## Special Relativity: Tensor Calculus and Four-Vectors

Looking ahead to *general* relativity, where such things are more important, we will now introduce the mathematics of tensors and four-vectors.

### The Mathematics of Spacetime

Let's start by defining some geometric objects. Bear with me for the first couple, which seem obvious but lay the groundwork for the less obvious sequels.

*Scalar.*—A scalar is a pure number, meaning that all observers will agree on its value. For example, the number "3" is a scalar. Okay, that one's trivial, but there are others that aren't so obvious, and we'll get to those in a bit.

*Event.*—Next we have an *event*. An event is effectively a "point" in spacetime. More generally, if you have an N-dimensional space, you need N numbers to label it uniquely. For example, in two dimensions you need two numbers; e.g., $x$ and $y$ for a plane, or $\theta$ and $\phi$ for the surface of a sphere. For spacetime, you need four numbers: e.g., $t$, $x$, $y$, and $z$. Naturally, the essence of the event isn't changed if you relabel the coordinates. I want to stress this, because something that is obvious for events but may not be obvious for some other geometric objects is that although when you finally calculate something you may choose a coordinate system and break things into components, there is also an independent reality (well, within the math at least!) of the objects. Going with the coordinate-free representation has proved very helpful in proving theorems about spacetime, but when doing astrophysics it is usually best to investigate components in some given system.

*Vector.*—Next, consider a *vector*. Normally we think about three-dimensional vectors, but here we have to consider four-vectors. This is something that has four numbers; for example, the location of an event can be written as $x^\mu = (ct, x, y, z)$. You can also imagine a line segment drawn between two events: this is a vector as well. Four-velocities are also vectors.

In the geometric sense, it is important to work with four-vectors rather than the more familiar purely spatial three-vectors. The reason is just as an event can be thought of as a purely geometric object with existence independent of the coordinate system, so can a four-vector. This is not the case with three-vectors. For example, consider an electric field at a given location. This is a three-vector. You know from electromagnetism that under many circumstances you can boost into a frame where the electric field vanishes. Three-vectors, then, are not coordinate-independent geometric objects.

*One-form.*—Events and vectors are pretty familiar. Not so with the next object. This is a *one-form*. Think again about a Euclidean plane, and two points very close to each

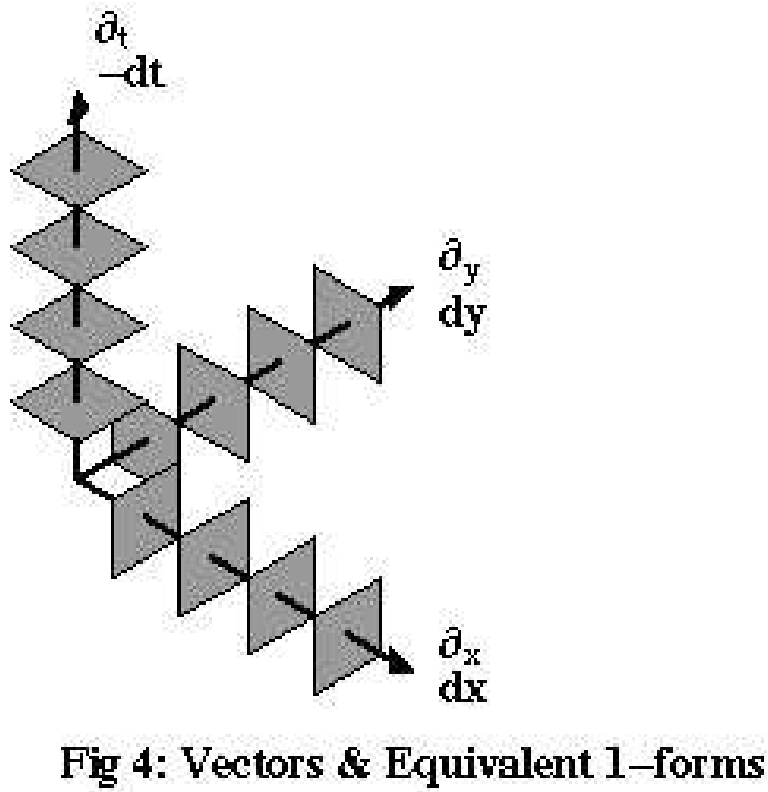


Fig. 1.— Vectors and their corresponding one-forms. From http://gregegan.customer.netspace.net.au/FOUNDATIONS/03/Fig04.gif

other. You can define a vector between them. But you could equally well define something perpendicular to that vector (see the figure). If you imagined two points in three dimensions, then after drawing a vector you could draw a plane perpendicular to it. In four dimensions (spacetime), you would have a three-dimensional thing perpendicular to the vector. This is called a one-form. One-forms are written differently than vectors. For example, consider the radial component of a velocity $\mathbf{v}$. It is written $v^r$, with a superscript. The $r$-component of the corresponding one-form would be written $v_r$, with a subscript. These are also called, respectively, *contravariant* (up) and *covariant* (down) components. In the GR lectures will get into how to transform between the two, by raising and lowering indicies.

*Tensor.*—One can generalize with the further concept of *tensors*. Think, for example, of the gradient of an electric field: $K \equiv \nabla\mathbf{E}$ (this isn't a tensor itself, because it involves just the spatial components, but we're using this as an analogy). At a given point, if you want to know the components of this you can't just specify its $x$ or $y$ component, for example. Instead, you need to specify things like the gradient of the $x$ component of $E$, in the $y$ direction. Then $K$ might have components like $K^{xy}$ or $K_{zz}$, depending on whether one wanted to go with a contravariant or covariant description. We'll get to how to manipulate tensors and their indices a little later. For now, it is also useful to think of a tensor as a machine that can take vectors or one-forms as inputs in its "slots", one slot per index, and return a number. It is a linear machine. As with vectors and one-forms, a tensor has a mathematical existence of its own, independent of coordinate systems, but when calculating it is usually convenient to select a particular coordinate system and compute with components.

The "rank" of the tensor is the number of separate indices it has. For example, $T^{\mu\nu}$ is a second-rank tensor and $R^{\alpha}_{\beta\gamma\delta}$ is a fourth-rank tensor. One especially important second-rank tensor is the metric tensor, which we'll talk about now.

**Metrics**

Now let's move a little from those basic definitions to how they are used in curved spacetime (yes, we're dealing with flat spacetime now but we want an easy generalization to other spacetimes). A great way to characterize the curvature is through the use of a *metric*. This effectively tells you the "distance" between two events. For example, in two dimensions, what is the distance between two events separated by $dx$ in the $x$ direction and $dy$ in the $y$ direction? The distance $ds$ is just given by $ds^2 = dx^2 + dy^2$. In polar coordinates we would write $ds^2 = dr^2 + r^2 d\theta^2$, but it's the same thing. One point (obvious here, not so obvious later) is that this distance is the same for two given events even if the coordinates are redefined (e.g., by rotation of the coordinate system). Therefore, as we did in the last lecture, we call this the invariant *interval* between the events.

In four dimensions and flat space, special relativity tells us that the invariant interval is defined as

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \ . \tag{1}$$

For two arbitrary events, $ds^2$ can be positive, negative, or zero. **Ask class:** consider just $dt$ and $dx$, so that $ds^2 = -c^2 dt^2 + dx^2$. What is the condition that $ds^2 = 0$? This is the condition that the two events could be connected by a photon going from one to the other. Therefore, $ds^2 = 0$ is the path of a light ray.

We can represent this more compactly, using the *metric tensor* $g_{\alpha\beta}$, as

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta \ . \tag{2}$$

Here we have introduced two conventions. The first is the use of greek indices to represent indices that might be any of the four in spacetime (for example, $t$, $x$, $y$, and $z$). The second is the Einstein summation convention. Whenever you see a symbol used as an up and a down index in the same expression, you are supposed to sum over the four possibilities. For example, $v^\alpha u_\alpha = v^t u_t + v^x u_x + v^y u_y + v^z u_z$. This also means that if you sum over indices, you don't count them in the rank of the tensor. For example, $v^\alpha u_\alpha$ is a scalar, $T^{\alpha\beta} u_\alpha$ is a rank one tensor, and $R^\alpha_{\beta\alpha\gamma}$ is a rank two tensor. This means that $ds^2$ is a scalar, so for example it transforms as a scalar does under Lorentz transformations, i.e., it's unchanged (that's why it is an "invariant" interval!).

Let's take a moment to emphasize that point a bit more. *Scalars are unchanged under Lorentz transformations!* At a given "point" (i.e., event) in spacetime, *everyone* will agree on the value of a scalar. If this isn't clear, just think about a trivial example of a scalar: a pure number. For example, suppose a box contains 3 particles. The number 3 is a scalar. Obviously everyone will agree that the number of the count is 3: countest thou not to 2 unless thou countest also to 3; 5 is right out; and so on. **Ask class:** what are other examples of scalars? Since scalars are invariant, it often gives insight to construct scalars out of the geometrical quantities of interest.

Of course, the particular indices used are dummy indices; we could as well have written $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$. Therefore, $dx^\mu$ (or $dx^\nu$ or $dx^\alpha$ or whatever, it's the same thing) is the $\mu$th component of the separation between the events, in this particular coordinate system.

The metric tensor is symmetric: $g_{\alpha\beta} = g_{\beta\alpha}$.

Going back to the particular metric we wrote earlier: $g_{\alpha\beta} = (-1, 1, 1, 1)$ down the diagonal. This is called the Minkowski metric, and is usually given a special symbol: $\eta_{\alpha\beta}$. The Minkowski metric is of special importance. It describes flat spacetime, which is spacetime without gravity. Any metric that can be put in the Minkowski form by a transformation also describes flat spacetime. Here, by the way, is a place to point out the difference between *spacetime* and a particular metric used to describe it. Spacetime has some particular geometric characteristics (for example, it's flat). A metric is what you get when you pick a coordinate system to use with that spacetime. Seems trivial, but as we'll see in a later lecture, sometimes a change in the coordinates (which therefore does not change the spacetime) has made a big difference in how things are perceived.

So, in flat spacetime one could easily write $ds^2 = -c^2dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)$, and recover the Minkowski metric by the usual spherical to Cartesian transformation. However, there are some metrics for which a global transformation to Minkowski is not possible, e.g., those applicable outside a massive object.

**Important Four-Vectors**

There are several especially useful four-vectors to consider. One, which we already encountered, is the label of an event: $x^\mu = (ct, x, y, z)$. How about the four-velocity? You might be tempted to write this as $U^\mu = dx^\mu/dt$, where $t$ is the coordinate time, but remember that since $t$ is specific to a given coordinate system, this would not be a geometric object. Instead, we write

$$U^\mu = \frac{dx^\mu}{d\tau} \tag{3}$$

where $\tau$ is the *proper time*, defined in terms of the invariant interval as

$$-c^2d\tau^2 = ds^2\ . \tag{4}$$

The negative sign here is because we've used a metric "signature" in which the sign of the $dt$ term is negative. The interpretation of this is the following: suppose that a freely falling (or freely drifting) observer moves with a particle that goes from one event to another. That observer's subjective time (measured, e.g., on a wristwatch) equals the proper time. You can see this by noting that for such an observer, $dx = dy = dz = 0$.

One can then define, analogously, the four-momentum

$$p^\mu \equiv mU^\mu \tag{5}$$

(for particles of non-zero rest mass $m$) and the four-acceleration

$$a^\mu \equiv \frac{dU^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2}\ . \tag{6}$$

Note that the squared four-velocity has a constant magnitude:

$$U_\mu U^\mu = \frac{dx_\mu}{d\tau}\frac{dx^\mu}{d\tau} = \frac{\eta_{\mu\nu}dx^\nu dx^\mu}{d\tau^2} = \frac{ds^2}{d\tau^2} = -c^2\ . \tag{7}$$

The square of the four-momentum is thus $p_\mu p^\mu = -m^2c^2$. Note that neither $U_\mu U^\mu$ nor $p_\mu p^\mu$ have a free index (because of the summing), hence these expressions are scalars and therefore are measured to have the same value by all observers.

What does this mean? The 0th (or time) component of the four-momentum is just $p^0 = E/c$, and the spatial components are the normal linear momentum, so we find

$$p_\mu p^\mu = -m^2c^2 = -E^2/c^2 + \vec{p}^2 \Rightarrow E^2 = p^2c^2 + m^2c^4 \ , \tag{8}$$

which may be a familiar equation. For particles of zero rest mass, the concept of proper time is no longer useful because since $ds^2 = 0$ along the path of a photon or anything else traveling at the speed of light, $d\tau^2 = 0$ as well. For such purposes, though, we can still define the four-velocity if we define a parameter along the trajectory. Then, for example, the four-momentum of a photon of energy E is $p^\mu = (E/c, \vec{p})$, where $\vec{p}$ is the three-momentum of the photon. Note that for a photon, $p_\mu p^\mu = 0$.

**Application: Single-photon absorption by electron?**

Can a free electron absorb a single photon? If so, then the conservation of four-momentum implies

$$p_\gamma^\mu + p_e^\mu = p_e^{\prime\mu} \ . \tag{9}$$

Here the $\gamma$ subscript means the photon, $e$ means the electron, and the prime is for after the absorption. We square both sides and use $p_{e\mu}p_e^\mu = -m_e^2c^2$ to get

$$p_{\gamma\mu}p_e^\mu = 0 \ . \tag{10}$$

Note that because there are no free indices left, this is a scalar and thus valid in all reference frames. We are therefore free to choose a convenient reference frame, specifically the one in which the electron is initially at rest. In this frame, the electron had no three-momentum, so $p_e^\mu = (m_ec, 0, 0, 0)$. The four-momentum of the photon was $p_\gamma^\mu = (E/c, p_x, p_y, p_z)$, so the dot product between them gives the condition

$$-m_eE = 0 \ . \tag{11}$$

This requires a photon of zero energy, meaning that this process is *not* possible. You need something else around to absorb momentum, e.g., a charged nucleus or a magnetic field.

One last comment before we conclude: as is shown in the text (see the book for a derivation), it turns out that there is another Lorentz invariant (i.e., scalar) around. This is the "phase space volume". Consider a collection of particles that inhabits a small region in 3-D space, $d^3r$, and has a small spread in three-momenta, $d^3p$. It turns out that $d^3rd^3p$ is invariant. This has many applications, some of which we will explore when we look at radiation processes in the next class.

## Intuition Builder

Can a photon spontaneously split into two photons, assuming that the system is isolated and in flat spacetime?