PHYS 480/581: General Relativity
Covariant Derivative

Prof. Cyr-Racine
(Dated: February 21, 2024)

I. PROBLEM WITH PARTIAL DERIVATIVES

[**Need for a tensorial derivative operator**] One issue that we have encountered so far is that partial derivatives of tensors in general spacetime are not tensors. This is a real problem: we want to write physical laws that are independent of any coordinate systems and only proper tensor equations can do that (non-tensor equations take different forms in different coordinate system). Since our physical laws are expressed as *differential* equations, having a derivative operator that provides bone fide tensors is very important.

[**Example with a constant vector field**] To illustrate the problem, take a vector field $\boldsymbol{A}$ that is constant across space and time. In flat cartesian coordinates, the components of this vector field are constant and $\partial_\alpha A^\mu = 0$, which makes sense since $\boldsymbol{A}$ is a constant vector field. But if I were to instead use curvilinear coordinates (like spherical coordinates), then the components of that vector fields would not be constant because they would have to compensate for the coordinate basis vectors changing direction and magnitude at different points in spacetime. For instance, imagine that my vector field is pointing purely in the x-direction, $\boldsymbol{A} = A^x\hat{\mathbf{e}}_x$ (no sum on x here), where $\hat{\mathbf{e}}_x$ is the unit vector in the cartesian x direction. In spherical coordinates, the components of this vector would be

$$\boldsymbol{A} = A^x\left(\sin\theta\cos\phi\hat{\mathbf{e}}_r + \cos\theta\cos\phi\hat{\mathbf{e}}_\theta - \sin\phi\hat{\mathbf{e}}_\phi\right), \tag{1}$$

which indeed vary with the spherical coordinates θ, ϕ, leading for instance to $\partial_\theta A^r \neq 0$ despite the vector $\boldsymbol{A}$ being constant across space and time. But this variation is artificial, it is caused by the basis vectors changing direction from point to point in this spacetime.

II. COVARIANT DERIVATIVE

To deal with such spurious coordinate artifacts, we would like to define a derivative operator that returns the absolute variation of a given tensor or vector, independent of which coordinate system we use. We shall call this operator the *covariant* derivative (or absolute gradient) $\boldsymbol{\nabla}$. For instance,

$$\boldsymbol{\nabla}\boldsymbol{A} = \boldsymbol{\nabla}(A^\mu\mathbf{e}_\mu) = \boldsymbol{\nabla}(A^\mu)\mathbf{e}_\mu + A^\mu\boldsymbol{\nabla}(\mathbf{e}_\mu), \tag{2}$$

where we have basically demanded that the covariant derivative obeys the product rule. Since A^μ is a component, which are scalar functions for a generic vector field, it is natural to demand that

$$\boldsymbol{\nabla}(A^\mu) = \partial_\gamma A^\mu\mathbf{e}^\gamma, \tag{3}$$

which are just the standard gradient of A^μ.

[**Impact on basis vectors**] Concerning the change to the basis vector themselves, we will assume that the basis vector change infinitesimally as we move a very small distance ϵ^μ in spacetime. For such small changes, we can write the rate of change of a given basis vector as a linear function of all basis vectors

$$\boldsymbol{\nabla}(\mathbf{e}_\mu) = (\partial_\gamma\mathbf{e}_\mu)\mathbf{e}^\gamma = (\Gamma^\nu_{\gamma\mu}\mathbf{e}_\nu)\mathbf{e}^\gamma. \tag{4}$$

We thus have

$$\boldsymbol{\nabla}\boldsymbol{A} = \left[\partial_\gamma A^\mu\mathbf{e}_\mu + A^\mu\Gamma^\nu_{\gamma\mu}\mathbf{e}_\nu\right]\mathbf{e}^\gamma = \left[\partial_\gamma A^\mu + \Gamma^\mu_{\gamma\nu}A^\nu\right]\mathbf{e}_\mu\mathbf{e}^\gamma. \tag{5}$$

The term proportional to Γ is the correction term accounting for the changes in the basis vectors as we take derivative along coordinate directions. $\Gamma^\nu_{\gamma\mu}$ is called a *connection coefficient* since it allows us to understand how basis vectors change as we move along certain direction in spacetime. **Specifically, $\Gamma^\nu_{\gamma\mu}$ describe the ν component of the change to the $\mathbf{e}_\mu$ basis vector as we move in the γ coordinate direction.** For spacetime that have a metric, the connection coefficient are derived from the metric and are usually referred to as the Christoffel symbols. Note that we use the word "symbol" here and not tensor since $\Gamma^{\nu}_{\gamma\mu}$ are not tensors. This is totally fine since the partial derivative of a vector is not a tensor either. But the *combination* of $\partial_\gamma A^\mu$ and $\Gamma^{\mu}_{\gamma\nu}A^\nu$ *is* a tensor since the non-tensorial part exactly cancels in this sum. In summary, in component notation, the covariant derivative of a vector $\boldsymbol{A}$ is

$$\nabla_\gamma A^\mu = \partial_\gamma A^\mu + \Gamma^{\mu}_{\gamma\nu}A^\nu. \tag{6}$$

[**Dual vectors**] What about dual vectors? We could repeat the exercise performed in Eq. (2) above for higher rank objects and realize that we would pick up extra terms, one for each of the coordinate basis (or dual basis) vectors need to write the higher order object in component form. For instance, for a dual vector $\boldsymbol{B}$, it's covariant derivative is

$$\nabla_\gamma B_\mu = \partial_\gamma B_\mu - \Gamma^{\nu}_{\gamma\mu}B_\nu. \tag{7}$$

[**Other tensors**] For higher order tensor, we basically pick up a Γ term for each index that the tensor components have, with its sign depending whether it is an upper or lower index. For instance, for a $(1,1)$ tensor we would get

$$\nabla_\gamma T^{\alpha}{}_{\beta} = \partial_\gamma T^{\alpha}{}_{\beta} + \Gamma^{\alpha}_{\gamma\nu}T^{\nu}{}_{\beta} - \Gamma^{\nu}_{\gamma\beta}T^{\alpha}{}_{\nu} \tag{8}$$

[**Properties of covariant derivatives**]:

- Torsion free: $\Gamma^{\gamma}_{\mu\nu} = \Gamma^{\gamma}_{\nu\mu}$.
- Metric compatibility: $\nabla_\rho g_{\mu\nu} = 0$.

III. DIRECTIONAL DERIVATIVE

Imagine I have a vector field $\boldsymbol{U}$. I want to compute the rate of change of this vector field in the direction of another vector field $\boldsymbol{X}$. Basically, we want to compute how $\boldsymbol{U}$ changes as we move along the field lines of vector field $\boldsymbol{X}$. We shall denote this directional derivative $\boldsymbol{\nabla_X U}$. In terms of its components, this takes the form

$$(\boldsymbol{\nabla_X U})^\mu = X^\lambda \nabla_\lambda U^\mu. \tag{9}$$

This form is familiar from, say, when we want to compute the rate of change of a scalar function f in multivariate calculus along a given 3-vector direction $\vec{v}$. This is given by $\vec{v}\cdot\vec{\nabla}f$, where $\vec{\nabla}$ is the standard 3-dimensional gradient.

IV. PARALLEL TRANSPORT

Imagine an observer is traveling along a worldline $x^\mu(\tau)$. What does it mean for a tensorial object $\boldsymbol{T}$ to be constant along that path? It means that covariant rate of change of $\boldsymbol{T}$ in the direction of the worldline is zero. Now, the tangent vector to the path is the observer four-velocity $dx^\mu/d\tau$. Thus the directional derivative of $\boldsymbol{T}$ along $x^\mu(\tau)$ is

$$\frac{dx^\lambda}{d\tau}\nabla_\lambda T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} \tag{10}$$

A tensor $\boldsymbol{T}$ is said to be "parallel transported" along the curve $x^\mu(\tau)$ if

$$\frac{dx^\lambda}{d\tau}\nabla_\lambda T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = 0. \tag{11}$$

Physical meaning of covariant derivative: it quantifies the instantaneous rate of change of a tensor field in comparison to what the tensor would be it was parallel transported.

V. LINK TO GEODESICS

In flat spacetime, geodesics are straight lines. Since the tangent vector to a straight line is always the same, we could say that a geodesic in flat spacetime is always "parallel transporting" its tangent vector. In fact, this definition of geodesics as curves that always parallel transport their tangent vector is correct. You can indeed show that in any spacetime, a geodesic is a curve that always parallel transports it's own tangent vector. Since the tangent vector to a worldline $x^\mu(\tau)$ is the four-velocity $u^\mu = dx^\mu/d\tau$, then we can show that this equation

$$\frac{dx^\mu}{d\tau}\nabla_\mu \frac{dx^\nu}{d\tau} = 0 \tag{12}$$

is entirely equivalent to the geodesic equation. Also, since the four-momentum of a particle is $p^\mu = mu^\mu$, this can also be written in the form:

$$p^\mu \nabla_\mu p^\nu = 0. \tag{13}$$

So, a geodesic is a trajectory that always parallel transports its own four-momentum.