GRADIENTS AND DIRECTIONAL DERIVATIVES OF TENSORS

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

References: Kip S. Thorne & Roger D. Blandford, *Modern Classical Physics*, Princeton University Press (2017). Section 1.7.

Post date: 18 Sep 2020.

In Section 1.7, T&B introduce the idea of a *directional derivative* of a tensor T, which they define as

$$\nabla_{\boldsymbol{A}}\mathsf{T} \equiv \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{A}\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)\right] \tag{1}$$

The tensor T may be any rank, but the slots into which one inserts the vectors have been suppressed in the notation. The vector $\boldsymbol{x}_{\mathcal{P}}$ indicates the location of a point $\mathcal{P}$, and the vector $\boldsymbol{A}$ indicates the direction along which the derivative is taken. The notation $\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)$ indicates that the tensor T is a general (not necessarily linear) function of position, and this dependence is not the same as that for the vectors which are inserted into the tensor's slots. A tensor is linear with respect to its dependence on the vectors which go into its slots (see T&B Sec 1.3), so that, for example

$$\mathsf{T}\left(e\boldsymbol{E}+f\boldsymbol{F},\boldsymbol{B},\boldsymbol{C}\right)=e\mathsf{T}\left(\boldsymbol{E},\boldsymbol{B},\boldsymbol{C}\right)+f\mathsf{T}\left(\boldsymbol{F},\boldsymbol{B},\boldsymbol{C}\right) \tag{2}$$

However the notation $\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)$ can represent *any* type of function of position, not necessarily a linear one.

T&B mention that 1 is linear in the vector $\boldsymbol{A}$ along which the derivative is taken. We can see this by starting with 1. Suppose we want the derivative along a direction $\boldsymbol{A}+\boldsymbol{B}$. Then we have

$$\nabla_{\boldsymbol{A}+\boldsymbol{B}}\mathsf{T} = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\left(\boldsymbol{A}+\boldsymbol{B}\right)\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)\right] \tag{3}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\left(\boldsymbol{A}+\boldsymbol{B}\right)\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{A}\right)+\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{A}\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)\right] \tag{4}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\left(\boldsymbol{A}+\boldsymbol{B}\right)\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{A}\right)\right]+ \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{A}\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)\right] \tag{5}$$

$$= \nabla_{\boldsymbol{B}}\mathsf{T}+\nabla_{\boldsymbol{A}}\mathsf{T} \tag{6}$$

Here, we've set

$$\nabla_{\boldsymbol{B}}\mathsf{T}=\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\left(\boldsymbol{A}+\boldsymbol{B}\right)\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{A}\right)\right] \tag{7}$$

which might bother you. However, the infinitesimal displacement in the $\boldsymbol{A}$ direction is the same in both terms, and goes to zero in the limit $\epsilon\to 0$, so it really is the same as the derivative in the $\boldsymbol{B}$ direction at the point $\mathcal{P}$.

Since the directional derivative is linear with respect to the directional vector $\boldsymbol{A}$, it defines a tensor with a rank one greater than the rank of T. In geometric notation, we have, for a rank-3 tensor T

$$\nabla_{\boldsymbol{A}}\mathsf{T}\left(\_,\_,\_\right)=\nabla\mathsf{T}\left(\_,\_,\_,\boldsymbol{A}\right) \tag{8}$$

The tensor on the RHS is defined to be the *gradient.* The last slot in the gradient is the 'differentiation slot', and defines the direction in which the gradient is taken.

To get the components of the gradient, we insert the basis vectors into the slots, so we have

$$\left(\nabla\mathsf{T}\right)_{abcj}=\nabla\mathsf{T}\left(\boldsymbol{e}_a,\boldsymbol{e}_b,\boldsymbol{e}_c,\boldsymbol{e}_j\right) \tag{9}$$

From 1 we have (no sum over j):

$$\begin{aligned}
\left(\nabla\mathsf{T}\right)_{abcj} &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}+\epsilon\boldsymbol{e}_j\right)-\mathsf{T}\left(\boldsymbol{x}_{\mathcal{P}}\right)\right]_{abc} && (10)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(x_j\boldsymbol{e}_j+\epsilon\boldsymbol{e}_j+\boldsymbol{x}'_{\mathcal{P}}\right)-\mathsf{T}\left(x_j\boldsymbol{e}_j+\boldsymbol{x}'_{\mathcal{P}}\right)\right]_{abc} && (11)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathsf{T}\left(\left(x_j+\epsilon\right)\boldsymbol{e}_j+\boldsymbol{x}'_{\mathcal{P}}\right)-\mathsf{T}\left(x_j\boldsymbol{e}_j+\boldsymbol{x}'_{\mathcal{P}}\right)\right]_{abc} && (12)
\end{aligned}$$

The suffix abc indicates that we've inserted the basis vectors $\boldsymbol{e}_a,\boldsymbol{e}_b,\boldsymbol{e}_c$ into 9 to get the corresponding components.

Here, $\boldsymbol{x}'_{\mathcal{P}}$ is defined as

$$\boldsymbol{x}'_{\mathcal{P}}\equiv\boldsymbol{x}_{\mathcal{P}}-x_j\boldsymbol{e}_j \tag{13}$$

That is, it's that component of $\boldsymbol{x}_{\mathcal{P}}$ that is perpendicular to $\boldsymbol{e}_j$.

From 12, we see that we're effectively taking the derivative of T_{abc} with respect to the coordinate x_j. That is

$$\left(\nabla_{\boldsymbol{A}}\mathsf{T}\right)_{abcj}=\frac{\partial T_{abc}}{\partial x_j} \tag{14}$$

The usual shorthand notation for the RHS is to write the derivative component after a comma, so we have

$$T_{abc,j} \equiv \frac{\partial T_{abc}}{\partial x_j} \tag{15}$$