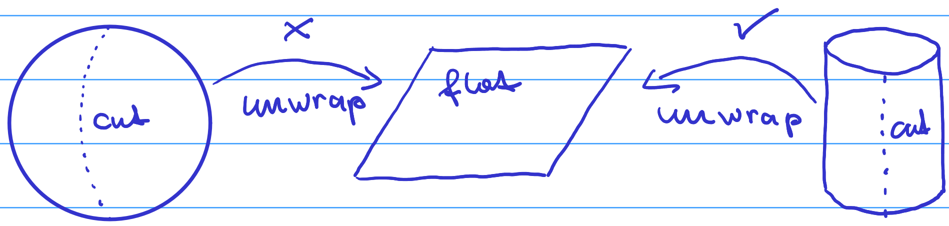
Curvature

- After introducing the concept of a manifold and various tensors and tensor operations on it, we are ready to introduce the measure for its curvature.

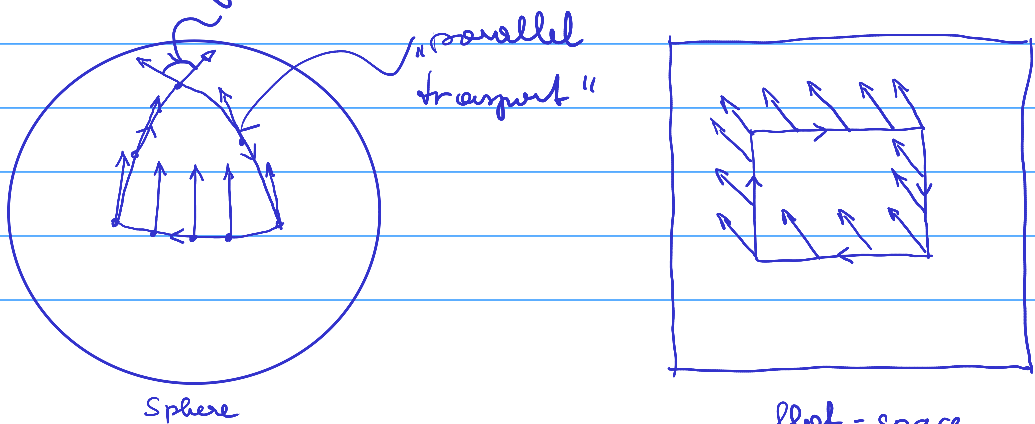
- Naively, one could expect that the metric can do this job and we will indeed show that such a measure can be derived from the metric.

- However, the metric can not directly be used to measure curvature, simply because one can always transform to Riemann normal coordinates where the metric is locally flat!

- Intuitively, a surface is curved if it can't be cut open and unwrapped to a plane:



- The main idea is simple: a space is curved if there is a „deficite angel" after parallel transport on closed loops:



•) Let's first take stock what we need for formalize this:

i) Need a mathematical tool called connection that allows to compare vectors in tangent spaces of nearby points. Unique connection given by Christoffel symbol:

$$\boxed{\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu} \right)}$$

Called „symbol" because it is not a tensor as we will see.

ii) Use of Γ is that it is required to define a generalization of the partial derivative (∂_μ) called covariant derivative:

$$\boxed{\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^{\nu}_{\mu\sigma} V^\sigma \text{, etc.}}$$

iii) Γ also appears in the definition of geodesics, i.e., the generalization of the notion of a straight line:

$$\boxed{\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^{\mu}_{\rho\sigma} \frac{dx^\rho}{d\lambda} \frac{dx^\sigma}{d\lambda} = 0 .}$$

iv) Finally, the mathematical object for the intrinsic measure of curvature is the Riemann tensor defined in terms of Γ:

$$\boxed{R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda} \Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda} \Gamma^{\lambda}_{\mu\sigma} .}$$

$R^{\rho}{}_{\sigma\mu\nu}$ encodes all information about curvature of a manifold. A manifold is flat iff $R^{\rho}{}_{\sigma\mu\nu} = 0$.

•) In the following we will develope the formulas i) - iv).

Covariant derivatives

•) On curved manifolds one needs to generalize the concept of partial derivatives to covariant derivatives:

e.g.: $\partial_\mu V^\nu$ is in general not a covariant expression

$$\partial_{\mu'} V^{\nu'} \neq \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \partial_\mu V^\nu .$$

•) Like any derivative operator, the covariant derivative must satisfy the following rules:

1. linearity: $\nabla(T+S) = \nabla T + \nabla S$
2. Leibnitz rule: $\nabla(T \otimes S) = \nabla T \otimes S + T \otimes \nabla S$
(product rule)

•) The tensor transformation law can be fixed by adding a linear transformation to the partial derivative:

e.g.: $\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\sigma} V^\sigma$

linear trasformation called connection.

•) Since $\Gamma^\nu_{\mu\sigma}$ by definition has to compensate the non-tensorial behavior of ∂_μ, it cannot be a tensor.

•) In order to figure out how $\Gamma^\nu_{\mu\sigma}$ transforms, we demand the desired transformation law for $\nabla_{\mu'} V^{\nu'}$:

$$\nabla_{\mu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \nabla_\mu V^\nu . \quad \text{„(lhs) = (rhs)“}$$

•) Then we write out explicitly the l.h.s:

$$\nabla_{\mu'} V^{\nu'} = \partial_{\mu'} V^{\nu'} + \Gamma^{\nu'}_{\mu'\sigma'} V^{\sigma'} ,$$

and perform the transformation in all terms except $\Gamma^{\nu'}_{\mu'\sigma'}$:

$$= \frac{\partial x^{\mu}}{\partial x^{\mu'}} \partial_{\mu} \left(\frac{\partial x^{\nu'}}{\partial x^{\nu}} V^{\nu} \right) + \Gamma^{\nu'}_{\mu'\sigma'} \frac{\partial x^{\sigma'}}{\partial x^{\sigma}} V^{\sigma}$$

$$= \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \partial_{\mu} V^{\nu} + \frac{\partial x^{\mu}}{\partial x^{\mu'}} V^{\nu} \frac{\partial}{\partial x^{\mu}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} + \Gamma^{\nu'}_{\mu'\sigma'} \frac{\partial x^{\sigma'}}{\partial x^{\sigma}} V^{\sigma} .$$

•) Then we expand the r.h.s:

$$\frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} \nabla_{\mu} V^{\nu} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \partial_{\mu} V^{\nu} + \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \Gamma^{\nu}_{\mu\sigma} V^{\sigma}$$

•) Since lhs = rhs we get:

$$\frac{\partial x^{\mu}}{\partial x^{\mu'}} V^{\nu} \frac{\partial}{\partial x^{\mu}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} + \Gamma^{\nu'}_{\mu'\sigma'} \frac{\partial x^{\sigma'}}{\partial x^{\sigma}} V^{\sigma} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \Gamma^{\nu}_{\mu\sigma} V^{\sigma} ,$$

where the first term on the lhs and rhs cancel.

•) After renaming some dummy indices we get:

$$\Gamma^{\nu'}_{\mu'\lambda'} \frac{\partial x^{\lambda'}}{\partial x^{\lambda}} V^{\lambda} + \frac{\partial x^{\mu}}{\partial x^{\mu'}} V^{\lambda} \frac{\partial}{\partial x^{\mu}} \frac{\partial x^{\nu'}}{\partial x^{\lambda}} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \Gamma^{\nu}_{\mu\lambda} V^{\lambda}$$

•) Since this equation has to hold for each vector we can drop V^{λ}:

$$\frac{\partial x^{\lambda'}}{\partial x^{\lambda}} \Gamma^{\nu'}_{\mu'\lambda'} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \Gamma^{\nu}_{\mu\lambda} - \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^{\mu} \partial x^{\lambda}}$$

- Finally, we isolate the transformed T' on the l.h.s by multiplying with the inverse trafo $\frac{\partial x^\lambda}{\partial x^{\lambda'}}$:

$$T^{\nu'}{}_{\mu'\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^{\nu'}}{\partial x^\nu} T^\nu{}_{\mu\lambda} - \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\lambda} .$$

- This is of course not the tensor transformation law, because of the last term.

- Note that one recovers the tensor transformation law for LT on flat space:

$$\frac{\partial x^\mu}{\partial x^{\mu'}} \to \Lambda^\mu{}_{\mu'} \quad , \text{ since } \frac{\partial}{\partial x^\mu} \underbrace{\Lambda^{\nu'}{}_\lambda}_{\text{const.}} = 0 \quad ,$$

which is why we don't need a connection there.

Covariant derivative of a one-form:

•) A priori there is no reason why connections for vectors and one-forms should be related:

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu + \tilde{\Gamma}^\lambda_{\mu\nu}\omega_\lambda \,.$$

in general $\tilde{\Gamma}^\lambda_{\mu\nu} \neq \Gamma^\lambda_{\mu\nu}$

•) Demand two additional properties to connect $\tilde{\Gamma}$ to Γ:

3) commute with contraction: $\nabla_\mu (T^\lambda{}_{\lambda\rho}) = (\nabla T)_\mu{}^\lambda{}_{\lambda\rho}$,

4) identical to partial der. on scalars: $\nabla_\mu \phi = \partial_\mu \phi \Rightarrow \nabla_\mu \delta^\lambda_\nu = 0$.

•) 3, 4, can not be derived, but are part of the definition of a covariant derivative.

•) Implications of 3, 4:

consider covariant derivative of a scalar $\omega_\lambda V^\lambda$

$$\begin{aligned}\nabla_\mu \omega_\lambda V^\lambda &= V^\lambda \nabla_\mu \omega_\lambda + \omega_\lambda \nabla_\mu V^\lambda \\ &= V^\lambda \partial_\mu \omega_\lambda + V^\lambda \tilde{\Gamma}^\sigma_{\mu\lambda}\omega_\sigma + \omega_\lambda \partial_\mu V^\lambda + \omega_\lambda \Gamma^\lambda_{\mu\sigma}V^\sigma\end{aligned}$$

since $\omega_\lambda V^\lambda$ is a scalar $\nabla_\mu \omega_\lambda V^\lambda = \partial_\mu(\omega_\lambda V^\lambda)$

$$\Rightarrow \cancel{V^\lambda \partial_\mu \omega_\lambda + \omega_\lambda \partial_\mu V^\lambda} = \cancel{V^\lambda \partial_\mu \omega_\lambda + \omega_\lambda \partial_\mu V^\lambda} + V^\lambda \tilde{\Gamma}^\sigma_{\mu\lambda}\omega_\sigma + \omega_\lambda \Gamma^\lambda_{\mu\sigma}V^\sigma$$

$$0 = V^\lambda \tilde{\Gamma}^\sigma_{\mu\lambda}\omega_\sigma + \omega_\lambda \Gamma^\lambda_{\mu\sigma}V^\sigma$$

relabel: $V^\lambda \tilde{\Gamma}^\sigma_{\mu\lambda}\omega_\sigma = -\omega_\sigma \Gamma^\sigma_{\mu\lambda}V^\lambda \Rightarrow \tilde{\Gamma}^\sigma_{\mu\lambda} = -\Gamma^\sigma_{\mu\lambda}$.

•) This means 3,4, establish a relation between $\tilde{\Gamma}$ and Γ:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \,, \quad \nabla_\mu W_\nu = \partial_\mu W_\nu - \Gamma^\lambda_{\mu\nu} W_\lambda$$

•) General rule: upper indices $(+\Gamma)$, lower indices $(-\Gamma)$

$$\nabla_\sigma T^{\mu_1 \mu_2 \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l} = \partial_\sigma T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \mu_l}$$

$$+ \Gamma^{\mu_1}_{\sigma\lambda} T^{\lambda \mu_2 \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l} + \Gamma^{\mu_2}_{\sigma\lambda} T^{\mu_1 \lambda \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l} + \dots$$

$$- \Gamma^\lambda_{\sigma\nu_1} T^{\mu_1 \mu_2 \dots \mu_k}{}_{\lambda \nu_2 \dots \nu_l} + \Gamma^\lambda_{\sigma\nu_2} T^{\mu_1 \mu_2 \dots \mu_k}{}_{\nu_1 \lambda \dots \nu_l} \,.$$

•) Side remark: Some books and papers use comma notation:

$$\nabla_\sigma T^{\mu_1 \mu_2 \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l} = T^{\mu_1 \mu_2 \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l ; \sigma} \,.$$

•) $\Gamma^\lambda_{\mu\nu}$ is called connection since it „transports" vectors from one tangent space to another.

•) The connection is not unique and each one defines its own covariant derivative.

•) However, each metric defines a unique connection, as we will show.

•) The diffrence between two connections is a tensor:

$$\begin{aligned}
S_{\mu'\nu'}{}^{\lambda'} &= \Gamma^{\lambda'}_{\mu'\nu'} - \tilde{\Gamma}^{\lambda'}_{\mu'\nu'} \\
&= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^{\lambda'}}{\partial x^\lambda} \Gamma^\lambda_{\mu\nu} \cancel{- \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial^2 x^{\lambda'}}{\partial x^\mu \partial x^\nu}} - \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^{\lambda'}}{\partial x^\lambda} \tilde{\Gamma}^\lambda_{\mu\nu} \cancel{+ \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial^2 x^{\lambda'}}{\partial x^\mu \partial x^\nu}} \\
&= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^{\lambda'}}{\partial x^\lambda} \left(\Gamma^\lambda_{\mu\nu} - \tilde{\Gamma}^\lambda_{\mu\nu} \right) \\
&= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^{\lambda'}}{\partial x^\lambda} S_{\mu\nu}{}^\lambda \,. \quad \square
\end{aligned}$$

•) Alternatively one can show this via:

$$\nabla_\mu V^\lambda - \tilde{\nabla}_\mu V^\lambda = \cancel{\partial_\mu V^\lambda} + \Gamma^\lambda_{\mu\sigma} V^\sigma - \cancel{\partial_\mu V^\lambda} - \tilde{\Gamma}^\lambda_{\mu\sigma} V^\sigma$$

tensor

$$= (\Gamma^\lambda_{\mu\sigma} - \tilde{\Gamma}^\lambda_{\mu\sigma}) V^\sigma = S_{\mu\sigma}{}^\lambda V^\sigma .$$

$\Rightarrow$ tensor

•) This means one can build from a given connection a new one by adding a (1,2) tensor:

$$\Gamma^\lambda_{\mu\nu} = \tilde{\Gamma}^\lambda_{\mu\nu} + S^\lambda_{\mu\nu}$$

•) Also, one can immediately build a new connection by permuting the lower indices:

$$\tilde{\Gamma}^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu} .$$

•) One can now define a new tensor called torsion tensor:

$$T^\lambda_{\mu\nu} = \Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu} = 2\,\Gamma^\lambda_{[\mu\nu]} .$$

2 Remarks: 1) $T^\lambda_{\mu\nu} = -T^\lambda_{\nu\mu}$ antisymmetric by construction.

2) $T^\lambda_{\mu\nu} = 0 \Leftrightarrow \Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu}$ torsion free connections are symmetric.

•) Now we can define a unique connection on a manifold with a metric from the two conditions:

1) torsion-free: $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu} = \Gamma^\lambda_{(\mu\nu)}$.

2) metric compatibility: $\nabla_\rho g_{\mu\nu} = 0$.

•) Using 1, 2, we can construct the connection in terms of $g_{\mu\nu}$

$$\nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\lambda_{\rho\nu} g_{\mu\lambda} - \Gamma^\lambda_{\rho\mu} g_{\lambda\nu} = 0 \quad (1)$$

$$\nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\lambda_{\mu\nu} g_{\lambda\rho} - \Gamma^\lambda_{\mu\rho} g_{\nu\lambda} = 0 \quad (2)$$

$$\nabla_\nu g_{\rho\mu} = \partial_\nu g_{\rho\mu} - \Gamma^\lambda_{\nu\rho} g_{\lambda\mu} - \Gamma^\lambda_{\nu\mu} g_{\rho\lambda} = 0 \quad (3)$$

(1) - (2) - (3) and using symmetry of $\Gamma^\lambda_{\mu\nu}$ $\Rightarrow$

$$\partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu} + 2\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} = 0 \quad \Big| \cdot g^{\sigma\rho}$$

$$\boxed{\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}\left(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}\right).}$$

This is one of the most important formulas in GR!

•) This specific connection has different names: Levi-Cività connection, Christoffel connection, Christoffel symbol, ...

•) The study of manifolds with metrics and this specific connection is called Riemannian geometry, or Pseudo-Riemannian geometry if the metric has Lorentzian signature.

•) Before putting the covariant derivative to work, let's discuss some of its properties first.

•) You should realize that non-vanishing Γ does not mean that the corresponding space curved!

i) Consider for example the plane in polar coordinates:

$$ds^2 = dr^2 + r^2 d\theta^2 \quad , \quad g^{rr} = 1 \, , \, g^{\theta\theta} = r^{-2} \, .$$

ii) As you will show in the exercises, not all Christoffel components are zero:

$$\Gamma^r{}_{\theta\theta} = -r \, , \, \Gamma^\theta{}_{r\theta} = \Gamma^\theta{}_{\theta r} = \frac{1}{r} \, , \text{ rest is zero.}$$

iii) This is a consequence of using curvi-linear coordinates, but doesn't mean the space is curved, as you can easily see by using (flat) Cartesian coordinates.

•) Another useful property is that the divergence of a vector has a simplified formula for the Christoffel con:

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu{}_{\mu\lambda} V^\lambda = \frac{1}{\sqrt{|g|}} \partial_\mu \left(\sqrt{|g|} \, V^\mu \right) \quad . \text{ (exercise!)}$$

•) Finally, there are derivatives that satisfy the tensor transformation law without the need of a Christoffel connection:

i) The exterior derivative d is a well defined tensor, because Γ would drop out by anti-symmetry:

$$\nabla_{[\mu} \omega_{\nu]} = \partial_{[\mu} \omega_{\nu]} - \Gamma^\lambda_{[\mu\nu]} \omega_\lambda = \partial_{[\mu} \omega_{\nu]} = d\omega \, .$$

ii) The commutator (Lie bracket) of derivatives is a tensor:

$$\mathcal{L}_X (Y^\mu) = [X, Y]^\mu = X^\lambda \partial_\lambda Y^\mu - Y^\lambda \partial_\lambda X^\mu \, . \quad \text{(later!)}$$

Parallel transport

•) Parallel transport is the curved-spacetime generalization of „keeping a tensor constant" as we move it along a path.

•) Flat space: Given a curve $x^\mu(\lambda)$, keeping the components of a tensor constant along the curve means that its directional derivative is zero:

$$\frac{d}{d\lambda} T^{\mu_1 \mu_2 \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l} = \frac{dx^\mu}{d\lambda} \frac{\partial}{\partial x^\mu} T^{\mu_1 \mu_2 \dots \mu_k}{}_{\nu_1 \nu_2 \dots \nu_l} = 0 \quad .$$

•) Curved space: we have to replace ∂_μ by ∇_μ:

$$\frac{D}{d\lambda} := \frac{dx^\mu}{d\lambda} \nabla_\mu \quad .$$

•) Definition of parallel transport of a tensor T along $x^\mu(\lambda)$:

$$\left(\frac{D}{d\lambda} T \right)^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} := \frac{dx^\sigma}{d\lambda} \nabla_\sigma T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = 0 \quad .$$

•) For a vector V^μ it takes the form

$$\frac{d}{d\lambda} V^\mu + \Gamma^\mu{}_{\sigma\rho} \frac{dx^\sigma}{d\lambda} V^\rho = 0 \quad . \qquad (1)$$

•) This first-order differential equation defines an initial value problem:

i) Given a path $x^\mu(\lambda)$, $\lambda \in (\lambda_0, \lambda_1)$ and some initial $V^\mu(x^\mu(\lambda_0))$, a solution of (1) gives a unique continuation of V^μ along $x^\mu(\lambda)$.

•) Notion of parallel transport depends on the connection: different connections lead to different answers.

•) If the connection is metric compatible, then the metric is always parallel transported with respect to it:

$$\frac{D}{d\lambda} g_{\mu\nu} = \frac{dx^\sigma}{d\lambda} \underbrace{\nabla_\sigma g_{\mu\nu}}_{0} = 0 .$$

•) From this follows that the inner product of vectors is preserved:

$$\frac{D}{d\lambda}\left(g_{\mu\nu}V^\mu W^\nu\right) = \left(\frac{D}{d\lambda} g_{\mu\nu}\right)V^\mu W^\nu + g_{\mu\nu}\left(\frac{D}{d\lambda}V^\mu\right)W^\nu + g_{\mu\nu}V^\mu \frac{D}{d\lambda}W^\nu = 0 .$$

$\Rightarrow$ Parallel transport w. metric compatible connection preserves norm of vectors.

Geodesics

•) A geodesic is the curved-space generalization of a straight line in Euclidean space.

•) Naively: „shortest path between two points"
! Not true in general on curved space:
e.g.: two equally straight great circle segments of different length connect any two points on a sphere.

•) Better definition of a straight line in curved space:

A straight line is a path that parallel transports its own tangent vector.

Geodesic equation: $\frac{D}{d\lambda}\frac{dx^\mu}{d\lambda} = \frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0$

•) Alternative notation: $V^\mu \nabla_\mu V^\nu = 0$, $V^\mu = \frac{dx^\mu}{d\lambda}$

•) The geodesic equation can also be derived from extremizing the proper time functional:

$$\tau[x^\mu] = \int \sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\, d\lambda \, .$$

•) For simplicity we assume time like paths $x^2 < 0$, which correspond to minima of the proper time functional and which is why we put the $-$ sign under the square root.

•) In order to vary the fuctional we consider the change of proper time under infinitesimal variations of the path:

$$x^\mu \to x^\mu + \delta x^\mu \, ,$$

and its induced variations of the metric:

$$g_{\mu\nu} \to g_{\mu\nu} + \delta x^\sigma \partial_\sigma g_{\mu\nu} \, .$$

•) Plugging this into the fuctional we obtain:

$$\tau + \delta\tau = \int \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - \partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - 2 g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{d(\delta x^\nu)}{d\lambda}\right)^{1/2} d\lambda$$

$$= \int \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{1/2}\left[1 + \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1}\right.$$

$$\left.\times\left(-\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\delta x^\sigma - 2 g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{d(\delta x^\nu)}{d\lambda}\right)\right]^{1/2} d\lambda$$

•) Since δx^σ is small we can expand the expression in brackets:

$$\delta\tau = \int \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1/2}\left(-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\delta x^\sigma - g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{d(\delta x^\nu)}{d\lambda}\right) d\lambda$$

•) At this point it is useful to change the parametrization from λ to the proper time τ itself, after all this is a free choice and we can pick whatever is convenient!

$$d\lambda = \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1/2} d\tau \ ,$$

$$\Rightarrow \quad \frac{d\tau}{d\lambda} = \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{1/2}, \quad \frac{dx^\mu}{d\lambda} = \frac{dx^\mu}{d\tau}\frac{d\tau}{d\lambda}$$

•) This results in the following simplified form:

$$\delta\tau = \int \left[-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\delta x^\sigma - g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{d(\delta x^\nu)}{d\tau}\right] d\tau$$

•) Next we have to pull out δx^ν from the derivative $\frac{d}{d\tau}$, which we do as usual by integration by parts:

$$\delta\tau = \int \left[-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} + \frac{d}{d\tau}\left(g_{\mu\sigma}\frac{dx^\mu}{d\tau}\right)\right]\delta x^\sigma d\tau + g_{\mu\sigma}\frac{dx^\mu}{d\tau}\delta x^\sigma\Big|_{\tau_1}^{\tau_2}$$

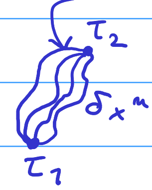
Reminder: $\displaystyle\int_{\tau_1}^{\tau_2} d\tau\, u\cdot v' = -\int_{\tau_1}^{\tau_2} d\tau\, u'v + uv\Big|_{\tau_1}^{\tau_2}$

or: $\displaystyle uv\Big|_{\tau_1}^{\tau_2} = \int_{\tau_1}^{\tau_2} d\tau (uv)' = \int_{\tau_1}^{\tau_2} d\tau\,(u'v + uv')$

$$\Rightarrow \int d\tau\, u'v = uv\Big| - \int d\tau\, uv' \ .$$

•) This is where we assume that the variation at the endpoints of the path vanishes $\delta x^\sigma\Big|_{\tau_1}^{\tau_2} = 0$, and where we can drop the boundary term.

consistent variations



$$\delta\tau = \int \left[-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} + \frac{d}{d\tau}\left(g_{\mu\nu}\frac{dx^\mu}{d\tau}\right)\right]\delta x^\sigma d\tau \ .$$

•) Since we look for stationary points of the functional, we want $\delta\tau$ to vanish for any (BC preserving) variations:

$$\Rightarrow -\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}+\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\partial_\nu g_{\mu\sigma}+g_{\mu\sigma}\frac{d^2x^\mu}{d\tau^2}=0,$$

where we have used $\frac{dg_{\mu\sigma}}{d\tau}=\frac{dx^\nu}{d\tau}\partial_\nu g_{\mu\sigma}$.

•) Reorganizing dummy indices and using the symmetry of $g_{\mu\nu}$, we arrive at the equation:

$$g_{\mu\sigma}\frac{d^2x^\mu}{d\tau^2}+\frac{1}{2}\left(-\partial_\sigma g_{\mu\nu}+\partial_\nu g_{\mu\sigma}+\partial_\mu g_{\nu\sigma}\right)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}=0.$$

•) Finally, multiplying with the inverse metric we get:

$$\frac{d^2x^\rho}{d\tau^2}+\underbrace{\frac{1}{2}g^{\rho\sigma}\left(\partial_\mu g_{\nu\sigma}+\partial_\nu g_{\sigma\mu}-\partial_\sigma g_{\mu\nu}\right)}_{\Gamma^\rho_{\mu\nu}}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}=0,$$

where we now recognize the definition of the Christoffel symbol.

$$\Rightarrow \frac{d^2x^\mu}{d\tau^2}+\Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau}=0.$$

•) In this derivation we made a specific choice for the parametrization, where we use the proper time as parameter. However this is not the most general choice!

•) There is a set of parameters called affine parameters that lead to the same form of the geodesic equation and which are related to τ by:

$$\tau\to\lambda=a\tau+b,\quad a,b=\text{const}.$$

•) The most general (non-affine) form of the geodesic equation is given by :

$$\frac{d^2x^\mu}{d\alpha^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\alpha}\frac{dx^\sigma}{d\alpha} = f(\alpha)\frac{dx^\mu}{d\alpha} \, ,$$

where $f(\alpha)$ is some function. (see exercise)