Physics 464/511 Lecture Q Fall, 2016

1 Parallel Transport

Consider a manifold with a vector or a 1-form defined at each point. Such an object might be a physical field, which would have field equations involving derivatives of this vector quantity. How can we tell on a curved manifold whether a vector $\mathbf{V} = V^\mu(x_A)\partial_\mu$ at the point A is the same or different from $\mathbf{V}$ at B, $V^\mu(x_B)\partial_\mu$? The naïve thing would be if $V^\mu(x_A) = V^\mu(x_B)$. But the V^μ's are chart dependent and such a statement of equality of components at different points can be true for one chart and not another, and has no real meaning for the manifold. It is also the *wrong* requirement even for the simple example of a two dimensional Euclidean space in polar coordinates, for a vector in the ρ direction for $\phi = 0$ is completely different from one of the same magnitude in the ρ direction at $\phi = \pi/2$.

This is true even for A and B very near each other. By definition, all charts agree on whether two directions differ by a finite angle as $A \to B$, but not on the rate. Thus $\lim_{x_B \to x_A} V^\mu(x_B)$ is well defined, but not $\partial_\nu V^\mu(x_B)$, in the sense that it is chart-dependent.

Thus an arbitrary manifold has no means of comparing vectors at different points, unless there is an extra structure placed on the manifold telling how to move a coordinate system from point A to a nearby point B.

Let the equivalence principle help us out, by giving us a chart[1] $C' = \{\xi^\alpha\}$ of a neighborhood of the event A which is cartesian and inertial at A. Then $g'_{\alpha\beta}(\mathbf{x}_A) = \eta_{\alpha\beta}$ and $g'_{\alpha\beta,\gamma}(\mathbf{x}_A) = 0$. A vector $\mathbf{V} = V^\mu\partial_\mu = V'^\alpha\partial'_\alpha$ defined at A is parallel transported an infinitesimal distance from A by holding its coordinates fixed, because that's how one parallel transports in flat space cartesian coordinates. Thus if we have a vector field $\mathbf{V}$ and we ask what the "physical change" in $\mathbf{V}$ is along the ∂'_α direction, it is $\mathbf{V}(B) - \mathbf{V}(A)_{\text{transported}} \approx \Delta\xi^\alpha \left(\partial'_\alpha V'^\beta\right)\partial'_\beta$. Let us define that to be $\Delta\xi^\alpha$ times the **covariant derivative** $D_{\alpha'}$ in the $\Delta\xi^\alpha$ direction. With Δx^μ the corresponding change in chart C's coordinates, we have

$$\begin{aligned}
\Delta x^\mu D_\mu \mathbf{V} &= \Delta\xi^\alpha \left(\partial'_\alpha V'^\beta\right)\partial'_\beta \\
&= \Delta x^\mu \frac{\partial \xi^\alpha}{\partial x^\mu}\partial'_\alpha \left(V^\nu \frac{\partial \xi^\beta}{\partial x^\nu}\right)\frac{\partial x^\rho}{\partial \xi^\beta}\partial_\rho \\
\text{or}\quad D_\mu \mathbf{V} &= \partial_\mu \left(V^\nu \frac{\partial \xi^\beta}{\partial x^\nu}\right)\frac{\partial x^\rho}{\partial \xi^\beta}\partial_\rho \\
&= \left[\partial_\mu V^\rho + \underbrace{\frac{\partial^2 \xi^\beta}{\partial x^\mu \partial x^\nu}\frac{\partial x^\rho}{\partial \xi^\beta}}_{\Gamma^\rho{}_{\nu\mu}} V^\nu\right]\partial_\rho
\end{aligned}$$

Thus the components of the vector $D_\mu \mathbf{V} = (D_\mu v)^\rho\, \partial_\rho$ are

$$(D_\mu V)^\rho = \partial_\mu V^\rho + \Gamma^\rho{}_{\nu\mu} V^\nu.$$

Again, D_μ is known as the **covariant dervivative**.

The chart C was an arbitrary chart. In some other chart with coordinates x'^μ, would we have $(D_\mu \mathbf{V})^\rho$ behave like a suitable tensor, with one covariant and one contravariant index? That is, is

$$\left(D'_\mu \mathbf{V}\right)^{\rho'} \stackrel{?}{=} \frac{\partial x'^\rho}{\partial x^\sigma}\frac{\partial x^\nu}{\partial x'^\mu}\left(D_\nu \mathbf{V}\right)^\sigma$$

The left hand side is

$$\begin{aligned}
\left(D_{\mu'} \mathbf{V}\right)^{\rho'} &= \partial'_\mu V'^\rho + \Gamma'^\rho{}_{\mu\kappa} V'^\kappa \\
&= \underbrace{\frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu\left(\frac{\partial x'^\rho}{\partial x^\sigma}V^\sigma\right)}_{\frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial x'^\rho}{\partial x^\sigma}\partial_\nu V^\sigma + \left(\frac{\partial}{\partial x'^\mu}\frac{\partial x'^\rho}{\partial x^\lambda}\right)V^\lambda} + \Gamma'^\rho{}_{\mu\kappa}\frac{\partial x'^\kappa}{\partial x^\lambda}V^\lambda
\end{aligned}$$

The right hand side is

$$\frac{\partial x'^\rho}{\partial x^\sigma}\frac{\partial x^\nu}{\partial x'^\mu}\left(\partial_\nu V^\sigma + \Gamma^\sigma{}_{\nu\lambda}V^\lambda\right).$$

We find covariance if it is true that

$$\Gamma'^\rho{}_{\mu\kappa}\frac{\partial x'^\kappa}{\partial x^\lambda} \stackrel{?}{=} \frac{\partial x'^\rho}{\partial x^\sigma}\frac{\partial x^\nu}{\partial x'^\mu}\Gamma^\sigma{}_{\nu\lambda} - \frac{\partial}{\partial x'^\mu}\frac{\partial x'^\rho}{\partial x^\lambda}.$$

[1] Some awkward notation here: the coordinates of C' will be called ξ^α without a prime, but the derivative will be $\partial'_\alpha = \partial/\partial\xi^\alpha$. The indices μ, ν, ρ, σ will refer to the coordinates x^μ of the chart C, while indices α, β will refer to the chart C'.

or

$$\Gamma'^{\rho}{}_{\mu\kappa} \stackrel{?}{=} \frac{\partial x'^{\rho}}{\partial x^{\sigma}}\frac{\partial x^{\nu}}{\partial x'^{\mu}}\frac{\partial x^{\lambda}}{\partial x'^{\kappa}}\Gamma^{\sigma}{}_{\nu\lambda} - \frac{\partial x^{\lambda}}{\partial x'^{\kappa}}\frac{\partial}{\partial x'^{\mu}}\frac{\partial x'^{\rho}}{\partial x^{\lambda}}.$$

$$\text{Note} \qquad \frac{\partial}{\partial x'^{\mu}}\left(\frac{\partial x^{\lambda}}{\partial x'^{\kappa}}\frac{\partial x'^{\rho}}{\partial x^{\lambda}}\right) = \frac{\partial}{\partial x'^{\mu}}\delta^{\rho}_{\kappa} = 0$$
$$= \frac{\partial^2 x^{\lambda}}{\partial x'^{\mu}\partial x'^{\kappa}}\frac{\partial x'^{\rho}}{\partial x^{\lambda}} + \frac{\partial x^{\lambda}}{\partial x'^{\rho}}\frac{\partial}{\partial x'^{\mu}}\frac{\partial x'^{\rho}}{\partial x^{\lambda}}$$

The second term matches our above questionable identity, which becomes

$$\Gamma'^{\rho}{}_{\mu\kappa} \stackrel{?}{=} \frac{\partial x'^{\rho}}{\partial x^{\sigma}}\frac{\partial x^{\nu}}{\partial x'^{\mu}}\frac{\partial x^{\lambda}}{\partial x'^{\kappa}}\Gamma^{\sigma}{}_{\nu\lambda} + \frac{\partial x'^{\kappa}}{\partial x^{\lambda}}\frac{\partial^2 x^{\lambda}}{\partial x'^{\mu}\partial x'^{\kappa}}.$$

But this is true, as verified from

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{\partial x^{\lambda}}{\partial \xi^{\alpha}}\frac{\partial^2 \xi^{\alpha}}{\partial x^{\mu}\partial x^{\nu}}$$

which is Eq. 1 from Lecture P.

We have gone to great lengths to define the covariant derivative of a vector, which is nontrivial because the basis vectors may change from point to point. There were no such difficulties for a scalar, as the scalar did not require basis vectors. Thus $D_\mu f = \partial_\mu f$. For forms we must again worry about a basis, but we can take a shortcut if we use Leibniz product rule, $D_\mu(AB) = D_\mu(A)B + AD_\mu(B)$ which must hold for any derivative (and in particular it holds for $\partial/\partial\xi^\alpha$). Let $\mathbf{A} = A_\mu \mathbf{d}x^\mu$ be a 1-form which we wish to covariantly differentiate. With $\mathbf{V}$ and arbitrary vector,

$$\begin{aligned} D_\mu\langle\mathbf{A}|\mathbf{V}\rangle = \partial_\mu\langle\mathbf{A}|\mathbf{V}\rangle &= \langle D_\mu\mathbf{A}|\mathbf{V}\rangle + \langle\mathbf{A}|D_\mu\mathbf{V}\rangle \\ = \partial_\mu\left(A_\nu V^\nu\right) &= = \left(D_\mu\mathbf{A}\right)_\nu V^\nu + A_\rho\left(\partial_\mu V^\rho + \Gamma^{\rho}{}_{\nu\mu}V^\nu\right). \end{aligned}$$

Thus

$$\left(D_\mu\mathbf{A}\right)_\nu = \partial_\mu A_\nu - \Gamma^{\rho}{}_{\nu\mu}A_\rho.$$

The rules for an arbitrary tensor can be found by considering tensor products of vectors and 1-forms. We find

$$\begin{aligned} \left(D_\mu T\right)^{\nu_1\ldots\nu_r}{}_{\rho_1\ldots\rho_s} &= \partial_\mu T^{\nu_1\ldots\nu_r}{}_{\rho_1\ldots\rho_s} + \sum_{i=1}^{r}\Gamma^{\nu_i}{}_{\alpha\mu}T^{\nu_1\ldots\nu_{i-1}\alpha\nu_{i+1}\ldots\nu_r}{}_{\rho_1\ldots\rho_s} \\ &\quad - \sum_{i=1}^{s}\Gamma^{\alpha}{}_{\rho_i\mu}T^{\nu_1\ldots\nu_r}{}_{\rho_1\ldots\rho_{i-1}\alpha\rho_{i+1}\ldots\rho_s}. \end{aligned}$$

The relationship between forms and vectors we just preserved in our definition of D on a form has nothing to do with the metric. But another connection we would like to have is that parallel transport of a pair of vectors should not change their inner product $\mathbf{g}(\mathbf{u},\mathbf{v})$. Thus $\partial_\mu\mathbf{g}(\mathbf{u},\mathbf{v}) = D_\mu\mathbf{g}(\mathbf{u},\mathbf{v}) = 0$ if $D_\mu\mathbf{u} = 0$ and $D_\mu\mathbf{v} = 0$. But

$$D_\mu\mathbf{g}(\mathbf{u},\mathbf{v}) = \left(\mathbf{D}_\mu\mathbf{g}\right)(u,v) + \mathbf{g}\left(\mathbf{D}_\mu\mathbf{u},v\right) + \mathbf{g}\left(\mathbf{u},\mathbf{D}_\mu\mathbf{v}\right),$$

and the last two terms are zero, so we must have

$$D_\mu\mathbf{g} = 0.$$

To check this, evaluate

$$\left(D_\mu\mathbf{g}\right)_{\rho\sigma} = g_{\rho\sigma,\mu} - \Gamma^{\lambda}{}_{\rho\mu}g_{\lambda\sigma} - \Gamma^{\lambda}{}_{\sigma\mu}g_{\lambda\rho}$$

by our general relation for a tensor, so

$$\left(D_\mu\mathbf{g}\right)_{\rho\sigma} = g_{\rho\sigma,\mu} - \Gamma_{\sigma\rho\mu} - \Gamma_{\rho\sigma\mu} = 0,$$

which is true (see Lecture P p. 7), so all is well.

Note: As $\mathbf{D}g = 0$, $\mathbf{D}$ commutes with raising and lowering indices! That is important, *e.g.*

$$g^{\mu\nu}\Big(\mathbf{D}_\rho \underbrace{\mathbf{A}}_{1\text{–form}}\Big)_\nu = \Big(\mathbf{D}_\rho \underbrace{\mathbf{A}}_{\text{vector}}\Big)^\mu.$$

Our definition of covariant derivative assumed the vector or scalar or 1-form was a field defined in the neighborhood of the event. Sometimes there are quantities only defined on, for example, a path. The velocity of a particle as it moves along its world-line is an example. $\mathbf{u}$ is simply not defined except along the path, and neither is, say, the spin of the particle $\mathbf{S}$. But we can define a covariant derivative along the path as it would be were $\mathbf{S}$ defined everywhere

$$\left(\frac{D}{D\lambda}\mathbf{S}\right)^\nu = \text{“}\left(\frac{dx^\mu}{d\lambda}D_\mu S\right)^\nu\text{”} = \text{“}\frac{dx^\mu}{d\lambda}\left(\frac{\partial S^\nu}{\partial x^\mu} + \Gamma^{\nu}{}_{\rho\mu}S^\rho\right)\text{”} = \frac{dS^\nu}{d\lambda} + \Gamma^{\nu}{}_{\rho\mu}S^\rho\frac{dx^\mu}{d\lambda}.$$

The last expression is well-defned entirely along the path of the particle, even though the expressions in quotes are not.

Recall from the end of Lecture D that $\vec{\nabla} f \sim \mathbf{d}f$ without requiring any knowledge of $\mathbf{g}$. Similarly $\vec{\nabla}\times\vec{A} \sim \mathbf{dA}$ doesn't depend on $\mathbf{g}$ or Γ. But $\vec{\nabla}\cdot\vec{A} = {}^*\mathbf{d}\,{}^*\mathbf{A}$, and the * requires the use of[2] $\varepsilon_{\mu\nu\rho\sigma} = \sqrt{g}\,\epsilon_{\mu\nu\rho\sigma}$, where it is $\epsilon_{\mu\nu\rho\sigma}$, not $\varepsilon_{\mu\nu\rho\sigma}$, which is a constant ($\pm 1$ or 0). Thus if $\mathbf{A} = A_\mu \mathbf{d}x^\mu$,

$$\begin{aligned} {}^*\mathbf{A} &= \frac{1}{3!}A^\mu\sqrt{g}\,\epsilon_{\mu\rho\sigma\kappa}\,\mathbf{d}x^\rho\wedge\mathbf{d}x^\sigma\wedge\mathbf{d}x^\kappa \\ \mathbf{d}\,{}^*\mathbf{A} &= \frac{1}{3!}\partial_\nu\left(A^\mu\sqrt{g}\right)\,\epsilon_{\mu\rho\sigma\kappa}\,\mathbf{d}x^\nu\wedge\mathbf{d}x^\rho\wedge\mathbf{d}x^\sigma\wedge\mathbf{d}x^\kappa \\ \text{and}\qquad {}^*\mathbf{d}\,{}^*\mathbf{A} &= \frac{1}{3!}\partial_\nu\left(A^\mu\sqrt{g}\right)g^{-1/2}\,\epsilon^{\nu\rho\sigma\kappa}\,\epsilon_{\mu\rho\sigma\kappa} = g^{-1/2}\partial_\mu\left(A^\mu\sqrt{g}\right) \end{aligned}$$

That is perhaps not what you expected ($\partial_\mu A^\mu$?). But it is the covariant derivative of the vector $\mathbf{A}$, contracted to form a divergence,

$$D_\mu A^\mu = \partial_\mu A^\mu + \Gamma^\mu{}_{\nu\mu}A^\nu,$$

$$\begin{aligned} \text{as}\qquad \Gamma^\mu{}_{\nu\mu} &= \frac{1}{2}g^{\mu\rho}\left(g_{\nu\rho,\mu}+g_{\mu\rho,\nu}-g_{\nu\mu,\rho}\right) = \frac{1}{2}g^{\mu\rho}g_{\mu\rho,\nu} = \frac{1}{2}\,\mathrm{Tr}\,G^{-1}\partial_\nu G \\ &= \frac{1}{2}\,\mathrm{Tr}\,\partial_\nu\,\ln G = \frac{1}{2}\,\partial_\nu\,\ln\det G = \partial_\nu\,\ln\left(g^{1/2}\right) = g^{-1/2}\partial_\nu\,g^{1/2} \end{aligned}$$

(where the matrix $G = g_{\cdot\cdot}$) .

$$\text{Thus}\qquad D_\mu A^\mu = g^{-1/2}\partial_\mu\left(A^\mu\sqrt{g}\right).$$

If we used D_μ for the divergence, why not for the curl? We did, but it made no difference,

$$\vec{\nabla}\times\vec{A} \sim D_\mu A_\nu - D_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu - \left(\underbrace{\Gamma^\rho{}_{\nu\mu} - \Gamma^\rho{}_{\mu\nu}}_{0}\right)A_\rho,$$

so the Γ dependence falls out of the antisymmetric part of the covariant derivative of a 1-form. Define the tensor product of two 1-forms and therein lies (if $\mathbf{A}$ is a 1-form)

$$\mathbf{d}x^\mu\otimes D_\mu\mathbf{A} = \mathbf{d}x^\mu\otimes\mathbf{d}x^\nu\left(\partial_\mu A_\nu - \Gamma^\rho{}_{\nu\mu}A_\rho\right).$$

[2]Recall from Lecture E that we defined $\epsilon_{\mu\nu\rho\sigma}$ to be the totally antisymmetric set of constants with $\epsilon_{0123} = 1$, but this was not covariant, and we also defined $\varepsilon_{\mu\nu\rho\sigma} = \sqrt{g}\,\epsilon_{\mu\nu\rho\sigma}$ which is covariant.

The antisymmetric part is just $\mathbf{dA}$, but the symmetric part is dependent on the connection coefficients. Similarly[3] $\mathbf{dF} \sim D_{[\mu}F_{\nu\rho]}$, and the Γ drops out.

Generalization to internal vector spaces:

Let us suppose we have a physical system involving a field ψ which takes values in an internal vector space, so that in some particular basis we have $\psi^a, a = 1,\cdots,N$. Let us also suppose that the physics is invariant under a group of transformations generated by $L^a{}_b$ of the basis, or under

$$\psi'^{\,a} = \left(e^{-i\theta L}\right)^a{}_b\,\psi^b \tag{1}$$

made independently at each spacetime point. Then if any derivatives are to enter the theory at all, there must be some additional structure. Let us assume a kind of equivalence principle: at any one point $\mathcal{P}$ of spacetime it is possible to find a set of bases $e_a(x)$ of the internal vector space such that, at $\mathcal{P}$, the physics is described by a Lagrangian $\mathcal{L}(\psi,\partial_\mu\psi)$ with no other fields (analogous to the laws of special relativity with no gravitational fields). Then in any other basis, the Lagrangian must be described by

$$\mathcal{L}(\psi', D'_\mu\psi')$$

where the relationship between the bases (1) also holds for

$$\begin{aligned} D'_\mu\psi' &= \left(\left(e^{-i\theta L}\right)^a{}_b\,\partial_\mu\psi^b = \left(e^{-i\theta L}\right)^a{}_b\,\partial_\mu\left(e^{i\theta L}\psi'\right)^b \right. \\ &= \partial_\mu\psi'^{\,a} + \left(e^{-i\theta L}\right)^a{}_b\,\partial_\mu\left(e^{i\theta L}\right)^b{}_c\,\psi'^{\,c}, \end{aligned}$$

or $D_\mu = \mathbb{I}\,\partial_\mu + e^{-i\theta L}\partial_\mu e^{i\theta L}$ is a matrix acting on the vector space of the ψ's. Define

$$A^a{}_{c\,\mu} = e^{-i\theta L}\,\partial_\mu\,e^{i\theta L} = \mathbf{A}_\mu.$$

Note that although $e^{i\theta L}$ connects two bases at the same point, the one for which the "inertial" frame has no A, and the other an arbitrary, general basis, the $\mathbf{A}$ refers only to the general basis, but in a sense at neighboring points. It defines parallel transport in the vector space of the ψ's.

[3]More neglected notation: a term with a bunch of lower indices enclosed in square brackets means to antisymmetrize in those indices, so $D_{[\mu}F_{\nu\rho]} := \frac{1}{3}\left(D_\mu F_{\nu\rho} - D_\nu F_{\mu\rho} - D_\rho F_{\nu\mu}\right)$ (where I already made use of the antisymmetry of $F_{\mu\nu}$). Enclosing in curly brackets means symmetrize ($F_{\{\mu,\nu\}} = 0$), and the same applies to a bunch of contravariant indices.