4-Vector Notation

Chris Clark

September 5, 2006

1 Lorentz Transformations

We will assume that the reader is familiar with the **Lorentz Transformations** for a boost in the x direction

$$\bar{x} = \gamma(x - vt)$$

$$\bar{y} = y$$

$$\bar{x} = z$$

$$\bar{t} = \gamma\left(t - \frac{v}{c^2}x\right)$$

(where $\gamma = 1/\sqrt{1-\beta^2}$ and $\beta = v/c$) and the thought experiments used to derive them. These equations contain all the geometrical information in the Special Theory of Relativity[1] We will take an axiomatic approach in this paper with these equations as our axioms. We will use the convention that $\mathbf{v}$ is the velocity of a moving reference frame and $\mathbf{u}$ is the velocity of a particle under consideration (except in the next section where they are the same).

2 Proper Time

Suppose an object is traveling at a velocity v along the x-axis in our reference frame, starting from the origin at time $t = 0$. Then in our coordinates the location of the object at time t is given by $x = vt, y = 0, z = 0$, and $t = t$. By performing a Lorentz transformation to the object's coordinates we get $\bar{x} = 0, \bar{y} = 0, \bar{z} = 0$, and

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

This means that during a time interval of length t in our reference frame, the object only traverses a smaller time interval $\bar{t} = t/\gamma$. So the rate of flow of time in a moving reference frame is diminished by a factor of γ with respect to the

[1] This statement is from Griffiths page 496.

stationary reference frame.[2] For an infinitesimal interval of time dt, we write the corresponding interval in the moving frame as

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

where τ is the **proper time**, which measures time in the moving reference frame.

3 Events

Lorentz transformation transform four distinct components, three space and one time. So if we wish to represent the transformations as operators on a vector space, then we will have to use a 4-D vector space, called Minkowski space, which is a 4-D pseudo-Riemann manifold. A **manifold** is basically a space that is locally Euclidean and a **Riemann manifold** is characterized by having a smoothly varying inner product. The "pseudo" here is included to indicate that we will not be using a true inner product, even though it is called one by physicists. In order to have all dimensions of the space have the same dimensions, we multiply the time dimension by the speed of light, c. Minkowski space can be thought of as a simple 4-D Cartesian space – there is nothing special about the space itself, it is the operations that we define on it that are special.

An **event** is a vector element of this space. This term entices one away from the abstraction, perhaps "Minkowski vector" would be a better term. But this is not the same thing as a 4-vector because 4-vectors are more general. An event can be written as

$$(x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \mathbf{x})$$

The Lorentz transformation equations in Minkowski space are obtained by replacing all occurrences of t and $\bar{t}$ with x^0/c and $\bar{x}^0/c$ respectively to account for the rescaling of the time axis.

$$\begin{aligned}
\bar{x}^0 &= \gamma \left(x^0 - \beta x^1 \right) \\
\bar{x}^1 &= \gamma \left(x^1 - \beta x^0 \right) \\
\bar{x}^2 &= x^2 \\
\bar{x}^3 &= x^3
\end{aligned}$$

The equations take on a more symmetrical form, reflecting the added symmetry in the dimensions of the four dimensions.

[2] The rate of flow must be uniformly diminished because nothing is changing during the interval, so there is no cause for the scale factor to vary.

4 4-Vectors and Contraction

We now define a **4-vector** to be any set of four components that transform in the same manner as an event under Lorentz transformations. [3] For an event, the components correspond directly to the components of a point in Minkowski space, but the components of a 4-vector can be anything, like velocity, energy, current, electric potential, and so on. We may write a 4-vector as

$$a^\mu = (a^0, a^1, a^2, a^3) = (a^0, \mathbf{a})$$

We now define the **contraction** of two 4-vectors a^μ and b^μ as

$$a^\mu \cdot b^\mu = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3$$

This is what we use to replace the inner product on our space, so we might call it a pseudo-inner product, but it is not a true inner product because it does not satisfy the requirement of positivity. For notational convenience we define the **covariant** 4-vector a_μ which differs from the **contravariant** 4-vector a^μ only in the sign of the zeroth component

$$a_\mu = (a_0, a_1, a_2, a_3) \equiv (-a^0, a^1, a^2, a^3)$$

Then Einstein summation notation allows us to write

$$a^\mu \cdot b^\mu = a^\mu b_\mu = a_\mu b^\mu$$

This definition of contraction is used because it has the very special property that every contraction of two 4-vectors is invariant under Lorentz transformations. Let's check that now.

$$\begin{aligned}
\bar{a}^\mu \bar{b}_\mu &= -\bar{a}^0 \bar{b}^0 + \bar{a}^1 \bar{b}^1 + \bar{a}^2 \bar{b}^2 + \bar{a}^3 \bar{b}^3 \\
&= -\gamma\left(a^0 - \beta a^1\right)\gamma\left(b^0 - \beta b^1\right) + \gamma\left(a^1 - \beta a^0\right)\gamma\left(b^1 - \beta b^0\right) + a^2 b^2 + a^3 b^3 \\
&= \gamma^2 \left(-a^0 b^0 + \cancel{\beta a^1 b^0} + \cancel{\beta a^0 b^1} - \beta^2 a^1 b^1 + a^1 b^1 - \cancel{\beta a^0 b^1} - \cancel{\beta a^1 b^0} + \beta^2 a^0 b^0\right) \\
&\quad + a^2 b^2 + a^3 b^3 \\
&= \gamma^2\left(\beta^2 - 1\right)\left(a^0 b^0 - a^1 b^1\right) + a^2 b^2 + a^3 b^3 \\
&= -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3 \\
&= a^\mu b_\mu
\end{aligned}$$

This only proves it for boosts in the x direction, but since no other directional assumptions were made, the x axis can be taken to point in any arbitrary direction.

Question: What is the concept behind the reason why this works out? Why should the time component be negated?

[3] Definition from Griffiths page 501

5 Specific 4-Vectors

Let's try to find some 4-vectors. The easiest way to do this is to start from the most trivial example of a 4-vector, an event. An event

$$x^\mu = (ct, \mathbf{x})$$

is clearly a 4-vector because an event must transform like an event under Lorentz transformations. Since we are in a vector space,

Any linear combination of 4-vectors is a 4-vector.

(we did not prove that it is a vector space here). Be careful not to think that γ or β can be used as the constant coefficients in the linear combination because they are not constant with respect to Lorentz transformations. Recall that Minkowski space is just a normal Cartesian space, so when we add 4-vectors the components just add.

Now notice that

$$d\tau^2 = dt^2/\gamma^2 = (1-\beta^2)dt^2 = dt^2 - d\mathbf{x}^2/c^2 = -x^\mu x_\mu/c^2$$

so the square of the proper time is a Lorentz invariant. But then we also have

$$d\tau' = d\tau$$

since it cannot flip its sign since that would mean time was going backwards, which doesn't happen. Therefore the proper time $d\tau$ is a Lorentz invariant and therefore it is a constant for a given situation.[4] This means that we can multiply and divide 4-vectors by $d\tau$ to obtain new 4-vectors. Now, da^μ is a 4-vector since it is a linear combination of two 4-vectors so if we divide this by $d\tau$ we obtain the proper-time derivative. Therefore

The proper-time derivative of a 4-vector is a 4-vector.

But the normal-time derivative of a 4-vector is not a 4-vector.

So we automatically know that the **proper velocity 4-vector** or **4-velocity**

$$U^\mu = \frac{dx^\mu}{d\tau} = \gamma\frac{dx^\mu}{dt} = \gamma\left(\frac{d(ct)}{dt}, \frac{d\mathbf{x}}{dt}\right) = (\gamma c, \gamma\mathbf{u})$$

is a 4-vector.[5]

Now what we really want to find is the relativistic momentum and energy. But how are these quantities defined? We can't really get the answer from classical mechanics, but we can take a look at classical mechanics to see how they did it back then. Newton defined momentum as mass times velocity. His

[4] I would like a better conceptual explanation for why all observers measure the same proper time.

[5] Note the sloppy use of differentials. This is known as non-standard analysis and it has been shown to be correct.

laws implies that momentum was conserved, but they gave no reason for the definition of momentum, that was based on experimental evidence. So if the definition of momentum is not exactly right, then its not obvious how to fix it. What really is the definition of momentum? Rather than defining a quantity by its expression, a more abstract and versatile method is to define it in terms of the properties that you want it to have. Therefore momentum should be defined as a conserved dynamical vector quantity proportional to the velocity of a particle. Here, dynamical means that it can only depend on dynamical quantities such as position, velocity, acceleration, mass, energy, and so on. So using some arguments based on an elastic collision between two particles viewed from different reference frames, one can find that

$$\mathbf{p} = \gamma m \mathbf{u}$$

is the proper relativistic definition of momentum. The only assumptions that go into this are the Lorentz transformation equations (I think).[6] Feynman points out in Volume I that after you realize that the mass m is replaced by the relativistic mass γm_o in the classical equations then you are all set with relativistic dynamics.

Now, the classical formalism proves that energy is conserved when momentum is defined as $\mathbf{p} = m\mathbf{u}$. But the derivation never requires that mass is constant (you never have to use $F = ma$, you can just use $F = dp/dt$), so the equivalent relativistic energy defined from this relativistic momentum will also be conserved. Again, we define energy abstractly as the conserved scalar dynamical quantity. So now we derive the expression for the relativistic energy. For this section we will use the relativistic mass

$$m = \gamma m_0$$

Now consider motion in one dimension. Notice that

$$\frac{dm}{dv} = \frac{m_0 v/c^2}{(1 - v^2/c^2)^{3/2}} = \frac{v}{c^2 - v^2} m$$

$$\begin{aligned} F = \frac{dp}{dt} = \frac{d(mv)}{dt} &= \frac{dm}{dt} v + m\frac{dv}{dt} = \frac{dm}{dt} v + \frac{c^2 - v^2}{v} \frac{dm}{dv} \frac{dv}{dt} \\ &= \frac{dm}{dt} v + \frac{c^2 - v^2}{v} \frac{dm}{dt} \\ \Rightarrow dE = F dx &= v^2 dm + (c^2 - v^2) dm = dmc^2 \\ &\Rightarrow E = mc^2 + C \end{aligned}$$

But when $m_0 = 0$, $m = 0$ and $E = 0$ since there is no particle there, so $C = 0$. Therefore

$$E = mc^2$$

[6] See `http://www.geocities.com/physics_world/sr/inertial_mass.htm`

or in our normal notation

$$E = \gamma mc^2$$

When you see the famous $E = mc^2$, the m is supposed to be the relativistic mass. [7]

Since we can multiply by a constant and still get a 4-vector, we multiply the 4-velocity by the rest mass m to get the **momentum 4-vector**

$$p^\mu = (\gamma mc, \gamma m\mathbf{u}) = \left(\frac{E}{c}, \mathbf{p}\right)$$

This has been derived for massive particles and the expressions $\mathbf{p} = \gamma m\mathbf{u}$ and $E = \gamma mc^2$ only apply to massive particles, but the last expression for p^μ is true even for photons. However, we haven't derived this, it is just fortunate that it works out to be correct for massless particles also.

Question: Does invariance of pseudo-length (contraction with itself) under arbitrary Lorentz transformations imply an object is a 4-vector? It seems that many four vectors are defined by a correspondence between the expression for their pseudo-length and some universal law. It would be nice if the answer to this question was yes and we had a proof of it because then we could derive

$$J^\mu = (c\rho, \mathbf{J}) \quad \text{and} \quad A^\mu = (\Phi, \mathbf{A})$$

Question: Is there an algorithm for checking whether a given four component set is a 4-vector? So far the best I can tell you is to first cancel off all constants and Lorentz invariants, including the rest mass m, c, and $d\tau$ since they won't affect the answer. Then if the rest is a linear combination of 4-vectors you have a 4-vector. Otherwise I don't know what to do next.

6 Invariant Quantities

$$U^\mu = (\gamma c, \gamma\mathbf{u}) \Rightarrow U^\mu U_\mu = -\gamma^2 c^2 + \gamma^2 u^2 = \gamma^2 c^2 (u^2/c^2 - 1) = -c^2$$

$$\Rightarrow U^\mu U_\mu = -c^2$$

$$p^\mu = \left(\frac{E}{c}, \mathbf{p}\right) \Rightarrow p^\mu p_\mu = -\frac{E^2}{c^2} + p^2$$

$$p^\mu = mU^\mu \Rightarrow p^\mu p_\mu = m^2 U^\mu U_\mu = -m^2 c^2$$

$$\Rightarrow E^2 = p^2 c^2 + m^2 c^4$$

So relativistic momentum and energy themselves are not Lorentz invariant, but the particular combination of them in $p^\mu p_\mu$ is equal to $-m^2c^2$, which is a constant, so that quantity is Lorentz invariant.

[7] Derivation of $E = mc^2$ from the rest mass relation was found at `http://www.karlscalculus.org/einstein.html`