Lecture VI: Tensor calculus in flat spacetime

(Dated: September 6, 2019)

I. OVERVIEW

This lecture covers tensor calculus, thus finishing up the material in Chapter 3. Some of the material on integrals is not in the book. We will do derivatives first, then integrals, and finally the relation between the two (the Fundamental Theorem of Calculus). In D dimensions, there are D variants of the Fundamental Theorem of Calculus; in 3 dimensions, for example, there was the gradient theorem, Stokes's theorem, and Gauss's divergence theorem.

II. DERIVATIVES

A. Gradients of scalars

You know in standard multivariable calculus, if one moves at velocity $\boldsymbol{v} = \mathrm{d}\boldsymbol{x}/\mathrm{d}\sigma$ (where σ is a parameter – I don't want to assume it is time!), then the scalar function $f(\boldsymbol{x})$ has derivative

$$\frac{\mathrm{d}f(\boldsymbol{x}(\sigma))}{\mathrm{d}\sigma} = \frac{\partial f}{\partial x^\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\sigma} = \frac{\partial f}{\partial x^\alpha}v^\alpha. \tag{1}$$

This same construction exists in special relativity. We will use the shorthand that a partial derivative is denoted with a comma:

$$f_{,\alpha} \equiv \partial_\alpha f \equiv \frac{\partial f}{\partial x^\alpha}, \tag{2}$$

and the chain rule guarantees that it transforms as a 1-form:

$$f_{,\bar\alpha} = \frac{\partial f}{\partial x^{\bar\alpha}} = \frac{\partial x^\beta}{\partial x^{\bar\alpha}}\frac{\partial f}{\partial x^\beta} = [\boldsymbol{\Lambda}^{-1}]^\beta{}_{\bar\alpha} f_{,\beta}. \tag{3}$$

This 1-form is called the *gradient* of f and is often denoted $\boldsymbol{\nabla} f$. The derivative along a path gets a special symbol:

$$\boldsymbol{\nabla}_{\boldsymbol{v}} f \equiv \frac{\mathrm{d}f(\boldsymbol{x}(\sigma))}{\mathrm{d}\sigma} = f_{,\alpha}v^\alpha. \tag{4}$$

B. Gradients of tensors

In flat spacetime, vectors or tensors are not tied to any position in spacetime: a vector $\boldsymbol{a}$ at event $\mathcal{P}$ is in the same vector space as a vector $\boldsymbol{b}$ at event $\mathcal{Q}$. Therefore, everything that we did above for a scalar f applies to a general vector or tensor. For example, we may write

$$\boldsymbol{\nabla}_{\boldsymbol{v}}\mathbf{T} = \frac{\mathrm{d}\mathbf{T}(\boldsymbol{x}(\sigma))}{\mathrm{d}\sigma}, \tag{5}$$

and the gradient of a rank $\binom{m}{n}$ tensor $\mathbf{T}$ is the rank $\binom{m}{n+1}$ tensor with components

$$T^{\beta_1...\beta_m}{}_{\nu_1...\nu_n,\alpha} = \frac{\partial T^{\beta_1...\beta_m}{}_{\nu_1...\nu_n}}{\partial x^\alpha}. \tag{6}$$

This tensor $\boldsymbol{\nabla}\mathbf{T}$ is called the *gradient* of $\mathbf{T}$.

When we go to curved spacetime, vectors and tensors at different events in spacetime are no longer directly comparable (they are not in the same vector space) – think of drawing a vector on a map of Columbus and a map of Beijing, and ask how you could possibly subtract them. Thus the construction for scalars

$$\frac{\mathrm{d}f(\boldsymbol{x}(\sigma))}{\mathrm{d}\sigma} = \lim_{\epsilon\to 0}\frac{f(\boldsymbol{x}(\sigma+\epsilon)) - f(\boldsymbol{x}(\sigma))}{\epsilon} \tag{7}$$

will have to be modified when we go to curved spacetime. In fact, "curvature" as it appears in the Einstein equations is defined in terms of how one transforms vectors between different points. But that isn't our challenge yet.

C. Exterior derivatives

One special type of derivative will turn out **not** to need modification when we go to curved spacetime (we will see why later). That is the fully antisymmetric derivative of a k-form. For a k-form $\boldsymbol{w}$, we define its *exterior derivative* $\tilde{\boldsymbol{d}}\boldsymbol{w}$ to be the $k+1$-form

$$[\tilde{\boldsymbol{d}}\boldsymbol{w}]_{\alpha_1\alpha_2...\alpha_{k+1}} = (k+1)\partial_{[\alpha_1} w_{\alpha_2...\alpha_{k+1}]}. \tag{8}$$

(In a formal sense, you could write this as $\tilde{\boldsymbol{d}}\boldsymbol{w} = \boldsymbol{\nabla} \wedge \boldsymbol{w}$.) For example, if $\boldsymbol{w}$ is a 1-form, then

$$[\tilde{\boldsymbol{d}}\boldsymbol{w}]_{\alpha\beta} = 2\partial_{[\alpha} w_{\beta]} = \partial_\alpha w_\beta - \partial_\beta w_\alpha. \tag{9}$$

Note that because partial derivatives are symmetric – $\partial_\mu\partial_\nu = \partial_\nu\partial_\mu$, and so $\partial_{[\mu}\partial_{\nu]} = 0$ – the second exterior derivative of a form is always zero:

$$[\tilde{\boldsymbol{d}}\tilde{\boldsymbol{d}}\boldsymbol{w}]_{\alpha_1\alpha_2...\alpha_{k+2}} = (k+2)(k+1)\partial_{[\alpha_1}\partial_{\alpha_2} w_{\alpha_3...\alpha_{k+2}]} = 0. \tag{10}$$

D. Divergences and curls

In electrodynamics, you learned that the divergence and curl were important vector operations for describing electric and magnetic fields. You might want to know how these extend to 4 dimensions. The divergence is most straightforward: if a vector has components v^α, then you know that

$$\boldsymbol{\nabla} \cdot \mathbf{v} = \frac{\partial v^\alpha}{\partial x^\alpha} = v^\alpha{}_{,\alpha}. \tag{11}$$

(At this point, I will skip the straightforward but in principle necessary exercise of proving this transforms as a scalar.) That is, "divergence" is a gradient followed by a contraction. In general, taking a divergence reduces the rank of a tensor by 1: it turns a vector into a scalar, and it can turn a rank 2 tensor into a vector, e.g., $T^{\alpha\beta}{}_{,\beta}$ has 1 index.

The notion of a curl is a bit different. You know that in 3 Euclidean dimensions, the curl of a vector $\boldsymbol{v}$ could be formally written as $\boldsymbol{\nabla} \times \boldsymbol{v}$. In Lecture V, we wrote the cross product as the dual of the wedge product. The same is true here: in 3 dimensions,

$$\boldsymbol{\nabla} \times \boldsymbol{v} = {}^\star(\tilde{\boldsymbol{d}}\boldsymbol{v}) \quad \text{or} \quad [\boldsymbol{\nabla} \times \boldsymbol{v}]^i = \varepsilon^{ijk}\partial_j v_k. \tag{12}$$

In a general number of dimensions, then, one replaces the concept of the curl with the exterior derivative. The exterior derivative of a vector (converted to a 1-form by index lowering) is a 2-form (rank 2 antisymmetric tensor). Only in 3 dimensions does this correspond to a vector through the notion of a dual.

It is less obvious, but the notion of a dual also allows us to write the divergence in terms of an exterior derivative. In D dimensions, one could write

$$[\tilde{\boldsymbol{d}}({}^\star\boldsymbol{v})]_{\alpha_1...\alpha_D} = D\partial_{[\alpha_1}(\varepsilon_{\alpha_2...\alpha_D]\mu} v^\mu). \tag{13}$$

Now since this is a fully antisymmetric tensor, let's consider the $1,2,...D$ component of this (or the 0123 component in special relativity); the remaining components are then determined by antisymmetrization and the fact that if there are repeated indices then the component is zero. This is

$$[\tilde{\boldsymbol{d}}({}^\star\boldsymbol{v})]_{1...D} = D\partial_{[1}(\varepsilon_{2...D]\mu} v^\mu) = D\frac{1}{D!}\sum_{\mu=1}^{D}\sum_{\text{perms}\,\pi} \pm_\pi \varepsilon_{\pi_2...\pi_D\mu}\partial_{\pi_1} v^\mu, \tag{14}$$

where the sum is over permutations of $1...D$ and the $\pm$ indicates even (+) or odd (−), and I have explicitly written all sums (it helps at this stage). Since $\boldsymbol{\varepsilon}$ is fully antisymmetric, the only non-zero components of $\boldsymbol{\varepsilon}$ are those where $\mu = \pi_1$. Thus:

$$[\tilde{\boldsymbol{d}}({}^\star\boldsymbol{v})]_{1...D} = \frac{1}{(D-1)!}\sum_{\mu}\sum_{\text{perms}\,\pi\text{ with }\pi_1=\mu} \pm_\pi \varepsilon_{\pi_2...\pi_D\pi_1}\partial_\mu v^\mu. \tag{15}$$

Now $\pi_2...\pi_D\pi_1$ is a permutation of $\pi_1\pi_2...\pi_D$; this permutation is odd (requires an odd number of swaps) if D is even and is even (requires an even number of swaps) if D is odd. Thus

$$\varepsilon_{\pi_2...\pi_D\pi_1} = (-1)^{D-1}\varepsilon_{\pi_1\pi_2...\pi_D} = \pm_\pi(-1)^{D-1}. \tag{16}$$

Substituting this in, and noting that there are $(D-1)!$ permutations with $\pi_1 = \mu$, we find

$$[\tilde{\boldsymbol{d}}(^\star\boldsymbol{v})]_{1\ldots D} = (-1)^{D-1}\sum_\mu \partial_\mu v^\mu = (-1)^{D-1}\partial_\mu v^\mu \quad (\text{E}\Sigma\text{C turned back on}). \tag{17}$$

One concludes that

$$\tilde{\boldsymbol{d}}(^\star\boldsymbol{v}) = (-1)^{D-1}(\boldsymbol{\nabla}\cdot\boldsymbol{v})\boldsymbol{\varepsilon}. \tag{18}$$

Thus the 3 fundamental operations of 3-dimensional vector calculus – the gradient, the curl, and the divergence – are actually exterior derivatives, taking us from 0-forms (scalars) to 1-forms, 1-forms to 2-forms, and 2-forms to 3-forms. Because of the dual correspondence in 3 dimensions, we can think of these as going from scalars to vectors to vectors and back to scalars. The rule that second exterior derivatives are zero corresponds to what you learned in undergraduate mathematical methods – that

$$\operatorname{curl}\operatorname{grad} f = 0 \quad \text{and} \quad \operatorname{div}\operatorname{curl}\boldsymbol{v} = 0. \tag{19}$$

In 4 dimensions, there are of course 4 such operations, but aside from grad and div (which exist in any number of dimensions) we will write these either with exterior derivatives or (more commonly for computational purposes) with the Levi-Civita symbol.

III. INTEGRALS

In 3D vector calculus, you probably did three types of integrals – integrals over 1D curves ("line integrals"), over 2D surfaces, or over 3D volumes. The same concepts exist in any number of dimensions.

The line integral really is the same concept in any number of dimensions: if we have a 1-form $\tilde{\boldsymbol{k}}$, then we can integrate over a path:

$$\int_{\mathcal{P}}^{\mathcal{Q}} \tilde{\boldsymbol{k}}\cdot \mathrm{d}\boldsymbol{x} = \int_{\mathcal{P}}^{\mathcal{Q}} k_\mu\,\mathrm{d}x^\mu = \int_{\mathcal{P}}^{\mathcal{Q}} k_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\,\mathrm{d}\sigma. \tag{20}$$

Surface integrals are trickier: in vector calculus, you took a surface $\mathcal{S}$ and wrote the flux of a vector field $\boldsymbol{B}$ as

$$\int_{\mathcal{S}} \boldsymbol{B}\cdot\mathrm{d}\boldsymbol{a}, \tag{21}$$

where $\mathrm{d}\boldsymbol{a}$ is a surface area element, written as a vector perpendicular to the surface. Unfortunately the concept of "vector perpendicular to a 2D surface" is unique to 3 dimensions. What we can do is think about area as an antisymmetric tensor: the area of a small bit of surface has some component projected into the xy, xz, or yz-planes (and, if area carries a sign, $a_{yx} = -a_{xy}$). To be rigorous about this, if we have a parallelogram defined by two vectors along its sides $\Delta\boldsymbol{u}$ and $\Delta\boldsymbol{v}$, we can define an area element:

$$\Delta\boldsymbol{a} = \Delta\boldsymbol{u}\wedge\Delta\boldsymbol{v} \quad \text{or} \quad \Delta a^{\mu\nu} = \Delta u^\mu\Delta v^\nu - \Delta u^\nu \Delta v^\mu. \tag{22}$$

In 3 dimensions, one could think about the vector $^\star\Delta\boldsymbol{a}$ as easily as $\Delta\boldsymbol{a}$, but from a more fundamental point of view, area is an antisymmetric rank 2 tensor. This point of view suggests that a 2-form field $\boldsymbol{F}$ could have a flux integral defined by a Riemann sum:

$$\int_{\mathcal{S}} \boldsymbol{F} = \lim \sum \frac{1}{2} F_{\mu\nu}\,\Delta a^{\mu\nu}, \tag{23}$$

where the $\frac{1}{2}$ removes the double-counting of $F_{xy}\,\Delta a^{xy}$ and $F_{yx}\,\Delta a^{yx}$ (which in 3D vector calculus you would write as $F_z\,\Delta a_z$). The flux integral of Eq. (21) could thus be written as $\int_{\mathcal{S}} {}^\star\boldsymbol{B}$. However, the formulation of Eq. (23) generalizes to arbitrary dimensions, and – if $\boldsymbol{F}$ has all lower indices – we don't even need to do any raising or lowering.

Volume integrals can be treated in much the same way, except now there is a volume element that is a 3D antisymmetric tensor: a parallelepiped with sides $\Delta\boldsymbol{u}$, $\Delta\boldsymbol{v}$, and $\Delta\boldsymbol{w}$ has volume

$$\Delta\boldsymbol{V} = \Delta\boldsymbol{u}\wedge\Delta\boldsymbol{v}\wedge\Delta\boldsymbol{w} \tag{24}$$

or

$$\Delta V^{\alpha\beta\gamma} = \Delta u^\alpha \Delta v^\beta \Delta w^\gamma + \Delta u^\beta \Delta v^\gamma \Delta w^\alpha + \Delta u^\gamma \Delta v^\alpha \Delta w^\beta - \Delta u^\alpha \Delta v^\gamma \Delta w^\beta - \Delta u^\beta \Delta v^\alpha \Delta w^\gamma - \Delta u^\gamma \Delta v^\beta \Delta w^\alpha. \quad (25)$$

We can define a similar Riemann sum to Eq. (23),

$$\int_{\mathcal{V}} \boldsymbol{H} = \lim \sum \frac{1}{3!} H_{\alpha\beta\gamma}\, \Delta V^{\alpha\beta\gamma}, \quad (26)$$

and then the traditional volume integral $\int_{\mathcal{V}} \rho\, d^3\boldsymbol{x}$ of multivariable calculus can be written as $\int_{\mathcal{V}} {}^\star\rho = \int_{\mathcal{V}} \rho\boldsymbol{\varepsilon}$.

In general, a k-form can be integrated over a k-dimensional surface, where $k \leq D$.

IV. THE FUNDAMENTAL THEOREM OF CALCULUS

We have seen that in D dimensions, there are D types of exterior derivatives (that turn k-forms into $k+1$-forms, for $k = 0 \ldots D-1$) and D types of integrals (that integrate k-forms over k-dimensional surfaces, for $k = 1 \ldots D$). You further know that in 3 dimensions, there are 3 vector calculus theorems:

- the *gradient theorem*, relating the line integral of grad f to the values of f on the endpoints of the curve;
- *Stokes's theorem*, relating the surface integral of curl $\boldsymbol{v}$ to the line integral of $\boldsymbol{v}$ around the boundary; and
- *Gauss's divergence theorem*, relating the volume integral of $\boldsymbol{\nabla}\cdot\boldsymbol{v}$ to the surface integral of $\boldsymbol{v}$ around the enclosing surface.

This concept generalizes to k-forms in D dimensions, and is known as the generalized Stokes's theorem.

I won't prove the generalized Stokes's theorem in full mathematical rigor, but it is instructive to do it when the region is a very small hypercube. (You could build a proof by building a general region out of many such hypercubes.) Let's consider, for definiteness, a $k = 3$-dimensional hypercube $\mathcal{H}$ of side length h on coordinate axes $1 \ldots 3$ with its lower-left corner at position $\boldsymbol{X}$. Let's imagine that $\boldsymbol{\alpha}$ is a 2-form, and $\boldsymbol{\beta} = \tilde{\boldsymbol{d}}\boldsymbol{\alpha}$ is its exterior derivative. Then, writing the integral as a Riemann sum with one term,

$$\int_{\mathcal{H}} \boldsymbol{\beta} = \frac{1}{3!}\beta_{\mu_1\mu_2\mu_3}\, \Delta V^{\mu_1\mu_2\mu_3} = \beta_{\mu_1\mu_2\mu_3}\, (h\boldsymbol{e}_1)^{[\mu_1}(h\boldsymbol{e}_2)^{\mu_2}(h\boldsymbol{e}_3)^{\mu_3]} = h^3\beta_{123}. \quad (27)$$

On the other hand, each component of $\boldsymbol{\beta}$ is itself a derivative:

$$\beta_{123} = 3\partial_{[1}\alpha_{23]} = \partial_1\alpha_{23} + \partial_2\alpha_{31} + \partial_3\alpha_{12}. \quad (28)$$

If we write $h\partial_1\alpha_{23} = \alpha_{23}(X^1+h, X^2, X^3) - \alpha_{23}(X^1, X^2, X^3)$, and use antisymmetry of $\boldsymbol{\alpha}$, then we find

$$\begin{aligned} \int_{\mathcal{H}} \boldsymbol{\beta} &= h^2\alpha_{23}(X^1+h, X^2, X^3) + h^2\alpha_{32}(X^1, X^2, X^3) + h^2\alpha_{31}(X^1, X^2+h, X^3) \\ &\quad + h^2\alpha_{13}(X^1, X^2, X^3) + h^2\alpha_{12}(X^1+h, X^2, X^3+h) + h^2\alpha_{21}(X^1, X^2, X^3) \\ &= \oint_{\partial\mathcal{H}} \boldsymbol{\alpha}, \end{aligned} \quad (29)$$

where in the last line we used the fact that the 6 faces of the cube each contribute to the closed-surface integral. The same type of argument works for any value of k; in general,

$$\int_{\mathcal{H}} \tilde{\boldsymbol{d}}\boldsymbol{\alpha} = \oint_{\partial\mathcal{H}} \boldsymbol{\alpha}, \quad (30)$$

where $\boldsymbol{\alpha}$ is any $k-1$-form, $\mathcal{H}$ is any k-dimensional region, and $\partial\mathcal{H}$ is its boundary.

The same argument works if $k = 1$, and if we formally define the "boundary" of a curve to be its two endpoints, with opposite polarity (the starting point gets a $-$ sign, just as an area element can have a negative component).

We can now highlight a few special cases:

- For $k = 1$: if $\boldsymbol{\alpha}$ is a scalar f, and $\mathcal{H} = \mathcal{L}$ is a line from point $\mathcal{P}$ to $\mathcal{Q}$, then the integral $\oint_{\partial\mathcal{L}} \boldsymbol{\alpha}$ is a trivial sum: it is $f(\mathcal{Q}) - f(\mathcal{P})$. Thus:

$$\int_{\mathcal{L}} \tilde{\boldsymbol{d}}f = \int_{\mathcal{P}}^{\mathcal{Q}} f_\alpha\, \mathrm{d}x^\alpha = f(\mathcal{Q}) - f(\mathcal{P}). \quad (31)$$

- For $k = D$: if $\boldsymbol{\alpha} = {}^{\star}\boldsymbol{v}$ for some vector field $\boldsymbol{v}$, and $\mathcal{H}$ is a D-dimensional volume $\mathcal{V}$, then

$$\int_{\mathcal{V}} (\boldsymbol{\nabla} \cdot \boldsymbol{v})\boldsymbol{\varepsilon} = \int_{\mathcal{V}} \tilde{\boldsymbol{d}}{}^{\star}\boldsymbol{v} = \oint_{\partial\mathcal{V}} {}^{\star}\boldsymbol{v}; \tag{32}$$

  the left-hand side is the volume integral of $\boldsymbol{\nabla} \cdot \boldsymbol{v}$, and the right-hand side is the $D-1$-dimensional surface integral of ${}^{\star}\boldsymbol{v}$ (just like a flux integral, but written with the dual).

- For $k = 2$ and $D = 3$: now if we take a 2-dimensional surface $\mathcal{S}$ and vector $\boldsymbol{v}$, we get

$$\int_{\mathcal{S}} {}^{\star}(\boldsymbol{\nabla} \times \boldsymbol{v}) = \int_{\mathcal{S}} {}^{\star\star}\tilde{\boldsymbol{d}}\boldsymbol{v} = \oint_{\partial\mathcal{S}} \boldsymbol{v}. \tag{33}$$

  This is the usual Stokes's theorem: the left-hand side is the surface integral of curl $\boldsymbol{v}$, and the right-hand side is the line integral of $\boldsymbol{v}$ around the boundary of $\mathcal{S}$.

Again, the generalized Stokes's theorem (Eq. 30), if written with $\boldsymbol{\alpha}$ as a form with down indices, does not raise or lower indices and hence is valid in curved spacetime as well as flat spacetime. (There **are** changes to this theorem if you integrate over regions with non-trivial topology, however we won't deal with that in this class.)