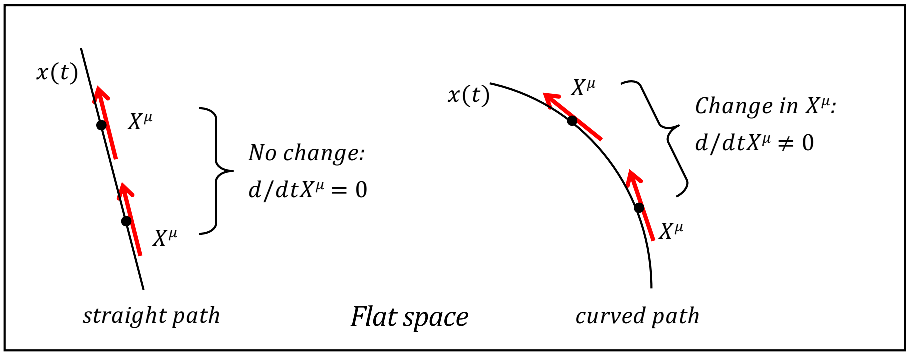
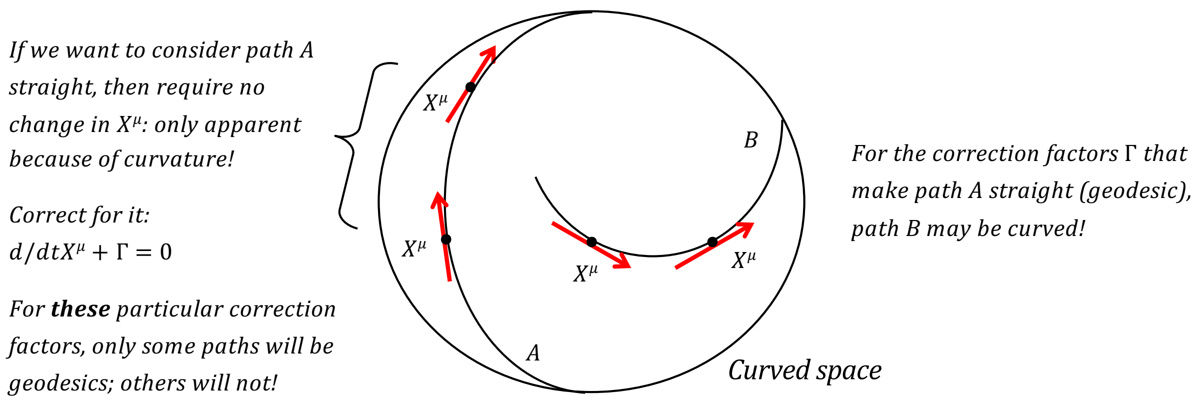
1. The Covariant Derivative (the general derivative for a curved space).

(a) $d^2x^\mu/dt^2 = d/dt(dx^\mu/dt)$ is the change of the **tangent vector** dx^μ/dt, call it X^μ, to the path $x(t)$ at different values of t:



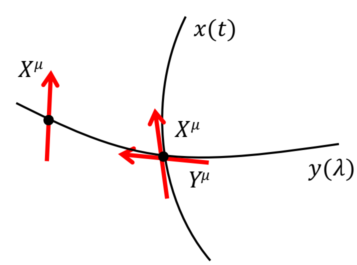


<u>Now</u>: What is the particular form of the correction Γ?

<u>First Note</u>:

$d/dtX^\mu + \Gamma = (X^\nu d/dx^\nu)X^\mu + \Gamma$
Change of X^μ in its own direction; *i.e.*, along the path $x(t)$ it's tangent to. Or, the "*directional derivative* of X^μ in the direction X^μ".

$d/dt = (dx^\nu/dt)(d/dx^\nu)$



$d/d\lambda X^\mu + \Gamma = (Y^\nu d/dy^\nu)X^\mu + \Gamma$
Change of X^μ in direction of $Y^\nu = dy^\nu/d\lambda$; *i.e.*, along the path $y(\lambda)$. Or, the "*directional derivative* of X^μ in the direction Y^ν".

$d/d\lambda = (dy^\nu/d\lambda)(d/dy^\nu)$

Moving X^μ in the direction Y^μ and measuring how it changes.

<u>So</u>: In a curved space, the *directional derivative* of a vector X^μ in the direction Y^ν must be "corrected". Write it as:

$$Y^\nu \nabla_\nu X^\mu = (Y^\nu \partial_\nu)X^\mu + \Gamma^\mu_{\nu\sigma} Y^\nu X^\sigma$$

← *Require the correction to depend on both X^μ and Y^ν.*

For $Y^\nu = X^\nu$, and setting the directional derivative to zero, we get the geodesic equation!

This motivates the definition for the "covariant derivative" ∇_ν (the general derivative for a curved space): →

$$\nabla_\nu X^\mu = \partial_\nu X^\mu + \Gamma^\mu_{\nu\sigma} X^\sigma$$

The derivative of X^μ in a curved space = *The derivative of X^μ in flat space* + *Correction factors*

(b) For any curved space, there may be many ways to construct "correction factors" Γ (technically called "connection coefficients"). But if the space has a metric $g_{\mu\nu}$, there is one particular set of correction factors associated with $g_{\mu\nu}$ (the "metric compatible connection coefficients"). These are defined by the condition $\nabla_\sigma g_{\mu\nu} = 0$. You can solve for the Γ's in this equation explicitly in terms of the values of $g_{\mu\nu}$. (Intuitively, when the space has a metric, you can measure the length of paths; hence you can "minimize" path lengths to find paths that are the "shortest distance" between points, which you can then identify as geodesics.)

(c) If the correction factors $\Gamma = 0$, then we have the flat space case. But even in flat spaces, there may be correction factors; and this may arise because of the use of kooky coordinates to label paths (*e.g.*, spherical coordinates instead of Cartesian). So just because $\Gamma \neq 0$, doesn't necessarily mean our space is curved.

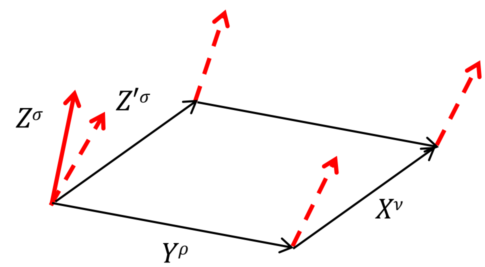
2. The Curvature Tensor

One can define a 4-indexed quantity, the **curvature tensor** $R^\sigma_{\mu\nu\rho}$ that acts on three vectors X^ν, Y^ρ, Z^μ and outputs the amount of change experienced by Z^μ upon parallel-transport around an infinitesimal curve defined by X^ν and Y^ρ:

$$R^\sigma_{\mu\nu\rho} X^\nu Y^\rho Z^\mu = \delta Z^\sigma$$

$$\delta Z^\sigma = Z^\sigma - Z'^\sigma$$

$= \textit{change in } Z^\sigma \textit{ upon parallel transport around loop defined by } X^\nu \text{ and } Y^\rho.$



(a) $R^\sigma_{\mu\nu\rho} = 0$ *if and only if* the space is flat.
(b) $R^\sigma_{\mu\nu\rho}$ depends explicitly on the metric $g_{\mu\nu}$.

The curvature tensor is officially defined by:

$$R^\sigma_{\mu\nu\rho} X^\nu Y^\rho Z^\mu = \underbrace{X^\nu \nabla_\nu (\underbrace{Y^\rho \nabla_\rho Z^\mu}_{\text{change in } Z^\mu \text{ along } Y^\rho})}_{\text{then change in the result along } X^\nu} - \underbrace{Y^\nu \nabla_\nu (\underbrace{X^\rho \nabla_\rho Z^\mu}_{\text{change in } Z^\mu \text{ along } X^\nu})}_{\text{then change in the result along } Y^\rho}$$

For "metric-compatible connections", $R^\sigma_{\mu\nu\rho}$ depends explicitly on first derivatives of the metric. So, since the Minkowski metric is constant ($= diag(-1, 1, 1, 1)$), the curvature tensor for Minkowski spacetime is zero: Minkowski spacetime is flat!

For convenience sake, write $X^\nu \nabla_\nu$ as ∇_X (the deriviative in the X^μ direction). Then the above becomes:

$$R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z$$

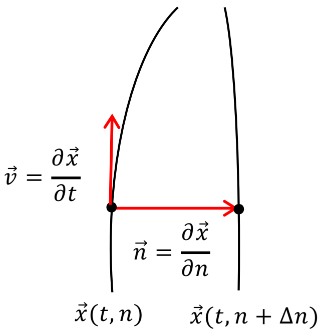
3. Motivating the Einstein Equations.

First, compare:

(a) The relative acceleration between 2 massive objects, according to Newtonian gravity.

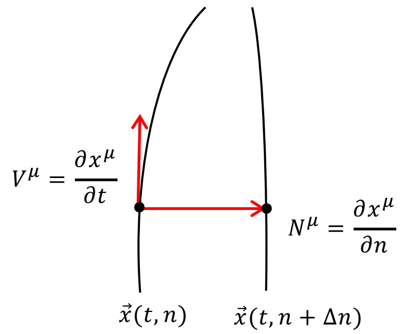
(b) The rate at which parallel geodesics deviate in a curved space.

(a)



$\vec{n}$ = *separation distance*

(b)



$$\begin{pmatrix} \text{rate of change in } \vec{n} \\ \text{in } \vec{v}\text{-direction} \end{pmatrix} = \left(\vec{v} \cdot \frac{\partial}{\partial \vec{x}}\right)\left(\vec{v} \cdot \frac{\partial}{\partial \vec{x}}\right)\vec{n} = \frac{\partial^2 \vec{n}}{\partial t^2}$$

Then:

$$\frac{\partial^2 \vec{n}}{\partial t^2} = \frac{\partial^2}{\partial t^2}\frac{\partial \vec{x}}{\partial n} = \frac{\partial}{\partial n}\frac{\partial^2 \vec{x}}{\partial t^2} = \frac{\partial}{\partial n}\left(-\frac{\partial \Phi}{\partial \vec{x}}\right)$$

$$= \left(\frac{\partial \vec{x}}{\partial n} \cdot \frac{\partial}{\partial \vec{x}}\right)\left(-\frac{\partial \Phi}{\partial \vec{x}}\right)$$

$$= -(\vec{n} \cdot \vec{\partial})\vec{\partial}\Phi$$

$$\begin{pmatrix} \text{rate of change in } N^\mu \\ \text{in } V^\mu \text{ direction} \end{pmatrix} = \nabla_V(\nabla_V N)$$

Require:

(1) $\nabla_V N = \nabla_N V$ (the "angle" between V, N remains constant).

(2) $\nabla_V V = 0$ (*i.e.*, the paths are geodesics).

Then:

$$\begin{aligned} \nabla_V(\nabla_V N) &= \nabla_V(\nabla_N V) \\ &= \nabla_N(\nabla_V V) - R(N, V, V) \\ &= -R(V, N, V) \end{aligned}$$

compare!

Suggests:

$R^\sigma_{\mu\nu\rho} V^\mu V^\nu \leftrightarrow V^\mu \nabla_\mu \Phi = 4\pi G T_{\mu\nu} V^\mu V^\nu$ where $V^\mu \nabla_\mu \Phi = 4\pi G\rho$ ← *Poissons's field equation for Newtonian gravity, ρ = mass density*

and $T_{\mu\nu} V^\mu V^\nu = \rho$

To cancel the $V^\mu V^\nu$ terms on both sides, and to get a $\mu\nu$-indexed object on the left, take the partial trace of $R^\sigma_{\mu\nu\rho}$. Define the "Ricci tensor" $R_{\mu\nu} \equiv R^\sigma_{\mu\nu\rho} = R^0_{\mu\nu 0} + R^1_{\mu\nu 1} + R^2_{\mu\nu 2} + R^3_{\mu\nu 3}$.

Then we have: $R_{\mu\nu} = 4\pi G T_{\mu\nu}$ $(*)$ ← *proposed by Einstein in 1915*

But! It's a mathematical property of $R^\sigma_{\mu\nu\rho}$ that,

$$\nabla_\sigma(R_{\mu\nu} - ½R g_{\mu\nu}) = 0$$

← *where the "Ricci scalar" $R \equiv \mathrm{Trace}(R_{\mu\nu}) = R_{00} + R_{11} + R_{22} + R_{33}$*

So: Since $\nabla_\sigma T_{\mu\nu} = 0$, equation $(*)$ then entails $\nabla_\sigma R = \nabla_\sigma T = 0$, where $T = \mathrm{Trace}(T_{\mu\nu})$; and this means mass-density is constant, which is unphysical!

Solution: Instead of equation $(*)$, write

$$R_{\mu\nu} - ½R g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

or ← *Take trace of both sides to get $R = -8\pi G T$*

$$R_{\mu\nu} = 8\pi G(T_{\mu\nu} - ½T g_{\mu\nu})$$

For very slow speeds (Newtonian limit), $T_{\mu\nu}$ and $T \approx \rho$, hence we get back Poisson's equation!