Lecture VII: Gravity gradients, the Ricci tensor, and the field equations

Christopher M. Hirata
Caltech M/C 350-17, Pasadena CA 91125, USA[*]
(Dated: October 29, 2012)

I. OVERVIEW

This lecture will complete our description of curvature tensors. In particular, we will build the Einstein tensor $G_{\mu\nu}$ that occupies the left-hand side of the field equations of general relativity. We will also prove two very important results:

- Any spacetime with zero Riemann tensor is (at least locally) equivalent to flat spacetime.
- The Einstein tensor is divergenceless, $G^{\mu\nu}{}_{;\nu} = 0$.

The recommended reading for this lecture is:

- MTW §8.7 (for geodesic deviation).
- MTW §§17.1–17.3 (for the discussion of the Ricci tensor and field equations).

II. SPACETIMES OF ZERO RIEMANN TENSOR

Thus far, we have referred to $\boldsymbol{Riemann}$ as the curvature tensor. However, while it is obvious that it vanishes in flat spacetime, it is not obvious that a spacetime with $\boldsymbol{Riemann} = 0$ is flat. We will prove this now – with an important condition, that the spacetime must have trivial topology (be simply connected). Otherwise, we can only say that any local neighborhood of a point coincides with Minkowski spacetime.

First, we observe that if a spacetime has zero Riemann tensor, the parallel transport of a vector from $\mathcal{A} \rightarrow \mathcal{B}$ is independent of the path taken. To see this, consider a continuous sequence of paths $\mathcal{P}(\lambda, n)$, where $n \in \mathbb{R}$ and $\mathcal{P}(0, n) = \mathcal{A}$, $\mathcal{P}(1, n) = \mathcal{B}$. These will *not* necessarily be geodesics. Now suppose we build the vector $\boldsymbol{w}(\lambda, n)$ at the point $\mathcal{P}(\lambda, n)$ with $\boldsymbol{w}(0, n) = \boldsymbol{W}$ and defined by parallel-transport along a curve $\mathcal{P}(\lambda, n)$ of constant n. In this case, $\boldsymbol{w}(1, n)$ is the value of the vector parallel-transported to $\mathcal{B}$.

Define, as in the previous lecture for a sequence of paths, $\boldsymbol{v} = \partial\mathcal{P}/\partial\lambda$ and $\boldsymbol{\xi} = \partial\mathcal{P}/\partial n$; then $[\boldsymbol{\xi}, \boldsymbol{v}] = 0$. Now parallel transport tells us that

$$\nabla_{\boldsymbol{v}} \boldsymbol{w} = 0. \tag{1}$$

It follows that

$$\nabla_{\boldsymbol{v}} \nabla_{\boldsymbol{\xi}} \boldsymbol{w} = \nabla_{\boldsymbol{\xi}} \nabla_{\boldsymbol{v}} \boldsymbol{w} + \nabla_{[\boldsymbol{v},\boldsymbol{\xi}]} \boldsymbol{w} + \boldsymbol{Riemann}(\_, \boldsymbol{w}, \boldsymbol{v}, \boldsymbol{\xi}) = 0 + 0 + 0 = 0. \tag{2}$$

But the left-hand side can be written as

$$\frac{D}{\partial\lambda} \frac{D\boldsymbol{w}}{\partial n} = 0. \tag{3}$$

Now since the initial condition is that $\boldsymbol{w}(0, n) = \boldsymbol{W}$ for all n, we have $D\boldsymbol{w}/\partial n = 0$ at $\lambda = 0$, hence also $D\boldsymbol{w}/\partial n = 0$ at $n = 1$. It follows that $\boldsymbol{w}(1, n)$ is independent of n, and hence the parallel transport of $\boldsymbol{W}$ from $\mathcal{A}$ to $\mathcal{B}$ is independent of path. [The ability to continuously deform any path into any other, which we have used here, is present only for simply connected spacetimes. So for more complex topologies our proof only works if we restrict to a simply connected neighborhood of the point $\mathcal{A}$.]

[*]Electronic address: `chirata@tapir.caltech.edu`

Let's take an orthonormal basis $\{\boldsymbol{e}_{\hat\mu}\}$ at any point $\mathcal{A}$; by the above construction it can be turned into a vector field by parallel transport to any other point, and moreover we have $\nabla_{\boldsymbol{v}}\boldsymbol{e}_{\hat\mu} = 0$ for all $\hat\mu$ and all $\boldsymbol{v}$. It follows from the definition that the commutator $[\boldsymbol{e}_{\hat\mu}, \boldsymbol{e}_{\hat\nu}] = 0$. Also parallel transport properties give $\boldsymbol{e}_{\hat\mu}\cdot\boldsymbol{e}_{\hat\nu} = \eta_{\hat\mu\hat\nu}$.

Now define the n functions $x_{\hat\mu}(\mathcal{B})$ (so far just a set of functions : $\mathcal{M}\to\mathbb{R}$) by taking a path $\mathcal{P}(\lambda)$ from $\mathcal{A}$ at $\lambda = 0$ to $\mathcal{B}$ at $\lambda = 1$ and setting

$$x_{\hat\mu}(\mathcal{B}) \equiv \int_0^1 \boldsymbol{e}_{\hat\mu}\cdot\boldsymbol{v}\,d\lambda, \tag{4}$$

where $\boldsymbol{v}$ is the velocity vector of the path. This is well-defined since as we vary the path,

$$\begin{aligned}
\frac{dx_{\hat\mu}(\mathcal{B})}{dn} &= \frac{d}{dn}\int_0^1 \boldsymbol{e}_{\hat\mu}\cdot\boldsymbol{v}\,d\lambda \\
&= \int_0^1 \frac{\partial}{\partial n}(\boldsymbol{e}_{\hat\mu}\cdot\boldsymbol{v})\,d\lambda \\
&= \int_0^1 \boldsymbol{e}_{\hat\mu}\cdot\nabla_{\boldsymbol{\xi}}\boldsymbol{v}\,d\lambda \\
&= \int_0^1 \boldsymbol{e}_{\hat\mu}\cdot\nabla_{\boldsymbol{v}}\boldsymbol{\xi}\,d\lambda \\
&= \int_0^1 \frac{\partial}{\partial\lambda}(e_{\hat\mu}\cdot\boldsymbol{\xi})\,d\lambda \\
&= (e_{\hat\mu}\cdot\boldsymbol{\xi})|_{\lambda=0}^1 \\
&= 0,
\end{aligned} \tag{5}$$

since $\boldsymbol{\xi} = 0$ at both $\mathcal{A}$ and $\mathcal{B}$ (both points are independent of n so $\boldsymbol{\xi} = \partial\mathcal{P}/\partial n = 0$). Thus we may define a coordinate system by $x^{\hat\mu} = \eta^{\hat\mu\hat\nu}x_{\hat\nu}$. It is easily seen that

$$\nabla_{\boldsymbol{v}}x^{\hat\mu} = \eta^{\hat\mu\hat\nu}\nabla_{\boldsymbol{v}}x_{\hat\nu} = \eta^{\hat\mu\hat\nu}\boldsymbol{e}_{\hat\nu}\cdot\boldsymbol{v} = \eta^{\hat\mu\hat\nu}v_{\hat\nu} = v^{\hat\mu}, \tag{6}$$

where the components of $\boldsymbol{v}$ are expressed in the orthonormal basis $\{\boldsymbol{e}_{\hat\mu}\}$. It follows from this that $\{\boldsymbol{e}_{\hat\mu}\}$ is the basis defined by the coordinate system; since it is orthonormal, $g_{\hat\mu\hat\nu} = \eta_{\hat\mu\hat\nu}$ and the spacetime is truly flat.

III. RICCI TENSOR

Now that we have established the Riemann tensor as the notion of "curvature," we may want to relate it to the stress-energy tensor. Unfortunately a 4th rank tensor cannot equal a 2nd rank one, so we are motivated to consider 2nd rank tensors built from ***Riemann***.

A. Gravity gradients

To do so, let's first consider the notion of gravity gradients. Going back to the geodesic deviation equation, we find that the apparent relative acceleration of a geodesic displaced by $\boldsymbol{\xi}$ from another with velocity $\boldsymbol{v}$ is:

$$\frac{D^2\xi^\gamma}{d\tau^2} = -R^\gamma{}_{\mu\alpha\beta}v^\mu\xi^\alpha v^\beta. \tag{7}$$

Thus if we build an orthonormal frame for the observer with velocity $\boldsymbol{v} = \boldsymbol{e}_{\hat 0}$, then there is a relative acceleration given by

$$a^{\hat i} = -R^{\hat i}{}_{\hat 0\hat j\hat 0}\xi^{\hat j}, \tag{8}$$

and so $R^{\hat i}{}_{\hat 0\hat j\hat 0}$ is the 3×3 gravity gradient seen by the observer. It is the relativistic analogue of the Newtonian $-\partial g^i/\partial x^j = \partial^2\Phi/(\partial x^i\,\partial x^j)$. [In fact, for a stationary spacetime that is a small perturbation around Minkowski and $g_{00} = -1 - 2\Phi$, $R^{\hat i}{}_{\hat 0\hat j\hat 0}$ is to first order the second derivative matrix of the potential Φ.] Note that because of the symmetry properties of ***Riemann***, the gravity gradient is symmetric: this is the equivalent of the Newtonian "curl $\boldsymbol{g} = 0$."

You know that in Newtonian physics the PDE for the gravitational potential was

$$\nabla^2\Phi = -\mathrm{div}\,\boldsymbol{g} = 4\pi\rho. \tag{9}$$

That is, positive density implies that the gravity gradient is converging, or that its trace is positive. (The shearing or traceless-symmetric part corresponds to tidal fields produced remotely by other matter.) These considerations motivate constructing the trace of the gravity gradient seen by an observer, $R^{\hat{i}}{}_{\hat{0}\hat{i}\hat{0}}$. In full tensor language, we write the *Ricci tensor*

$$R_{\beta\delta} = R^{\alpha}{}_{\beta\alpha\delta}, \tag{10}$$

so that in the frame of an observer of 4-velocity $\boldsymbol{v}$ we have

$$\mathbf{R}(\boldsymbol{v},\boldsymbol{v}) = R^{\hat{i}}{}_{\hat{0}\hat{i}\hat{0}}. \tag{11}$$

By construction, the Ricci tensor is symmetric. It takes in a 4-velocity (in both slots) and returns a positive number if gravity is converging as measured by that observer and a negative number if it is diverging.

The energy density seen by an observer is $\rho = \mathbf{T}(\boldsymbol{v},\boldsymbol{v})$. It is tempting therefore to suggest a relativistic generalization of Newtonian dynamics:

$$R_{\mu\nu} \stackrel{?}{=} 4\pi T_{\mu\nu}. \tag{12}$$

Unfortunately this doesn't quite work: we want a relativistic field equation that automatically implies a conserved source, $T^{\mu\nu}{}_{;\nu} = 0$, just as the electrodynamic field equations implied $J^{\mu}{}_{;\mu} = 0$. We therefore want to compute the divergence of the Ricci tensor, and decide whether it always vanishes. It doesn't, but the calculation will tell us how to fix Eq. (12).

B. Divergences

We would like to prove a general theorem about the divergence of the Ricci tensor. A good starting point is the only nontrivial theorem on gradients of ***Riemann*** that we have encountered – the Bianchi identity:

$$R^{\alpha\beta}{}_{[\gamma\delta;\epsilon]} = 0. \tag{13}$$

We can get the divergence of the Ricci tensor by contracting the index pairs $\alpha\gamma$ and $\beta\epsilon$:

$$3R^{\alpha\beta}{}_{[\alpha\delta;\beta]} = R^{\alpha\beta}{}_{\alpha\delta;\beta} + R^{\alpha\beta}{}_{\delta\beta;\alpha} + R^{\alpha\beta}{}_{\beta\alpha;\delta} = 0. \tag{14}$$

Using the symmetries of the Riemann tensor, each of the three terms can be turned into a Ricci tensor:

$$R^{\beta}{}_{\delta;\beta} + R^{\alpha}{}_{\delta;\alpha} - R_{;\delta} = 0, \tag{15}$$

where R is the *Ricci scalar*, defined as the contraction of the Ricci tensor: $R = R^{\alpha}{}_{\alpha}$. The first two terms are identical, so

$$R^{\beta}{}_{\delta;\beta} = \frac{1}{2}R_{;\delta}. \tag{16}$$

It can easily be shown that $R_{;\delta}$ may be anything (try your favorite messy metric), so the Ricci tensor need not vanish.

But Eq. (16) also tells us how to fix the problem. Let us define the *Einstein tensor* $\mathbf{G}$

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \tag{17}$$

which is obviously a rank 2 symmetric tensor. We then have

$$G^{\beta}{}_{\delta;\beta} = R^{\beta}{}_{\delta;\beta} - \frac{1}{2}R_{;\beta}\delta^{\beta}{}_{\delta} = \frac{1}{2}R_{;\delta} - \frac{1}{2}R_{;\delta} = 0. \tag{18}$$

IV. THE EINSTEIN FIELD EQUATIONS

A. The equations

The Einstein tensor $\mathbf{G}$ has the correct properties to be on the left-hand side of Einstein's equations:

- It is rank 2 and symmetric.
- It vanishes in flat spacetime.
- It is divergenceless, $\nabla \cdot \mathbf{G} = 0$.

Thus we expect Einstein's equations to take the form $\mathbf{G} = \kappa\mathbf{T}$ for some constant κ – but what should this constant be? To gain insight, let's re-write this in terms of the Ricci tensor. Recall that the trace of the metric tensor is $g^{\alpha}{}_{\alpha} = n$ (the number of dimensions), and then the trace of the Einstein tensor is

$$G = R - \frac{1}{2}nR = -\frac{n-2}{2}R. \tag{19}$$

It follows that the Ricci tensor is expressible in terms of the Einstein tensor:

$$\mathbf{R} = \mathbf{G} + \frac{1}{2}R\mathbf{g} = \mathbf{G} - \frac{1}{n-2}G\mathbf{g}. \tag{20}$$

In particular, operating on the 4-velocity $\boldsymbol{v}$ of any observer, and assuming that $\mathbf{G} = \kappa\mathbf{T}$, we find

$$\mathbf{R}(\boldsymbol{v},\boldsymbol{v}) = \kappa\left[\mathbf{T}(\boldsymbol{v},\boldsymbol{v}) - \frac{1}{n-2}T\mathbf{g}(\boldsymbol{v},\boldsymbol{v})\right] = \kappa\left[\rho - \frac{1}{n-2}[-\rho + (n-1)p](-1)\right] = \frac{1}{n-2}\kappa[(n-3)\rho + (n-1)p], \tag{21}$$

where $p = \frac{1}{n-1}T^{\hat{i}}{}_{\hat{i}}$ is the isotropic part of the stress tensor (the pressure) and $T \equiv T^{\alpha}{}_{\alpha}$. In $n = 4$ dimensions, this simplifies to

$$\mathbf{R}(\boldsymbol{v},\boldsymbol{v}) = \frac{\kappa}{2}(\rho + 3p). \tag{22}$$

Now the left-hand side is the trace of the gravity gradient, and is $4\pi\rho$ in Newtonian physics. So we need $\kappa = 8\pi$ to reproduce Newtonian gravity in the limit of $p \ll \rho$. So we see that

$$\mathbf{G} = 8\pi\mathbf{T}. \tag{23}$$

Note that the choice of 8π really corresponds to a choice of units that makes the Newtonian gravitation constant $G_{\rm N} = 1$, and does not alter the physical content of the theory. We will use such units throughout the course.

B. Some consequences

The first consequence of the Einstein equations actually comes from Eq. (22): it tells us that the local convergence of the gravitational field is $4\pi(\rho + 3p)$. This means that **not only does energy density pull, so does pressure**. Under ordinary circumstances this is irrelevant: in the case of the Earth, the ratio of the pressure to the density is of order $gR \sim M/R \sim v_{\rm c}^2$ where $v_{\rm c}$ is the speed of an orbiting satellite (8 km/s $= 10^{-4.5}c$), so we have $p/\rho \sim 10^{-9}$. Even for the Sun $p/\rho \sim 10^{-6}$. But in the case of a neutron star, the gravitational attraction provided by the pressure in the core of the star is significant, and plays a key role in establishing a maximum possible mass for the neutron star.

A second consequence concerns the nature of gravity gradients, or in general of the Riemann tensor. Of the 20 independent components, 10 are set locally by the stress-energy tensor. The 10 remaining components, including the 5 components of tidal shear seen by an observer ($R_{\hat{i}\hat{0}\hat{j}\hat{0}} - \frac{1}{3}R_{\hat{0}\hat{0}}\delta_{\hat{i}\hat{j}}$), are determined remotely. These components are collected into the *Weyl tensor*:

$$C_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} - \frac{1}{n-2}(g_{\alpha\gamma}R_{\beta\delta} - g_{\alpha\delta}R_{\beta\gamma} - g_{\beta\gamma}R_{\alpha\delta} + g_{\beta\delta}R_{\alpha\gamma}) + \frac{1}{(n-1)(n-2)}R(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}). \tag{24}$$

It's straightforward to see that $\mathbf{C}$ obeys the same symmetry relations as ***Riemann***, and is traceless ($C^{\alpha}{}_{\beta\alpha\delta} = 0$), so it has only 10 linearly independent components. Five of these – those of the traceless symmetric tensor $C_{\hat{i}\hat{0}\hat{j}\hat{0}}$ – are the tidal field seen by an observer. However, since the background spacetime is dynamic, it is usually *not* correct to think of $\mathbf{T}$ as being "given" in the sense that one may give charges and currents in an electromagnetics problem: rather one must *simultaneously* solve for the behavior of the matter and the metric. The exception is for solving the behavior in vacuum where $\mathbf{T} = 0$ and hence $\mathbf{G} = \mathbf{R} = 0$. The study of spacetimes of zero Ricci tensor is incredibly rich.

V. COSMOLOGICAL CONSTANT

Einstein postulated that the Universe, despite being filled with matter, was static. This caused a serious problem for his theory, since the static assumption forced gravity gradients to vanish, contradicting Eq. (22).

He resolved the contradiction by modifying the left-hand side of his equation. Instead of putting **G** there, it is possible to try another tensor. Clearly having something of 2nd rank, symmetric, and divergenceless is non-negotiable. The complete structure of the theory would have to be changed if higher than 2nd derivatives of the metric (e.g. $R_{;\mu\nu}$) or something nonlinear in the second derivative (e.g. $R_{\mu\beta}R^{\beta}{}_{\nu}$) were introduced. Such theories are legal, and have even been seriously suggested as alternative explanations of cosmic acceleration, but Einstein chose a less radical alternative: he could replace the field equation with

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = 8\pi T^{\mu\nu} \tag{25}$$

or

$$\mathbf{G} + \Lambda\mathbf{g} = 8\pi\mathbf{T}, \tag{26}$$

where Λ is a small number called the *cosmological constant*. If Λ is constant then the source is still automatically conserved, and the left-hand side is a symmetric 2nd rank tensor. However, in this case, flat spacetime with no matter is no longer a solution. The cosmological constant forces even empty space to be curved!

We will solve cosmology including the Λ term in a few weeks, but we can immediately see what effect it will have. By moving it to the right-hand side of Einstein's equations, we see that it is equivalent to adding a contribution to the stress-energy tensor of $(\Lambda/8\pi)\mathbf{g}$, i.e. it is equivalent to adding $\Lambda/8\pi$ to the density and $-\Lambda/8\pi$ to the pressure. Thus Eq. (22) becomes

$$\mathbf{R}(\boldsymbol{v}, \boldsymbol{v}) = 4\pi(\rho + 3p) - \Lambda. \tag{27}$$

Remember that the left-hand side is the measured converging part of the gravity gradient (relativistic analogue of "$-\mathrm{div}\,\boldsymbol{g}$"). So $\Lambda > 0$ implies a form of repulsive gravity even in the absence of matter, whereas $\Lambda < 0$ would imply attractive gravity. In particular, if $\Lambda > 0$, then in vacuum in an isotropic system the gravity gradient tensor is $R_{\hat{i}\hat{0}\hat{j}\hat{0}} = (\Lambda/3)\delta_{\hat{i}\hat{j}}$, and the geodesic deviation equation reads

$$\frac{D^2\xi^{\hat{i}}}{d\tau^2} = \frac{\Lambda}{3}\xi^{\hat{i}}; \tag{28}$$

then over a timescale $\sqrt{3/\Lambda}$, nearby objects tend to drift apart exponentially!

We now know that the Universe is expanding and there is no such balance as Einstein envisioned. For many years it was also believed that the expanding Universe did away with the need for the Λ term. However, observations of the expansion history of the Universe (which we will cover later) in fact show positive acceleration, which requires that the Λ term (or some alternative) be re-introduced. So far as we can tell, cosmological observations are consistent with Λ as the single modification to gravity, with a value [1]

$$\Lambda = (1.27 \pm 0.07) \times 10^{-56}\,\mathrm{cm}^{-2} = (1.14 \pm 0.06) \times 10^{-35}\,\mathrm{s}^{-2} = (2.76 \pm 0.15) \times 10^{-46}\,M_{\odot}^{-2}. \tag{29}$$

This is such a small value that it has no significant effect on laboratory physics, or stellar astrophysics, or the observable properties of black holes. Only in the regions of lowest density, i.e. in the setting of cosmology, does Λ play a role.

---

[1] WMAP 7-year result, including BAO and H_0 determinations; $\Omega_\Lambda h^2 = .36 \pm 0.02$. Note that $\Lambda = 3\Omega_\Lambda H_0^2$. `http://lambda.gsfc.nasa.gov/product/map/dr4/best_params.cfm`