Lecture # 9 : Covariant Derivatives & curvature

Note Title 9/27/2011

I. The covariant Derivative
II. Maxwell & Differential Forms
III. Geodesics, Again
IV. Curvature
V. Tidal Forces & The equation of geodesic Deviation.

I. The covariant Derivative

One of the main items on our wish list is some means of building tensors from Derivatives of tensors.

Recall the problem:

$$\partial_c g_{ab} \text{ is not a tensor!}$$

In most coordinate systems $\partial_c g_{ab} \neq 0$ $(\& \ \Gamma^a{}_{bc} \neq 0)$, but at any pt. p we can always change to locally inertial coordinates in which

$$\partial_{\tilde{c}} g_{\tilde{a}\tilde{b}}\Big|_p = 0, \quad \Gamma^{\tilde{a}}{}_{\tilde{b}\tilde{c}}\Big|_p = 0$$

$$\Rightarrow \quad \partial_{\tilde{c}}\, g_{\tilde{a}\tilde{b}} \neq \frac{\partial x^c}{\partial x^{\tilde{c}}}\frac{\partial x^a}{\partial x^{\tilde{a}}}\frac{\partial x^b}{\partial x^{\tilde{b}}}\partial_c g_{ab}$$

$$\frac{\partial x^c}{\partial x^{\tilde{c}}}\frac{\partial}{\partial x^c}\left[\frac{\partial x^a}{\partial x^{\tilde{a}}}\frac{\partial x^b}{\partial x^{\tilde{b}}}\, g_{ab}\right]$$

The problem clearly comes from "acceleration terms" involving $\frac{\partial}{\partial x^c}\frac{\partial x^a}{\partial x^{\tilde{a}}}$.

Note that a similar term arises for any vector, co-vector, or tensor field which is not a scalar field.

So, what do we do about this?

One answer is to use the equivalence principle! This tells us that the derivative that will give us the right physics is the one computed in a locally inertial frame!

I.e, let's try to define a "covariant derivative ∇T" in any coordinate system via the following procedure:

i) Let p be the space time point where we wish to compute ∇T. Find new coordinates $y^{\tilde{a}}$ that are locally inertial at p.

ii) Compute ∂T in the locally inertial coordinates $y^{\tilde{a}}$.

iii) Define ∇T to be the tensor that agrees w/ (ii) at p in the coords $y^{\tilde{a}}$. I.e, for a vector

$$\nabla_a V^b \Big|_p := \frac{\partial x^b}{\partial y^{\tilde{b}}} \frac{\partial y^{\tilde{a}}}{\partial x^a} \partial_{\tilde{a}} V^{\tilde{b}} \Big|_p, \text{ etc.}$$

(while for scalars $\nabla_a \phi = \frac{\partial y^{\tilde{a}}}{\partial x^a} \underbrace{\partial_{\tilde{a}} \phi}_{\text{locally inertial}} = \partial_a \phi$)

Note 1: Recall that one must find new coords $y^{\tilde{a}}$ at each pt p! I.e, the choice of $y^{\tilde{a}}$ depends on p!

Note 2: Since the problem came only from acceleration terms, it does not matter which set of locally inertial coords one chooses at p. I.e., any two sets of locally inertial coordinates $y^{\hat{a}}$ & $y^{\tilde{a}}$ at p satisfy

$$\frac{\partial^2 y^{\hat{}}}{\partial y^{\tilde{}2}} = 0 = \frac{\partial^2 y^{\tilde{}}}{\partial y^{\hat{}2}}$$

& thus lead to the <u>same</u> $\nabla T|_p$.

In fact, we are free to weaken the notion of "locally inertial" as on your last HW to require only that

$$\partial_{\tilde{c}}\, g_{\tilde{a}\tilde{b}}\Big|_p = 0$$

<u>without</u> necessarily requiring $g_{\tilde{a}\tilde{b}}\Big|_p = \eta_{\tilde{a}\tilde{b}}$, since one may pass back & forth btwn $\eta_{\tilde{a}\tilde{b}}$ & any other (constant) symmetric matrix using a linear coordinate transformation.

Now, this definition sounds like a real pain. But luckily we have already worked out how to obtain locally inertial coords $y^{\tilde{a}}$ from any set of coordinates x^a. From HW #3, the answer is

christoffels in the x-coords

$$y^{\tilde{a}} = \delta^{\tilde{a}}_a \left(\Delta x^a + \tfrac{1}{2}\Gamma^a{}_{bc}\Delta x^b \Delta x^c\right)$$

$$\Rightarrow \quad \Delta x^a = \delta^a_{\tilde{a}}\left(y^{\tilde{a}}\right) - \tfrac{1}{2}\Gamma^a{}_{bc}\,\delta^b_{\tilde{b}}\,\delta^c_{\tilde{c}}\, y^{\tilde{b}} y^{\tilde{c}}$$

$$+ O(y^3)$$

where $\Delta x^a = x^a - x^a(p)$

So, $\dfrac{\partial^2 y^{\tilde{a}}}{\partial x^b \partial x^c} = \delta^{\tilde{a}}_a \Gamma^a{}_{bc}$ & $\dfrac{\partial^2 x^a}{\partial y^{\tilde{b}} \partial y^{\tilde{c}}} = -\Gamma^a{}_{bc}\,\delta^b_{\tilde{b}}\,\delta^c_{\tilde{c}}$

Note: Γ is not a tensor, so with the natural definition of $\Gamma^{\tilde a}{}_{\tilde b \tilde c}$ (where these are the christoffel symbols in $y^{\tilde a}$ coordinates) we have

$$\Gamma^{\tilde a}{}_{\tilde b \tilde c} = 0.$$

This is why I have written $\delta^{\tilde a}{}_{a} \Gamma^{a}{}_{bc}$ above instead of $\Gamma^{\tilde a}{}_{bc}$, which might have been confusing. [And writing $y^{\tilde a} = \cdots \Gamma^{\tilde a}{}_{\cdot\cdot}$ would make it look like the indices don't match up, so that might lead to confusion as well.]

This means that we can write a "simple" formula for ∇T directly in the original x-coordinates! E.g., for a vector we have

$$\nabla_a V^b \Big|_p := \frac{\partial y^{\tilde a}}{\partial x^a} \frac{\partial x^b}{\partial y^{\tilde b}} \nabla_{\tilde a} V^{\tilde b} \Big|_p$$

$$= \frac{\partial y^{\tilde a}}{\partial x^a} \frac{\partial x^b}{\partial y^{\tilde b}} \partial_{\tilde a} V^{\tilde b} \Big|_p$$

$$= \left(\underbrace{\frac{\partial y^{\tilde a}}{\partial x^a} \partial_{\tilde a}}_{\partial_a} \Big[\underbrace{\frac{\partial x^b}{\partial y^{\tilde b}} V^{\tilde b}}_{V^b} \Big] - \frac{\partial y^{\tilde a}}{\partial x^a} V^{\tilde b} \frac{\partial^2 x^b}{\partial y^{\tilde a} \partial y^{\tilde b}} \right) \Bigg|_p$$

$$= \left(\partial_a V^b - \frac{\partial y^{\tilde a}}{\partial x^a} V^{\tilde b} \left(-\Gamma^{b}{}_{cd} \, \delta^{c}{}_{\tilde a} \delta^{d}{}_{\tilde b} \right) \right) \Bigg|_p$$

But $\displaystyle \frac{\partial y^{\tilde a}}{\partial x^a}\Big|_p = \delta^{\tilde a}{}_{a}$ & $\displaystyle \delta^{a}{}_{\tilde a} = \frac{\partial x^a}{\partial y^{\tilde a}}\Big|_p$

$$\Rightarrow \quad \nabla_a V^b \Big|_p = \left(\partial_a V^b + \Gamma^{b}{}_{ad} \frac{\partial x^d}{\partial y^{\tilde b}} V^{\tilde b} \right) \Bigg|_p$$

$$= \left(\partial_a V^b + \Gamma^b{}_{ad} V^d\right)_p$$

But now that both sides are written in the original coordinates (so that we don't have to recall that the def. of $g_{\tilde{a}}$ depends on the choice of p), we can conclude

$$\nabla_a V^b = \partial_a V^b + \Gamma^b{}_{ac} V^c$$

everywhere.

Note: this is a tensor by construction!

Similarly, given any co-vector field ω_a one finds (from an analogous calculation)

$$\nabla_a \omega_b = \partial_a \omega_b - \Gamma^c{}_{ab} \omega_c$$

comes from "opposite" coord trans. of ω_a vs. V?

I'll let you work through the details on your own, but one can quickly see that this is the right answer from the following argument: ∇_a is just $\partial_{\tilde{a}}$ in locally inertial coordinates. So it must satisfy the Leibnitz rule:

$$\nabla (T_1 T_2) = (\nabla T_1) T_2 + T_1 \nabla T_2$$

Thus we have $\nabla_a (V^b \omega_b) = \partial_a (V^b \omega_b)$

Since $V^b \omega_b$ is a scalar.

So for all v^b we have

$$v^b(\nabla_a \omega_b) = \partial_a(v^b \omega_b) - \omega_b \nabla_a v^b$$

$$= v^b \partial_a \omega_b + \cancel{\omega_b \partial_a v^b} - \cancel{\omega_b \partial_a v^b} - \omega_b \Gamma^b_{ac} v^c$$

(Renaming b as c & c as b)

$$= v^b(\partial_a \omega_b - \Gamma^c_{ab}\omega_c)$$

Similarly, for any tensor $T^{a_1 \cdots a_n}{}_{b_1 \cdots b_m}$, we have

$$\nabla_c T^{a_1 \cdots a_n}{}_{b_1 \cdots b_n} = \partial_c T^{a_1 \cdots a_n}{}_{b_1 \cdots b_m} + \sum_{i=1}^{n} \Gamma^{a_i}_{cd} T^{a_1 \cdots d \cdots a_n}{}_{b_1 \cdots b_m} - \sum_{j=1}^{m} \Gamma^d_{cb_j} T^{a_1 \cdots a_n}{}_{b_1 \cdots d \cdots b_m}$$

Note: while it is natural to be confused by all of these incides, it is actually easy to ~~get~~ the index contractions right. E.g., in

$$\nabla_a v^b = \partial_a v^b + \Gamma^b_{ac} v^c$$

So long as you recall that this index is contracted w/ Γ, there is only one way to match up the indices! (Since $\Gamma^b_{ac} = \Gamma^b_{ca}$).

O.K., now for the all important question: what is
$\nabla_a g_{bc}$?

Direct calculation yields

Have the students do this in class as practice

$$\nabla_a g_{bc} = \partial_a g_{bc} - \Gamma^d{}_{ab}\, g_{dc} - \Gamma^d{}_{ac}\, g_{bd}$$

$$= \partial_a g_{bc} - \left(\Gamma_{cab} + \Gamma_{bac}\right)$$

But $\Gamma_{cab} = \frac{1}{2}\left(\partial_a g_{bc} + \cancel{\partial_b g_{ac}} - \cancel{\partial_c g_{ab}}\right)$

$+ \; \Gamma_{bac} = \frac{1}{2}\left(\partial_a g_{bc} + \cancel{\partial_c g_{ba}} - \cancel{\partial_b g_{ac}}\right)$

$$= \partial_a g_{bc}$$

So $\boxed{\nabla_a g_{bc} = 0\,!}$ ($\Rightarrow \nabla^{(n)} g = 0$ as well!)

This should not really be a surprise. After all, we defined $\nabla_a g_{bc}$ in terms of $\partial_{\tilde{a}} g_{\tilde{b}\tilde{c}}$ in locally inertial coordinates at each pt. p.

But in locally inertial coordinates we have

$$\partial_{\tilde{a}} g_{\tilde{b}\tilde{c}}\Big|_p = 0 \quad \text{by definition!}$$

Said differently, we have seen that there is no coordinate-invariant info in 1st derivatives of the metric. So how could we have found any non-zero result? The surprise perhaps is that

$\nabla_a \nabla_b g_{cd} = 0$ as well, since there is invariant info in 2nd derivatives of g_{cd}.

But don't worry. We'll see soon how we can extract this info from ∇_a any way.

Remark #1: The covariant derivative of a tensor density can be defined in precisely the same way.

E.g., consider a scalar density $\tilde{\phi} = \sqrt{-g}\,\phi$ for some scalar ϕ. Clearly $\nabla_a \sqrt{-g} = 0$ (since $\partial_a \sqrt{-g} = 0$ in locally inertial coordinates).

$$\text{so } \nabla_a \tilde{\phi} = \nabla_a (\sqrt{-g}\,\phi) = \sqrt{-g}\,\nabla_a \phi = \sqrt{-g}\,\partial_a \phi$$

$$= \partial_a (\sqrt{-g}\,\phi) - \phi\,\partial_a \sqrt{-g}$$

$$= \partial_a \tilde{\phi} - \tilde{\phi}\,\partial_a \ln \sqrt{-g}$$

$$= \partial_a \tilde{\phi} - \tfrac{1}{2}\,\partial_a \ln \det g \quad (\text{since } \partial_a \sqrt{-1} = 0)$$

But recall that $\ln \det M = \operatorname{Tr} \ln M$ for any matrix M, so

$$\partial_a \ln \det g = \operatorname{Tr}(g^{-1} \partial_a g) \leftarrow$$ uses cyclic property of the trace

$$= (g)^{bc}\,\partial_a g_{bc}$$

Note: $\Gamma^b{}_{ab} = g^{bc}\,\Gamma_{bac} = \frac{1}{2} g^{bc} (\partial_a g_{bc} + \cancel{\partial_c g_{ba}} - \cancel{\partial_b g_{ac}})$

$$= \tfrac{1}{2} g^{bc}\,\partial_a g_{bc}$$

So $\nabla_a \tilde{\varphi} = \partial_a \tilde{\varphi} - \Gamma^b{}_{ab} \tilde{\varphi}$

Remark #2

One difference btwn the operators ∇_a & ∂_a is that of course $\partial_{[a} \partial_{b]} T = 0$ (since derivatives commute!

But it turns out that covariant derivatives do not generally commute! You might ask how can this be? Doesn't $\nabla_a \nabla_b V^c|_p$ become just $\partial_{\tilde{a}} \partial_{\tilde{b}} V^{\tilde{c}}$ in some locally inertial coordinate system at p? No!, this works for the first derivative $\nabla_{\tilde{b}} V^{\tilde{c}}|_p = \partial_{\tilde{b}} V^{\tilde{c}}|_p$. But to compute $\partial_{\tilde{a}} \partial_{\tilde{b}} V^{\tilde{c}}$ we need to know $\partial_{\tilde{b}} V^{\tilde{c}}|_p$ and also at nearby points $p' \neq p$ where our current coordinates are not locally inertial! So 2nd covariant derivatives know abt the way that locally inertial coordinates change from pt. to pt. I.e., they know that the spacetime is not flat & they encode spacetime curvature!

We'll discuss this more next time, but I want to quickly point out a special case first.

Let ϕ be a scalar. Then $\nabla_a \phi = \partial_a \phi$

& $\nabla_b \nabla_a \phi = \partial_b \nabla_a \phi - \Gamma^c{}_{ba} \nabla_c \phi$

$= \partial_b \partial_a \phi - \Gamma^c{}_{ba} \partial_c \phi$

So $\nabla_{[b} \nabla_{a]} \phi = \cancel{\partial_{[b} \partial_{a]}}^{0} \phi - \cancel{\Gamma^c{}_{[ba]}}^{0} \partial_c \phi = 0$

Since $\Gamma^a{}_{bc}$ is symmetric in b & c.

So covariant Derivatives do commute when acting on Scalars. This is known as the torsion-free property. (And any ∇_a defined via its $\Gamma^a{}_{bc}$ w/ $\Gamma^a{}_{[bc]} \neq 0$ is said to have torsion.)

So, our covariant Derivative satisfies the following useful properties:

1) $\nabla_a \phi = \partial_a \phi$ for Scalars

2) The Leibnitz rule $\nabla T_1 T_2 = T_2 \nabla T_1 + T_1 \nabla T_2$

3) Linearity: For $A, B \in \mathbb{R}$,

$$\nabla (AT_1 + BT_2) = A\nabla T_1 + B\nabla T_2$$

4) commutes w/ contractions:

i.e. $\nabla_d \left(T^{a_1 \ldots c \ldots a_m}{}_{b_1 \ldots c \ldots b_n}\right)$
$= (\nabla T)_d{}^{a_1 \ldots c \ldots a_m}{}_{b_1 \ldots c \ldots b_n}$

In general, any operator w/ these properties is said to be a covariant Derivative. Wald proves that any such $\tilde{\nabla}$ satisfying (1-4) is of the form

$$\tilde{\nabla}_a v^b = \partial_a v^b + C^b{}_{ac} v^c$$

for some "connection coeff's $C^b{}_{ac}$" that play the role of our Γ's (but which in general are not related to locally inertial coordinates.)

Note that while neither of $\Gamma^a{}_{bc}$, $C^a{}_{bc}$ are tensors, $\Gamma^a{}_{bc} - C^a{}_{bc}$ <u>is</u> a tensor since for any vector field V^c we have

$$(\nabla_a - \tilde{\nabla}_a) V^c = (\Gamma^c_{ab} - C^c_{ab}) V^b$$

(The $\partial_a V^c$ terms cancelled out).

But our ∇ has two more properties:

4) torsion-free

& 5) "metric-compatible" $\nabla_a g_{bc} = 0$.

Wald & Carroll take (1–5) as an axiomatic definition of ∇ & essentially reverse our logic above to show that this implies that the Christoffel symbols $\Gamma^a{}_{bc}$ are the connection coefficients for ∇ (from which the relation to locally inertial coordinates then follows from the fact that they make Γ vanish).

II. Maxwell & Differential Forms

Let's just check on how much progress we are making toward our goal.

We now have a way of making a tensor ∇T from ∂T. Furthermore, $\nabla T = \partial T$ in locally inertial coords.

So, if we know the action for any system in special relativity, the equivalence principle suggests that the action in G.R. should be of the same form w/

1) η replaced by g

2) $d^4x \rightarrow \sqrt{-g}\, d^4x$

3) ∂_a replaced by ∇_a

The resulting action is known as the minimally-coupled action.

E.g., for a scalar $S = -\int d^4x \left(\frac{1}{2}\partial_a \phi\, \partial_b \phi\, \eta^{ab} + V(\phi)\right)$

$$\downarrow$$

$$S = -\int d^4x \left(\frac{1}{2} \nabla_a \phi \nabla_b \phi\, g^{ab} + V\right)\sqrt{-g}$$

Note: A "non-minimally-coupled action" would contain some additional term that explicitly vanishes in Mink space like $|\nabla_{[a}\nabla_{b]}\nabla_c \phi|^2$.

(Also note if the original action already contains $\partial_a \partial_b \partial_c \phi$ then the notion of minimal coupling is ill-defined as one has to worry abt how to order the covariant Derivatives.)

Of course, this will not yet give us an action for gravity itself since we did not have a good S.R. action for the metric!

But we can do (e.g.) maxwell Theory!

$$S = -\frac{1}{4}\int d^4x\, F_{ab}F^{ab} + \int d^4x\, A_a J^a$$

$$\downarrow \qquad \text{w/ } F_{ab} = \partial_a A_b - \partial_b A_a$$

$$S = -\frac{1}{4}\int d^4x \sqrt{-g}\, F_{ab}F^{ab} + \int d^4x\, A_a J^a$$

$$\text{w/ } F_{ab} = \nabla_a A_b - \nabla_b A_a$$

$$= \partial_a A_b - \partial_b A_a + \cancel{\Gamma^c_{ab} A_c - \Gamma^c_{ba} A_c}$$

Aha! So we did not really need covariant Derivatives here after all!

It turns out that the symmetry of $\Gamma^a{}_{bc}$ makes it cancel out of many such calculations. Given any totally anti-symmetric covariant (all indices down) tensor $\omega_{a_1 \dots a_n} = \omega_{[a_1 \dots a_n]}$ one finds

$$\nabla_{[a_{n+1}}\omega_{a_1 \dots a_n]} = \partial_{[a_{n+1}}\omega_{a_1 \dots a_n]}$$

This observation is clearly useful, so such totally anti-sym tensors are given the special name "differential form" & the totally anti-symmetric derivative is given the special name "exterior derivative". It is called "d."

I.e. for any differential form $\omega_{a_1 \cdots a_n}$,

$$(d\omega)_{a_1 \cdots a_{n+1}} := (n+1)\, \partial_{[a_{n+1}} \omega_{a_1 \cdots a_n]}$$

$$\Rightarrow \quad d^2\omega = 0 \qquad \text{since} \qquad \partial_{[a_{n+1}} \partial_{a_{n+2}]} = 0.$$

While we won't distract ourselves further right now w/ the mathematics of differential forms, they are very useful. So we will return to them soon.

III. Geodesics, Again

Let us return to the (affinely parameterized) geodesic equation

$$0 = \frac{d^2 x^a}{d\lambda^2} + \Gamma^a{}_{bc} \frac{dx^b}{d\lambda} \frac{dx^c}{d\lambda}$$

In terms of the 4-velocity $U^a = \frac{dx^a}{d\lambda}$

(not normalized [i.e., U^2 may not be -1] for $\lambda \neq \tau$

=> allows null case w/ $U^2 = 0$)

This is

$$0 = \underbrace{\frac{dU^a}{d\lambda}}_{\frac{\partial x^c}{\partial \lambda}\underbrace{\frac{\partial}{\partial x^c}U^a}} + \Gamma^a{}_{bc}\, U^b U^b$$

note: $\frac{\partial}{\partial x^c} U^a$ is not really well-defined for a single worldline because it requires info abt how U^a changes in the direction transverse to the worldline. On the other hand, this may be supplied by considering a sufficiently large family of smooth curves containing ours & $\frac{\partial x^c}{\partial \lambda}\frac{\partial}{\partial x^c} U^a$ is independent of how this family is defined. It depends only on our worldline.

I.e., the geodesic Eqn is $0 = U^b\left(\partial_b U^a + \Gamma^a{}_{bc} U^c\right)$

$$= U^b \nabla_b U^a$$

We immediately see that, in locally inertial coordinates at the pt. P, we have

$$U^{\hat{b}}\, \partial_{\hat{b}}\, U^{\hat{a}}\Big|_P = 0$$

I.e. the tangent along the curve is locally constant along the geodesic (as described in appropriate local inertial frames).

IV. Curvature

This brings up an interesting question. Does it make any sense to say that U^a is globally constant? Not really. Note that $\partial_b U^a = 0$ cannot hold in all coordinate systems (unless $U^a = 0$). One could try to solve

$$0 = \nabla_a U^b = \partial_a U^b + \Gamma^b_{ac} U^c \qquad (*)$$

but in a general spacetime metric this has no solutions!

To see why, note that any solution must satisfy an integrability condition given as usual by taking ∂_d of the above eqn & anti-symmetrizing:

$$0 = \partial_{[d} \nabla_{a]} U^b = \partial_{[d} \Gamma^b_{a]c} U^c + \Gamma^b_{[a|c|} \partial_{d]} U^c$$

& using $(*)$ to replace $\partial_d U^c = -\Gamma^c_{de} U^e$

I.e., unless

$$\frac{1}{2} R_{adc}{}^b = \partial_{[d} \Gamma^b_{a]c} - \Gamma^b_{[a|f|} \Gamma^f_{d]c} \quad \text{has}$$

a zero eigenvalue eigenvector, there are no solutions!

This suggests that $R_{adc}{}^b$ has at least some invariant info and, indeed, with a bit of work one may check that it transforms as a tensor. It is called the Riemann tensor.

Related objects that come up often are the

Ricci tensor $\quad R_{ab} = R_{acb}{}^c$

& the Ricci scalar $\quad R = R_a{}^a$.

Note that The Riemann Tensor has the following

Symmetries: R_{abcd} is

1) anti sym under $(a \leftrightarrow b)$ [manifest from above]
2) anti sym under $(c \leftrightarrow d)$
3) sym under $(a,b) \leftrightarrow (c,d)$

4) A "cyclic sym"
$$0 = R_{abcd} + R_{adbc} + R_{acdb}$$
(w/ #2, #4 $\Leftrightarrow$ $R_{a[bcd]} = 0$)

It turns out that it also satisfies a so-called "Bianchi identity"
$$\nabla_{[a} R_{bc]de} = 0,$$

which is somewhat analogous to $d_{[a} F_{bc]} = 0$ in Maxwell Theory.

Let's postpone the derivations for a bit. We will discuss them shortly

Before doing so, I'd like to mention that the easiest way to verify that this is in fact a tensor is to study the commutator $\nabla_{[a}, \nabla_{b]}$ that we noted before should be generally non-zero.

A straight forward computation shows

$$[\nabla_a, \nabla_b] V^c := \nabla_{[a} \nabla_{b]} V^c = -R_{abc}{}^d V^c$$

and $\quad [\nabla_a, \nabla_b] \omega_c = + R_{abc}{}^d \omega_d$

& similarly $\quad [\nabla_a, \nabla_b] T^{c_1 \cdots c_n}{}_{d_1 \cdots d_m} = \sum$ similar terms for each index

This shows that $R_{abc}{}^d$ is a tensor.

BTW, the key point here (as emphasized by Wald & Carroll) is that in computing

$[\nabla_a, \nabla_b] T$ one finds that all terms involving derivatives (either 1st or 2nd!) of T cancel out.

The easiest way to see that this must happen is to note that, for any scalar f,

$$[\nabla_a, \nabla_b](fT) = (\nabla_a \nabla_b - \nabla_b \nabla_a)(fT)$$

$$= \nabla_a (f \nabla_b T + T \nabla_b f) - \nabla_b (f \nabla_a T - T \nabla_a f)$$

$$= (\underbrace{[\nabla_a, \nabla_b] f}_{0}) T + (\nabla_a f \cancel{\nabla_b T} + \cancel{\nabla_b f} \nabla_a T - \cancel{\nabla_b f} \nabla_a T - \nabla_a f \cancel{\nabla_a T})$$
$$+ f [\nabla_a, \nabla_b] T$$

One may use the fact that this holds for any $T^{c_1 \cdots c_n}{}_{d_1 \cdots d_n}$ to argue that $[\nabla_a, \nabla_b] T$ is in fact algebraic in T.

I can now say just a few brief words about how the above symmetry properties are derived

#2) $R_{ab[cd]}$ Follows From

$$0 = \nabla_{[a} \nabla_{b]} g_{cd} = R_{abc}{}^{e} g_{ed} + R_{abd}{}^{e} g_{ce}$$

$$= R_{abcd} + R_{abdc}$$

#4) $R_{[abc]}{}^{d}$ Follows From

$$R_{[abc]}{}^{d} w_d = \nabla_{[[a} \nabla_{b]} w_{c]} = \nabla_{[a} \nabla_{b} w_{c]} = \partial_{[a} \partial_b w_{c]} = 0$$

I.e., follows from $\Gamma^{a}{}_{[bc]} = 0$

Cyclic sym then follows from algebra & #2

#3) $R_{abcd} = R_{cdab}$ Then follows by algebra from the above

(though it is perhaps easier just to see this from the formula in terms of Γ's, see Carroll.)

#5) $\nabla_{[a} R_{bc]d}{}^{e} = 0$

Follows From $\nabla_{[a} \nabla_b \nabla_{c]} w_d = 0$

$$= \nabla_{[a} \nabla_{[b} \nabla_{c]]} w_d$$

$$= \nabla_{[a} (R_{bc]d}{}^{e} w_e)$$

$$= (\nabla_{[a} R_{bc]d}{}^{e}) w_e + \nabla_{[a|} w_e R_{|bc]d}{}^{e}$$

But also $0 = \nabla_{[a} \nabla_b \nabla_{c]} w_d = \nabla_{[[a} \nabla_{b]} \nabla_{c]} w_d = R_{[ab|c|}{}^{e} \nabla_{e]} w_d$

$$+ R_{[ab|d}{}^{e} \nabla_c w_{|e]}$$

$\nabla_{[a|} w_e R_{|bc]d}{}^e = R_{[bc|d}{}^e \nabla_{|a]} w_e$ Trivial rewriting

$= R_{[ab|d}{}^e \nabla_{|c]} w_e$ cyclic

& $R_{[ab|d}{}^e \nabla_c w_{|e]} = R_{[ab|d|e]} \nabla_c w^e$

$= -R_{[abe]d} \nabla_c w^e = 0$

by # 4.

$\Rightarrow \nabla_{[a} R_{bc]d}{}^e = 0$

Note that since $\nabla_a g_{bc} = 0$ it is easy to contract a & e to find

$0 = \nabla_{[a} R_{bc]d}{}^a = \frac{1}{3}\left(\nabla_a R_{bcd}{}^a + \nabla_b R_{cad}{}^a + \nabla_c R_{abd}{}^a\right)$

(each term appears twice due to $R_{[bc]d}{}^a$)

$= \frac{1}{3}\left(\nabla_a R_{bcd}{}^a + \nabla_b R_{cd} - \nabla_c R_{bd}\right)$

messy, but let's do one more:

$0 = 3\nabla_{[a} \ {}_{bc]}{}^{ba}$

$= \nabla_a R_{bc}{}^{ba} + \nabla_b R_c{}^b - \nabla_c R$

$= 2\nabla_a R_c{}^a - \nabla_c R$

To clean this up, define $G_{ab} = R_{ab} - \frac{1}{2} g_{ab} R$ (symmetric!)

$\Rightarrow \nabla_a G_c{}^a = 0$

This gives a kind of conservation law for the "Einstein Tensor" G_{ab}.

There is one other final key point that I will mention: we have seen that there is coordinate invariant information in 2nd derivatives of the metric, & that at least some of this info is encoded in R_{abcd}. In fact, one may show that all such info is encoded in R_{abcd} (& similarly the info in higher derivativatives is encoded in

$\nabla_{e_1} \cdots \nabla_{e_n} R_{abcd}$) & that g_{ab} may be completely reconstructed (at least locally, & up to certain choices of coords) from R_{abcd}. We'll see that directly in the next lecture or two, but it is also instructive to go through the counting argument given in Carroll (& Wald??) which shows that there are the same # of independent components in R_{abcd} & in $\partial_a \partial_b g_{cd}$.

V. Tidal Forces & the equation of geodesic deviation

Mathematically: $[\nabla_a, \nabla_b] V^d = - R_{abc}{}^d V^c$

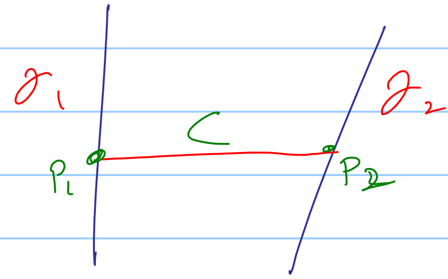
on vectors, & similarly on other tensors defines curvature.

But what does it mean physically?

we know that it must be associated w/ the relative acceleration of geodesics. In particular, two geodesics that are "initially at rest wrt one another" will not remain at relative rest.

But, what precisely does this mean? How do we define the notion of two geodesics initially being at relative rest?

In flat space, we would construct a spacelike geodesic (straight line) btwn the 1st geodesic & the 2nd & which is orthogonal to the 1st geodesic.



we would then compare the tangent vectors U_1 & U_2 to γ_1 & γ_2 at the corresponding pts P_1 & P_2 shown above.

But comparing these 2 vectors is non-trivial in curved space — we need to specify some way of moving a vector from one pt to the other.

A natural prescription is to move the vector (say U_1) along the connecting curve C in such a way that, at each pt $P \in C$,

$$\left.\frac{d\tilde{U}_1^{a}(\lambda)}{d\lambda}\right|_P = 0 \qquad \text{as described by a locally inertial observer at } P$$

Here is any parameter along the curve C.

we have seen that this condition may be written

$$C^b \nabla_b U_1^{a}(\lambda) = 0 \qquad \text{where } C^b = \frac{dx^b}{d\lambda} \text{ is the tangent to } C.$$

Here we should view this as a diff EQ to be solved for $U_1{}^a(\lambda)$. Since $c^a \nabla_b = \frac{d}{d\lambda}$, we may write

$$\frac{d}{d\lambda} U_1{}^a + c^b \Gamma^a_{bc} U_1{}^c = 0$$

This is an eqn of the form

$$\frac{d}{d\lambda} \vec{v} = M \vec{v}$$

vector matrix

where $M^a{}_c = c^b \Gamma^a_{bc}$. [Note: True "matrizes" have one index up & one down so that they can be multiplied together.]

So the solution is $U_1{}^a(\lambda) = \mathcal{P}\exp\left(-\int c\cdot\Gamma\right)^a{}_c V^c$.

Moving a vector along a curve in this way is called "parallel transporting the vector along the curve" because $U_1{}^a$ remains "locally parallel to itself."

Note that parallel transport keeps the length of the vector constant, & in fact preserves the inner product of any two vectors.

I.e., for

$$c^b \nabla_b U_1{}^a = 0$$

$$\& \quad c^b \nabla_b U_2{}^a = 0$$

Then since $\nabla_c g^{ab} = 0$ we have

$$c^b \nabla_b \left(U_1 \cdot U_2\right) = c^b \nabla_b (U_1{}^a U_2{}^b g_{ab}) = 0.$$

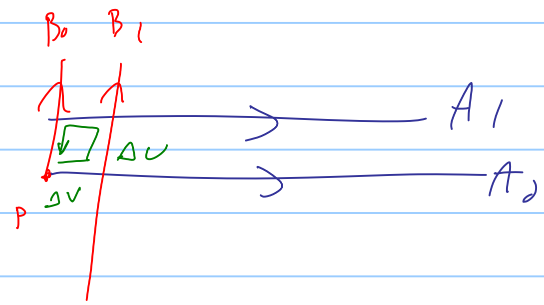
One can also parallel transport co-vectors, tensors, etc. Note that "parallel transporting" a scalar would just say that the scalar is constant along c.

Note that this means that we get the same result for $U_1 \cdot U_2$ whether we parallel transport U_1 to P_2 or U_2 to P_1 (so long as we use the same curve C in either case.)

However, the result does depend on which curve C we choose to use.

In particular, moving a vector around a closed loop can change the vector (by a Lorentz trans, since the length does not change).

This is easiest to see for an infinitesimal loop. Consider some 2-surface & two (suitably non-degenerate) smooth families of curves $A_{(u)}$ & $B_{(v)}$ in that surface.



To leading order in $\Delta u, \Delta v$ the tangents to A_0 & A_1 agree (call them A^a) as do those of B_0 & B_1 (call them B^b).

And similarly $\Gamma^a{}_{bc}$ is constant to leading order in Δu, Δv. But generally $M_u{}^a{}_c = \Gamma^a{}_{bc} A^b$
$\neq \Gamma^a{}_{bc} B^b = M_v{}^a{}_c$

So the result for parallel transporting some vector U^a around this square is (using v as the parameter along B_v & u along A_u)

$$U^a_{final} = \left(e^{-M_u \Delta u} e^{-M_v \Delta v} e^{M_u \Delta u} e^{M_v \Delta v}\right)^a{}_c \, U^c$$

$$= (1 + (\cancel{M_u} - \cancel{M_u})\Delta u + (M_v - M_v)\Delta v$$

$$+\Delta v^2\left(\frac{\cancel{M_v}^2}{2}+\frac{(-\cancel{M_v})^2}{2}-\cancel{M_v^2}\right)+\Delta u^2\left(\frac{\cancel{M_u}^2}{2}+\frac{(-\cancel{M_u})^2}{\cancel{2}}-\cancel{M_u^2}\right)$$

$$+\frac{\cancel{2}\Delta u\Delta v}{\cancel{2}}\left[M_u, M_v\right]+O(\Delta^3)\Big)^a{}_c\, U^c$$

where

$$[M_u, M_v]^a{}_c = \Gamma^a{}_{de}B^d\Gamma^e{}_{fc}A^f - \Gamma^a{}_{de}A^d\Gamma^e{}_{fc}B^f$$

Of course, since the lowest order terms cancelled out completely, we should be more careful about the higher order terms. In particular, we neglected terms of quadratic order associated w/ changes in A^a, B^a, & $\Gamma^a{}_{bc}$.

Some of these (e.g., $\frac{dA^a}{dv}$) give terms of order Δv^2 (or Δu^2). But it is easy to see that such terms must cancel among themselves since clearly $U^a_{final}=U^a_{initial}$ if either $\Delta u=0$ or $\Delta v=0$. So only "cross terms" involving both Δu & Δv can survive.

Such terms come from $\frac{\partial M_v}{\partial u}$ & $\frac{\partial M_u}{\partial v}$.

It turns out that the $\frac{\partial A^b}{\partial v}$ & $\frac{\partial B^b}{\partial v}$ terms (all multiplying the same $\Gamma^a{}_{bc}$!) cancel against each other because the square must close

$$\Rightarrow \cancel{A^a\Delta v}+\cancel{B^a\Delta u}-\left(\cancel{A^a}+\frac{\partial A^a}{\partial u}\Delta u\right)\Delta v - \left(\cancel{B^a}+\frac{\partial B^a}{\partial v}\Delta v\right)\Delta u=0$$

So the only additional terms come from the fact that on A_1 we need

$$\Gamma^a{}_{bc}(0) + \frac{\partial \Gamma^a{}_{bc}}{\partial u}\Delta u = \Gamma^a{}_{bc}(0) + \partial_d \Gamma^a{}_{bc} B^d \Delta u$$

& on β_1 we need

$$\Gamma^a{}_{bc}(0) + \frac{\partial \Gamma^a{}_{bc}}{\partial v}\Delta v = \Gamma^a{}_{bc}(0) + \partial_d \Gamma^a{}_{bc} A^d \Delta v$$

So we obtain

$$U^a_{final} - U^a_{initial} = \Big([M_u, M_v]^a{}_c + B^d A^b \partial_d \Gamma^a{}_{bc} \Delta u \Delta v - B^b A^d \partial_d \Gamma^a{}_{bc} \Delta u \Delta v \Big) U^c_{initial} + O(\Delta^3)$$

$$= - R_{bdc}{}^a \, U^c_{initial} A^d B^b$$

since

$$R_{bdc}{}^a = \partial_d \Gamma_b{}^a{}_c - \partial_b \Gamma^a{}_{dc} + \Gamma^a{}_{df} \Gamma^f{}_{bc} - \Gamma^a{}_{bf} \Gamma^f{}_{dc}$$

I.e. The Riemann tensor at p is a direct measure of the Lorentz transformation induced by parallel transport around infinitesimal closed curves at p.

In particular, parallel transport around the parallelogram defined by A^d B^b induces

$$- A^d B^b R_{bdc}{}^a$$

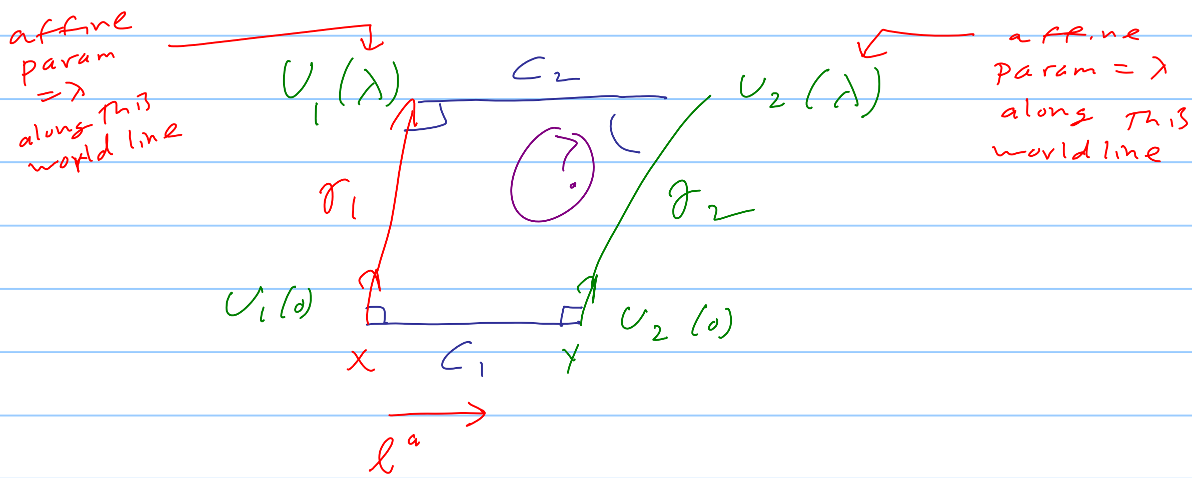
This is indeed a Lorentz Transformation (as it must be!) since R_{bdca} is anti-sym in ca.

Lesson: There is no canonical way to identify tangent spaces at distinct points!

Let's now tie this back in to our discussion about the relative acceleration of two nearby geodesics.

Note that the above picture $\#$ looks much like our relative acceleration problem. We use parallel transport to set up two "initially parallel geodesics."

So we take u_1 & u_2 at ("time zero") to be related by parallel transport along C_1. We took C_1 to be a geodesic $\perp$ u_1 (& thus $\perp$ u_2 as well, but we will shortly see that this is irrelevant.



We now want to choose some C_2 to call "$t = \lambda$" and use parallel propagation along this curve to compare $u_1(\lambda)$ & $u_2(\lambda)$. We could think hard about the best curve to choose, but we will shortly see that it makes no difference (so long as it is a smooth deformation of C_1). Note, however, that we <u>cannot</u> choose C_2 <u>both</u> a geodesic <u>&</u> to be $\perp$ to both $u_1(\lambda)$ & $u_2(\lambda)$. [Consider the 2d case. If we could, then since $u_1(\lambda)$ & $u_2(\lambda)$ are normalized, it would imply that parallel propagating $u_2(\lambda)$ along C_2 yields $u_1(\lambda)$. But this would mean that parallel propagating around the full loop gives the trivial Lorentz transformation, so we must have $U^a C^b R_{abc}{}^d = 0$.]

This failure is precisely the effect we wish to measure! It means that the observers are <u>no longer</u> at relative rest!

So all we need to do is to quantify it properly.

Perhaps the most intuitively obvious thing to do is to parallel propagate $u_2^a(0)$ to γ_1 along c_1 & $u_2^a(\lambda)$ to γ_1 along c_2. Let's call the results $\hat{u}_2^a(0) = u_1^a(0)$ & $\hat{u}_2^a(\lambda)$. Then we can take $\frac{d}{d\lambda}$ of $\hat{u}_2^a(\lambda)$ as computed in a locally inertial reference frame (e.g., the frame of our observer on γ_1)!

I.e., the relative 4-acceleration is

$$a^a(0) := \frac{d}{d\lambda} \hat{u}_2^a(\lambda)\Big|_{\lambda=0} = u_1^b \nabla_b \hat{u}_2^a\Big|_{\lambda=0}$$

(again using the fact that $\nabla_b = \partial_b$ in a locally inertial frame).

Since $\hat{u}_2(0) = u_1(0)$, it is useful to consider

$$\hat{u}_2^a(\lambda) - u_1^a(\lambda) = \mathcal{O}(\lambda) = \lambda \Delta^a + \mathcal{O}(\lambda^2)$$

$\mathcal{O}(1)$. Doesn't actually matter if we choose Δ^a "constant"; as we will see

Note that $a^a(0) = u_1^b \nabla_b \hat{u}_2^a(\lambda)\Big|_{\lambda=0}$

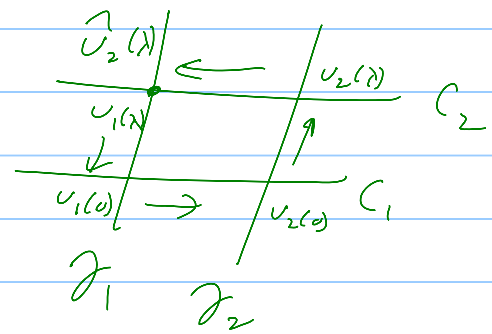
$$= u_1^b \nabla_b (\hat{u}_2^a - u_1^a)\Big|_{\lambda=0}$$ By the geodesic Eqn!

$$= u_1^b \nabla_b (\lambda \Delta^a + \mathcal{O}(\lambda^2))\Big|_{\lambda=0}$$

$$= \Delta^a$$ since $u_1^b \nabla_b \lambda = u_1^b \partial_b \lambda = \frac{d\lambda}{d\lambda} = 1$

& all other terms vanish at $\lambda = 0$.

On the other hand, note that $\tilde{U}_2(\lambda)$ is the result of parallel transporting $U_1(\lambda)$ around the following loop:



So, by our previous result

$$\tilde{U}_2^{\,a}(\lambda) - U_1^{\,a}(\lambda) = U_{final} - U_{initial}$$

$$= -R_{bdc}{}^{a}\, U^c_{initial}\,(-U^d)\, C^b\, \Delta\lambda_c \Delta\lambda_c$$

$$= +\underbrace{R_{bcd}}_{\text{swapped}}{}^{a}\, C^b U^c U^d \Delta\lambda_c \,\Delta\lambda_c$$

swapped since contracted w/ $U^c U^d$

I.e.,
$$a^a = R_{bcd}{}^{a} C^b U^c U^d \Delta\lambda_c$$

Note: your textbooks do not have the $\Delta\lambda_c$ factor in their equations because they normalize C^b by setting $\Delta\lambda_c = 1$; i.e., the think of C^b as <u>the</u> displacement vector btwn the two worldlines (as opposed to merely specifying the direction). This is a fine convention (& the one

I prefer myself, but it is sometimes confusing to students (so I don't insist on it).

Note that the relative acceleration is linear in $\Delta\lambda$ (for small $\Delta\lambda$). Of course, we should really say "$+ o(\Delta\lambda^2)$".

We are now almost finished with our story, but I want to emphasize that the above definition of relative acceleration as

$$a^a = U^b \nabla_b \hat{U}_2$$

coincides with another physical notion given by the 2nd derivative of the separation vector

$$\ell^a = C^a \, \Delta\lambda_c$$

Note that $\ell^a = C^a$ if we simply choose λ_c along the curves $C(\lambda)$ so that $\Delta\lambda_c = 1$. This is a vector defined at each point of γ_1, so we can ask for it's 1st derivative as computed by a locally inertial observer (i.e., by observer #1!)

$$V^a = \frac{d\ell^a}{d\lambda_U} = U^b \nabla_b \ell^a$$

& the 2nd derivative

$$a^a = \frac{dV^a}{d\lambda_U} = U^b \nabla_b V^a$$

To see this, we simply expand

$$x_2^a(\lambda_v) - x_1^a(\lambda_v) = C^a_{\lambda_v}\,\Delta\lambda_c + \mathcal{O}(\Delta\lambda_c^2)$$

$$= \ell^a(\lambda_v) + \mathcal{O}(\Delta\lambda_c^2)$$

Let us compute $\frac{d\ell^a}{d\lambda_v}$ in locally inertial coordinates.

Then clearly $\frac{d\ell^a}{d\lambda^v} = \left. u_2^a - u_1^a \right\}$ not a tensor, since involves vectors at different pts!

$= \underbrace{\hat{u}_2^a}_{\text{since parallel propagation leaves locally inertial coordinates unchanged}} - u_1^a \leftarrow$ A Tensor: vectors at same pt!

But the LHS defines the tensor v^a, so we have

$$v^a = u_1^a - \hat{u}_2^a$$

while we have derived this result using locally inertial coords, since both sides are tensors (at the same spacetime point), the eqn must in fact hold in all coordinate systems.

But we saw above that

$$a^a = u_1^b \nabla_b (u_2^a - \hat{u}_1^a)$$

so $a^a = u^b \nabla_b v^a$ as claimed.

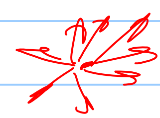
Remarks:

#1) The same argument works for null & spacelike geodesics as well. (Note that we defined a^a in terms of $\frac{d}{d\lambda}$)

#2) A clarification: The Eq. of geodesic deviation is an exact mathematical result about the vector ξ^a & its derivatives along a worldline. This describes the infinitesimal displacement relating the two worldlines or, equivalently, it describes the displacement to first order in the separation.

#3) We can use this result to show that R_{abcd} contains all of the (local) coordinate invariant info in any metric. I.e., we can reconstruct the metric (in a sufficiently small but finite-sized region surrounding any point p) from R_{abcd}.

To do so, consider any pt p & fire geodesics outward from p in all (spacetime) directions



Choose some affine parameter λ along each geodesic that varies smoothly from one geodesic to the next. [In Riemanian (all +) signature, just use proper distance.]. Now, if a point q lies on the geodesic w/ tangent T^a (at p) at affine parameter λ, assign it coordinates

$$x^a(q) = \lambda T^a$$

This will be a non-singular system of coordinates until the geodesics begin to cross each other. (I.e., in some finite-sized region.) They are called "Riemann normal coordinates."

Note that the distance btwn any two points on the same geodesic is just $\Delta\lambda$. And as we have seen, the distance btwn neighboring geodesics satisfies a 2nd order ODE built from $R_{abc}{}^{d}$. So the metric in these coordinates can be completely constructed by solving 2nd order ODEs w/ source $R_{abc}{}^{d}$.

Some physics:

So, how big is the Riemann Tensor near the earth, & what does it look like?

We know empirically that relative accelerations match those given by Newton's laws.

$$\vec{a}_{Newton} = -\frac{MG}{r^2}\hat{r} = -\frac{MG\vec{r}}{r^3}$$

$$= \vec{a}_{Newton}(x_0) - \frac{MG(\vec{r}-\vec{x}_0)}{r^3} + \frac{3MG\vec{r}}{r^4}(|\vec{r}|-|\vec{x}_0|) + \text{2nd order}$$

(in a coord system where the metric is approximately $g_{ab}=\text{diag}(-1,1,1,1)$ @ x_0)

Define $r^a = (0,\vec{r})$ & $\ell^a = (0,\vec{\ell})$ w/ $\vec{\ell} = \vec{r}-\vec{x}_0$

(note: $\ell^0 = 0$ since $t = t_0$)

$$\Rightarrow \quad a^a_{relative} = -\frac{MG}{r^3}\ell^a + \frac{3MG\,r^a}{r^4}\partial_b r\,\ell^b$$

On the other hand, defining $t^a = (1,0,0,0)$,

The eqn of geodesic deviation says

$$a^a = R_{bcd}{}^a \ell^b t^c t^d = R_{btt}{}^a \ell^b$$

Thus we identify

$$R_{btt}{}^a = -\frac{MG}{r^3} \underbrace{\delta_b{}^a}_{\delta \text{ in space}} + \frac{3MG r^a}{r^4} \partial_b r$$

vanishes for $a=t$ or $b=t$

Note that this allows us to easily compute the R_{tt} component of the Ricci tensor:

$$R_{cd} = R_{bcd}{}^b$$

i.e., $R_{tt} = -\frac{3MG}{r^2} + \frac{3MG}{r^4} \underbrace{r^a \partial_a r}_{r} = 0$

(easy in spherical coords!)

(at least in the Newtonian approximation).

It is interesting to repeat this calculation inside the earth (using, say, a constant density model of the earth w/ $M = \frac{4\pi}{3} r^3 \rho$).

This adds an extra term to

$$\vec{a}_{Newton} = ---- \frac{-4\pi\rho G \vec{r}}{r^3} (r^2 \Delta r)$$

$$a^a_{newton} = ... -8\pi \rho G r^a \ell^b \frac{\partial_b r}{r}$$

$$\Rightarrow \quad R_{btt}{}^{a} \rightarrow \dots - \frac{8\pi \rho G \, r^a \partial_b r}{r}$$

$$\Rightarrow \quad R_{tt} = -8\pi \rho G$$

This looks like it might be part of a set of equations of motion for the metric! Einstein worked rather hard to find a "good" relativistic formula that reproduces the above non-relativistic limit. But we will make our lives somewhat easier by working at the level of the action. This will guarantee that the EOMs have all possible "good" properties. We'll discuss these good properties after finding the right action.

BTW, we find no measureable curvature of space, so, to the accuracy of our experiments we also find

$$R_{ij} = 0.$$

But note that these experiments are intrinsically much harder. Recall that observing a timelike geodesic for 1 second is like monitoring a spacelike geodesic over $3 \times 10^7 m = 3 \times 10^4 km$!

$$\approx r_{earth}/3 \, ,$$