A Neat Mathematical Description of Units and The Geometry of Moving Bodies

Herkus Strasunskas

April 13, 2026

1 Units

1.1 Defining Units

When we measure things in the real world, we measure them using units. For this section it will be important to pay attention to the fact that not only do units tell you the "magnitude" of the unit that is being used, but they also tell you what it is that you're measuring in the first place. The big question is: what exactly is it that we are doing when we multiply units together? As an example, let's look at

$$\mathrm{N} = \mathrm{kg} \cdot \frac{\mathrm{m}}{\mathrm{s}^2} = \mathrm{kg} \cdot \mathrm{m} \cdot \mathrm{s}^{-2}. \tag{1}$$

Here, not only are we multiplying units together, we are also dividing by units as well. Simultaneously, it makes no sense to add units that are not "the same"?

$$\mathrm{m} + \mathrm{kg} = (\text{error}). \tag{2}$$

What does it mean for units to "not be the same"? What are these mysterious objects we call "units" really? Well perhaps a way we can tackle this question is by introducing some **common unit space** $(\mathcal{U}_c, r)$. We want some kind of map that we will call the **ratio map** $r : \mathcal{U}_c \times \mathcal{U}_c \to \mathbb{R}$ to act as something that compares how much one unit "fits" into another (think $\frac{1\ \mathrm{in}}{1\ \mathrm{cm}} = 2.54$). We can now complete the set of axioms for r:

1. $\forall a, b \in \mathcal{U}_c, r(a, b) = 1 \implies a = b$. If the "ratio of a, and b" is 1, they are the same unit (think $\frac{\mathrm{m}}{\mathrm{m}} = 1$, or $\frac{100\mathrm{cm}}{\mathrm{m}} = 1$).
2. $\forall a, b \in \mathcal{U}_c, r(a, b) = \frac{1}{r(b,a)}$ (think $\frac{\mathrm{in}}{\mathrm{cm}} = \frac{1/\mathrm{cm}}{1/\mathrm{in}} = 2.54 = \frac{1}{1/2.54}$).
3. $\forall a, b, c \in \mathcal{U}_c, r(a, c) = r(a, b) \cdot r(b, c)$ (think $\frac{\mathrm{yd}}{\mathrm{in}} = \frac{\mathrm{yd}}{\mathrm{ft}} \cdot \frac{\mathrm{ft}}{\mathrm{in}} = 3 \cdot 12 = 36$).

So now that we have properly defined the notion of "common units," like meters, miles, feet, etc., or kilograms, pounds (mass), ounces (mass), etc., we can start thinking about what it is that we're actually doing when we multiply units together. It will prove useful to view the act of multiplying units together as the addition of elements from some group. Let's define the **unit group** $(\mathcal{U}, +)$, where $\mathcal{U}$ contains every single possible common unit space $(\mathcal{U}_c, r)$. This group will actually turn out to fulfill every requirement we would want it to fulfill, given that we provide the proper axioms for the addition operation $+$ that we would normally expect from any old addition operation:

1. Commutativity: $\forall a, b \in \mathcal{U}, a + b = b + a \in \mathcal{U}$.
2. Associativity: $\forall a, b, c \in \mathcal{U}, (a + b) + c = a + (b + c) \in \mathcal{U}$.
3. Additive identity: $\exists \mathbf{0} \in \mathcal{U}, s.t. \forall a \in \mathcal{U}, a + \mathbf{0} = a$.
4. Dual element: $\forall a \in \mathcal{U}, \exists(-a) \in \mathcal{U}, s.t.\ a + (-a) = \mathbf{0}$.

So now we have a way of expressing what we mean by $\mathrm{N} \cdot \mathrm{m} = \mathrm{J}$. What we're actually doing is we're taking two separate common unit spaces $[\text{Force}], [\text{Length}] \in \mathcal{U}$, and using them to create another space:

$$[\text{Force}] + [\text{Length}] = [\text{Work}] \in \mathcal{U}. \tag{3}$$

Notice that the space of real numbers is the additive identity element of the unit group:

$$(\mathbb{R}, r) = \mathbf{0} \in \mathcal{U}, \tag{4}$$

since for any common unit space $a \in \mathcal{U}, a + (\mathbb{R}, r) = a$, as in $2 \cdot \mathrm{ft} = 2\mathrm{ft}$ (you're still "measuring the same thing"). But when we speak of elements of the sets from these spaces, we **multiply** them together, instead of adding them (as in our example of $\mathrm{N} \cdot \mathrm{m} = \mathrm{J}$). This means that, loosely speaking the difference between individual units ($\mathrm{N} \cdot \mathrm{m} = \mathrm{J}$) and the spaces of those units ($[\text{Force}] + [\text{Length}] = [\text{Work}]$) is the same as the difference between numbers $\in \mathbb{R}$ and the logarithms of those numbers. This means we can define a "logarithm map" $\log_{\mathcal{U}} : \mathcal{U}_c \to \mathcal{U}$, and use it to very elegantly "zoom out" to see what it is that we're measuring when we use some unit $a \in \mathcal{U}_c$. For instance,

$$\log_{\mathcal{U}}(\mathrm{m}) = [\text{Length}]. \tag{5}$$

The fact that the dual element exists actually explains division of units. For instance, the inverse second s^{-1} is simply part of the **dual time space**! If we use our **unit group logarithm** and treat it as any other logarithm, we get

$$\log_{\mathcal{U}}(\mathrm{s}^{-1}) = -\log_{\mathcal{U}}(\mathrm{s}) = -[\text{Time}] \implies \mathrm{s} \cdot \mathrm{s}^{-1} \in \mathbb{R}. \tag{6}$$

Hence we can finally–very elegantly–explain why we can't add two separate, nonzero units together (by a nonzero unit $a \in \mathcal{U}_c$, I mean $\log_{\mathcal{U}}(a)$ exists; the unit whose unit group logarithm does not exist is equivalent to what we know as the number 0): it is because they are different elements from completely different common unit spaces. This means if there exists an equation $a + b = c$, where a, b, and c are units, these units must all satisfy $\log_{\mathcal{U}}(a) = \log_{\mathcal{U}}(b) = \log_{\mathcal{U}}(c)$.

We can now see how nicely different units are related to one-another using this new formalism. If you're a runner, for instance, you might be used to measuring how fast you're moving with something called "pace," being the reciprocal–or in this formal model, the negative–of speed. We can show how all of these units are related using the unit group:

$$\begin{aligned} &[\text{Length}] - [\text{Time}] = [\text{Speed}] \\ \implies &-[\text{Speed}] + [\text{Length}] = [\text{Pace}] + [\text{Length}] = [\text{Time}]. \end{aligned} \tag{7}$$

1.2 Further Analysis

All of this seems to point to the fact that units of length, for instance, are not the same unit in every direction. If I have some length space in the x-direction, $[\text{Length}]_x$, and one in the y-direction, $[\text{Length}]_y$, I cannot add these units together, as they are from completely separate spaces. This explains why you have squares of lengths in the Pythagorean theorem: it is because, although we cannot add two units "from two separate directions," we can pull them out of an area space:

$$[\text{Length}]_x + [\text{Length}]_y = [\text{Area}]_{xy}. \tag{8}$$

This new space $[\text{Area}]_{xy}$ forms a plane that contains a set of directions, which then allows us to go from knowing measurements of two lengths from two different common unit spaces to a measurement in a common unit space that intersects these two length spaces. This means if we are using meters to measure length in both of these directions, we would end up with units of m^2 for $[\text{Area}]_{xy}$, when something like $\text{m}_x \cdot \text{m}_y$ would be more descriptive. This is essentially the distinction between "square meter" and "meter squared"; the former is actually referring to the common unit space

$$[\text{Area}]_{xy} = [\text{Length}]_x + [\text{Length}]_y, \tag{9}$$

whereas the latter is part of the unit space

$$2[\text{Length}]. \tag{10}$$

Really, if we wanted to be more descriptive, the Pythagorean theorem simply utilizes the fact that $2[\text{Length}]_x$, $2[\text{Length}]_y$, and $2[\text{Length}]_s$ (where the "direction" of s is within the same plane as x and y) are "collapsed" spaces of $[\text{Area}]_{xy}$ to transform between them. This also reveals what fractional dimensions mean: they are exactly like integer dimensions. The reason why you can't imagine a $\text{m}^{1/2}$ is because you're thinking too hard! if you could gaze upon this unit, you'd just see a regular line (or an area if we "decolapse" it further, or a volume, etc.). The reason why we can't create a $\text{m}^{1/2}$ is because the space $[\text{half-Length}]_1 + [\text{half-Length}]_2 = [\text{Length}]_{12}$ collapses into the same length space ($2[\text{half-Length}]_i = [\text{Length}]_{12}$), and we just end up seeing a single length, as opposed to some area or something else.

2 The Metric Tensor

2.1 Introduction

Let's suppose we have two maps $\gamma : X \subseteq \mathbb{R} \to L \subseteq U$ (where U is an open set on the manifold with a map ϕ that defines the coordinate system), and $s : L \to S \subseteq \mathbb{R}$. Let's look at $\gamma(\lambda) = p \in U$, where λ is a measured distance with respect to the coordinate system. We are interested in the limit of the quotient $\frac{ds}{d\lambda}$, since we want to know how the coordinate system gets "stretched" as you go along the manifold. Now, we have to be a little bit careful when using the chain rule to rewrite it in terms of the coordinate system x^μ, because the map s is only defined for points that are along a single line, meaning $\frac{\partial(s \circ \phi^{-1})}{\partial x^\mu}$ doesn't make any sense.

The problem is fixed, however, if we consider maps $s^{(\mu)} : \phi^{-1}(x^\mu) \to S^{(\mu)} \subseteq \mathbb{R}$, where the parenthesized (x^μ) indicates that all the coordinates are 0, except for x^μ, and $s^{(\mu)} \circ \phi^{-1}(0) = 0$. It is here where we have to also introduce a new convention, so as to not cause confusion: if an index is parenthesized, it shall be assumed that there is no sum over that index. Using these paths, we can expand our original derivative:

$$\frac{d(s \circ \gamma)}{d\lambda} = \frac{d(s^{(\mu)} \circ \phi^{-1}(x^\mu))}{dx^\mu} \frac{dx^\mu}{d\lambda}. \tag{11}$$

This implies this is a tensor $\frac{ds^{(\mu)}}{dx^\mu} dx^\mu$ acting on a vector $\frac{d}{d\lambda}$. But we're not entirely sure what we are to do with this tensor. Is it even useful? What does it mean? This is where we try our luck by squaring our original expression:

$$\left(\frac{ds}{d\lambda}\right)^2 = \left(\frac{ds}{d\lambda}\right)\left(\frac{ds^{(\mu)}}{dx^\mu}\frac{dx^\mu}{d\lambda}\right) = \left(\frac{ds}{d\lambda}\right)\frac{ds^{(\mu)}}{dx^\mu} dx^\mu \left(\frac{d}{d\lambda}\right). \tag{12}$$

We have to, again, be careful when expanding the other $\frac{ds}{d\lambda}$. If we were to expand it the same way we expanded the other factor, and sloppily let it "join" the other dual vector to form a tensor that's acting on $\left(\frac{d}{d\lambda}, \frac{d}{d\lambda}\right)$, we would have a problem, because, for the cross-terms, the derivatives would be acting on two different paths, $s^{(\mu)}$ and $s^{(\nu)}$:

$$\left(\frac{ds}{d\lambda}\right)^2 = \frac{ds^{(\mu)}}{dx^\mu}\frac{ds^{(\nu)}}{dx^\nu} dx^\mu dx^\nu \left(\frac{d}{d\lambda}, \frac{d}{d\lambda}\right). \tag{13}$$

This would result in an inconsistent "focus" for those terms, which is not what should be happening. Here's the fix: for every path $s^{(\mu)}$, we define a scalar field $s_\phi^{(\mu)} : U \to S_\phi^{(\mu)} \subseteq \mathbb{R}$, such that

$$\frac{\partial s_\phi^{(\mu)}}{\partial x^\mu} = \frac{ds^{(\mu)}}{dx^\mu}, \tag{14}$$

and

$$s_\phi^{(\mu)} \circ \phi^{-1}(x^\mu) = s^{(\mu)} \circ \phi^{-1}(x^\mu). \tag{15}$$

This allows us to expand Eq. 5, and write

$$\left(\frac{ds}{d\lambda}\right)^2 = \left(\frac{ds}{d\lambda}\right)\frac{\partial s_\phi^{(\mu)}}{\partial x^\nu}\frac{\partial x^\nu}{\partial x^\mu} dx^\mu \left(\frac{d}{d\lambda}\right) = \left(\frac{ds}{d\lambda}\right)\frac{\partial s_\phi^{(\mu)}}{\partial x^\nu} dx^\nu \left(\frac{d}{d\lambda}\right). \tag{16}$$

This allows us to "separate" the two indices, μ and ν, and thus we can write

$$\left(\frac{ds}{d\lambda}\right)^2 = \frac{ds^{(\mu)}}{dx^\mu}\frac{\partial s_\phi^{(\mu)}}{\partial x^\nu} dx^\mu dx^\nu \left(\frac{d}{d\lambda}, \frac{d}{d\lambda}\right). \tag{17}$$

So now, we have successfully "separated" the $(0, 2)$ tensor $\frac{ds^{(\mu)}}{dx^\mu}\frac{\partial s_\phi^{(\mu)}}{\partial x^\nu} dx^\mu dx^\nu$ from the $(2, 0)$ tensor $\left(\frac{d}{d\lambda}, \frac{d}{d\lambda}\right)$. Now we can think of $\left(\frac{ds}{d\lambda}\right)^2$ as a $(0, 2)$ tensor that acts on two vectors and spits out a real number. So let's make it act on two different vectors:

$$\begin{aligned}\frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}dx^{\mu}dx^{\nu}\left(\frac{d}{d\lambda},\frac{d}{d\lambda_1}\right) &= \frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}\frac{dx^{\mu}}{d\lambda_1}\frac{dx^{\nu}}{d\lambda} \\ &= \frac{ds^{(\nu)}}{dx^{\nu}}\frac{\partial s_{\phi}^{(\nu)}}{\partial x^{\mu}}\frac{dx^{\nu}}{d\lambda_1}\frac{dx^{\mu}}{d\lambda} \\ &= \frac{ds^{(\nu)}}{dx^{\nu}}\frac{\partial s_{\phi}^{(\nu)}}{\partial x^{\mu}}dx^{\mu}dx^{\nu}\left(\frac{d}{d\lambda_1},\frac{d}{d\lambda}\right)\end{aligned} \tag{18}$$

(I am able to swap $\mu \to \nu$ and $\nu \to \mu$, because we are summing over all of the possible indices anyway, so it doesn't matter what we call them). Now, the two sets of paths $s^{(\mu)}$ and $s^{(\nu)}$ have to be the same, because we're dealing with the exact same coordinate system. And since we're still summing over all indices, it means the two tensors have to be the same:

$$\frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}dx^{\mu}dx^{\nu} = \frac{ds^{(\nu)}}{dx^{\nu}}\frac{\partial s_{\phi}^{(\nu)}}{\partial x^{\mu}}dx^{\mu}dx^{\nu}. \tag{19}$$

Meaning this $(0,2)$ tensor has this special commutative property:

$$\frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}dx^{\mu}dx^{\nu}\left(\frac{d}{d\lambda},\frac{d}{d\lambda_1}\right) = \frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}dx^{\mu}dx^{\nu}\left(\frac{d}{d\lambda_1},\frac{d}{d\lambda}\right). \tag{20}$$

So, no matter the order in which I let the tensor act on the two vectors, I will get the same result. This means I can replace $\frac{d}{d\lambda}$ with another vector (meaning this tensor would be acting on two vectors that are not related to the original path) and still end up with that same commutative property. It therefore follows that no matter what two vectors we pick on this tangent vector space, we will still get this exact same commutative property. This is groundbreaking, because we can now define the dot product of any two vectors that lie on the tangent space T_p as simply this tensor acting on the two vectors. It means we can "relate" two vectors together using this tensor. Now notice that, in general, there has to be some different $\frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}$ for virtually every point on the n-manifold (if we don't constrain $n = 1$). This is because there are infinitely many paths running through those same points, meaning (if we're not constraining $n = 1$) we would have to consider infinitely many paths all over the manifold in order to calculate the coefficients of the tensor. Usually, for efficiency, we refer to these coefficients as $g_{\mu\nu}$. This means we can define, what is referred to as the **metric tensor**, as

$$g = g_{\mu\nu}dx^{\mu}dx^{\nu}, \tag{21}$$

where the coefficients $g_{\mu\nu}$ are defined as

$$g_{\mu\nu} = \frac{ds^{(\mu)}}{dx^{\mu}}\frac{\partial s_{\phi}^{(\mu)}}{\partial x^{\nu}}. \tag{22}$$

We can also look at our original path again and calculate Δs over an interval $\Delta\lambda$:

$$\left(\frac{ds}{d\lambda}\right)^2 = g_{\mu\nu}\frac{dx^{\mu}}{d\lambda}\frac{dx^{\nu}}{d\lambda}. \tag{23}$$

If we take the square root on both sides (assuming $\frac{ds}{d\lambda} > 0$), and integrate with respect to λ,

$$\begin{aligned}\Delta s &= \int_{\lambda_0}^{\lambda_0+\Delta\lambda} \frac{ds}{d\lambda}\, d\lambda = \int_{\lambda_0}^{\lambda_0+\Delta\lambda} d(s \circ \gamma(\lambda)) \\ &= \int_{\lambda_0}^{\lambda_0+\Delta\lambda} \sqrt{g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\, d\lambda. \end{aligned} \tag{24}$$

2.2 Applications of The Metric Tensor

It may also be useful to remember Eq. 15, as it can be used when wanting to measure the coefficients for real-life geometries. For instance, say we wanted to measure those coefficients for a flat piece of paper. We can draw our coordinate system with the familiar, perpendicular, x- and y-axes. If we were to define our two paths as ones that are going along those axes, we will notice that, if we align our coordinate system so that a change in x or y = a change in the "number of units" on our "ruler," $\frac{ds^{(x)}}{dx} = \frac{ds^{(y)}}{dy} = 1$ (for all x and y). The other derivatives, $\frac{\partial s_\phi^{(x)}}{\partial y}, \frac{\partial s_\phi^{(y)}}{\partial x}$, are a little bit more tricky; what exactly are these scalar fields? Well, using our earlier definitions of s^μ and s_ϕ^μ, we know

$$s_\phi^{(x)} = x + f(y),\ s_\phi^{(y)} = y + g(x),\ f(0) = g(0) = 0. \tag{25}$$

Now, due to the commutativity of the metric tensor, it also follows that

$$g_{\mu\nu} = g_{\nu\mu}. \tag{26}$$

Since $\frac{ds^{(x)}}{dx} = \frac{ds^{(y)}}{dy} = 1$, this would mean

$$\frac{\partial s_\phi^{(x)}}{\partial y} = \frac{\partial s_\phi^{(y)}}{\partial x} = C \tag{27}$$

(a constant, since that's the only way these derivatives can be equal). Knowing this, we can use Eq. 17, and the fact that all of the derivatives are constant to get:

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + 2C(\Delta x)(\Delta y). \tag{28}$$

And, after a bit of basic algebra, we get a formula for calculating C for measured values of Δs, Δx, and Δy:

$$C = \frac{(\Delta s)^2 - (\Delta x)^2 - (\Delta y)^2}{2(\Delta x)(\Delta y)}. \tag{29}$$

If we do this on our *orthogonal* coordinate system, we will get that $C = 0$. This means the coefficients of our metric tensor are just 1 (diagonal) and 0 (off-diagonal). Thus, we have arrived at the famous **Pythagorean theorem**:

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2. \tag{30}$$

If we were to choose a set of coordinates that weren't orthogonal, we would have defined $C = \cos\theta$, where θ would be the familiar angle between the axes. From that, we would've ended up with the **law of cosines**:

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + 2(\Delta x)(\Delta y)\cos\theta. \tag{31}$$

I would like to take a moment to mention a very dangerous abuse of notation. Sometimes, when people get introduced to the metric tensor, they are told that it can be written as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{32}$$

But this is a very slippery slope. It does seem like it kind of makes sense, since it defines the notion of "length" between two points, and the notation ds^2 seems like it defines a tensor product of two dual vectors, $ds \otimes ds$. But that's not it at all! It makes no sense to define a dual vector using a path, because a dual vector is supposed to exist at a point and needs no path.

3 Viewing Relativistic Effects as Geometry

In relativity, it is taken as an axiom that all bodies part of the spacetime manifold travel at the speed of light:

$$g_{\mu\nu}\frac{dx^\mu}{d\tau'}\frac{dx^\nu}{d\tau'} = g_{\mu\nu}U^\mu U^\nu = c^2, \tag{33}$$

where U is the 4-velocity of a body that's parameterized by τ' (τ is reserved for the x^μ frame for the sake of clarity), and c is the speed of light. Now we can use the parameter τ of the x^μ frame to parameterize this path:

$$g_{\mu\nu}\frac{dx^\mu}{d\tau'}\frac{dx^\nu}{d\tau'} = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\left(\frac{d\tau}{d\tau'}\right)^2 = g_{\mu\nu}v^\mu v^\nu \gamma^2 = c^2, \tag{34}$$

where v is the "Newtonian velocity," in the sense that $v^0 = c$, while the rest of the components are the familiar "regular" velocities from good ol' clasical mechanics that can be measured using the clock of the x^μ frame (their magnitude, $\sqrt{v_i v^i}$, will be abbreviated with v). Now, the metric of flat spacetime (in the mostly-minuses signature) looks like this:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{35}$$

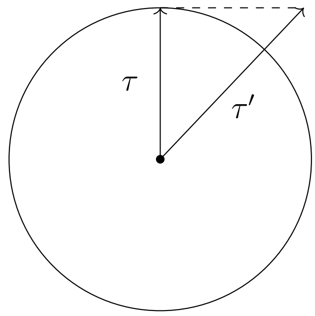

Figure 1: The projection of the path results in a shorter length (in lorentzian geometry the lengths flip in size)

so, aligning the path onto the x^0-x^1-plane, we have

$$\left[(v^0)^2 - (v^1)^2\right]\gamma^2 = \left[c^2 - v^2\right]\gamma^2 = c^2. \tag{36}$$

And hence we have solved for the **Lorentz factor**, γ:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{37}$$

Notice that we have not assumed anything about how light travels forward and back (the way it's usually derived). The Lorentz factor is actually just a choice of "reparameterization" using the observer's frame x^μ. It's not that we cannot know the one-way speed of *light*, more so the one-way speed of the object that is actually traveling in some direction. The speed of light is actually 0, as far as the spacetime manifold is concerned; since a light path is one that has a spacetime interval of 0, taking its derivative with respect to any parameter τ' will also result in 0. The weird thing about light is that it does not travel at all; it is a particle that you can leave behind in a stationary position until your friend perhaps runs into it. what's weirder is that, due to the geometry, your friend can also leave behind a particle of light in the other direction which you will run into. In this sense, both of you are "ahead" of one-another, because the Lorentzian geometry turns one point on the manifold into two *divergent* points. Really, the phenomenon of time dilation is simply a result of a projection from the $x^{\mu'}$ frame to the x^μ frame, which is why clocks always seem to tick slower for "the other observer" (Fig. 1).