# PHY 420 Spring 2021
## Class Summary—Week 7, Day 1—Tuesday, May 11

## 4-vectors

Having learned the Lorentz transformations, it is natural to anticipate that there are numerous physical quantities that transform under Lorentz transformations in the same manner as the time and space coordinates of a point. With this in mind, we introduce **4-vectors**. The coordinate 4-vector is $(x_0, x_1, x_2, x_3)$; we designate the components of an arbitrary 4-vector as $(A_0, A_1, A_2, A_3)$, where $A_1, A_2, A_3$ are the components of a 3-vector $\vec{A}$. Take care you understand the terminology here: a 4-dimensional vector is any vector with four coordinates, whereas **a 4-vector is a special 4-dimensional vector that obeys the Lorentz transformations** (written below).

Now, recall that the Lorentz transformations were based on the invariance of the squared quantity

$$c^2t^2 - (x^2 + y^2 + z^2) = c^2t'^2 - (x'^2 + y'^2 + z'^2) \tag{11.15}$$

Similarly, we can write for any 4-vector the invariance

$$A_0'^2 - \left|\vec{A}'\right|^2 = A_0^2 - \left|\vec{A}\right|^2 \tag{11.23}$$

where the components $(A_0', \vec{A}')$ and $(A_0, \vec{A})$ refer to any two inertial reference frames, and hence the **Lorentz transformation law for an arbitrary 4-vector**, equivalent to equation (11.16) for position-time coordinates that we wrote in the previous class, is

$$\left.\begin{aligned} A_0' &= \gamma\,(A_0 - \vec{\beta}\cdot\vec{A}) \\ A_\parallel' &= \gamma\,(A_\parallel - \beta A_0) \\ A_\perp' &= A_\perp \end{aligned}\right\} \tag{11.22}$$

where the parallel and perpendicular signs indicate components relative to the velocity $\vec{v} = c\vec{\beta}$.

Now, the addition of velocities in relativity is given by

$$u_\parallel = \frac{u_\parallel' + v}{1 + \dfrac{\vec{v}\cdot\vec{u}'}{c^2}} \qquad\qquad \vec{u}_\perp = \frac{\vec{u}_\perp'}{\gamma_v\left(1 + \dfrac{\vec{v}\cdot\vec{u}'}{c^2}\right)} \tag{11.31}$$

where $u_\parallel$ and $\vec{u}_\perp$ refer to the components of velocity parallel and perpendicular, respectively, to $\vec{v}$, and the subscript on $\gamma_v$ explicitly identifies the relationship to be $\gamma_v = (1 - v^2/c^2)^{-1/2}$.

The structure of equation (11.31) for the addition of velocities makes it clear that the law of transformation of velocities is not that of 4-vectors. There is, however, a 4-vector related to ordinary velocity. To find it, we will need the **proper time** $d\tau$. Recall that if a clock is at rest in frame $K'$, then it measures the shortest time interval in that frame, and this is known as the *proper time* $\Delta\tau$.

Suppose frame $K'$ is moving with velocity $v$ along the $x_1$ axis of frame $K$, so that $\gamma = (1-v^2/c^2)^{-1/2}$. To find a relation between $\Delta\tau$ and the time interval $\Delta t$ measured in frame $K$, recall that proper time is measured in the rest frame of the clock, so it must be measured at the same position $x_1'$ in $K'$. The Lorentz transformation equation for the time coordinate $x_0 = ct$ given by

$$x_0 = \gamma\Big(x_0' + \beta x_1'\Big)$$

can be written at the instant $t_1'$ in frame $K'$, corresponding to $t_1$ in frame $K$ as

$$ct_1 = \gamma\Big(ct_1' + \beta x_1'\Big)$$

and, at the instant $t_2'$ in frame $K'$, corresponding to $t_2$ in frame $K$ as

$$ct_2 = \gamma\Big(ct_2' + \beta x_1'\Big)$$

Notice that I've written the same $x_1'$ in both equations, because the proper time is measured in the rest frame of the clock, i.e., the clock is at the same position $x_1'$ in frame $K'$, in order for it to measure the proper time interval. Subtracting, we get

$$c\Big(t_2 - t_1\Big) = \gamma c\Big(t_2' - t_1'\Big)$$

and therefore

$$\Delta t = \gamma \Delta\tau$$

*as you showed in Question 1 on today's worksheet.*

To obtain the **4-velocity**, we will have to divide $(cdt, dx, dy, dz)$ by $\Delta\tau$. Before doing that, let us put our notation in place. The object is moving in an inertial frame with velocity $u$, so the proper time is measured by a clock fixed to the object, and so now the $\gamma$-factor is actually

$$\gamma_u = (1 - u^2/c^2)^{-1/2}$$

and $dt = \gamma_u d\tau$, so that

$$\frac{dt}{d\tau} = \gamma_u$$

Then, if we write the 4-velocity as $U = (U_0, \vec{U})$, we have

$$U_0 = c\frac{dt}{d\tau} = c\Big(\gamma_u\Big)$$

Meanwhile

$$U_x = \frac{dx}{d\tau} = \frac{dx}{dt}\,\frac{dt}{d\tau} = u_x\Big(\gamma_u\Big)$$

Likewise

$$U_y = u_y\Big(\gamma_u\Big), \qquad U_z = u_z\Big(\gamma_u\Big)$$

so that

$$\vec{U} = \gamma_u\,\vec{u}$$

Therefore, the 4-velocity is

$$U = (U_0, \vec{U}) = \Big(\gamma_u\,c, \gamma_u\,\vec{u}\Big)$$

*as you showed in Question 2(a) on today's worksheet.*

Using the conventional definition of the momentum, we can use the 4-velocity defined on the previous page to define the **4-momentum**

$$P = mU = m\,(U_0, \vec{U})$$

and since $(U_0, \vec{U}) = \Big(\gamma_u\, c, \gamma_u\, \vec{u}\Big)$, this means that

$$P = \Big(\gamma_u mc, \gamma_u m\vec{u}\Big)$$

Now, recall the following expressions for the relativistic energy and momentum:

$$\vec{p} = \gamma_u m\vec{u} \tag{11.46}$$

and

$$E = \gamma_u mc^2 \tag{11.51}$$

where $\gamma_u = (1 - u^2/c^2)^{-1/2}$. You used these to demonstrate that $\vec{u} = c^2\,\vec{p}/E$ *in Question 2(b) on today's worksheet.*

From equation (11.51), we have $\gamma_u mc = E/c$, and since $\gamma_u m\vec{u} = \vec{p}$ from equation (11.46), *you showed in Question 3(a) on today's worksheet* that the 4-momentum can also be written as

$$P = \Big(\gamma_u mc, \gamma_u m\vec{u}\Big) = \left(\frac{E}{c}, \vec{p}\right)$$

which explains why the 4-momentum is also known as the **energy-momentum 4-vector**.

Next, we need to find the invariant quantity for the energy-momentum 4-vector ($p_0{=}E/c$, $\vec{p}$). To do so, evaluate

$$\begin{aligned} p_0^2 - \vec{p}\cdot\vec{p} &= \Big(\gamma_u mc\Big)^2 - \Big(\gamma_u m\vec{u}\Big)\cdot\Big(\gamma_u m\vec{u}\Big) \\ &= \gamma_u^2 m^2 c^2 - \gamma_u^2 m^2 u^2 \\ &= \gamma_u^2 m^2 c^2 \left[1 - \frac{u^2}{c^2}\right] \end{aligned}$$

But since $\gamma_u = (1 - u^2/c^2)^{-1/2}$, we have $\gamma_u^{-2} = (1 - u^2/c^2)$, so that the above modifies to

$$p_0^2 - \vec{p}\cdot\vec{p} = \gamma_u^2 m^2 c^2\,\gamma_u^{-2}$$

Therefore, the invariant for the energy-momentum 4-vector is

$$p_0^2 - \vec{p}\cdot\vec{p} = (mc)^2 \tag{11.54}$$

*as you showed in Question 3(b) on today's worksheet.*

It is easy to see that equation (11.54) permits the energy $E$ to be expressed in terms of the momentum as

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{11.55}$$

*as you showed in Question 3(c) on today's worksheet.*

## Mathematical properties of space-time

The transformations that we've discussed above, and all aspects of the kinematics of special relativity discussed by Jackson on pages 524-539, can be discussed in a more profound and elegant manner in terms of group theory.

If you're unaware of groups, it's a good thing to brush up on if you have the time (e.g., the Wolfram Mathworld page). Very briefly, a group is a set of elements $(a, b, c, \ldots)$ along with a binary operation called the group operation that together satisfy the four fundamental properties of closure, associativity, identity and inverse. That is, if $*$ denotes the group operation, then

- Closure: If $a$ and $b$ are any two elements in a group, then $c = a * b$ must also be in the group.
- Associativity: For any elements $a, b, c$ in the group, $a * (b * c) = (a * b) * c$.
- Identity: The group must contain a special element $I$ called the identity element such that $a * I = I * a = a$.
- Inverse: Every element in the group must contain an inverse that is also in the group, where the inverse is defined such that if $a^{-1}$ is the inverse of $a$, then $a * a^{-1} = a^{-1}a = I$, where $I$ is the identity element.

For an example of groups, three-dimensional rotations in classical and quantum mechanics can be discussed in terms of the group of transformations of the coordinates that leave the norm (i.e., the length) of the vector $\vec{x}$ invariant.

In the Special Theory of Relativity, Lorentz transformations of the four-dimensional coordinates $(x_0, \vec{x})$ follow from the invariance of

$$s^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2 \tag{11.59}$$

Therefore, the kinematics of relativity can be developed in terms of the group of all transformations that leave $s^2$ invariant. This group is called the *homogenous Lorentz group*, and it contains ordinary rotations as well the Lorentz transformations that we wrote in equation (11.16).

Meanwhile, the group of transformations that leave invariant the quantity

$$s^2(x, y) = (x_0 - y_0)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2$$

is called the *inhomogenous Lorentz group* or the *Poincaré group*. It contains translations and reflections in both space and time, as well as the transformations of the homogenous Lorentz group. The quantity $s(x, y)$ is defined to be the norm of relativistic space-time, and may be thought of as the distance squared between two events $x$ and $y$.

## Tensors

In special relativity, we are concerned with transformations from one inertial frame into another. From the first postulate of relativity, we know that the laws of physics must be the same in all inertial frames. In technical language, this is stated by saying that the mathematical equations expressing the laws of nature must be covariant, that is, invariant in form under the transformations of the Lorentz group, which involve displacements of coordinate axes, rotations, Lorentz transformations, and parity reversals (i.e., from $x, y, z$ to $-x, -y, -z$). Tensors, about which we will now learn, have the property that if they possess a certain relationship in one coordinate system, that same relationship holds in any other coordinate system that is related to the first by a certain class of transformations. Therefore, the laws of physics can be expressed as these relationships between tensors.

To motivate this discussion, let us look at it by way of Griffiths (although with symbols modified to conform to the ones we're using).

Recall that the dot (or scalar) product of a 3-dimensional vector

$$\vec{A} \cdot \vec{B} \equiv A_x B_x + A_y B_y + A_z B_z$$

is invariant (unchanged) under rotations.

The analog to this in 4-dimensions is

$$c^2 t^2 - (x^2 + y^2 + z^2)$$

because no matter what frame of reference we are in, this quantity is invariant, i.e.,

$$c^2 t'^2 - (x'^2 + y'^2 + z'^2) = c^2 t^2 - (x^2 + y^2 + z^2)$$

This motivates us to establish a set of rules for the scalar product of any two 4-vectors to be invariant under Lorentz transformations.

So, think of designating a 4-vector $B_\alpha$ as a **covariant vector**, where covariant implies that its dot (or scalar) product is *invariant* under Lorentz transformations. We'll see later how this can be written as a row vector.

Then, to form a dot product, we need a column vector. So, we'll define such a vector $A^\alpha$ and designate it as a **contravariant vector** so that the dot (or scalar) product of two vectors is defined as the product of the components of a covariant and a contravariant vector:

$$B \cdot A \equiv \sum_{\alpha=0}^{3} B_\alpha A^\alpha$$

Starting with this equation, we will begin using a convention developed by Einstein himself. Known as **Einsteinian summation**, it invokes the rule that repeated indices indicate summation, so that we write

$$B \cdot A \equiv B_\alpha A^\alpha \tag{11.66}$$

where $\alpha = 0, 1, 2, 3$. The summation is implied by the repeated index $\alpha$. We will frequently use this summation convention for repeated indices.

Now, recall that the space-time continuum is defined in terms of a 4-dimensional space for which we've designated the coordinates as

$$(ct, z, x, y) \equiv (x^0, x^1, x^2, x^3)$$

Previously, we had written the indices as subscripts, but henceforth, we will write them as superscripts so that they constitute a contravariant vector.

We suppose that there is a well-defined transformation that yields new coordinates $x'^0, x'^1, x'^2, x'^3$ according to some rule:

$$x'^\alpha = x'^\alpha\,(x^0, x^1, x^2, x^3) \qquad \text{where } \alpha = 0, 1, 2, 3 \tag{11.60}$$

For example, the rule could be the Lorentz transformations; writing these for a general contravariant vector $(A^0, A^1, A^2, A^3)$, we have

$$A'^{\,0} = \gamma\,(A^0 - \beta A^1) \qquad\qquad A'^{\,2} = A^2$$

$$A'^{\,1} = \gamma\,(A^1 - \beta A^0) \qquad\qquad A'^{\,3} = A^3$$

with the invariant

$$(A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2$$

We say that the *contravariant vector* $A^\alpha$ has four components $A^0, A^1, A^2, A^3$. Then, we have that the **transformation rule for *contravariant* vectors** is

$$A'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta}\,A^\beta \tag{11.61}$$

where the repeated index $\beta$ implies a summation over $\beta = 0, 1, 2, 3$. So, explicitly, we have

$$A'^\alpha = \frac{\partial x'^\alpha}{\partial x^0}\,A^0 + \frac{\partial x'^\alpha}{\partial x^1}\,A^1 + \frac{\partial x'^\alpha}{\partial x^2}\,A^2 + \frac{\partial x'^\alpha}{\partial x^3}\,A^3$$

*as you wrote in Question 4 on today's worksheet.*

Let's look at an example to see how this works. Let's take $\alpha = 0$, so that

$$A'^{\,0} = \frac{\partial x'^{\,0}}{\partial x^0}\,A^0 + \frac{\partial x'^{\,0}}{\partial x^1}\,A^1 + \frac{\partial x'^{\,0}}{\partial x^2}\,A^2 + \frac{\partial x'^{\,0}}{\partial x^3}\,A^3 \tag{11.61.a}$$

Suppose $x'^\alpha$ also transform according to Lorentz transformations so that

$$x'^{\,0} = \gamma\,(x^0 - \beta x^1) \qquad\qquad x'^{\,2} = x^2$$

$$x'^{\,1} = \gamma\,(x^1 - \beta x^0) \qquad\qquad x'^{\,3} = x^3$$

so that

$$\frac{\partial x'^{\,0}}{\partial x^0} = \gamma, \qquad \frac{\partial x'^{\,0}}{\partial x^1} = -\gamma\beta, \qquad \frac{\partial x'^{\,0}}{\partial x^2} = 0, \qquad \frac{\partial x'^{\,0}}{\partial x^3} = 0$$

Substituting these in equation (11.61.a), we get

$$A'^{\,0} = \gamma\,A^0 + (-\gamma\beta)\,A^1 + (0)\,A^2 + (0)\,A^3 = \gamma(A^0 - \beta A^1)$$

which is consistent with the Lorentz transformation relation for the component $A'^{\,0}$ of the general covariant vector, *as you showed in Question 5 on today's worksheet.* I'll leave it to you to check for the other three components.

Having defined equation (11.61) as the transformation for a contravariant vector, we can now define the **transformation rule for *covariant* vectors**:

$$B'_\alpha = \frac{\partial x^\beta}{\partial x'^\alpha} B_\beta \tag{11.62}$$

where, now, the coordinate indices are on top while the new primed coordinate is on the bottom. Writing it out explicitly

$$B'_\alpha = \frac{\partial x^0}{\partial x'^\alpha} B_0 + \frac{\partial x^1}{\partial x'^\alpha} B_1 + \frac{\partial x^2}{\partial x'^\alpha} B_2 + \frac{\partial x^3}{\partial x'^\alpha} B_3$$

The partial derivative in equation (11.62) is to be calculated from the inverse of equation (11.60) with $x^\beta$ expressed as a function of $x'^0, x'^1, x'^2, x'^3$, that is, from

$$x^\beta = x^\beta\left(x'^0, x'^1, x'^2, x'^3\right) \qquad \text{where } \beta = 0, 1, 2, 3$$

Note again that contravariant vectors have superscripts and covariant vectors have subscripts, corresponding to the presence of $\partial x'^\alpha / \partial x^\beta$ and its inverse in the rule of transformation.

And so, we have started our discussion of tensors, because vectors are tensors of rank one. The **rank of a tensor** is determined by the number of indices it possesses, and you've seen above how we've designated vectors as $A^\alpha$ or $B_\beta$, that is, with one index (more on this below).

From equation (11.61) and equation (11.62), we see that tensors are defined by their transformation properties relative to a transformation of the underlying coordinate system $x \to x'$.

A **scalar** is a single quantity (an ordinary number, which may be real or complex) whose value is unchanged by such a transformation (of the underlying coordinate system), and is said to be a **tensor of rank zero**. Examples of scalars are numbers like $\pi$, and the inner product $\vec{A} \cdot \vec{B}$.

For vectors, or tensors of rank 1, we have identified two types. *Contravariant vectors* $A^\alpha \equiv (A^0, A^1, A^2, A^3)$ transform according to equation (11.61), where all the indices are in the superscript, as is the new primed coordinate. *Covariant vectors* $B_\alpha \equiv (B_0, B_1, B_2, B_3)$ transform according to equation (11.62).

Next, consider **tensors of rank two** (i.e., with two indices).

- A contravariant tensor of rank two, $F^{\alpha\beta}$, consists of 16 quantities that transform according to

$$F'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x'^\beta}{\partial x^\delta} F^{\gamma\delta} \tag{11.63}$$

- A covariant tensor of rank two, $G_{\alpha\beta}$ transforms as

$$G'_{\alpha\beta} = \frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\delta}{\partial x'^\beta} G_{\gamma\delta} \tag{11.64}$$

- The mixed second-rank tensor $H^\alpha_\beta$ transforms as

$$H'^\alpha_\beta = \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x^\delta}{\partial x'^\beta} H^\gamma_\delta \tag{11.65}$$

*all of which you wrote in Question 6 on today's worksheet.*

The **generalization** to contravariant, covariant, and mixed tensors of arbitrary rank should be obvious from the previous examples. We can form higher rank tensors by doing an outer product, in which we just take two tensors of some rank and multiply them component by component. For example, we can form the contravariant tensor of rank two, $F^{\alpha\beta}$, by doing

$$F^{\alpha\beta} = A^\alpha B^\beta$$

Going the other way, we can also reduce the rank of tensors by a process known as *contraction*. To contract two tensors, we set two of the indices to be equal and carry out the Einstein summation over the common range of indices. One index is contravariant and the other covariant.

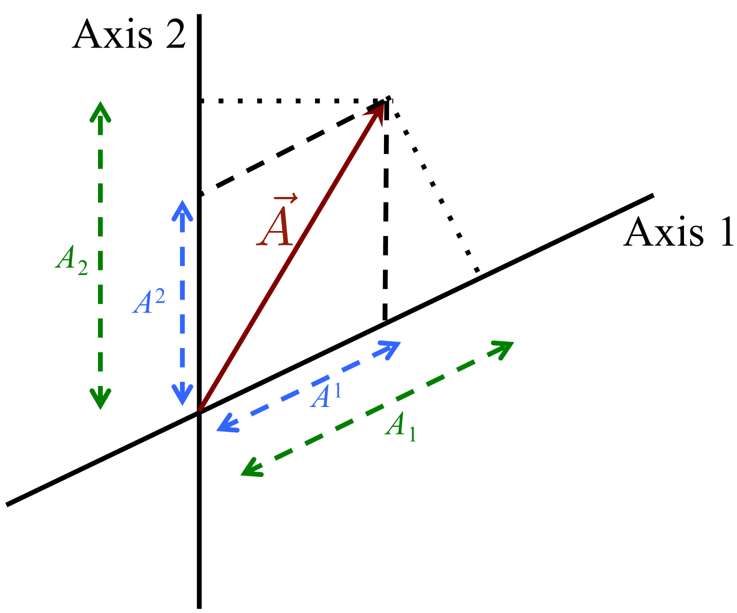
An example of such a contraction is the dot (or scalar) product of two vectors that we discussed previously in equation (11.66); recall that we defined the dot product as the product of the components of a covariant and a contravariant vector:

$$B \cdot A \equiv B_\alpha A^\alpha \tag{11.66}$$

where $A$ is a contravariant vector defined in equation (11.61), and $B$ is a covariant vector defined in equation (11.62), and the Einsteinian summation over repeated indices is implied.

## Appendix

The terms contravariant and covariant can be confusing. The terms are just used to identify two different conventions for interpreting components with respect to a coordinate system (keeping in mind that vector and tensor properties must be independent of the coordinate system). The following figure should help in distinguishing between the two.




From the figure, we see that contravariant components are projections of the vector onto an axis such that the projection (indicated by the dashed black line) is parallel to the other axis, whereas covariant components are a perpendicular projection of the vector onto an axis.