GR lecture 5

Covariant derivatives, Christoffel connection, geodesics, electromagnetism in curved spacetime, local conservation of 4-momentum

I. PARALLEL TRANSPORT, AFFINE CONNECTIONS, COVARIANT DERIVATIVES

The Lie derivative $\mathcal{L}_u$ is all fine and well, but it requires a vector field $u^\mu(x)$ to define the relevant flow. It would be nice to have a tensorial derivative operator similar to $\partial_\mu v^\nu$ – let us call it $\nabla_\mu v^\nu$ – from which we can construct directional derivatives $u^\mu \nabla_\mu v^\nu$ by simply contracting with a vector u^μ, without needing to know its derivatives $\partial_\mu u^\nu$. As we've seen, such an operator would require some prescription for relating the vector bases at adjacent points, or, equivalently, for deciding how to move the vector's "head", given some motion (along u^μ) of its "tail". In other words, such an operator ∇_μ requires some extra geometric structure on our spacetime. When it exists, ∇_μ is called a covariant derivative. The corresponding geometric structure is called parallel transport, because it generalizes the flat notion of moving a vector's head in parallel to the motion of its tail.

Let us then postulate this extra structure! What we need is a basis transformation matrix $M_\mu{}^\nu$ corresponding to motion along an infinitesimal vector u^μ. Since we're dealing with infinitesimals, this basis transformation should be linear in u^μ:

$$M_\mu{}^\nu = \delta_\mu^\nu + u^\rho \Gamma^\nu_{\rho\mu} \; ; \quad (M^{-1})_\mu{}^\nu = \delta_\mu^\nu - u^\rho \Gamma^\nu_{\rho\mu} \; . \tag{1}$$

Thus, the recipe for parallel-transporting vectors and covectors reads:

$$v^\mu_{\text{transported}} = v^\mu - u^\nu \Gamma^\mu_{\nu\rho} v^\rho \; ; \tag{2}$$

$$w_\mu^{\text{transported}} = w_\mu + u^\nu \Gamma^\rho_{\nu\mu} w_\rho \; . \tag{3}$$

The object $\Gamma^\mu_{\nu\rho}$ is known as an affine connection. As we'll see, it is not a tensor, but can be used in the construction of tensors. We normally don't try to raise and lower its indices. Let us work out how $\Gamma^\mu_{\nu\rho}$ should transform under a change of coordinates. It is slightly easier to do this starting from the covector transport rule (3). Under a change of coordinates, a covector w_μ at a point x^μ transforms via the matrix:

$$w'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} w_\nu \; . \tag{4}$$

The RHS of the transport rule (3) transforms as:

$$w'^{\text{transported}}_\mu = w'_\mu - u'^\nu \Gamma'^\rho_{\nu\mu} w'_\rho \ , \tag{5}$$

where $\Gamma'^\mu_{\nu\rho}$ is the sought-after transformation of $\Gamma^\mu_{\nu\rho}$. To find it, we should notice that the parallel-transported covector $w^{\text{transported}}_\mu$ should actually transform according to the $\partial x/\partial x'$ matrix evaluated not at x^μ (or x'^μ), but at the new point $x^\mu + u^\mu$ (or $x'^\mu + u'^\mu$):

$$\begin{aligned} w'^{\text{transported}}_\mu &= \left(\frac{\partial x^\nu}{\partial x'^\mu} + u'^\rho \frac{\partial^2 x^\nu}{\partial x'^\rho \partial x'^\mu}\right) w^{\text{transported}}_\nu \\ &= \left(\frac{\partial x^\nu}{\partial x'^\mu} + u'^\rho \frac{\partial^2 x^\nu}{\partial x'^\rho \partial x'^\mu}\right) \left(w_\nu + u^\rho \Gamma^\sigma_{\rho\nu} w_\sigma\right) \\ &\approx \frac{\partial x^\nu}{\partial x'^\mu} w_\nu + \frac{\partial x^\nu}{\partial x'^\mu} u^\rho \Gamma^\sigma_{\rho\nu} w_\sigma + u'^\rho \frac{\partial^2 x^\nu}{\partial x'^\rho \partial x'^\mu} w_\nu \\ &= w'_\mu + \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\lambda} \frac{\partial x'^\kappa}{\partial x^\sigma} u'^\lambda \Gamma^\sigma_{\rho\nu} w'_\kappa + u'^\rho \frac{\partial^2 x^\nu}{\partial x'^\rho \partial x'^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} w'_\sigma \ . \end{aligned} \tag{6}$$

where we neglected a piece quadratic in the inifinitesimal u^μ. Comparing (5) and (6), we obtain the transformation rule:

$$\Gamma'^\mu_{\nu\rho} = \frac{\partial x'^\mu}{\partial x^\sigma} \frac{\partial x^\kappa}{\partial x'^\nu} \frac{\partial x^\lambda}{\partial x'^\rho} \Gamma^\sigma_{\kappa\lambda} + \frac{\partial x'^\mu}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x'^\nu \partial x'^\rho} \ . \tag{7}$$

The first term is the standard tensor transformation rule; the second tells us that $\Gamma^\mu_{\nu\rho}$ is not a tensor. Nevertheless, we are now ready to use $\Gamma^\mu_{\nu\rho}$ to construct the covariant derivative ∇_μ. To take the covariant derivative $u^\mu \nabla_\mu v^\nu$ of e.g. a vector, we must simply take the transported value (2), and subtract it from the actual value $v^\nu + u^\mu \partial_\mu v^\nu$ at the new point:

$$u^\mu \nabla_\mu v^\nu \equiv u^\mu \partial_\mu v^\nu + u^\mu \Gamma^\nu_{\mu\rho} v^\rho \ , \tag{8}$$

and similarly for a covector w_μ. Since everything is just proportional to u^μ, we can get rid of it. We end up with the following expressions for the covariant derivative:

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\mu\rho} v^\rho \ ; \quad \nabla_\mu w_\nu = \partial_\mu w_\nu - \Gamma^\rho_{\mu\nu} w_\rho \ . \tag{9}$$

When performing a coordinate transformation, the "unwanted" terms in the transformation of e.g. $\partial_\mu v^\nu$ and $\Gamma^\mu_{\nu\rho}$ cancel. Thus, the covariant derivative $\nabla_\mu v^\nu$ (as well as $\nabla_\mu w_\nu$) is an honest tensor! The covariant derivative can now be defined for tensors with any number of indices. For scalars, we define simply $\nabla_\mu f \equiv \partial_\mu f$.

Exercise 1. *Demonstrate the Leibniz rules:*

$$\nabla_\mu (f v^\nu) = v^\nu \partial_\mu f + f \nabla_\mu v^\nu \ ; \quad \partial_\mu (u_\nu v^\nu) = v^\nu \nabla_\mu u_\nu + u_\nu \nabla_\mu v^\nu \ . \tag{10}$$

Just as we did for Lie derivatives, we use the Leibniz rule to define the covariant derivative of arbitrary tensors, e.g.:

$$\nabla_\mu T_{\nu\rho\sigma}{}^\lambda = \partial_\mu T_{\nu\rho\sigma}{}^\lambda - \Gamma^\kappa_{\mu\nu} T_{\kappa\rho\sigma}{}^\lambda - \Gamma^\kappa_{\mu\rho} T_{\nu\kappa\sigma}{}^\lambda - \Gamma^\kappa_{\mu\sigma} T_{\nu\rho\kappa}{}^\lambda + \Gamma^\lambda_{\mu\kappa} T_{\nu\rho\sigma}{}^\kappa \ . \tag{11}$$

The construction above may be familiar if you've encountered electromagnetism at a sufficiently high level. The wavefunction $\psi(x^\mu)$ of a charged non-spinning particle is a complex number with a phase. The physics (i.e. the Schrodinger equation and Born's rule) is invariant under rotating this phase by a constant $\psi \to e^{i\theta}\psi$. However, once we allow spacetime-dependent phase transformations $\theta(x^\mu)$, the Schrodinger equation is no longer invariant, since the derivative $\partial_\mu\psi$ now has the messy transformation law:

$$\partial_\mu\psi \ \to \ \partial_\mu(e^{i\theta}\psi) = e^{i\theta}(\partial_\mu\psi + i\psi\partial_\mu\theta) \ . \tag{12}$$

The wavefunction ψ in this story is analogous to our tensors. The phase transformation θ is analogous to the basis transformation matrix $\partial x^\mu/\partial x'^\nu$. The second term in (12) is analogous to the unwanted $\partial^2 x^\mu/\partial x'^\nu \partial x'^\rho$ term in the transformation of a partial derivative $\partial_\mu v^\nu$. In electromagnetism, the way to make the local phase transformation (12) a symmetry after all is to introduce the <u>electromagnetic potential</u> A_μ, analogous to $\Gamma^\mu_{\nu\rho}$ in the GR story. We then postulate the gauge transformation:

$$A_\mu \ \to \ A_\mu - \partial_\mu\theta \ , \tag{13}$$

which is analogous to the second term in (7). The presence of the first term in (7) signals that the relevant "force field" is self-interacting; this is true for GR (gravitational waves have energy) and Yang-Mills (gluons have color), but not for electromagnetism (photons have no charge). Anyway, now that we have the new field A_μ, we can use it to replace the partial derivative (12) with a covariant derivative:

$$\nabla_\mu\psi \equiv \partial_\mu\psi + iA_\mu\psi \ , \tag{14}$$

which transforms nicely as $\nabla_\mu\psi \to e^{i\theta}\psi$, and is analogous to (9) in the GR story. Note that this modern understanding of electromagnetism came only <u>after</u> GR, and was inspired by it! In the geometric jargon, both A_μ and $\Gamma^\mu_{\nu\rho}$ are called <u>connections</u>. The more specific name "affine connection" for $\Gamma^\mu_{\nu\rho}$ indicates its more specialized role in transforming the tangent space (rather than some internal space, such as the space of wavefunction values).

II. COMPARING CONNECTIONS; TORSION

Our discussion so far of the connection $\Gamma^\mu_{\nu\rho}$ has been very abstract, and completely disconnected from the fact that spacetime has a metric $g_{\mu\nu}$. As we will see, the metric actually chooses a connection for us: parallel transport, just like all other geometric structures, is already encoded in $g_{\mu\nu}$. Nevertheless, it's often useful to consider parallel transport as an independent geometric concept, and the connection as $\Gamma^\mu_{\nu\rho}$ as an independent variable. One useful observation is that if we have two different connections, then the difference $\Gamma^\mu_{\nu\rho} - \tilde{\Gamma}^\mu_{\nu\rho}$ between them is a tensor: the "unwanted" second transformation term in (7) cancels. This also makes sense from the point of view of the corresponding covariant derivatives ∇ and $\tilde{\nabla}$: the explicitly tensorial quantity $(\nabla_\mu - \tilde{\nabla}_\mu)v^\nu$ can be written also as $(\Gamma^\nu_{\mu\rho} - \tilde{\Gamma}^\nu_{\mu\rho})v^\rho$, since the partial-derivative pieces in (9) cancel.

A related observation is that the "unwanted" second term in (7) is symmetric in the two lower indices. Thus, we can obtain a tensor from $\Gamma^\mu_{\nu\rho}$ by simply antisymmetrizing these two indices:

$$T^\mu_{\nu\rho} \equiv 2\Gamma^\mu_{[\nu\rho]} = \Gamma^\mu_{\nu\rho} - \Gamma^\mu_{\rho\nu} \ . \tag{15}$$

This tensor is known as the torsion of the connection $\Gamma^\mu_{\nu\rho}$. A slightly fancier way to define it is through the commutator of covariant derivatives:

$$[\nabla_\mu, \nabla_\nu] \equiv \nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu \ . \tag{16}$$

Note that partial derivatives always commute: $[\partial_\mu, \partial_\nu] = 0$, but covariant derivatives don't have to: in fact, we'll gradually see how their non-commutativity essentially defines the curvature of spacetime.

Exercise 2. *Show that the torsion $T^\mu_{\nu\rho}$ defines the commutator of covariant derivatives on a scalar field $f(x^\mu)$:*

$$[\nabla_\mu, \nabla_\nu]f = -T^\rho_{\mu\nu}\partial_\rho f \ . \tag{17}$$

For yet another definition of torsion, recall that the "exterior derivative" $\partial_{[\mu}v_{\nu]}$ is a tensor, which doesn't require a connection to define it, and can be compared against $\nabla_{[\mu}v_{\nu]}$:

Exercise 3. *Show that the torsion tensor can be defined via:*

$$\nabla_{[\mu}v_{\nu]} - \partial_{[\mu}v_{\nu]} = -\frac{1}{2}T^\rho_{\mu\nu}v^\rho \ , \tag{18}$$

and that, more generally, when torsion vanishes, we have:

$$\nabla_{[\mu_1} f_{\mu_2 \ldots \mu_n]} = \partial_{[\mu_2} f_{\mu_2 \ldots \mu_n]} \ . \tag{19}$$

Clearly, life is simpler when the torsion vanishes: the antisymmetrized ∇ then agrees with the antisymmetrized ∂, and inherits its interpretation of measuring circulations around closed loops. We then accordingly have $[\nabla_\mu, \nabla_\nu] f = 0$, which expresses the fact that, upon traveling around a closed loop, f returns to itself. In fact, even if we have a connection with torsion, we can always just replace it by one without:

Exercise 4. *If $\Gamma^\mu_{\nu\rho}$ is a connection, show that $\Gamma^\mu_{(\nu\rho)}$ is also a connection, and is torsionless.*

From now on, we will assume by default that our connection is torsion-free, i.e. that $\Gamma^\mu_{\nu\rho} = \Gamma^\mu_{\rho\nu}$.

III. GEODESICS AND THE CHRISTOFFEL CONNECTION

It is time to show how the connection $\Gamma^\mu_{\nu\rho}$ can be derived from the metric $g_{\mu\nu}$. To do this, let's consider the notion of a straight line in flat space. A straight line has two different-sounding but equivalent definitions:

- It is the shortest (or, for timelike lines in spacetime, the longest) path between two points.
- It is the line you get by parallel-transporting a tangent vector along itself.

The first definition is an "integral one", referring to entire paths – just like the action principle. The second is a "differential one", referring to how the line twists and turns (or rather doesn't) at each successive point – just like equations of motion. Crucially, note that the first definition is about the metric, but the second is about the connection! We can use both definitions to define "straight lines", or geodesics, in curved spacetime. The requirement that both definitions agree will provide us with the desired relation between $g_{\mu\nu}$ and $\Gamma^\mu_{\nu\rho}$. We've already done most of the work. When discussing the free-falling particle, we already defined a geodesic, via the action principle, as the longest path between two points in a curved metric. We also derived the Euler-Lagrange equations of motion, which define the geodesic motion differentially:

$$\frac{du^\mu}{d\tau} + g^{\mu\lambda}\left(\partial_\nu g_{\rho\lambda} - \frac{1}{2}\partial_\lambda g_{\nu\rho}\right)u^\nu u^\rho = 0 \ , \tag{20}$$

where $d\tau = \sqrt{-g_{\mu\nu}dx^\mu dx^\nu}$ is the proper time (i.e. length) parameter along the worldline, and $u^\mu = dx^\mu/d\tau$ is the unit tangent vector to the worldline, a.k.a. the 4-velocity. On the other hand, a geodesic should be defined as the line obtained by parallel-transporting u^μ along itself. In a proper-time interval $d\tau$, the line advances by $dx^\mu = u^\mu d\tau$, so we expect the parallel transport to change the components of u^μ by $-\Gamma^\mu_{\nu\rho}dx^\nu u^\rho = -\Gamma^\mu_{\nu\rho}u^\nu u^\rho d\tau$. Thus, we get the geodesic equation in the form:

$$\frac{du^\mu}{d\tau} + \Gamma^\mu_{\nu\rho}u^\nu u^\rho = 0 \ . \tag{21}$$

Comparing eqs. (20) and (21) for arbitrary u^μ, and demanding the torsion-free condition $\Gamma^\mu_{\nu\rho} = \Gamma^\mu_{\rho\nu}$, can we obtain a unique connection $\Gamma^\mu_{\nu\rho}$? Yes! If we focus on the $\nu\rho$ indices, the question is essentially whether a symmetric matrix $A_{\mu\nu}$ is completely defined by its products $A_{\mu\nu}u^\mu u^\nu$ with arbitrary unit vectors u^μ. To see that the answer is yes, we proceed in two steps. First, note that knowing the products $A_{\mu\nu}u^\mu u^\nu$ with unit vectors implies also knowing the products with arbitrary vectors: we can just rescale $u^\mu \to \alpha u^\mu$. Second:

Exercise 5. *Let $A_{\mu\nu}$ be a symmetric matrix 4×4 matrix, written in an arbitrary basis. Find an appropriate set of 10 vectors u^μ, and explicitly express the matrix elements of $A_{\mu\nu}$ in terms of the products $A_{\mu\nu}u^\mu u^\nu$.*

In fact, we've already been using this trick for a while, for the particular case of the metric: we've been equating $g_{\mu\nu}$ with knowing the squared lengths $g_{\mu\nu}dx^\mu dx^\nu$ for arbitrary dx^μ.

To sum up, we can read off from (20)-(21) a unique torsion-free connection by symmetrizing the expression in (20) over $\nu\rho$:

$$\Gamma^\mu_{\nu\rho} = g^{\mu\lambda}\left(\partial_{(\nu}g_{\rho)\lambda} - \frac{1}{2}\partial_\lambda g_{\nu\rho}\right) = \frac{1}{2}g^{\mu\lambda}(\partial_\nu g_{\rho\lambda} + \partial_\rho g_{\nu\lambda} - \partial_\lambda g_{\nu\rho}) \ . \tag{22}$$

The unique torsion-free connection (22) derived from the metric $g_{\mu\nu}$ is known as the Christoffel connection. We will always use this connection unless stated otherwise. When the metric is constant in spacetime, i.e. when the coordinate axes are flat, the connection $\Gamma^\mu_{\nu\rho}$ vanishes. Thus, it measures the curvature of the coordinate axes. For a timelike axis, this curvature can be interpreted as acceleration. Note that this is completely consistent with (21): for a free-falling particle traveling along a geodesic, the acceleration $du^\mu/d\tau$ with respect to the coordinates must cancel against the acceleration $\Gamma^\mu_{\nu\rho}u^\nu u^\rho$ of the coordinates.

Exercise 6. *Find the elements of the Christoffel connection $\Gamma^\mu_{\nu\rho}$ for polar, spherical and Rindler coordinates.*

Exercise 7. *Find the Christoffel connection $\Gamma^\mu_{\nu\rho}$ for the cosmological metric $ds^2 = -dt^2 + a(t)^2\mathbf{dx}^2$.*

For a general worldline, the LHS of (21) can be used to define the 4-acceleration α^μ, which vanishes in the particular case of geodesic motion. Thus, for a charge q in an electromagnetic field $F_{\mu\nu}$ (in addition to the gravitational one), we will have:

$$\frac{du^\mu}{d\tau} + \Gamma^\mu_{\nu\rho}u^\nu u^\rho = \frac{q}{m}F^\mu{}_\nu u^\nu \ . \tag{23}$$

Exercise 8. *Consider motion along a $\rho = const$ worldline in Rindler coordinates. Working completely in Rindler coordinates, find the 4-velocity (u^ρ, u^τ) and 4-acceleration $(\alpha^\rho, \alpha^\tau)$.*

Exercise 9. *Consider circular motion with angular velocity ω along a circle of radius ρ. Working completely in polar coordinates, find the velocity $(v^\rho, v^\phi) = (\dot\rho, \dot\phi)$ and the acceleration (a^ρ, a^ϕ), where dots represent time derivatives. Note that this isn't a relativistic exercise!*

The geodesic equation (21) retains its form if we replace τ by some rescaled parameter $\lambda = \text{const} \times \tau$, and replace $u^\mu = dx^\mu/d\tau$ by $dx^\mu/d\lambda \equiv \dot{x}^\mu$:

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\rho}\dot{x}^\nu\dot{x}^\rho = 0 \ . \tag{24}$$

In this form, the equation is said to describe an affine parameterization $x^\mu(\lambda)$ of the worldline: $d\lambda$ is no longer the proper length, but equal $d\lambda$'s do represent equal intervals. Conversely, $\dot{x}^\mu$ is not necessarily a unit tangent, but it is an affine one: it is constant under parallel transport along the geodesic. When understood in this way, the geodesic equation applies equally well to null geodesics, i.e. to lightrays. There, the length of the line is zero, but an affine parameterization still makes sense: it generalizes the flat notion of using the same lightlike tangent vector throughout the worldline. An important example of an affine vector, which applies both to the timelike and null cases, is the particle's 4-momentum p^μ.

Exercise 10. *Rewrite the equation of motion* (23) *for a charged particle in a way that applies also to the* $m = 0$ *case.*

Sometimes, it may be useful to generalize a bit further, and write the geodesic equation for an arbitrary parameterization $x^\mu(\lambda)$, not necessarily affine. In that case, the RHS of (24) is no longer zero, but must be proportional (i.e. parallel) to the tangent vector $\dot{x}^\mu$:

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\rho}\dot{x}^\nu\dot{x}^\rho \sim \dot{x}^\mu \ . \tag{25}$$

In other words, if the tangent $\dot{x}^\mu$ doesn't maintain its normalization, then it might not be constant under parallel transport, but it must still remain "parallel to itself".

Exercise 11. *Derive* (25)*, starting from the affine geodesic equation* (24) *and replacing* λ *with an arbitrary function* $\tilde{\lambda}(\lambda)$*.*

IV. METRIC COMPATIBILITY AND LIE DERIVATIVES

There is an alternative definition of the Christoffel connection $\Gamma^\mu_{\nu\rho}$: it is the unique torsion-free connection for which the covariant derivative of the metric vanishes: $\nabla_\mu g_{\nu\rho} = 0$. In other words, the metric is constant under parallel transport. This means that when we parallel-transport any vector (not necessarily along itself), it will retain its length. Equivalently, the inner product of any two vectors is unchanged under parallel transport. When manipulating equations, what this means is that the metric can be freely taken into or out of covariant derivatives, i.e. that indices inside covariant derivatives can be raised and lowered without worrying. The condition $\nabla_\mu g_{\nu\rho} = 0$ is known as metric compatibility.

Exercise 12. *Show that the connection* (22) *indeed satisfies* $\nabla_\mu g_{\nu\rho} = 0$*. Show its uniqueness by counting degrees of freedom.*

Exercise 13. *Show that the Lie derivative can be defined equally well with covariant derivatives* ∇_μ *in place of partials* ∂_μ *(note that this only works when torsion vanishes):*

$$\mathcal{L}_u v^\mu = u^\nu\partial_\nu v^\mu - v^\nu\partial_\nu u^\mu = u^\nu\nabla_\nu v^\mu - v^\nu\nabla_\nu u^\mu \ . \tag{26}$$

Exercise 14. *Show that the Lie derivative of the metric along a vector field* ξ^μ *can be written as:*

$$\mathcal{L}_\xi \, g_{\mu\nu} = 2\nabla_{(\mu}\xi_{\nu)} = \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu \ . \tag{27}$$

Note that eq. (27) can be treated as an upgraded version of the electromagnetic gauge transformation $\delta A_\mu = \partial_\mu \theta$. As we can see, A_μ is sometimes analogous to $g_{\mu\nu}$, and sometimes to $\Gamma^\mu_{\rho\sigma}$.

Exercise 15 (TRICKY). *Show that the Christoffel covariant derivative can be constructed using just exterior and Lie derivatives, without any "bare" partial derivatives as in* (22).

V. ELECTROMAGNETISM IN CURVED SPACETIME; DENSITIES AND CONSERVATION LAWS

We now understand the motion of free-falling particles in a curved metric, as well as the motion of charges subject to an electromagnetic field. To complete the picture of electromagnetism in curved spacetime, it remains to rewrite the dynamics of the electromagnetic field, i.e. the Maxwell equations. We begin with the Special Relativistic action:

$$S = -m \int \sqrt{-dx_\mu dx^\mu} + q \int A_\mu dx^\mu - \frac{1}{4} \int F_{\mu\nu} F^{\mu\nu}\, d^4x \ , \tag{28}$$

where we got lazy and set $\epsilon_0 = 1$. As we've seen, the only modification necessary in the first term is to write $dx_\mu dx^\mu$ more explicitly as $g_{\mu\nu} dx^\mu dx^\nu$, and remember that $g_{\mu\nu}$ can be x-dependent. Remarkably, the second term – the interaction between the charges and the EM field – requires no modifications at all: A_μ is a covector, and $A_\mu dx^\mu$ can be integrated just as well in curved spacetime. The third term will require a small amount of work. First, we note that the definition $F_{\mu\nu} = 2\partial_{[\mu} A_{\nu]}$ of the field strength in terms of the potential doesn't require any changes: the antisymmetrized derivative is a tensor. We do, however, need to include explicit $g^{\mu\nu}$ factors to raise the indices in $F_{\mu\nu}F^{\mu\nu}$. Finally, we should worry about the integration measure $d^4x = dx^0 dx^1 dx^2 dx^3$. In flat coordinates, this corresponds to 4d spacetime volume, which is a scalar. In curved coordinates, that's clearly not the case: we can just rescale the coordinates arbitrarily, and the measure will change. In fact, under a general coordinate transformation, the "coordinate volume" d^4x transforms as:

$$d^4x' = \det \left| \frac{\partial x'^\mu}{\partial x^\nu} \right| d^4x \ , \tag{29}$$

where the determinant of the transformation matrix $\partial x'^\mu / \partial x^\nu$ should be familiar from calculus courses as the Jacobian. It is possible to prove (29) directly using $\epsilon_{\mu\nu\rho\sigma}$. Alternatively, we can just notice that it gives the right answer for rescalings $x'^1 = \alpha x^1$ (the coordinate volume rescales by α), as well as for "slants" $x'^1 = x^1 + \beta x^2$ (the coordinate volume remains unchanged).

Thus, for the spacetime integral in (28) to make sense, we must replace d^4x with a version that's invariant under coordinate transformations, and really does measure spacetime volume, using the lengths and angles defined by the metric. It's easy to see that such a measure is actually given by $\sqrt{-g}\,d^4x$, where $g \equiv \det|g_{\mu\nu}|$ is the determinant of $g_{\mu\nu}$, and the minus sign is due to the time dimension. Indeed, for a diagonal metric, lengths along the coordinate axes are given by $\sqrt{-g_{00}}\,dx^0$, $\sqrt{g_{11}}\,dx^1$, $\sqrt{g_{22}}\,dx^2$ and$\sqrt{g_{33}}\,dx^3$, and the product $g_{00}g_{11}g_{22}g_{33}$ is the same as the determinant g. The transformation law for g is the inverse square of (29), so that $\sqrt{-g}\,d^4x$ is indeed invariant:

$$g' = \det\left|\frac{\partial x^\mu}{\partial x'^\nu}\right|^2 g \ ; \qquad \sqrt{-g'} = \det\left|\frac{\partial x^\mu}{\partial x'^\nu}\right|\sqrt{-g} \ . \tag{30}$$

Exercise 16. *Prove this transformation law, by rewriting the tensor transformation law for $g_{\mu\nu}$ in matrix notation.*

Putting everything together, the curved-spacetime version of the EM action (28) is:

$$S = -m\int\sqrt{-g_{\mu\nu}dx^\mu dx^\nu} + q\int A_\mu dx^\mu - \frac{1}{4}\int d^4x\sqrt{-g}\,g^{\mu\rho}g^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma} \ . \tag{31}$$

We will soon come back to the Maxwell equations defined from this action. First, let us extend the discussion of $\sqrt{-g}$ and its transformation law. Since $\sqrt{-g}\,d^4x$ is the spacetime volume, $\sqrt{-g}$ is known as the volume density. More generally, any quantity that satisfies a transformation law of the form (30) is called a "scalar density". Similarly, quantities with indices that have a Jacobian in their transformation in addition to the ordinary basis transformation matrices are known as "tensor densities". Given a metric, we can always create densities by multiplying ordinary scalars and tensors by $\sqrt{-g}$.

I'd rather not develop here the theory in tensor densities in detail. Very briefly, densities are needed whenever we want to integrate. Thus, integrals over 4d spacetime require a scalar density, as in $\sqrt{-g}\,d^4x$. Similarly, integrals over a 3d volume require a vector density. In particular, if we have a vector 4-current j^μ, then the vector density $\sqrt{-g}\,j^\mu$ can be integrated over 3d volume as $\sqrt{-g}\,j^\mu d^3V_\mu$ to produce a total charge. Similarly, the amount of charge produced in an infinitesimal 4d volume will be given by the divergence $\partial_\mu(\sqrt{-g}\,j^\mu)$. This is yet another case where a partial derivative is good enough – in this case, making a scalar density $\partial_\mu(\sqrt{-g}j^\mu)$ out of a vector density $\sqrt{-g}\,j^\mu$. For the (non-densitized) vector j^μ, the divergence $\partial_\mu j^\mu$ doesn't transform nicely under coordinate changes. However, we can replace the partial derivative with a covariant derivative, and talk about $\nabla_\mu j^\mu$:

$$\nabla_\mu j^\mu = \partial_\mu j^\mu + \Gamma^\mu_{\mu\nu} j^\nu \ , \tag{32}$$

where the trace $\Gamma^\mu_{\mu\nu}$ of the Christoffel connection (22) reads:

$$\Gamma^\mu_{\mu\nu} = \frac{1}{2} g^{\mu\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) = \frac{1}{2} g^{\mu\rho} \partial_\nu g_{\mu\rho} \ . \tag{33}$$

The last expression deserves some attention. In Lecture 1-2, we learned some fancy formulas for determinants. One can use them to derive a much simpler formula for the determinant's variation. For an arbitrary matrix $A_{\mu\nu}$ with inverse $(A^{-1})^{\mu\nu}$, we have:

$$d(\det A) = (\det A)(A^{-1})^{\nu\mu} dA_{\mu\nu} \ . \tag{34}$$

This can be easily verified for diagonal matrices, where $A_{\mu\nu} = \mathrm{diag}(A_{00}, A_{11}, A_{22}, A_{33})$, $(A^{-1})^{\mu\nu} = \mathrm{diag}(1/A_{00}, 1/A_{11}, 1/A_{22}, 1/A_{33})$, and $\det A = A_{00}A_{11}A_{22}A_{33}$. Applying (34) to (33), and noting that index order doesn't matter since $g_{\mu\nu}$ is symmetric, we get:

$$\Gamma^\mu_{\mu\nu} = \frac{\partial_\nu g}{2g} = \frac{\partial_\nu \sqrt{-g}}{\sqrt{-g}} \ . \tag{35}$$

Eq. (32) now becomes:

$$\nabla_\mu j^\mu = \partial_\mu j^\mu + \frac{j^\mu \partial_\mu \sqrt{-g}}{\sqrt{-g}} = \frac{1}{\sqrt{-g}} \partial_\mu\left(\sqrt{-g}\, j^\mu\right) \ . \tag{36}$$

Thus, up to the factor of $\sqrt{-g}$, the covariant divergence $\nabla_\mu j^\mu$ of the vector j^μ is the same as the ordinary divergence $\partial_\mu(\sqrt{-g}\,j^\mu)$ of the vector density $\sqrt{-g}\,j^\mu$. This is important, because ordinary divergences have a clear geometric meaning in terms of fluxes and conservation laws! In fact, the same pattern continues for all tensors with just antisymmetrized upper indices. For example, the covariant divergence $\nabla_\mu F^{\mu\nu}$ is proportional to $\partial_\mu(\sqrt{-g}\,F^{\mu\nu})$, where $\sqrt{-g}\,F^{\mu\nu}$ is a density that can be integrated over 2d surfaces as $\frac{1}{2}\sqrt{-g}\,F^{\mu\nu} d^2S_{\mu\nu}$ (the factor of $1/2$ is coming from the permutations of the $\mu\nu$ indices).

We are now ready to write the Maxwell equations in curved spacetime. The source-free half of the equations, which encodes the fact that $F_{\mu\nu}$ has the form $2\partial_{[\mu}A_{\nu]}$, is an exterior derivative, and thus remains unchanged:

$$\partial_{[\mu} F_{\nu\rho]} = 0 \ . \tag{37}$$

The source-dependent part is derived, as before, by varying the action (31). It is easy to see that the $\sqrt{-g}$ factor in the last term will modify the Euler-Lagrange equations from $\partial_\nu F^{\mu\nu} = j^\mu$ into:

$$\partial_\nu\left(\sqrt{-g}\,F^{\mu\nu}\right) = \sqrt{-g}\,j^\mu \quad \Longleftrightarrow \quad \nabla_\nu F^{\mu\nu} = j^\mu\ . \tag{38}$$

The partial-derivative form of the equations ensures that two features survive the upgrade into curved spacetime:

- The Maxwell equations still imply charge conservation, $\nabla_\mu j^\mu \sim \partial_\mu(\sqrt{-g}\,j^\mu) = 0$.
- The Maxwell equations still have the same integral interpretation. For example, one can integrate $\sqrt{-g}\,F^{\mu\nu}$ over 2d spatial surfaces, and obtain the charge inside via Gauss' law.

VI. THE LOCAL CONSERVATION OF 4-MOMENTUM

We've seen that even in curved spacetime, a scalar quantity like the electric charge can be defined as the integral of a local current j^μ (multiplied appropriately by $\sqrt{-g}$), and its conservation can be expressed locally as $\nabla_\mu j^\mu \sim \partial_\mu(\sqrt{-g}\,j^\mu) = 0$. For energy and momentum, both of this statements become problematic. One can still define the local density of 4-momentum via the stress-energy tensor $T^{\mu\nu}$. However, there is no coordinate-invariant way to integrate $T^{\mu\nu}$ over a 3d region and obtain a total 4-momentum. This is only to be expected: we should not be able to add vectors at different points. There is generally no such thing as the total 4-momentum in a region of curved spacetime: a vector needs a single point at which it is defined!

As for the conservation of 4-momentum, the situation is subtle and interesting. Briefly, there exists a local conservation law $\nabla_\mu T^{\mu\nu} = 0$, but it doesn't correspond to the conservation of any overall integrated quantity – indeed, as we've seen, a conservation law of "total 4-momentum" is generally an ill-defined concept! Another way to put this is that we expect a conserved momentum only when there's a corresponding translational symmetry. Since a general curved metric has no such symmetry, there should not be a conserved momentum.

To explore this terrain in some more detail, consider again the geodesic equation (21) for the motion of a free-falling particle. Let us multiply it by the mass m to obtain an equation for the 4-momentum $p^\mu = mu^\mu$:

$$\frac{dp^\mu}{d\tau} + \Gamma^\mu_{\nu\rho} u^\nu p^\rho = 0 \ . \tag{39}$$

Now let's multiply by $d\tau$, using $u^\mu d\tau = dx^\mu$ to obtain:

$$dp^\mu + \Gamma^\mu_{\nu\rho} dx^\nu p^\rho = 0 \ . \tag{40}$$

This equation tells us that the 4-momentum p^μ is constant under parallel transport along the displacement dx^μ, i.e. along the worldline. Note that in this form, the equation applies equally well to the lightlike motion of a massless particle. We already see here the essential subtlety: the 4-momentum is conserved locally, under infinitesimal displacements along the worldline, but we cannot compare its values at points that are finitely separated: the p^μ vectors at those points belong to different tangent spaces.

Let us now work out the corresponding statement for the local density and current of 4-momentum, i.e. for the stress-energy tensor $T^{\mu\nu}$. As before, to calibrate our expectations, we consider the $T^{\mu\nu}$ due to a distribution of free particles. Specifically, we imagine some identical particles that travel along geodesic worldlines, without being created or destroyed. Thus, the number of these particles is a conserved scalar, much like electric charge. Therefore, their volume density and current density can be arranged into a 4-current j^μ with vanishing divergence $\nabla_\mu j^\mu = 0$. If every particle has an electric charge q, then we will have a charge 4-current qj^μ, which of course also has zero divergence. We are however interested in the particles' 4-momentum. Let the particles at position x^μ have 4-momentum p^μ. Then the stress-energy tensor – the 4-current of 4-momentum – is $T^{\mu\nu} = j^\mu p^\nu$. Despite appearances, this tensor is symmetric in its two indices, as we've already seen before. One must simply notice that j^μ and p^μ point in the same direction – along the worldline. Thus, at each point, we can write:

$$j^\mu = \alpha p^\mu \ ; \quad T^{\mu\nu} = \alpha p^\mu p^\nu = \frac{1}{\alpha} j^\mu j^\nu \ , \tag{41}$$

for some scalar α. Consider now the covariant divergence of $T^{\mu\nu}$:

$$\nabla_\mu T^{\mu\nu} = \nabla_\mu (j^\mu p^\nu) = p^\nu \nabla_\mu j^\mu + j^\mu \nabla_\mu p^\nu \ , \tag{42}$$

where we used the Leibniz rule for derivatives of products (recall that it was a defining property in our construction of the covariant derivative (11) for arbitrary tensors!). Now, the first term in (42) vanishes due to the conservation of particle number. The second term also vanishes, thanks to the geodesic equation (40). Indeed, the derivative $j^\mu \nabla_\mu p^\nu$ precisely encodes the parallel transport of p^ν along j^μ, i.e. along the direction of the particle's worldline! We conclude that the stress-energy tensor $T^{\mu\nu}$ satisfies the "local conservation law" $\nabla_\mu T^{\mu\nu} = 0$. One crucial difference between this conservation law and that for j^μ is that it cannot be rewritten in terms of a partial derivative $\partial_\mu (\dots)^{\mu\nu} = 0$: since $T^{\mu\nu}$ is symmetric in its indices rather than antisymmetric, its covariant divergence doesn't get the same "special treatment" as $\nabla_\mu F^{\mu\nu}$ in the previous section. Our inability to write the local conservation law in terms of partial derivatives is just another way of saying that it doesn't encode the conservation of any integrated quantity: $\nabla_\mu T^{\mu\nu}$ is not the actual production rate of anything per unit spacetime volume.