Product Rule for Derivatives of Tensors

09/10/2020

In this note, I'll try to elaborate more on the discussion we had in problem session on the product rule for derivatives acting on tensors. In the following, we will consider a d-dimensional Minkowski spacetime with metric tensor $\eta_{\mu\nu}$. Additionally, we will let T, S, and R be tensors of rank n, m, and p, respectively, for some arbitrary n, m, p.

The derivative operator $\partial_\mu = \frac{\partial}{\partial x^\mu}$ denotes partial derivatives with respect to the coordinates x^μ. Importantly, even when these derivatives act on tensors instead of just scalars, we still have the product rule for differentiation:

$$\partial_\mu \left(T^{\nu_1 \dots \nu_n} S^{\rho_1 \dots \rho_m} \right) = S^{\rho_1 \dots \rho_m} \partial_\mu T^{\nu_1 \dots \nu_n} + T^{\nu_1 \dots \nu_n} \partial_\mu S^{\rho_1 \dots \rho_m} \ . \tag{1}$$

Note that we wrote this rule where all indices on S and T appear upstairs, but this positioning of the indices doesn't matter; the same product rule will apply when any of the indices are lowered. Similarly, if we had acted with ∂^μ instead of ∂_μ, we would also have a similar product rule. This is because indices are raised and lowered with the metric tensor $\eta_{\mu\nu}$, which is a constant and can thus always be moved around freely, even with the derivative operator present.

As a corollary of the product rule (1), we note the following useful identity:

$$\boxed{\partial_\mu \left(T^{\nu_1 \dots \nu_n} T_{\nu_1 \dots \nu_n} \right) = 2 T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n}} \ . \tag{2}$$

To see this, we first use the product rule (1) to write the left-hand side of this equation as follows:

$$\partial_\mu \left(T^{\nu_1 \dots \nu_n} T_{\nu_1 \dots \nu_n} \right) = T_{\nu_1 \dots \nu_n} \partial_\mu T^{\nu_1 \dots \nu_n} + T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n} \ . \tag{3}$$

From here, we know that the indices ν_1, ν_2, $\dots \nu_n$ are all dummy indices that are summed over, and so we should be able to freely raise and lower pairs of contracted indices. (See questions 5 and 6 on Problem Set 1 for more details). This is true even when derivatives are present, since the Minkowski metric is constant and so $\partial_\mu \eta_{\nu\rho} = 0$. For completeness sake, I will go through this raising/lowering process explicitly:

$$\begin{aligned} T_{\nu_1 \dots \nu_n} \partial_\mu T^{\nu_1 \dots \nu_n} &= T_{\nu_1 \dots \nu_n} \partial_\mu \left(\eta^{\nu_1 \rho_1} \dots \eta^{\nu_n \rho_n} T_{\rho_1 \dots \rho_n} \right) \\ &= T_{\nu_1 \dots \nu_n} \eta^{\nu_1 \rho_1} \dots \eta^{\nu_n \rho_n} \partial_\mu T_{\rho_1 \dots \rho_n} \\ &= \left(\eta^{\nu_1 \rho_1} \dots \eta^{\nu_n \rho_n} T_{\nu_1 \dots \nu_n} \right) \partial_\mu T_{\rho_1 \dots \rho_n} \\ &= T^{\rho_1 \dots \rho_n} \partial_\mu T_{\rho_1 \dots \rho_n} \\ &= T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n} \ , \end{aligned} \tag{4}$$

where in the last line we simply relabelled all dummy indices. Therefore the first term on the right-hand side of (3) is precisely the same as the second term on the right-hand side of (3), and they can be combined appropriately:

$$\begin{aligned} \partial_\mu \left(T^{\nu_1 \dots \nu_n} T_{\nu_1 \dots \nu_n} \right) &= T_{\nu_1 \dots \nu_n} \partial_\mu T^{\nu_1 \dots \nu_n} + T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n} \\ &= T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n} + T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n} \\ &= 2 T^{\nu_1 \dots \nu_n} \partial_\mu T_{\nu_1 \dots \nu_n} \ . \end{aligned} \tag{5}$$

This is precisely the identity (2) we wanted to establish.

We have so far only been discussing partial differentiation with respect to the spacetime coordinates x^μ. However, as discussed somewhat in Problem Set 1 and Problem Set 2, we may also be interested in differentation with respect to *tensors*, which is relevant for the Euler-Lagrange equations. Fortunately, the rules for differentiation with respect to tensors are the same as with respect to coordinates (coordinates are just a specific instance of a tensor!), and so we also have a product rule:

$$\frac{\partial}{\partial R^{\mu_1 \cdots \mu_p}} \left(T^{\nu_1 \ldots \nu_n} S^{\rho_1 \ldots \rho_m} \right) = S^{\rho_1 \ldots \rho_m} \frac{\partial}{\partial R^{\mu_1 \cdots \mu_p}} T^{\nu_1 \ldots \nu_n} + T^{\nu_1 \ldots \nu_n} \frac{\partial}{\partial R^{\mu_1 \cdots \mu_p}} S^{\rho_1 \ldots \rho_m} \ . \tag{6}$$

And, once again, the location (up or down) of indices doesn't matter, since we can move the Minkowski metric in and out of the derivative freely.[1] As a result, all of our hard work in deriving the identity (2) will also apply when differentiating with respect to tensors, and thus we arrive at another useful identity:

$$\boxed{\frac{\partial}{\partial R^{\mu_1 \cdots \mu_p}} \left(T^{\nu_1 \ldots \nu_n} T_{\nu_1 \ldots \nu_n} \right) = 2 T^{\nu_1 \ldots \nu_n} \frac{\partial}{\partial R^{\mu_1 \cdots \mu_p}} T_{\nu_1 \ldots \nu_n}} \ . \tag{7}$$

You may find such an identity very useful when deriving the Euler-Lagrange equations for various Lagrangians throughout this class.

For example, consider the Maxwell Lagrangian we discussed in Problem Set 3:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + 4\pi A_\mu J^\mu \ , \tag{8}$$

where the field strength is defined by $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$, and J^μ is a fixed background field that is independent from the dynamical vector field A_μ. In deriving the equations of motion for A_μ, we end up having to differentiate the Lagrangian $\mathcal{L}$ with respect to the tensor $\partial_\mu A_\nu$. This computation can get a little hairy, because you may try expanding out all of the field strengths explicitly and then differentiating. However, the identity (7) gives us a much quicker way to do the differentiation:

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{4} \frac{\partial}{\partial(\partial_\mu A_\nu)} \left(F_{\rho\sigma} F^{\rho\sigma} \right) = -\frac{1}{2} F^{\rho\sigma} \frac{\partial F_{\rho\sigma}}{\partial(\partial_\mu A_\nu)} \ . \tag{9}$$

The first equality in (9) came from noticing that the term $A_\mu J^\mu$ has no derivatives acting on A, and so it is not relevant for our differentiation. The second equality is then just a simple application of (7). Thus, our product rule identities help shorten the number of calculations you have to do when deriving Maxwell's equations from the Maxwell Lagrangian.

[1] The only exception to this would be if we were differentiating *with respect to* the metric tensor. In that case, we would have to keep careful track of how the metric tensor appears in all expressions, and thus the index positioning would matter.