Massachusetts Institute of Technology
Department of Physics
8.962 Spring 2024

Lecture 5
More on the stress-energy tensor; curvilinear coordinates

5.1 Stress-energy symmetry

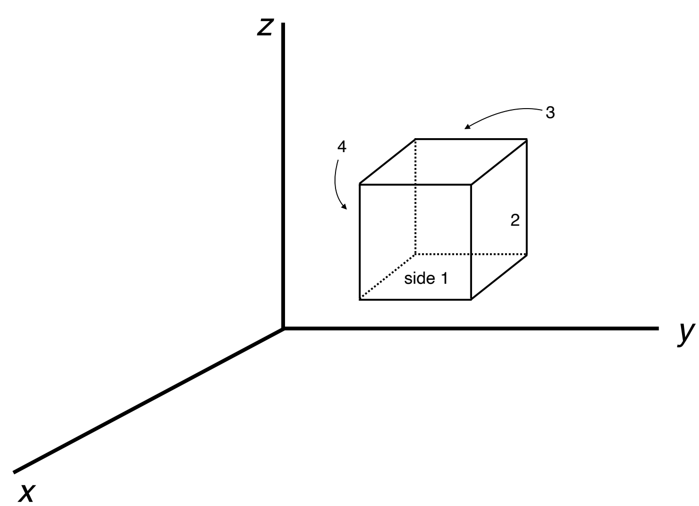
In the previous lecture, we introduced the stress-energy tensor, a quantity $T^{\alpha\beta}$ which describes the flux of p^α in the x^β direction. Let us look at it component by component for the case of energy and momentum of dust moving with 4-velocity $\vec{u} \doteq (\gamma, \gamma\mathbf{v})$ in some specified IRF:

$$T^{00} = \text{Energy density} = \gamma^2\rho_0 \ , \tag{5.1}$$
$$T^{0i} = \text{Energy flux} = \gamma^2\rho_0 v^i \ , \tag{5.2}$$
$$T^{i0} = \text{Momentum density} = \gamma^2\rho_0 v^i \ , \tag{5.3}$$
$$T^{ij} = \text{Momentum flux or stress} = \gamma^2\rho_0 v^i v^j \ . \tag{5.4}$$

(Here, ρ_0 is the energy density of the dust in its rest frame, i.e., the dust's rest-energy density.) Notice that $T^{\alpha\beta} = T^{\beta\alpha}$: the stress-energy tensor is symmetric under exchange of indices.

This symmetry is manifestly clear for the case of dust, but in fact holds much more generally. Picking a particular IRF, the symmetry $T^{0i} = T^{i0}$ follows from the fact that energy and momentum are the same thing reviewed in different reference frames. In other words, if energy is flowing, we must have momentum as well[1]. Energy flux is thus equivalent to momentum density. The symmetry $T^{ij} = T^{ji}$ is less obvious. However, one can quickly show that if this symmetry were violated, then a physically absurd situation could develop.

Consider a small cube with sides l embedded in some $T^{\alpha\beta}$:



Let's compute the 3-forces that act due to the stress-energy tensor upon faces 1 through 4 as illustrated here, and then use those forces to compute the torque about the z axis.

[1]It is worth noting that in "normal" units, we may need to insert factors of c to connect T^{0i} and T^{i0} with energy flux and momentum density, respectively.

The total 3-force that acts on face 1 is given by

$$F^i_{\text{face 1}} = \int_{\text{face 1}} T^{ix} dy\, dz \simeq T^{ix}\, l^2 \ . \tag{5.5}$$

The approximate equality follows because we assume l is very small (smaller than the scale over which the stress-energy tensor varies in this region). Repeating this calculation for the other three faces, we find

$$F^i_{\text{face 2}} \simeq T^{iy}\, l^2 \ , \qquad F^i_{\text{face 3}} \simeq -T^{ix}\, l^2 \ , \qquad F^i_{\text{face 4}} \simeq -T^{iy}\, l^2 \ . \tag{5.6}$$

(Our assumption of l small means that we assume the components T^{ij} are approximately constant over the cube.) Note that no net force acts on the cube, so our stress-energy tensor does not lead to weird accelerations.

Let us next examine the torques that these forces exert about an axis through the center of the cube, parallel to the z axis:

$$\tau^z_1 = (\text{torque due to } \mathbf{F}_{\text{face 1}})^z = -xF^y_{\text{face 1}} = -\frac{1}{2}T^{yx}\, l^3 \ . \tag{5.7}$$

The other 3-force components generate the following contributions to the torque:

$$\tau^z_3 = xF^y_{\text{face 3}} = -\frac{1}{2}T^{yx}\, l^3 \ , \tag{5.8}$$

$$\tau^z_2 = yF^x_{\text{face 2}} = \frac{1}{2}T^{xy}\, l^3 \ , \tag{5.9}$$

$$\tau^z_4 = -yF^x_{\text{face 4}} = \frac{1}{2}T^{xy}\, l^3 \ . \tag{5.10}$$

The total torque about the axis parallel to z exerted on this cube is thus

$$\tau^z = l^3\left(T^{xy} - T^{yx}\right) \ . \tag{5.11}$$

By itself, there's nothing very interesting here. We would like to understand how this behaves as $l \to 0$. Equation (5.11) seems to indicate that the torque strictly vanishes as $l \to 0$, independent of the properties of T^{xy} and T^{yx}. This is not wrong, but we should dig in a little deeper before concluding that everything is OK. Imagine that the cube is made of some material of density ρ, and thus has a momentum of inertia

$$I = \alpha(\rho l^3) l^2 \ . \tag{5.12}$$

(The quantity α is a number less than or of order unity; if ρ is constant, $\alpha = 1/6$.) The cube feels an angular acceleration, tending to make it spin about this axis, whose value is

$$\frac{d^2\theta}{dt^2} = \frac{\tau^z}{I} \propto \frac{\left(T^{xy} - T^{yx}\right)}{l^2} \ . \tag{5.13}$$

Unless this is strictly zero — i.e., unless $T^{xy} = T^{yx}$ — then the source of our stress-energy would tend to cause small regions to unstably rotate. Considering torques about the other axes of the cube leads us to conclude we need $T^{ij} = T^{ji}$ in order to avoid “weird spin ups” in small regions of our stress-energy tensor.

We emphasize that this discussion is *not* a rigorous proof that $T^{ij} = T^{ji}$; it is just a physical motivation that if this symmetry were violated, then weird and unphysical behavior would result. Later in the course, we will see a way to extract the stress-energy tensor from a Lagrangian density, and will see that the procedure for doing this guarantees that $T^{\alpha\beta} = T^{\beta\alpha}$.

5.2 Examples of stress-energy tensors

So far, we have discussed two examples of stress-energy tensors: that of dust, $T^{\alpha\beta} = \rho_0 u^\alpha u^\beta$; and that of the perfect fluid, $T^{\alpha\beta} = (\rho + P)u^\alpha u^\beta + P\eta^{\alpha\beta}$. Arguably, this is actually *one* example, since dust can be considered a perfect fluid with $P = 0$. This example is particularly important in gravity studies, but it is worth knowing of other examples:

5.2.1 A point particle moving through spacetime

A point body of mass m moving on some worldline $z^\mu(\tau)$ has a stress-energy tensor given by

$$T^{\mu\nu} = m \int d\tau\, u^\mu u^\nu\, \delta^{(4)}\left[\vec{x} - \vec{z}(\tau)\right] . \tag{5.14}$$

We have introduced here the 4-dimensional Dirac delta function; it can be decomposed into a product of 4 1-dimensional deltas:

$$\delta^{(4)}\left[\vec{x} - \vec{z}(\tau)\right] = \delta\left[t - z^0(\tau)\right]\delta\left[x - z^1(\tau)\right]\delta\left[y - z^2(\tau)\right]\delta\left[z - z^3(\tau)\right] . \tag{5.15}$$

Using the rule

$$\int f(x)\delta[g(x)]dx = \frac{f(x_0)}{|dg/dx|_{x=x_0}} , \tag{5.16}$$

where x_0 is the[2] zero of $g(x)$ in the domain of the integral, we can integrate up to find

$$T^{\mu\nu} = m\frac{u^\mu u^\nu}{u^t}\delta^{(3)}\left[\mathbf{x} - \mathbf{z}(\tau)\right] . \tag{5.17}$$

This form is used quite a bit in modern research as a tool for modeling the stress-energy of "small" bodies (i.e., bodies sufficiently pointlike that their internal structure can be ignored).

5.2.2 An electromagnetic field

If a region is filled with an electromagnetic field with field tensor $F^{\mu\nu}$, then

$$T^{\mu\nu} = \frac{1}{4\pi}\left[F^{\mu\lambda}F^\nu{}_\lambda - \frac{1}{4}\eta^{\mu\nu}F^{\lambda\sigma}F_{\lambda\sigma}\right] . \tag{5.18}$$

Going into a particular frame and using the association of components $F^{\mu\nu}$ with electric and magnetic field components, it is not hard to see that

$$T^{00} = \frac{\mathbf{E}\cdot\mathbf{E} + \mathbf{B}\cdot\mathbf{B}}{8\pi} , \tag{5.19}$$

$$T^{0i} = \frac{(\mathbf{E}\times\mathbf{B})^i}{4\pi} , \tag{5.20}$$

$$= T^{i0} , \tag{5.21}$$

$$T^{ij} = \frac{1}{8\pi}\left[(\mathbf{E}\cdot\mathbf{E} + \mathbf{B}\cdot\mathbf{B})\,\delta^{ij} - 2\left(E^iE^j + B^iB^j\right)\right] . \tag{5.22}$$

The terms for T^{00} and T^{0i} are hopefully familiar: they are the usual electromagnetic energy density (perhaps in funny units) and the Poynting vector, respectively. The result for T^{ij} is perhaps a little

[2] If there is more than one zero, then one must sum over all the zeros. This does not pertain to our analysis.

less familiar; it expresses a both an attractive tension associated with the field as well as a pressure that the field generates. A simple example can be found by putting $\mathbf{E} = E^x \mathbf{e}_x$, for which

$$T^{\mu\nu} \doteq \frac{(E^x)^2}{8\pi} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{5.23}$$

The term T^{xx} indicates that this electromagnetic field generates an attractive stress along the x direction (think of the plates of a capacitor being pulled toward one another), but also a pressure in the y and z directions.

5.3 Flat spacetime in curvilinear coordinates

Everything we have done so far has been done in *inertial* coordinates — that is, coordinates for which an inertial observer's worldline can be expressed as a line. These coordinates are spatially Cartesian, and really easy to work with. However, they are limiting. To prepare us for the calculations we will need to do when we start thinking about spacetimes that are curved, let us start developing tools to do flat spacetime physics in curvilinear coordinate systems. Let's begin with the simplest example beyond Cartesian/inertial coordinates: plane polar coordinates (t, r, ϕ, z), related to our inertial coordinate system by $x = r\cos\phi$, $y = r\sin\phi$.

We continue to work in a coordinate basis, so it remains the case that

$$d\vec{x} = dx^\alpha \vec{e}_\alpha = dt\,\vec{e}_t + dr\,\vec{e}_r + d\phi\,\vec{e}_\phi + dz\,\vec{e}_z\ . \tag{5.24}$$

Since $d\phi$ has dimensions of angle, $\vec{e}_\phi$ must have dimensions of length. This is not a "normal" basis: $\vec{e}_\phi \cdot \vec{e}_\phi \neq 1$. We transform between these bases using the transformation matrix $L^\alpha{}_{\bar\mu} \equiv \partial x^\alpha/\partial x^{\bar\mu}$ (reserving from now on the symbol Λ for transformation matrices that are just Lorentz transformations). Let us write x^α for the inertial coordinates, and $x^{\bar\mu}$ for the curvilinear coordinates. The components of this transformation matrix are then given by

$$L^\alpha{}_{\bar\mu} = \frac{\partial x^\alpha}{\partial x^{\bar\mu}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -r\sin\phi & r\cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{5.25}$$

The inverse transformation involves the matrix whose components are

$$L^{\bar\mu}{}_\alpha = \frac{\partial x^{\bar\mu}}{\partial x^\alpha} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi/r & 0 \\ 0 & \sin\phi & \cos\phi/r & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{5.26}$$

You should be able to show straightforwardly that these matrices are inverses of each other.

Moving forward, we will define objects as tensors if their components transform between different coordinate representations using the L matrix which expresses the Jacobian between those coordinates:

$$V^{\bar\mu} = L^{\bar\mu}{}_\alpha V^\alpha = \frac{\partial x^{\bar\mu}}{\partial x^\alpha} V^\alpha\ , \tag{5.27}$$

$$\vec{e}_{\bar\mu} = L^\alpha{}_{\bar\mu}\,\vec{e}_\alpha = \frac{\partial x^\alpha}{\partial x^{\bar\mu}}\,\vec{e}_\alpha\ , \tag{5.28}$$

$$T^{\bar\mu\bar\nu}{}_{\bar\sigma} = L^{\bar\mu}{}_\alpha L^{\bar\nu}{}_\beta L^\gamma{}_{\bar\sigma}\, T^{\alpha\beta}{}_\gamma = \left(\frac{\partial x^{\bar\mu}}{\partial x^\alpha}\right)\left(\frac{\partial x^{\bar\nu}}{\partial x^\beta}\right)\left(\frac{\partial x^\gamma}{\partial x^{\bar\sigma}}\right) T^{\alpha\beta}{}_\gamma\ . \tag{5.29}$$

These forms reduce to our previous rule when the two coordinates are inertial representations of special relativity. However, there is now in general a significant difference: $L^{\bar{\mu}}{}_{\alpha}$ is now a function, and can vary from point to point. Our previous transformation only involved matrices whose components $\Lambda^{\bar{\mu}}{}_{\alpha}$ are constants.

Let's briefly examine several important geometric entities in the new coordinate representation. Two of the basis vectors are different:

$$\vec{e}_r = L^{\alpha}{}_r\,\vec{e}_\alpha = \cos\phi\,\vec{e}_x + \sin\phi\,\vec{e}_y\ , \tag{5.30}$$

$$\vec{e}_\phi = L^{\alpha}{}_\phi\,\vec{e}_\alpha = -r\sin\phi\,\vec{e}_x + r\cos\phi\,\vec{e}_y\ , \tag{5.31}$$

Notice that $\vec{e}_\phi$ grows with distance. You can think of this because the basis vector subtends a particular amount of fixed angle.

We find the metric in the usual way, though we now introduce a new symbol:

$$g_{\alpha\beta} = \vec{e}_\alpha \cdot \vec{e}_\beta = \mathrm{diag}(-1, 1, r^2, 1)\ . \tag{5.32}$$

From now on, we reserve the symbol $\eta_{\alpha\beta}$ strictly for diag$(-1, 1, 1, 1)$, the inertial coordinate representation of the metric of special relativity. Notice that the components of the inverse metric are no longer the same as the components of the metric:

$$g^{\alpha\beta} = \mathrm{diag}(-1, 1, 1/r^2, 1)\ . \tag{5.33}$$

The line element we find with this metric is

$$ds^2 = g_{\alpha\beta}dx^\alpha\,dx^\beta = -dt^2 + dr^2 + r^2\,d\phi^2 + dz^2\ . \tag{5.34}$$

As with the basis vectors, two of the basis 1-forms are different:

$$\tilde{d}r = L^{r}{}_{\alpha}\,\tilde{d}x^\alpha = \cos\phi\,\tilde{d}x + \sin\phi\,\tilde{d}y\ , \tag{5.35}$$

$$\tilde{d}\phi = L^{\phi}{}_{\alpha}\,\tilde{d}x^\alpha = -\frac{\sin\phi}{r}\tilde{d}x + \frac{\cos\phi}{r}\tilde{d}y\ . \tag{5.36}$$

5.4 Christoffel symbols

A major reason why this detail matters is in calculating derivatives of tensorial objects: our basis objects are no longer constant, and we need to include how they vary when we take derivatives. For the case of plane polar coordinates,

$$\begin{aligned} \frac{\partial\vec{e}_r}{\partial r} &= 0\ , & \frac{\partial\vec{e}_r}{\partial\phi} &= \frac{\vec{e}_\phi}{r}\ , \\ \frac{\partial\vec{e}_\phi}{\partial r} &= \frac{\vec{e}_\phi}{r}\ , & \frac{\partial\vec{e}_\phi}{\partial\phi} &= -r\vec{e}_r\ . \end{aligned} \tag{5.37}$$

Let's think about this in the context of the derivative of a simple tensor object, a vector $\vec{V} = V^\alpha\vec{e}_\alpha$. We write its gradient in fully tensorial form using the basis 1-forms as follows:

$$\boldsymbol{\nabla}\vec{V} = \frac{\partial}{\partial x^\beta}\left[(V^\alpha\vec{e}_\alpha)\right]\tilde{d}x^\beta \equiv \partial_\beta\left[(V^\alpha\vec{e}_\alpha)\right]\tilde{d}x^\beta\ . \tag{5.38}$$

The derivative we need to evaluate here is

$$\partial_\beta\left(V^\alpha\vec{e}_\alpha\right) = \left(\partial_\beta V^\alpha\right)\vec{e}_\alpha + V^\alpha\partial_\beta\vec{e}_\alpha\ . \tag{5.39}$$

The last term involves the derivatives of the basis vectors themselves. They can be organized in a very helpful way as follows:

$$\partial_\beta\vec{e}_\alpha = \Gamma^\mu_{\beta\alpha}\vec{e}_\mu\ . \tag{5.40}$$

The quantity we introduced, $\Gamma^\mu_{\beta\alpha}$, is known as the *Christoffel symbol*.

Caution: Slightly different conventions exist for how the Christoffel symbol is defined. The way we have defined it (the x^β derivative of $\vec{e}_\alpha$) is how it is done in, for example, the textbooks by Carroll and Schutz. Some other sources (for example, Schutz and MTW) define this quantity to be $\Gamma^\mu_{\alpha\beta}$. This difference turns out to be irrelevant as long as we work in a coordinate basis, since the Christoffel symbol is symmetric on the lower two indices in this case. (We will soon prove this.)

For the plane polar coordinates we have introduced, there are three non-zero Christoffel symbols (two of which are equal due to this symmetry):

$$\Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = 1/r \;, \qquad \Gamma^r_{\phi\phi} = -r \;. \tag{5.41}$$

All the other Christoffel symbols are zero in this case.

Let us return to the derivative of the vector:

$$\begin{aligned}
\partial_\beta \vec{V} &= [(\partial_\beta V^\alpha)\vec{e}_\alpha + V^\alpha(\partial_\beta \vec{e}_\alpha)] \\
&= \left[(\partial_\beta V^\alpha)\vec{e}_\alpha + V^\alpha \Gamma^\mu_{\beta\alpha}\vec{e}_\mu\right] \\
&= \left[(\partial_\beta V^\alpha)\vec{e}_\alpha + V^\mu \Gamma^\alpha_{\beta\mu}\vec{e}_\alpha\right] \\
&= \left(\partial_\beta V^\alpha + V^\mu \Gamma^\alpha_{\beta\mu}\right)\vec{e}_\alpha \\
&\equiv \left(\nabla_\beta V^\alpha\right)\vec{e}_\alpha \;.
\end{aligned} \tag{5.42}$$

On the last line, we introduced the *covariant derivative* of the vector components V^α:

$$\nabla_\beta V^\alpha = \partial_\beta V^\alpha + V^\mu \Gamma^\alpha_{\beta\mu} \;. \tag{5.43}$$

The importance of the covariant derivative is that $\nabla_\beta V^\alpha$ represents the components of a tensor: if we change our coordinate representation, then the components change in the usual way:

$$\nabla_{\bar\nu} V^{\bar\mu} = \left(\frac{\partial x^{\bar\mu}}{\partial x^\alpha}\right)\left(\frac{\partial x^\beta}{\partial x^{\bar\nu}}\right)\nabla_\beta V^\alpha \;. \tag{5.44}$$

A corollary of this statement is that in general coordinates, the components $\partial_\beta V^\alpha$ do *not* represent the components of a tensor. If these components did in fact represent tensor components, then we would expect that

$$\partial_{\bar\nu} V^{\bar\mu} \stackrel{?}{=} \left(\frac{\partial x^{\bar\mu}}{\partial x^\alpha}\right)\left(\frac{\partial x^\beta}{\partial x^{\bar\nu}}\right)\partial_\beta V^\alpha \;. \tag{5.45}$$

Let's see what happens if we enact this transformation:

$$\begin{aligned}
\partial_{\bar\nu} V^{\bar\mu} &= \left(\frac{\partial x^\beta}{\partial x^{\bar\nu}}\right)\partial_\beta\left[\left(\frac{\partial x^{\bar\mu}}{\partial x^\alpha}\right)V^\alpha\right] \\
&= \left(\frac{\partial x^\beta}{\partial x^{\bar\nu}}\right)\left(\frac{\partial x^{\bar\mu}}{\partial x^\alpha}\right)\partial_\beta V^\alpha + \left(\frac{\partial x^\beta}{\partial x^{\bar\nu}}\right)\left(\frac{\partial^2 x^{\bar\mu}}{\partial x^\alpha \partial x^\beta}\right)V^\alpha \;.
\end{aligned} \tag{5.46}$$

The second term in the above equation spoils the tensorial nature of $\partial_\beta V^\alpha$, showing that the $\stackrel{?}{=}$ in (5.45) is not in fact an equality. On an upcoming problem set, you will show that that the Christoffel symbol $\Gamma^\alpha_{\beta\mu}$ is also not tensorial; but, that the manner in which it "fails" to transform as a tensor exactly compensates for how the partial derivative fails. As a consequence, the sum of the partial derivative and the Christoffel term transforms exactly as tensor components need to transform.