# LIE DERIVATIVES

Link to: physicspages home page.

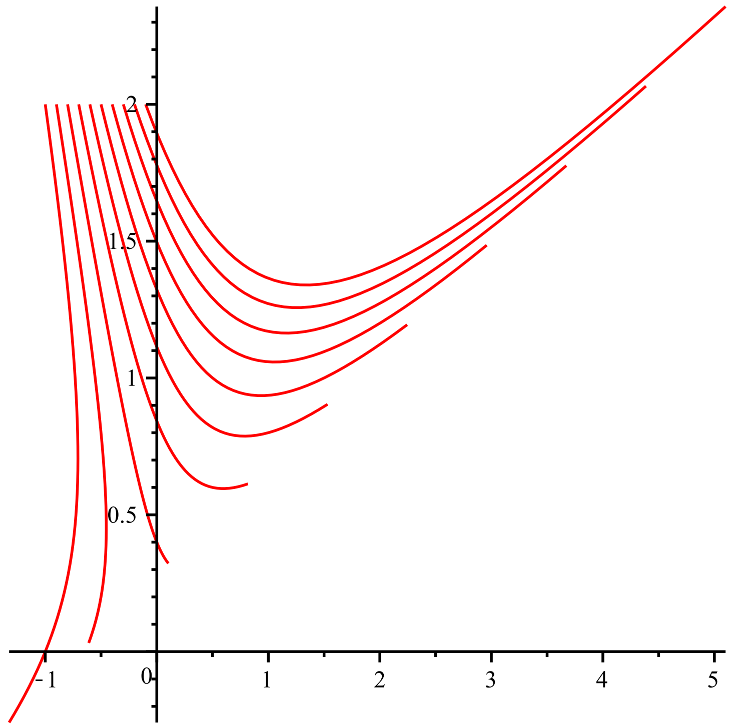
To leave a comment or report an error, please use the auxiliary blog.

Reference: d'Inverno, Ray, *Introducing Einstein's Relativity* (1992), Oxford Uni Press. - Section 6.2.

The straightforward derivative of a tensor is not itself a tensor, as we've seen in an earlier example. In this post, we'll examine one of the ways in which a tensor derivative that actually *is* a tensor can be defined: the Lie derivative.

The concept can be devilishly difficult to understand, and most of the derivations I've seen tend to gloss over the conceptual aspect of a Lie derivative. Probably the clearest way of introducing it is in the context of fluid flow.

In the last post, we examined the congruence of curves that can be derived from a vector field. A vector field can be used to derive such a set of curves, and we'll reproduce here the plot included in the last post as an example. (The actual equations used to derive these curves are not important here; all that matters is that you can visualize a set of curves.)




These curves are all defined by a parameter $t$, (we'll use $t$ instead of $u$ as the parameter here, since it fits in better with the interpretation that follows) which in the plot, begins at $t = 0$ at the starting point of each curve at the upper left, and increases up to $t = 1.5$ as you move along each curve away the starting point. Imagine now that these curves are flow lines in a fluid; that is, if you placed a particle on one of the curves, it would follow that curve as the fluid flowed. The first thing to notice is that the speed of the particle varies from one curve to the next. For the same range in $t$ (which could be thought of as time) some curves are longer than others, so the particle would move farther in the same time.

Let's begin with the simplest case. Suppose we have a scalar field $f$, such as the temperature or density of the fluid, which is defined everywhere in the fluid. If we look at the value of $f$ at some point, say the starting point of the first curve on the left, then as the fluid flows, this point gets dragged along the flow line defined by the first curve. We can follow this point and look at how $f$ changes as this point flows along. The rate of change is, using the chain rule:

$$\frac{d}{dt} f(\mathbf{r}(t)) = \frac{df}{dx}\frac{dx}{dt} + \frac{df}{dy}\frac{dy}{dt} + \frac{df}{dz}\frac{dz}{dt} \quad (1)$$

$$= \mathbf{v} \cdot \nabla f \quad (2)$$

This change is zero if $\mathbf{v}$ is perpendicular to $\nabla f$ (in other words, the path of flow is along a line of constant $f$).

This derivative is defined as the Lie derivative of a scalar field, and thus measures the change in the field as the observation point moves along a flow line.

Now suppose we consider two points $\mathbf{r}_1(t=0)$ and $\mathbf{r}_2(t=0)$. To make things definite, suppose $\mathbf{r}_1$ is the starting point of the first curve on the left, and $\mathbf{r}_2$ is the corresponding point on the second curve. Now suppose we define a vector $\lambda\mathbf{w} = \mathbf{r}_2(0) - \mathbf{r}_1(0)$. Here, $\lambda$ is some small quantity, so that the two curves are assumed to be very close to each other. The vector $\mathbf{w}$ could measure any vector property of the fluid, such as the temperature or density gradient. The units of $\lambda$ and $\mathbf{w}$ must be such that $\lambda\mathbf{w}$ has the units of length, but apart from that, there is no restriction.

As the fluid flows, the points $\mathbf{r}_1$ and $\mathbf{r}_2$ get dragged along with the fluid, and it's possible that $\mathbf{r}_1(t) + \lambda\mathbf{w}(\mathbf{r}(t))$ might drift away from $\mathbf{r}_2(t)$. If we require that *not* to be the case, that is, that the condition

$$\mathbf{r}_1(t) + \lambda\mathbf{w}(\mathbf{r}_1(t)) = \mathbf{r}_2(t) \tag{3}$$

remain true (for small $\lambda$ but for $t > 0$), then we are saying that the corresponding points on the two flow lines are connected by the same relationship as the flow proceeds. Note that this doesn't mean that $\mathbf{w}$ is constant; it merely means that as $\mathbf{w}$ is dragged along, $\lambda\mathbf{w}$ still connects the two points on the two flow lines.

Now take the derivative with respect to $t$ of this requirement, and use the chain rule as above. We get

$$\mathbf{v}(\mathbf{r}_1) + \lambda(\mathbf{v}(\mathbf{r}_1)\cdot\nabla)\mathbf{w}(\mathbf{r}_1) = \mathbf{v}(\mathbf{r}_2) \tag{4}$$

Substituting for $\mathbf{r}_2$ we can expand the RHS in a Taylor series:

$$\begin{aligned}\mathbf{v}(\mathbf{r}_2) &= \mathbf{v}(\mathbf{r}_1 + \lambda\mathbf{w}(\mathbf{r}_1)) && (5)\\ &= \mathbf{v}(\mathbf{r}_1) + \lambda(\mathbf{w}\cdot\nabla)\mathbf{v}(\mathbf{r}_1) + O(\lambda^2) && (6)\end{aligned}$$

In the limit $\lambda \to 0$, the condition becomes

$$(\mathbf{v}(\mathbf{r}_1)\cdot\nabla)\mathbf{w}(\mathbf{r}_1) = (\mathbf{w}(\mathbf{r}_1)\cdot\nabla)\mathbf{v}(\mathbf{r}_1) \tag{7}$$

In briefer notation, we can define the Lie derivative of a vector field $\mathbf{w}$ with respect to the vector field $\mathbf{v}$ as

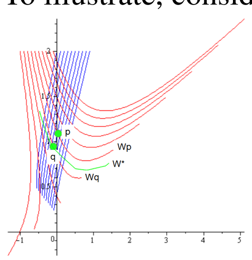
$$L_{\mathbf{v}}\mathbf{w} = (\mathbf{v}\cdot\nabla)\mathbf{w} - (\mathbf{w}\cdot\nabla)\mathbf{v} \tag{8}$$

What exactly does this *mean*? The most common explanation given in textbooks seems to be something like this. The Lie derivative looks at the change in the vector field **w** between two points and compares this with the change in **w** that would result by dragging the vector between these two points. If you're like me, this makes no sense whatsoever. In particular, how *can* there be a difference between these two things?

One way of thinking about it that may help goes like this. Suppose we had only the vector field **w**. If we were asked for its derivative, what would we do? Clearly, we have to know what we're taking the derivative with respect to, and therein lies the problem. We've seen earlier that we can generate a congruence of curves from **w**, so we might ask for the derivative as we make the transition between two adjacent curves from this congruence. But how are we going to make the transition? In the figure above, for example, there are many ways we could cross the set of curves to provide paths along which we might find the derivative. Without specifying these paths, we really can't do any calculations.

What we need, then, is *another* congruence of curves which provides these paths. That is where the flow lines described above come in: they provide the paths along which we move to find the derivative. The flow lines we've been considering all along are *not* from the congruence due to **w**; in fact, we haven't mentioned **w**'s congruence yet at all.

To illustrate, consider the following diagram.




The blue curves are the flow lines we've been considering above, which we can take as the congruence from **v**; the red curves are the congruence from **w**. A curve from a congruence is called an *integral curve*, since it's obtained by integrating vector field.

The blue curves are parametrized by $t$ since they are the flow curves, while the red curves are parametrized by a different parameter, say $u$. Consider point $p$ and suppose we want to find the derivative of **w** at this point. We need to use the blue curves to define the direction along which the derivative is to be taken. Point $p$ lies on curve $Wp$ from **w**'s congruence. We move a distance $\Delta t$ down the blue curve passing through $p$ until we reach point $q$ which lies on the curve $Wq$, also from **w**'s congruence. Having done this, we can now define a new curve $W*$ (in green) by starting on $Wp$ and moving a distance $\Delta t$ down *all* the blue curves. This green curve is *not* in the congruence of either **w** or **v** and in fact it intersects $Wq$ at point $q$. (Remember that no two curves in a congruence can intersect.) This green curve is the one resulting from 'dragging' $Wp$ along the flow lines of **v**, while curve $Wq$ is the integral curve passing through the same point $q$. That is, a tangent to $W*$ is where a vector from the field **w** would end up if it was dragged along with the flow, while $Wq$ is what the vector field **w** actually *is* along the integral curve passing through $q$. In general, these two curves are not the same.

Both $W*$ and all the $W$ integral curves are parametrized by $u$, so we can calculate the derivative along each of them. That is, we can calculate $dW/du$ and $dW*/du$. If we take the difference, we get

$$\Delta W \equiv \left.\frac{dW}{du}\right|_{t_p+\Delta t} - \frac{dW*}{du} \tag{9}$$

where we're interested in the behaviour at point $q$ which corresponds to $t_p+\Delta t$ ($t_p$ is the value of $t$ at point $p$).

Since $Wq$ is an integral curve of **w**, the derivative $dW/du$ is just the vector field back again. Since $W*$ is not an integral curve, we have to treat it differently.

We can also think of each of $W$ and $W*$ as being functions of $t$ since they are both determined by how far we've moved along the flow curve in the calculation. , we can write, using a Taylor series:

$$\frac{\partial W}{\partial u} = \left.\frac{\partial W}{\partial u}\right|_{t_p+\Delta t} \tag{10}$$

$$= \left.\frac{\partial W}{\partial u}\right|_{t_p} + \left.\frac{\partial}{\partial t}\frac{\partial W}{\partial u}\right|_{t_p}\Delta t + O\left(\Delta t^2\right) \tag{11}$$

We've now switched to using partial derivatives since we've got 2 different variables.

For $W*$, the entire curve was constructed by moving $Wp$ along by a distance $\Delta t$ so we can write

$$W* = W|_{t_p} + \left.\frac{\partial W}{\partial t}\right|_{t_p}\Delta t + O\left(\Delta t^2\right) \tag{12}$$

$$\frac{\partial W*}{\partial u} = \left.\frac{\partial W}{\partial u}\right|_{t_p} + \left.\frac{\partial}{\partial u}\frac{\partial W}{\partial t}\right|_{t_p}\Delta t + O\left(\Delta t^2\right) \tag{13}$$

Therefore

$$\Delta W=\left.\frac{\partial}{\partial t}\frac{\partial W}{\partial u}\right|_{t_p}\Delta t-\left.\frac{\partial}{\partial u}\frac{\partial W}{\partial t}\right|_{t_p}\Delta t+O\left(\Delta t^2\right) \tag{14}$$

We can now define a derivative for $W$ by

$$\begin{aligned}\mathfrak{L}_v W &= \lim_{\Delta t\to 0}\frac{\Delta W}{\Delta t} &(15)\\ &= \frac{\partial}{\partial t}\frac{\partial W}{\partial u}-\frac{\partial}{\partial u}\frac{\partial W}{\partial t} &(16)\end{aligned}$$

You might think this is always zero, since in standard calculus for continuous functions, derivatives commute. However, these aren't ordinary derivatives; they are directional derivatives, and in this case the order does matter. As we saw when discussing the tangent space, the directional derivative has the form, in a specific coordinate system:

$$\frac{\partial}{\partial t}=X^a\partial_a \tag{17}$$

where $X^a=dx^a/dt$. We therefore get

$$\begin{aligned}\mathfrak{L}_v W &= X^a\partial_a\left(Y^b\partial_b W\right)-Y^a\partial_a\left(X^b\partial_b W\right) &(18)\\ X^a &= \frac{dx^a}{dt} &(19)\\ Y^a &= \frac{dx^a}{du} &(20)\end{aligned}$$

Expanding this we get

$$\begin{aligned}\mathfrak{L}_v W &= X^a\left(\partial_a Y^b\right)\left(\partial_b W\right)+X^aY^b\partial^2_{a,b}W-Y^a\left(\partial_a X^b\right)\left(\partial_b W\right)-Y^aX^b\partial^2_{a,b}W &(21)\\ &= \left[X^a\left(\partial_a Y^b\right)-Y^a\left(\partial_a X^b\right)\right]\partial_b W &(22)\end{aligned}$$

To convert this to the notation used at the beginning of this post, the vector field **v** is the one associated with the parameter $t$, so $\mathbf{v}=X$. Similarly, the vector field **w** is associated with parameter $u$, so $\mathbf{w}=Y$, and we get

$$\mathfrak{L}_v W=\left[(\mathbf{v}\cdot\nabla)\,\mathbf{w}-(\mathbf{w}\cdot\nabla)\,\mathbf{v}\right]^b\partial_b W \tag{23}$$

## PINGBACKS

Pingback: Lie derivative: higher-rank tensors
Pingback: Covariant derivative and connections