The stress tensor as the canonical generator of spacetime symmetries

Don Marolf

For historical reasons, the following notes are presented in the form of a "HW problem" and solution (provided below). I am not asking you to try and work the HW problem. These notes are a supplement to the main material in class that you are welcome to study at your leisure. You are also welcome to ask me any questions about them during office hours.

The Peierls Bracket [Note: There is a lot of background material below to read for this problem. The actual problem starts after **The problem** below.]

Context: GR textbooks typically discuss T_{ab} in two ways. One is to explain T_{ab} as the covariant *local* description of energy and momentum (giving the flux through any surface element). However, we *defined* T_{ab} in terms of the variation of the action with respect to the metric. The attentive student might ask, "just how are these notions connected?" Of course, we saw in simple examples (such a the free relativistic particle) that they are indeed the same. But is there a general argument to this effect?

The answer is a resounding YES. In general, energy and momentum are the conserved quantities associated with time and space translation invariance. In GR, we know that for any Killing Field ξ there is a conserved current $T_{ab}\xi^a$ (with T_{ab} defined by varying the action with respect to the metric) which leads to a corresponding conserved charge. This strongly suggests that this conserved charge is energy (or momentum, angular momentum, etc., depending on the nature of the Killing field), and helps to build the connection we want.

In the (optional!) problem below, you will take this one step further. The point is to show not just that $T_{ab}\xi^a$ is conserved, but that integrals of this quantity in fact *generate* translations along ξ in the usual sense of classical mechanics. This allows us to conclusively identify these conserved charges as energy (momentum, angular momentum, etc...) and to complete the desired connection.

Background Material: The Peierls bracket is a covariant version of the Poisson bracket from classical mechanics. Though the definition may seem very different from that of the Poisson bracket, they can be shown to agree (by what is essentially a direct calculation; see R. E. Peierls, The commutation laws of relativistic field theory, Proc. Roy. Soc. (London) 214 (1952), 143-157.) The Peierls bracket is often used in quantum field theory (though for some reason it is just called the "commutator function.")

The Peierls bracket is defined for any system that 1) is described by an action S and 2) admits some kind of initial value formulation. To define the Peierls bracket $[A, B]$ of two (gauge-invariant) quantities A and B, one considers a 1-parameter family of new actions

$$S_\epsilon = S + \epsilon A \tag{1}$$

for (small) real ϵ. For example, in a scalar theory we may have $A = \phi(x)$ and $B = \phi(x')$. Then

$$S_\epsilon = S + \epsilon\phi(x) = S + \epsilon \int d^4y \delta(x, y)\phi(x), \tag{2}$$

(where here $\delta(x, y)$ is taken to be a density with respect to the argument y).

Now, consider some solution s_0 to the equations of motion defined by the *original* action $S = S_0$. Each new action leads to a new set of equations of motion with a 'source' proportional to ϵ, so that s_0 will no longer be a solution. However, it *will* be a solution to the modified equations of motion in the region R^- of spacetime to the past of A, and also in the region R^s which is spacelike separated from A, since no causal influences from A will reach into $R^- \cup R^s$. As a result, there is *some* solution s_ϵ^- to the modified equations of motion which agrees with s_0 in $R^- \cup R^s$. This s_ϵ^- is called the "retarded solution." Similarly, there is an "advanced solution" s_ϵ^+ which agrees with s_0 in the region $R^+ \cup R^s$, where R^+ denotes the region to the future of A.

Now, we may evaluate the quantity B on each solution $s_\epsilon^\pm$. In fact, we wish to compute the retarded (advanced) *change* in B due to A defined as:

$$D_A^\pm B(s_0) := \frac{1}{\epsilon} \lim_{\epsilon \to 0} B(s_\epsilon^\pm) - B(s_0). \tag{3}$$

We are now ready to define the Peierls Bracket $[A, B]$. Like A and B, it is some physical quantity which depends on the particular solution chosen:

$$[A, B] = D_A^- B - D_A^+ B. \tag{4}$$

The problem: Consider any matter theory described by an action S. (If you would like to work with an explicit example, you may take the matter theory to be a scalar field with standard kinetic and potential terms.) We take the spacetime metric g to be fixed (i.e., gravity is not dynamical) and to have a Killing field ξ.

Use the Peierls Bracket to show that $\int_\Sigma T_{ab}\xi^a n^b \sqrt{g_\Sigma}$ generates translations along ξ. Here Σ is a "Cauchy surface." (A surface on which initial data can be used to determine the entire evolution of the system for all time. Roughly, a surface of constant time.) The vector n^b is the future-pointing unit normal to Σ, and g_Σ is the determinant of the induced metric on Σ.

The proof is easiest if you break it down into parts as follows:

a) Consider any smooth function f such that i) $f = 0$ in the far past and ii) $f = 1$ in the far future. Define $\xi_f = f\xi$.

   Let the dynamical fields of the matter theory be called ϕ. [It doesn't really matter whether these are scalar fields or not.] Consider a small variation of these fields defined by

$$\delta\phi = \pounds_{\xi_f}\phi, \tag{5}$$

   where $\pounds_{\xi_f}$ is the so-called "Lie Derivative" along ξ_f. If you like, you can read about Lie Derivatives in the appendices of Wald or Carroll. But what is important is that the Lie Derivative of ϕ is defined to be the infinitesimal change in ϕ along the flow generated by ξ_f.

   Show that, when the fields are varied as in (5), the change in the action is:

$$\delta S = -\int \sqrt{-g}T^{ab}\nabla_a(f\xi)_b. \tag{6}$$

   Hint #1: The metric g_{ab} is not one of the dynamical fields being varied, but $\pounds_{\xi_f} g_{ab} = \nabla_a(\xi_f)_b + \nabla_a(\xi_f)_b$.

   Hint #2: When the equations of motion hold, the action is stationary with respect to compactly supported variations of the matter fields. But the above variation is not compactly supported (since $f\xi$ does not vanish in the far future). This is what makes the above δS non-zero.

   Solution: The key point here is that the δS defined through (5) above is the variation under just changing the fields, not a full diffeomorphism where we change both the fields and the metric. To emphasize this, let us for the moment refer to the above δS as $\delta_{\phi-\text{only}}S$. On the other hand, if we did a full diffeo we would find

$$\delta_{\text{diffeo}}S = \int \Big(\frac{\delta S}{\delta g_{ab}}\delta_{\text{diffeo}}g_{ab} + \frac{\delta S}{\delta\phi}\delta_{\text{diffeo}}\phi\Big) = \delta_{\phi-\text{only}}S + \int \frac{\delta S}{\delta g_{ab}}\delta_{\text{diffeo}}g_{ab} = 0 \tag{7}$$

Then, taking care to get the right sign for upper indices (recall $\frac{\delta}{\delta g_{ab}} = -g^{ac}g^{bd}\frac{\delta}{\delta g^{cd}}$),

$$T^{ab} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g_{ab}} \tag{8}$$

and since

$$\delta_{\text{diffeo}} g_{ab} = \nabla_a(\xi_f)_b + \nabla_b(\xi_f)_a \tag{9}$$

$$\delta_{\phi-\text{only}} S = \int \frac{\delta S}{\delta \phi}\delta\phi = -2\int \sqrt{-g}T^{ab}\delta g_{ab} = -\int \sqrt{-g}T^{ab}\nabla_a(\xi_f)_b \tag{10}$$

using the symmetry of T^{ab}. Below, I will return to referring to (10) as just δS without a subscript on the δ.

Let me comment on a subtlety: But wait, many of you may say, isn't $\frac{\delta S}{\delta \phi}$ zero by the equations of motion? So, if the above is true, doesn't it say that $T^{ab} = 0$? Actually, no, due to a subtlety. The point is that the equations of motion make S stationary under *compactly supported* variations of the matter fields. But here the variation is not compactly supported (or even close!) because $f\xi$ does not vanish in the far future. As a result, the desired variation δS has a non-zero boundary term. [Some of you may object that, due to this boundary term, the functional derivatives $\frac{\delta S}{\delta \phi}$ are not well-defined and I should not have used this symbol above. While one may take this point of view, is is just as useful to think of $\frac{\delta S}{\delta \phi}$ as containing an extra term proportional to a delta-function at future infinity which does not vanish by the equations of motion.]

b) Consider the action $S + \epsilon\delta S$, where δS was defined in (a). Consider some solution s_0 of the original theory (defined by S), and lest $s_\epsilon^\pm$ be the associated advanced and retarded solutions of the deformed theory as defined in the text at the beginning of the problem. That is, in the far past region (where $f = 0$) we have $s_0 = s_\epsilon^-$, while in the far future region (where $f = 1$) we have $s_0 = s_\epsilon^+$.

   Argue that (to first order in ϵ) in the far future region (where $f = 1$), s_ϵ^- can be obtained from s_0 by acting which the diffeomorphism which flows *backwards* along ξ with parameter ϵ.

Also argue that (to first order in ϵ) in the far past region (where $f = 0$), s_ϵ^+ can be obtained from s_0 by acting which the diffeomorphism which flows *forwards* along ξ with parameter ϵ.

Hint: recall that ξ is a Killing field, so that flow along ξ is a symmetry of the system.

Solution:

The action is generally covariant so if we modify it by performing a diffeomorphism along $f\xi$, the solutions to the resulting equations of motion will be the result of applying this same diffeomorphism to the solutions of the original theory. There is, however, a somewhat subtle sign:

$$s_0 \to s_0 - \epsilon \mathcal{L}_{f\xi} s_0 \tag{11}$$

This subtle sign is actually quite familiar: It is the same sign that appears in the statement that, on the real line, shifting any *function* $f(x)$ to the right is equivalent to shifting the *argument* to the right. E.g., the quadratic x^2 has a minimum at $x = 0$. But the quadratic $(1 + \partial_x)x^2 = x^2 + 2x = (x + 1)^2 - 1$ has a minimum at $x = -1$ (i.e., at the point obtained from $x = 0$ by moving along the vector field $-\partial_x$). For what follows, note that it suffices to work to first order in ϵ.

Now in the far past $f = 0$ and s_{ϵ^-} just matches s_0

$$s_0 \to s_0 \tag{12}$$

that is

$$s_{\epsilon^-} = s_0 \tag{13}$$

while the far future $f = 1$ and

$$s_0 \to s_0 - \epsilon \mathcal{L}_\xi s_0 \tag{14}$$

i.e.

$$s_{\epsilon^-} = s_0 - \epsilon \mathcal{L}_\xi s_0 \tag{15}$$

that is, the desired result.

Now for finding s_{ϵ^+} consider a solution s_1 which satisfies the S equations of motion (i.e. the unperturbed action) and for which $s_1 - \epsilon \mathcal{L}_{\xi f} s_1$ solves the $S + \epsilon\delta S$ equations of motion. Further we want to make $s_1 = s_0$ in the far future where $f = 1$. Then, to first order in ϵ, all we need do is take $s_1 = s_0 + \epsilon \mathcal{L}_\xi s_0$. That is in the far future ($f = 1$)

$$s_{\epsilon^+} = s_0 \tag{16}$$

while in the far past ($f = 0$)

$$s_{\epsilon^+} = s_0 + \epsilon \mathcal{L}_\xi s_0 \tag{17}$$

and we have the desired result.

c) Consider some quantity B built entirely from dynamical fields in the far future region where $f = 1$. (In particular, B does not depend on any non-dynamical background structure. You might focus on the example $B = \phi(x_0)$ where ϕ is a dynamical scalar field and x_0 is a spacetime point for which $f(x_0) = 1$.) Argue that the Peierls bracket yields

$$[\delta S, B] = -\pounds_\xi B, \tag{18}$$

i.e., that it is the infinitesimal change in B associated with flowing along the vector field ξ.

Solution:

Now in the far future $s_{\epsilon^+} = s_0$ so taking $A = \delta S$

$$D_A^+ B = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \Big(B(s_{\epsilon^+}) - B(s_0) \Big) = 0 \tag{19}$$

while

$$D_A^- B = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \Big(B(s_{\epsilon^-}) - B(s_0) \Big) = -\lim_{\epsilon \to 0} \frac{1}{\epsilon} (\epsilon \mathcal{L}_\xi B) = -\mathcal{L}_\xi B \tag{20}$$

so then

$$[\delta S, B] = D_A^- B - D_A^+ B = -\mathcal{L}_\xi B. \tag{21}$$

d) Finally, show that we may write

$$\delta S = \int_\Sigma T_{ab} \xi^a n^b \sqrt{g_\Sigma}, \tag{22}$$

where Σ is any surface in the far future region (where $f = 1$). Then use conservation of T_{ab} to conclude that (22) holds for any surface Σ. *Note:*

See appendix B of Wald for the correct sign of Stoke's theorem in curved spacetimes. Carroll's treatment (appendix E) has a typo that yields the wrong sign in Lorentz signature.

Solution:

Now we want to write the result from part A in terms of a surface integral. Recalling the conservation of stress energy

$$\nabla_a(T^{ab}) = 0 \tag{23}$$

$$\begin{aligned} \delta S &= -\int \sqrt{-g}T^{ab}\nabla_a(f\xi)_b = -\int \sqrt{-g}\nabla_a\Big(T^{ab}(f\xi)_b\Big) \\ &= -\int_\Sigma \Big((-n_a)T^{ab}f\xi_b\Big)\sqrt{g_\Sigma} = \int_\Sigma T_{ab}\xi^a n^b\sqrt{g_\Sigma} \end{aligned} \tag{24}$$

where Σ is any surface in the far future (where $f = 1$) and $\mathbf{n}$ is a future directed timelike normal to Σ. Note for this surface term we must recall that for timelike normals the unit normal is chosen to be inward pointing to satisfy Stoke's theorem (so in practice for timelike normals just add an extra sign). See appendix B of Wald. Carroll also discusses this point in appendix E, but he has a typo which yields the wrong overall sign in Lorentz signature (i.e., regardless of the spacetime signature, it is timelike normals that get the extra sign) . The other surface terms vanish since $f = 0$ at the boundary in the far past and I will assume T_{ab} vanishes at our spacelike boundaries (no matter or energy flowing in or out of our volume at large spatial distances).

Finally, note that (24) is independent of what surface we take since

$$\nabla_a(T^{ab}\xi_b) = \nabla_a(T^{ab})\xi_b + T^{ab}\nabla_a\xi_b = 0 \tag{25}$$

where the first term vanishes by conservation of T^{ab} and the second via Killing's equation and the symmetry of T^{ab}. That is, for any other surface Σ' we can use Stoke's theorem and the fact that the resulting volume integral vanishes to get the same answer for Σ' as for Σ.

e) Explain why (a-d) together prove the desired result; i.e., that $\int_\Sigma T_{ab}\xi^a n^b\sqrt{g_\Sigma}$ generates the symmetry associated with ξ. In particular, argue that (18) is true for *any* B.

Solution:

Note nothing in the above argument depends on precisely where in spacetime B is localized. In particular, for any B we can always choose a function f which vanishes in the far past and goes to one before B. Alternatively, one could work out the past story analogue of part C.