"GR is locally SR"

Canonical Form

The metric can, by a suitable change of coordinates, be brought to the form

(1.1) $$(g_{\mu\nu}) = \mathrm{diag}(-1, +1, +1, +1) .$$

This has "Lorentzian" signature, whereas our old friend δ_{ij} from $\mathbb{R}^3$ has "Euclidean" signature.

Locally Inertial Coords

At any particular point p, you can choose a reference frame in which

(1.2) $$\partial_{\hat\sigma} g_{\hat\mu\hat\nu}\Big|_p = 0$$

BUT this cannot hold beyond first derivatives. Let's analyze here by looking at a Taylor series expansion for the locally inertial coords (call them $\hat{x}^\mu$) in terms of the regular ones; choosing $\hat{x}^\mu\big|_p = x^\mu\big|_p$ for simplicity, and expanding in a Taylor series, gives

(1.3) $$x^\mu = \left(\frac{\partial x^\mu}{\partial x^{\hat\nu}}\right)\Bigg|_p x^{\hat\nu} + \text{higher-order terms}$$

↑ (4×4 = 16 independent components)

(2-1) We also know that $g_{\hat{\mu}\hat{\nu}} = \frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}} g_{\lambda\sigma}$.

So $g_{\hat{\mu}\hat{\nu}}(\hat{x})\big|_p$

(2-2) $= g_{\hat{\mu}\hat{\nu}}\big|_p + \frac{\partial}{\partial x^{\hat{\beta}}} g_{\hat{\mu}\hat{\nu}}\big|_p x^{\hat{\beta}} + \mathcal{O}(\hat{x}^2)$

$= \left(\frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}} g_{\lambda\sigma}\right)\Big|_p + \left(\frac{\partial}{\partial x^{\hat{\beta}}}\left(\frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}}\right) g_{\lambda\sigma} + \left(\frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}} \left(\frac{\partial}{\partial x^{\hat{\beta}}} g_{\lambda\sigma}\right)\right)\right)\Big|_p x^{\hat{\beta}} + \ldots$

Let's compare $\mathcal{O}(1)$ and $\mathcal{O}(\hat{x})$ terms here. We get two equations (to this next-to-lowest order)

(2-3) (a) $g_{\hat{\mu}\hat{\nu}}\big|_p = \left(\frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}} g_{\lambda\sigma}\right)\Big|_p$ and

(2-4) (b) $\left(\partial_{\hat{\beta}} g_{\hat{\mu}\hat{\nu}}\right)\big|_p = \left(\partial_{\hat{\beta}}\left(\frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}}\right) g_{\lambda\sigma} + \frac{\partial x^\lambda}{\partial x^{\hat{\mu}}} \frac{\partial x^\sigma}{\partial x^{\hat{\nu}}} \partial_{\hat{\beta}} g_{\lambda\sigma}\right)\Big|_p$

▷ For (a), having 16 functions (see (4.1)) <u>is</u> enough to specify $\frac{1}{2!}(4\times 5) = 10$ independent metric components.

The 6 left over are the 6 parameters of the Lorentz group in the locally inertial frame! (3 rotations & 3 boosts.) ☺

▷ For (b), for 2nd term, we have 4 derivatives of 10 cpts, i.e. 40 independent bits to worry about. Also, for 1st term,

(3.1) $$\partial_{\hat\beta}\left((\partial_{\hat\mu}x^\lambda)(\partial_{\hat\nu}x^\sigma)\right) = (\partial_{\hat\beta}\partial_{\hat\mu}x^\lambda)(\partial_{\hat\nu}x^\sigma) + (\partial_{\hat\mu}x^\lambda)(\partial_{\hat\beta}\partial_{\hat\nu}x^\sigma)$$

Notice that $\frac{\partial^2 x^\lambda}{\partial x^{\hat\beta}\partial x^{\hat\mu}}$ is symmetric in $(\hat\beta \leftrightarrow \hat\mu)$ so it also has 40 cpts.

So by adjusting ↑ as necessary, we can make $\partial_{\hat\beta} g_{\hat\mu\hat\nu} = 0$ at p (ONLY!), as required.

<u>Pattern</u>: at higher orders in Taylor expansion, there are <u>not</u> enough independent components of 3rd derivative of x to adjust 2nd deriv's of g to zero. [This is where a new construct, called the Riemann curvature tensor, which we will meet later, comes in.]

See Carroll p.75 for the proof.

Covariant Derivative

Remember when we tried to make a tensor in curved space out of ∂(tensor) but failed, $\because \partial_\lambda (\Lambda^{\mu'}{}_{\nu}) \neq 0$ ☹

Let's fix this by inventing a new type of derivative which **does** make a new tensor out of an old one.

We define the covariant derivative ∇ to satisfy

(4.1a) (1) linearity : $\nabla(T+S) = \nabla T + \nabla S$

(4.1b) (2) Leibnitz rule : $\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$

[like $\partial_\mu (ab) = (\partial_\mu a)\cdot b + a \cdot (\partial_\mu b)$ ...]

[not pronounced "Christ-awful"! ☺]

We define the Christoffel connection Γ by (in components)

(4.2)
$$\boxed{\Gamma^\lambda{}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma} (\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu})}$$
(σ is summed over)

and the covariant derivative on vectors by (again, in components)

(4.3)
$$\boxed{\nabla_\mu V^\lambda = \partial_\mu V^\lambda + \Gamma^\lambda{}_{\mu\nu} V^\nu}$$

The important feature of this Γ thingy is that $(\nabla_\mu V^\nu)$ transforms properly i.e. as a bona fide (1,1) tensor!

(*) ▷ The proper derivation is quite boring and lengthy; I do not believe it wise to grind through it here in-class. It is much better to figure it out on your own time – you will appreciate it more this way (although it does take some "sweat" – but so do most good things in life ☺)

N.B. → See e.g. pp 97–98 of Carroll.

For a one-form: in components

(5.1) $$\boxed{\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda{}_{\mu\nu}\,\omega_\lambda}$$

(N.B: Neither term on the RHS here or in (4.1) is a bona fide tensor.)

It is also straightforward (but boring) to work out the rule for a general (k,l) tensor :-

(5.2) $$\begin{aligned}\nabla_\sigma T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = {} & \partial_\sigma T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} \\ & + \Gamma^{\mu_1}{}_{\sigma\lambda} T^{\lambda\mu_2\cdots\mu_k}{}_{\nu_1\cdots\nu_l} + \Gamma^{\mu_2}{}_{\sigma\lambda} T^{\mu_1\lambda\mu_3\cdots\mu_k}{}_{\nu_1\cdots\nu_l} + \cdots \\ & - \Gamma^\lambda{}_{\sigma\nu_1} T^{\mu_1\cdots\mu_k}{}_{\lambda\nu_2\cdots\nu_k} - \Gamma^\lambda{}_{\sigma\nu_2} T^{\mu_1\cdots\mu_k}{}_{\nu_1\lambda\nu_3\cdots\nu_l} + \cdots\end{aligned}$$

This definition of ∇ has the other nice properties:

(6.1) (3) commutes with contraction, i.e.

e.g. $\nabla_\mu (T^\lambda{}_{\lambda\rho}) = (\nabla T)_\mu{}^\lambda{}_{\lambda\rho}$ or $\nabla_\mu(\delta^\nu{}_\lambda) = 0$

(6.2) (4) ∇ reduces to a action on scalars

i.e. $\nabla_\mu \varphi = \partial_\mu \varphi$.

There is a discussion in Carroll (somewhat advanced) showing that (4.1) is actually the unique $\Gamma^\lambda{}_{\mu\nu}$ satisfying axioms (1)–(4) [on pp4-5 above] provided that

(6.3) (a) $\boxed{\Gamma^\lambda{}_{\mu\nu} = \Gamma^\lambda{}_{\nu\mu}}$ ← "torsion-free"

and

(6.4) (b) $\boxed{\nabla_\sigma g^{\mu\nu} = 0}$ ← "metric-compatible"

Notice that the equation $\nabla_\sigma g^{\mu\nu} = 0$ is a bona fide tensor equation: it is built from the tensor $g^{\mu\nu}$ (the upstairs version of the metric tensor) and our new tensorial covariant derivative ∇_σ.

⊛ Let's now shift gears to a more geometrical picture of the meaning of ∇.

Parallel Transport

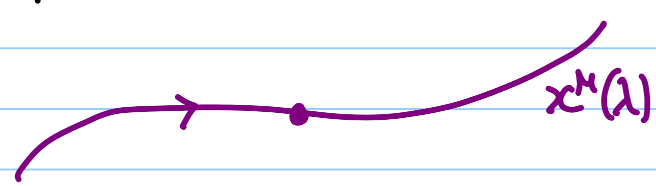
← really big, important idea!!

Introducing a covariant derivative was a great idea. ☺ It allows us to write tensor equations "wherever we go"; "all" we need to do is be sure to write D's (∇'s) rather than ∂'s. But... what rate of change does D measure?

A way to answer this question is to ask when ∇ of some tensor is zero, to get a handle on ∇.

For this, we actually have to specify what path along which we hope to compare tensors – because comparing tensors at two different points is, a priori, meaningless in GR: the spacetime metric differs!

So consider a path $x^\mu(\lambda)$



and define the directional covariant derivative

(7.1) $$\boxed{\frac{D}{d\lambda} \equiv \frac{dx^\mu}{d\lambda}\nabla_\mu}$$

(Carroll's funny notation)

This guy is defined only along the path $x^\mu(\lambda)$.

* We say that a tensor T is <u>parallel-transported</u> along the path $x^\mu(\lambda)$ if

(8.1)
$$\boxed{\left(\frac{D}{d\lambda}T\right)^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = \frac{dx^\sigma}{d\lambda}\nabla_\sigma T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = 0}$$

(a.k.a. "Equation of parallel transport")

(Notice that this is a nice proper tensor eqn because ∇T is a tensor and so is $\frac{dx}{d\lambda}$. ☺)

- Also, since $\nabla g = 0$, <u>parallel transport preserves the inner product of two tensors</u>

e.g. for 2 vectors V & W,

(8.2)
$$\frac{D}{d\lambda}\left(g_{\mu\nu}V^\mu W^\nu\right) = \underbrace{\left(\frac{D}{d\lambda}g_{\mu\nu}\right)}_{=0}V^\mu W^\nu + g_{\mu\nu}\left[\underbrace{\left(\frac{D}{d\lambda}V^\mu\right)}_{=0}W^\nu + V^\mu\underbrace{\left(\frac{D}{d\lambda}W^\nu\right)}_{=0}\right]$$
$$= 0 \quad (\checkmark).$$

This is a <u>crucial</u> property of parallel transport.

When we take an ordinary derivative, we do it

(9.1) by, e.g. taking $\lim_{\Delta x^\mu \to 0} \frac{f(x^\mu + \Delta x^\mu) - f(x^\mu)}{\Delta x^\mu} = \frac{\partial f}{\partial x^\mu}$

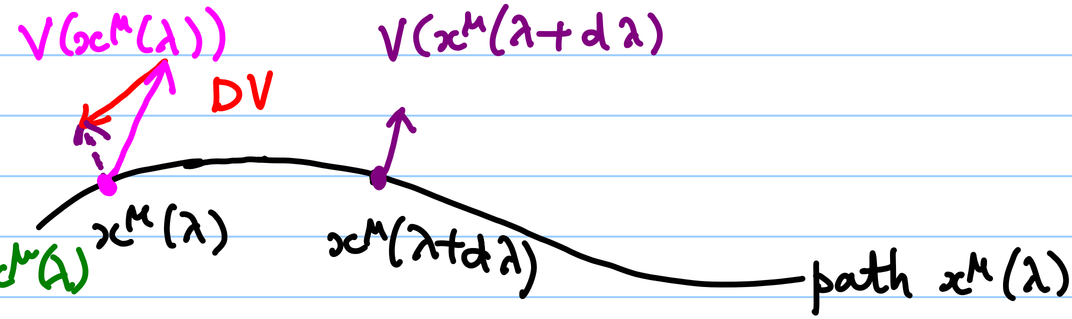
In curved space, the result of this isn't a tensor.
What we do to take the covariant derivative is:

(a) Take our tensor T at "$x^\mu + \Delta x^\mu$" where we measure Δx^μ along the path $x^\mu(\lambda)$: we take $x^\mu(\lambda + \Delta\lambda)$ as our "$x^\mu + \Delta x^\mu$" & find T there.

(b) We parallel-transport T back to $x^\mu(\lambda)$ along $x^\mu(\lambda)$

(c) We compare the parallel-transported-back T to the actual T at $x^\mu(\lambda)$, and we divide by $\Delta\lambda$.
The result is $\frac{D}{d\lambda} T$.

E.g. for a 4-vector V with components V^μ: & $\frac{DV}{d\lambda}$ is the directional covariant derivative along $x^\mu(\lambda)$



(9.2) i.e. $\frac{DV^\nu}{d\lambda} = \frac{dx^\mu}{d\lambda} \nabla_\mu V^\nu$

is the "covariant" rate of change of V^ν w.r.t. λ at the spacetime point $x^\mu(\lambda)$.