discussed above, the operator is not changed by its multiplication and division by another differential, e.g., $\partial x'^\beta$ (with the corresponding implied summation over all four values of β), so that

$$\frac{\partial}{\partial x^\alpha} = \frac{\partial x'^\beta}{\partial x^\alpha} \frac{\partial}{\partial x'^\beta}. \quad (9.102)$$

But, according to the first of Eqs. (99), this is exactly how the covariant vectors are Lorentz-transformed! Hence, we have to consider the derivative over a contravariant space-time interval as a covariant 4-vector, and vice versa.⁴³ (This result might be also expected from the index balance rule.) In particular, this means that the scalar product

$$\frac{\partial}{\partial x^\alpha} A^\alpha \equiv \frac{\partial A_0}{\partial(ct)} + \nabla \cdot \mathbf{A} \quad (9.103)$$

should be Lorentz-invariant for any legitimate 4-vector. A convenient shorthand for the covariant derivative, which complies with the index balance rule, is

$$\frac{\partial}{\partial x^\alpha} \equiv \partial_\alpha, \quad (9.104)$$

so that the invariant scalar product may be written just as $\partial_\alpha A^\alpha$. A similar definition of the contravariant derivative,

$$\partial^\alpha \equiv \frac{\partial}{\partial x_\alpha} = \left\{ \frac{\partial}{\partial(ct)}, -\nabla \right\}, \quad (9.105)$$

allows us to write the Lorentz-invariant scalar product (103) in any of the following two forms:

$$\frac{\partial A_0}{\partial(ct)} + \nabla \cdot \mathbf{A} = \partial^\alpha A_\alpha = \partial_\alpha A^\alpha. \quad (9.106)$$

Finally, let us see how does the general Lorentz transform change 4-tensors. A second-rank 4×4 matrix is a legitimate 4-tensor if the 4-vectors it relates obey the Lorentz transform. For example, if two legitimate 4-vectors are related as

$$A^\alpha = T^{\alpha\beta} B_\beta, \quad (9.107)$$

we should require that

$$A'^\alpha = T'^{\alpha\beta} B'_\beta, \quad (9.108)$$

where A^α and A'^α are related by Eqs. (97), while B_β and B'_β , by Eqs. (99). This requirement immediately yields

$$\text{Lorentz transform of 4-tensors} \quad T^{\alpha\beta} = \frac{\partial x^\alpha}{\partial x'^\gamma} \frac{\partial x^\beta}{\partial x'^\delta} T'^{\gamma\delta}, \quad T'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x'^\beta}{\partial x^\delta} T^{\gamma\delta}, \quad (9.109)$$

with the implied summation over two indices, γ and δ . The rules for the covariant and mixed tensors are similar.⁴⁴

Reference

³⁷ This compact notation may take some time to be accustomed to, but is very convenient (compact) and can hardly lead to any confusion, due to the following rule: the summation is implied when (and only when) an index is repeated twice, once on the top and another at the bottom. In this course, this shorthand notation will be used only for 4-vectors, but not for the usual (spatial) vectors.

³⁸ Note also that, by definition, for any two 4-vectors, $A_\alpha B^\alpha = B^\alpha A_\alpha$.

³⁹ These forms are 4-vector extensions of the notions of contravariance and covariance, introduced in the 1850s by J. Sylvester (who also introduced the term “matrix”) for the description of the change of the usual geometric (3-component) vectors at the

transfer between different reference frames – e.g., resulting from the frame rotation. In this case, the contravariance or covariance of a vector is uniquely determined by its nature: if the Cartesian coordinates of a vector (such as the non-relativistic velocity $\mathbf{v} = d\mathbf{r}/dt$) are transformed similarly to the radius-vector \mathbf{r} , it is called contravariant, while the vectors (such as ∇f) that require the reciprocal transform, are called covariant. In the Minkowski space, both forms may be used for any 4-vector.

⁴⁰ Just as the 4-vectors, 4-tensors with two top indices are called contravariant, and those with two bottom indices, covariant. The tensors with one top and one bottom index are called mixed.

⁴¹ Another way to write this relation is $(ds)^2 = g_{\alpha\beta} dx^\alpha dx^\beta = g^{\alpha\beta} dx_\alpha dx_\beta$, where double summation over indices α and β is implied, and g is the so-called metric tensor,

$$g^{\alpha\beta} \equiv g_{\alpha\beta} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

which may be used, in particular, to transfer a covariant vector into the corresponding contravariant one and back: $A^\alpha = g^{\alpha\beta} A_\beta$, $A_\alpha = g_{\alpha\beta} A^\beta$. The metric tensor plays a key role in general relativity, in which it is affected by gravity – “curved” by particles’ masses.

⁴² Note that in the index balance rule, the top index in the denominator of a fraction is counted as a bottom index in the numerator, and vice versa.

⁴³ As was mentioned above, this is also a property of the reference-frame transform of the “usual” 3D vectors.

⁴⁴ It is straightforward to check that transfer between the contravariant and covariant forms of the same tensor may be readily achieved using the metric tensor g : $T_{\alpha\beta} = g_{\alpha\gamma} T^{\gamma\delta} g_{\delta\beta}$, $T^{\alpha\beta} = g^{\alpha\gamma} T_{\gamma\delta} g^{\delta\beta}$.

This page titled [9.4: More on 4-vectors and 4-tensors](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Konstantin K. Likharev](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.