Massachusetts Institute of Technology
Department of Physics
8.962 Spring 2024

Lecture 2
Spacetime and geometry part II

2.1 More on the metric and basis vectors

The metric is the mathematical entity which attaches a notion of distance to the manifold: Δs^2 in the relation

$$\Delta s^2 = \Delta\vec{x} \cdot \Delta\vec{x} = \eta_{\alpha\beta}\Delta x^\alpha \Delta x^\beta \tag{2.1}$$

provides a lot of information about the distances between events that different observers measure. If $\Delta\vec{x}$ is spacelike, then we can perform a Lorentz transformation into a frame such that $\Delta\vec{x}$ has no timelike components; $\Delta s = \sqrt{\Delta s^2}$ is then the distance an observer in this frame measures between the two events which define the endpoints of $\Delta\vec{x}$. If $\Delta\vec{x}$ is timelike, we can perform a Lorentz transformation into a frame such that $\Delta\vec{x}$ has no spatial components; $\Delta\tau = \sqrt{-\Delta s^2}$ is the elapsed time an observer in this frame measures between the events which define the endpoints of $\Delta\vec{x}$.

A differential displacement vector is given by $d\vec{x} = dx^\alpha \vec{e}_\alpha$. This relationship defines $\vec{e}_\alpha$ as a *coordinate basis* vector. For rectilinear coordinates, this term doesn't mean very much. However, the distinction between "coordinate bases" and other kinds of bases becomes very important when we begin to examine curvilinear coordinates. Consider for example two ways of writing the spatial interval $d\mathbf{x}$ in spherical polar coordinates:

$$\begin{aligned} d\mathbf{x} &= dr\,\mathbf{e}_{\hat{r}} + r\,d\theta\,\mathbf{e}_{\hat{\theta}} + r\sin\theta d\phi\,\mathbf{e}_{\hat{\phi}} & (2.2)\\ &= dr\,\mathbf{e}_r + d\theta\,\mathbf{e}_\theta + d\phi\,\mathbf{e}_\phi\ . & (2.3) \end{aligned}$$

The form (2.2) is surely familiar to you if have studied electrodynamics; it uses what are called orthonormal polar basis vectors. The form (2.3) is perhaps new, and perhaps confusing. On first sight, you might wonder how this can possibly be dimensionally consistent — shouldn't every term on the right-hand side have the units of length? Indeed, each term on the right-hand side *does* have the units of length: the "missing" factors which make the units right are in the basis vectors. In a coordinate basis, the basis vectors are *not necessarily unit vectors*. We must in fact have $\mathbf{e}_\theta \cdot \mathbf{e}_\theta = r^2$, $\mathbf{e}_\phi \cdot \mathbf{e}_\phi = r^2 \sin^2\theta$ in order for Eq. (2.3) to define a sensible displacement vector. Although Eqs. (2.2) and (2.3) use somewhat different forms, they say exactly the same thing.

We will use coordinate basis vectors exclusively in this class. The reason is that as we move to more general transformations (going beyond Lorentz transformations), it will remain the case that to change the representation we use a matrix whose elements are $\partial x^{\bar\mu}/\partial x^\alpha$. That this works can be seen quite simply: $d\vec{x}$ is a geometric object that is the same in all representations, so

$$d\vec{x} = dx^\alpha \vec{e}_\alpha = dx^{\bar\mu}\vec{e}_{\bar\mu}\ . \tag{2.4}$$

It follows that the matrix elements $\partial x^{\bar\mu}/\partial x^\alpha$ produce the transformation of the components of $d\vec{x}$ — and all other 4-vector components — from those used in frame $\mathcal{O}$ to those used in frame $\bar{\mathcal{O}}$.

The transformation we need to use in the case of orthonormal bases is a little more complicated. Though not significantly so; interested students are strongly encouraged to read about the mathematical structure introduced to describe "noncoordinate" bases in Appendix J of Carroll. But this is enough of a diversion that in a single semester class, it is one of those interesting topics that we must unfortunately not include.

2.2 Other important 4-vectors

For us, a major purpose of constructing geometric objects like vectors is that they describe aspects of the physics that "lives" in the manifold of spacetime. Perhaps the simplest example is the 4-velocity. Suppose observer $\mathcal{O}$ follows a trajectory $x^\alpha(\tau)$, with the coordinates specified in some IRF (not necessarily the rest frame of $\mathcal{O}$; indeed, perhaps $\mathcal{O}$ is accelerating, and has different rest frames at different moments). A body's trajectory through spacetime is known as its *worldline*. We will define the parameter τ more precisely in a moment; for now, just imagine that it increases in some fashion as $\mathcal{O}$ moves along this worldline. Then, we define the 4-velocity as

$$\vec{u} = \frac{d\vec{x}}{d\tau} \ . \tag{2.5}$$

The vector interval $d\vec{x}$ is bracketed by $\mathcal{O}$'s location in spacetime at parameter τ and at parameter $\tau + d\tau$. For this parameter we use the *proper time* of $\mathcal{O}$ — the time that $\mathcal{O}$ measures on their own clock[1]. Different inertial observers will report different intervals of time for $\mathcal{O}$ to move from event $\mathcal{P}$ to event $\mathcal{Q}$, but all will agree that $\mathcal{O}$'s own clock records an interval $\Delta\tau$ for this interval. The proper time is thus an excellent way of defining a time interval that everyone can agree on. Notice that the 4-velocity is the tangent vector in spacetime to the worldline that $\mathcal{O}$ traces out.

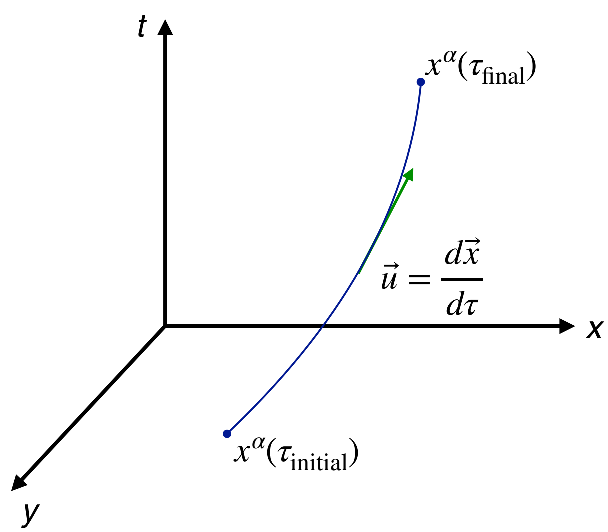

Figure 1: Illustration of the 4-velocity $\vec{u} = d\vec{x}/d\tau$ along a trajectory in spacetime.

In special relativity, we learn that the components of 4-velocity as measured by an observer who sees $\mathcal{O}$ move past with 3-velocity $\mathbf{v}$ is given by

$$\vec{u} \doteq (\gamma, \gamma\mathbf{v}) \ , \tag{2.6}$$

with $\gamma = 1/\sqrt{1 - |\mathbf{v}|^2}$. Notice that if $\mathcal{O}$ happens to be at rest with respect to that observer, then $\vec{u} \doteq (1, \mathbf{0})$. Notice also that

$$\vec{u} \cdot \vec{u} = -\gamma^2 + |\mathbf{v}|^2\gamma^2 = -1 \ . \tag{2.7}$$

The 4-velocity is *always* a timelike vector, with unit magnitude.

An important quantity we can construct from the 4-velocity is the 4-momentum:

$$\vec{p} = m\vec{u} \tag{2.8}$$

$$\doteq (E, \mathbf{p}) \ . \tag{2.9}$$

[1]Sometimes people ask what is "proper" about proper time. This term is taken from French or Latin, and really means "$\mathcal{O}$'s own time." I like the German term eigenzeit, but it doesn't seem to have caught on in English.

The form (2.8) is a definition of $\vec{p}$ that describes any object with mass[2] m and 4-velocity $\vec{u}$; the form (2.9) describes the timelike and spacelike components of $\vec{p}$ as measured in some particular IRF. The timelike component is the *energy* measured in that IRF; the spacelike components are the *3-momentum* (often simply "momentum") which that IRF measures. A few remarks about this very important object:

- All observers agree on $\vec{p} \cdot \vec{p}$; using the fact that $\vec{p} = m\vec{u}$ and $\vec{u} \cdot \vec{u} = -1$, we see that

$$\vec{p} \cdot \vec{p} = -m^2 \ . \tag{2.10}$$

  Because this is true in all IRFs, we can use it to understand the relation between an object's energy and 3-momentum as measured in some IRF:

$$\begin{aligned} &\vec{p} \cdot \vec{p} = -m^2 = -(p^0)^2 + |\mathbf{p}|^2 \\ \longrightarrow \qquad &E^2 - |\mathbf{p}|^2 = m^2 \qquad \text{or} \qquad E^2 = |\mathbf{p}|^2 + m^2 \ . \end{aligned} \tag{2.11}$$

  This final form should be familiar from studies of special relativity (albeit perhaps with factors of c hidden by our choice of units).

- A "body" can have a mass of zero; photons are well-known examples of this limit. In this case, the body still has energy and momentum, and thus has a well-defined 4-momentum. From the above relation, we see that in this case, any observer measures this zero-mass body to have energy equal to the magnitude of its 3-momentum: $E = |\mathbf{p}|$. It's worth noting one cannot assign a 4-velocity to such a body. If a body has a 4-velocity, then there exists a rest frame for that body at each moment on its worldline. You are hopefully familiar with the fact that zero mass objects cannot be at rest, consistent with the idea that they do not have a well-defined 4-velocity.

- 4-momentum is conserved in any interaction. If we have N particles that interact, then

$$\vec{p}_{\rm tot} \equiv \sum_{i=1}^{N} \vec{p}_i \tag{2.12}$$

  is conserved in the interaction. This is true even if N changes during the interaction (e.g., a particle decays, or two bodies collide inelastically to form a new body).

- Many calculations based on conservation of 4-momentum can be expedited by working in a special IRF in which

$$\vec{p}_{\rm tot} \doteq (E, \mathbf{0}) \ . \tag{2.13}$$

  This special IRF is known as the *center of momentum* frame. In this frame, the interacting system is at rest as a whole (although members of the system could very well be zipping around quite speedily).

2.3 Exploiting the invariant product

The invariance of the scalar product is often a way to rapidly deduce a result that is of great utility. The first example we'll look at concerns the energy of a body as measured by some particular observer. Let $\vec{p}_A$ be the 4-momentum of body A, and let $\vec{u}_{\mathcal{O}}$ be the 4-velocity of observer $\mathcal{O}$. What does $\mathcal{O}$ measure for the energy of body A?

[2]Notice that the mass is a Lorentz scalar — all observers agree on its value. It is often called the "rest mass," in recognition of the fact that it describes the object's energy when at rest (i.e., $E = mc^2$, restoring the c^2 hidden by our choice of units).

One way to do this is to go into the rest frame of $\mathcal{O}$. In that frame, $\vec{u}_{\mathcal{O}} \doteq (1, \mathbf{0})$. Also in that frame, $\vec{p} \doteq (E_{\mathcal{O}}, \mathbf{p}_{\mathcal{O}})$, where $E_{\mathcal{O}}$ and $\mathbf{p}_{\mathcal{O}}$ are the energy and 3-momentum of body A as measured by $\mathcal{O}$. This suggests that we should figure out the Lorentz transformation that puts us into the rest frame of $\mathcal{O}$, apply this transformation to $\vec{p}_A$, and from the resulting components read out $E_{\mathcal{O}}$.

That's not wrong, but the invariant product enables a tremendously effective shortcut. Because $\vec{u}_{\mathcal{O}} \doteq (1, \mathbf{0})$ in the rest frame of $\mathcal{O}$, it is clear that $E_{\mathcal{O}} = -\vec{p}_A \cdot \vec{u}_{\mathcal{O}}$. But this result involves an invariant *that holds in all IRFs.* Thus, no matter what frame one uses to construct $\vec{p}_A$ and $\vec{u}_{\mathcal{O}}$, the energy that $\mathcal{O}$ measures for the energy of body A is given by

$$\boxed{E_{\mathcal{O}} = -\vec{p}_A \cdot \vec{u}_{\mathcal{O}}} \tag{2.14}$$

This very powerful result gives us a totally frame-independent way to express what the energy of a body is according to some IRF. We are going to use this result several times at points in this class.

Another result concerns the 4-acceleration of a body. We described the fact that the 4-velocity $\vec{u} = d\vec{x}/d\tau$ is essentially the tangent vector to the worldline of a body moving through spacetime. If the body is not inertial, then it is accelerating, and we can define an acceleration 4-vector

$$\vec{a} = \frac{d\vec{u}}{d\tau} \,. \tag{2.15}$$

The 4-acceleration has a very important constrained relationship with respect to the 4-velocity. To see this, we use the fact that $\vec{u} \cdot \vec{u} = -1$. Differentiating this relationship, we have

$$\vec{a} \cdot \vec{u} = 0 \,. \tag{2.16}$$

The 4-acceleration is always orthogonal to the 4-velocity in spacetime. This is in stark contrast to physics based on 3-velocity and 3-acceleration, where these quantities need not have any particular relationship to one another at any moment. Equation (2.16) is an important sanity check when doing relativistic kinematics; a friend of mine was able to instantly reject a paper he was refereeing for a journal by noting that the author's proposed kinematics did not respect this relationship.

More importantly for us, Eq. (2.16) helps us understand the meaning of 4-acceleration. Let us go into the rest frame of the observer whose 4-velocity is $\vec{u}$. (If the body is accelerating, that frame will not be its rest frame for long. As such, this is often called its "instantaneous rest frame.") As mentioned many times, in this frame $\vec{u} \doteq (1, \mathbf{0})$. In order for $\vec{a} \cdot \vec{u} = 0$, we must have $a \doteq (0, \mathbf{a})$ in this frame. Here, $\mathbf{a}$ is the ordinary 3-acceleration in this frame. Further, we see from this that $\vec{a} \cdot \vec{a} = |\mathbf{a}|^2$ — a frame-independent statement about the 4-acceleration that nonetheless involves the 3-acceleration in the body's instantaneous rest frame. Putting all this together, we see that a body experiencing a 4-acceleration $\vec{a}$ feels a 3-acceleration of magnitude $|\mathbf{a}| = \sqrt{\vec{a} \cdot \vec{a}}$ in their own rest frame.

2.4 Tensors more generally

So far, we have introduced one object that we have labeled a "tensor," the metric $\eta_{\alpha\beta} = \vec{e}_\alpha \cdot \vec{e}_\beta$. More generally, tensors are geometric objects which act like functions which enact linear operations on vectors. Let us introduce a precise definition:

A tensor of type $\binom{0}{N}$ is a geometric object which enacts a functional mapping of N vectors into a Lorentz-invariant scalar, and which is linear in each of its N arguments.

(You might wonder: why the zero in $\binom{0}{N}$? Can't there be something else there? Hold that thought a bit: we're walking up to a fully general definition, but need to lay some groundwork first.)

Using this definition, we see that the invariant scalar product is actually a $\binom{0}{2}$ tensor:

$$\begin{aligned}
\vec{A}\cdot\vec{B} &= a && \text{2 vectors map to a Lorentz scalar.} && (2.17)\\
(\alpha\vec{A})\cdot\vec{B} &= \alpha a && \text{Linear in the 1st argument (part 1 of linearity).} && (2.18)\\
(\vec{A}+\vec{C})\cdot\vec{B} &= \vec{A}\cdot\vec{B}+\vec{C}\cdot\vec{B} && \text{Linear in the 1st argument (part 2 of linearity).} && (2.19)\\
\vec{A}\cdot(\beta\vec{B}) &= \beta a && \text{Linear in the 2nd argument (part 1 of linearity).} && (2.20)\\
\vec{A}\cdot(\vec{B}+\vec{C}) &= \vec{A}\cdot\vec{B}+\vec{A}\cdot\vec{C} && \text{Linear in the 2nd argument (part 2 of linearity).} && (2.21)
\end{aligned}$$

The "scalar product" tensor is in fact the metric. In very abstract form, we define the tensor $\boldsymbol{\eta}$ as a "2 slot" function that takes vectors as inputs, and yields the scalar product as output:

$$\boldsymbol{\eta}(\vec{A},\vec{B}) = \vec{A}\cdot\vec{B} = \eta_{\alpha\beta}A^\alpha B^\beta\ . \tag{2.22}$$

The usual linearity rules apply: $\boldsymbol{\eta}(\alpha\vec{A}+\vec{B},\vec{C}) = \alpha\boldsymbol{\eta}(\vec{A},\vec{C})+\boldsymbol{\eta}(\vec{B},\vec{C})$, etc.

Because tensors take vectors — frame-independent, geometric objects — as input, and yield Lorentz invariant scalars as output, they are themselves a kind of geometric object. We have intuition about vectors as "directed line segments" pointing in space or spacetime. It's harder to build such intuition about tensors; some can be thought of as an object that, in a meaningful way, points in multiple directions at once, but other tensors are harder to characterize. Because the tensors are themselves frame-independent geometric objects, they will have different components — i.e., different representations — in different reference frames. The scalar we get out when we fill up all their "slots" with vectors will, however, be invariant.

To get the components out, we simply put basis vectors in the slots. Our lone example so far is the metric:

$$\boldsymbol{\eta}(\vec{e}_\alpha,\vec{e}_\beta) = \vec{e}_\alpha\cdot\vec{e}_\beta = \eta_{\alpha\beta}\ . \tag{2.23}$$

Using this plus the linearity rules, we can see how to get tensor components in another frame of reference. Let's work this out: using $\vec{e}_{\bar\mu} = \Lambda^\alpha{}_{\bar\mu}\vec{e}_\alpha$ plus $\eta_{\bar\mu\bar\nu} = \boldsymbol{\eta}(\vec{e}_{\bar\mu},\vec{e}_{\bar\nu})$, we have

$$\begin{aligned}
\eta_{\bar\mu\bar\nu} &= \boldsymbol{\eta}(\Lambda^\alpha{}_{\bar\mu}\vec{e}_\alpha,\Lambda^\beta{}_{\bar\nu}\vec{e}_\beta)\\
&= \Lambda^\alpha{}_{\bar\mu}\Lambda^\beta{}_{\bar\nu}\boldsymbol{\eta}(\vec{e}_\alpha,\vec{e}_\beta)\\
&= \Lambda^\alpha{}_{\bar\mu}\Lambda^\beta{}_{\bar\nu}\eta_{\alpha\beta}\ .
\end{aligned}\tag{2.24}$$

The final rule we find aligns with a slogan introduced previously: "Just line up the indices." It should be noted that this is a fairly silly example, because it turns out that the components of the metric are just $\mathrm{diag}(-1,1,1,1)$ in all IRFs[3], at least as long as we use rectilinear coordinates. Nonetheless, the principle holds even in this silly case, and the need to enact this transformation law is far less trivial on all other tensors we will encounter in this class.

2.5 $\binom{0}{1}$ tensors: 1-forms

A 1-form is a mapping from a single vector to Lorentz invariant scalars:

$$\tilde{p}(\vec{A}) = \alpha\ . \tag{2.25}$$

Here, $\tilde{p}$ is a 1-form, and α is a Lorentz scalar. By inheriting tensor properties, it is easy to see that 1-forms satisfy all the rules needed to define a vector space: if $\tilde{s} = \tilde{p}+\tilde{q}$ and $\tilde{r} = \beta\tilde{p}$ then $\tilde{s}$ and $\tilde{r}$ are also 1-forms; $\tilde{s}(\vec{A}) = \tilde{p}(\vec{A})+\tilde{q}(\vec{A})$, $\tilde{r}(\vec{A}) = \beta\tilde{p}(\vec{A})$. We get their components with the basis vectors:

$$\tilde{p}(\vec{e}_\alpha) = p_\alpha\ . \tag{2.26}$$

[3] $\mathrm{diag}(a,b,c,d)$ means the matrix whose entries are a, b, c, d down the diagonal, and zero everywhere else.

This gives us a very nice way of understanding the Lorentz scalar that we get when we apply a 1-form to a vector:

$$\begin{aligned}\tilde{p}(\vec{A}) &= \tilde{p}(A^\alpha \vec{e}_\alpha) \\ &= A^\alpha\,\tilde{p}(\vec{e}_\alpha) \\ &= A^\alpha p_\alpha\ .\end{aligned} \tag{2.27}$$

The resulting Lorentz scalar is just what we get when we *contract* the (upstairs) vector indices on the components of $\vec{A}$ with the (downstairs) 1-form indices on the components of $\tilde{p}$. This operation is called *contraction*. Using a similar operation, it is straightforward to show that 1-form components transform by the "line up the indices" rule:

$$\begin{aligned}p_{\bar{\mu}} &= \tilde{p}(\vec{e}_{\bar{\mu}}) \\ &= \tilde{p}(\Lambda^\alpha{}_{\bar{\mu}}\vec{e}_\alpha) \\ &= \Lambda^\alpha{}_{\bar{\mu}}\tilde{p}(\vec{e}_\alpha) \\ &= \Lambda^\alpha{}_{\bar{\mu}}p_\alpha\ .\end{aligned} \tag{2.28}$$

In our discussion of basis vectors, we introduced a set of special vectors which were oriented along the 4 directions of spacetime. We similarly will find it useful to define a set of special 1-forms oriented (in a way to be described later) to pick out special directions. Let us define a set of 1-forms $\tilde{\omega}^\alpha$ such that any 1-form can be written using them and components as follows:

$$\tilde{p} = p_\alpha \tilde{\omega}^\alpha\ . \tag{2.29}$$

To actually define what these quantities are, we use the fact that $\tilde{p}(\vec{A}) = p_\alpha A^\alpha$ in the following way:

$$\begin{aligned}p_\alpha A^\alpha &= \tilde{p}(\vec{A}) \\ &= p_\alpha \tilde{\omega}^\alpha(A^\beta \vec{e}_\beta) \\ &= p_\alpha A^\beta \tilde{\omega}^\alpha(\vec{e}_\beta)\ .\end{aligned} \tag{2.30}$$

This last equation leads us to require that

$$\boxed{\tilde{\omega}^\alpha(\vec{e}_\beta) = \delta^\alpha{}_\beta} \tag{2.31}$$

An example representation that satisfies this is the very simple set

$$\tilde{\omega}^0 \doteq (1,0,0,0)\ , \tag{2.32}$$
$$\tilde{\omega}^1 \doteq (0,1,0,0)\ , \tag{2.33}$$
$$\tilde{\omega}^2 \doteq (0,0,1,0)\ , \tag{2.34}$$
$$\tilde{\omega}^3 \doteq (0,0,0,1)\ . \tag{2.35}$$

These look quite a lot like the basis vectors. They are not; the difference is akin to the difference between row vectors and column vectors in matrix manipulations. In this vein, it's worth remarking that 1-forms are sometimes called "dual vectors," especially in older texts. We are going to see very soon that 1-forms can be constructed from vectors, and vice versa, in a very simple way, making it clear that they are very closely related geometric entities.

We conclude this discussion with an example of a 1-form. Imagine that some observer follows a trajectory through spacetime with 4-velocity $\vec{u} = d\vec{x}/d\tau$ along their worldline, and that all of spacetime is filled with some scalar field which varies from event to event $\Phi(t,x,y,z)$. What is the rate of change of Φ along the worldline?

If this were simple Euclidean 3-space, you would surely assert that the answer is

$$\left.\frac{d\Phi}{dt}\right|_{\rm along\ traj} = \frac{\partial\Phi}{\partial t} + \frac{dx}{dt}\frac{\partial\Phi}{\partial x} + \frac{dy}{dt}\frac{\partial\Phi}{\partial y} + \frac{dz}{dt}\frac{\partial\Phi}{\partial z}$$

$$= \frac{\partial\Phi}{\partial t} + \mathbf{v}\cdot\boldsymbol{\nabla}\Phi\ . \tag{2.36}$$

Generalizing this to spacetime is easy: we change per unit time to per unit proper time, and "upgrade" certain other concepts in a similar manner:

$$\left.\frac{d\Phi}{d\tau}\right|_{\rm along\ traj} = \frac{dt}{d\tau}\frac{\partial\Phi}{\partial t} + \frac{dx}{d\tau}\frac{\partial\Phi}{\partial x} + \frac{dy}{d\tau}\frac{\partial\Phi}{\partial y} + \frac{dz}{d\tau}\frac{\partial\Phi}{\partial z}$$

$$= u^\alpha\frac{\partial\Phi}{\partial x^\alpha}$$

$$= u^\alpha\partial_\alpha\Phi\ . \tag{2.37}$$

On the last line, we have introduced the notation $\partial/\partial x^\alpha = \partial_\alpha$. We see that the components of the spacetime gradient are the components of a 1-form:

$$\tilde{d}\Phi \equiv \partial_\alpha\Phi\,\tilde{\omega}^\alpha\ . \tag{2.38}$$

We will use this in the next lecture to get some intuition about the geometry associated with 1-forms.