RELATIVITY & GRAVITATION

Lectured by Timothy Clifton

SECTION 1 - GEOMETRY AND SPACETIME

1.1 Manifolds

Our starting point for discussing GR is a 4D topological space known as a manifold. There is a technical mathematical definition for a manifold, but for our purposes we can think of it as a blank canvas. We will add structure on top until it can be described as a spacetime.

1.2 Coordinates

Coordinates are labels that we use to denote the positions of points in the manifold. In a 4D manifold we require the use of 4 coordinates to specify a single point $\{x, y, z, t\}$.

It is important to understand that coordinates, in general, have no physical significance. In particular an observable quantity should be independent of our choice of coordinates. We should be able to change coordinates, or redo a derivation in different coordinates, and any physical observable that results should be entirely unchanged.

1.3 Curves and surfaces

A curve is a 1D object that exists within a manifold. It can be described by

$$x^\mu = x^\mu(\lambda)$$

where $\mu = 0, 1, 2, 3$. The x^μ are the coordinates on the manifold and λ is a parameter that labels positions along the curve. Now consider a 3D subspace of the manifold, with coordinates $\{y^1, y^2, y^3\}$. The position of this space is then given by

$$x^\mu = x^\mu(y^1, y^2, y^3)$$

This is 4 equations of three variables. We can therefore eliminate y^1, y^2 and y^3 to leave us with a single equation

$$S(x^0, x^1, x^2, x^3) = 0$$

This one equation is sufficient to specify the position of the space, i.e. given any values for 3 of the x^μ, we can solve it to find the fourth.

1.4 Tangent vectors

Consider a point P on a curve C in a manifold M

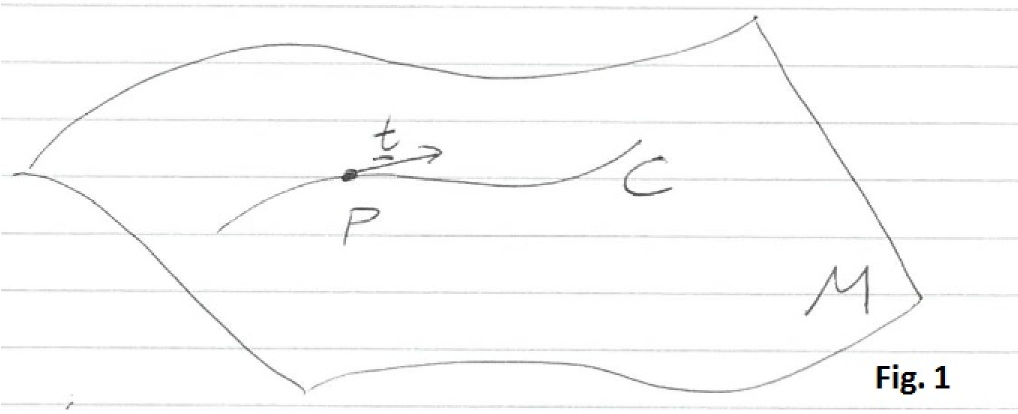

Fig. 1

The tangent vector $\vec{t}$ at P is defined by

$$\vec{t} \equiv \lim_{\delta\lambda \to 0} \frac{\delta\vec{s}}{\delta\lambda} = \frac{d\vec{s}}{d\lambda}$$

where $\delta\vec{s}$ is the infinitesimal separation vector between point P and some nearby point on the curve with parameter value $\lambda + \delta\lambda$

The tangent vector $\vec{t}$ only touches the manifold M at point P. Note that $\vec{t}$ exists independent of any choice of coordinates.

Example: in 2D Euclidean space consider the curve

$$x = \lambda\,, \quad y = \lambda^2$$

the tangent vector at $\lambda = 0$ is

$$\vec{t}_0 = \frac{dx}{d\lambda}\Big|_{\lambda=0}\hat{\vec{x}} + \frac{dy}{d\lambda}\Big|_{\lambda=0}\hat{\vec{y}} = 1 \times \hat{\vec{x}} + 0 \times \hat{\vec{y}} = \hat{\vec{x}}$$

At $\lambda = 1$

$$\vec{t}_1 = \frac{dx}{d\lambda}\Big|_{\lambda=1}\hat{\vec{x}} + \frac{dy}{d\lambda}\Big|_{\lambda=1}\hat{\vec{y}} = \hat{\vec{x}} + 2\hat{\vec{y}}$$

This curve could just as well exist in a curved 2D space, as illustrated in the figure below.

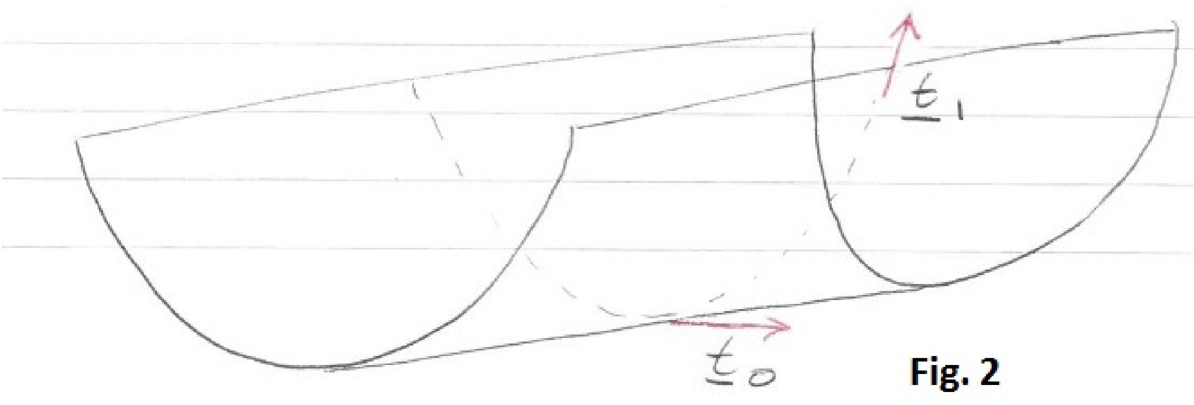

Fig. 2

1.5 Tangent spaces

At any point P in a curved space we can arrange a Euclidean plane such that it tangential to the curved space, e.g. imagine a flat solid sheet balanced on the surface of a ball. Any vector that lies in this plane is a tangent vector. In general the dimension of the tangent space at a point P which we call T_P must have the same dimension as the manifold $\mathcal{M}$. The tangent vector to any curve that passes through P can be drawn as an arrow in T_P.

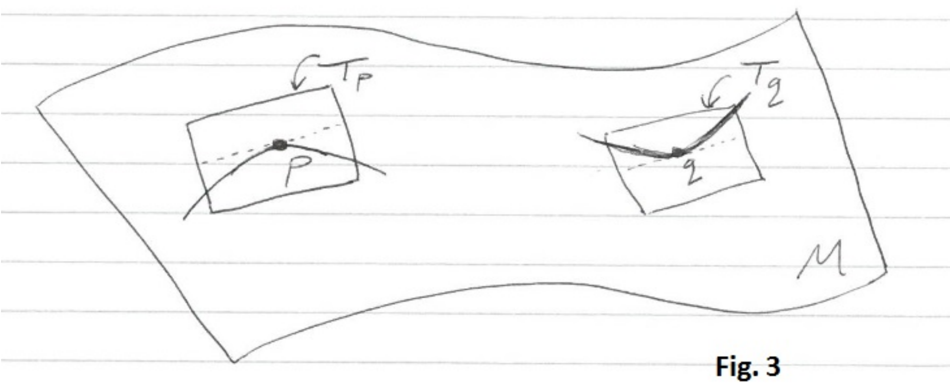

Fig. 3

If the geometry of $\mathcal{M}$ is sufficiently smooth (i.e. no sudden jumps) then the neighbourhood of space around the point P can be approximated by the Euclidean geometry of T_P at that point. Note that in GR the tangent spaces should be $4D$ Minkowski spaces, so that in the neighbourhood of P where the geometry of space is close to the geometry of the tangent space then we recover Special Relativity (approximately up to tidal effects).

1.6 Vectors and vector fields

Now that we have tangent spaces, we can define vectors and vector fields (two different but related concepts).

Let us give the definition of a vector. A vector $\vec{v}$ at point P is an object that exists in the tangent space T_P and that obeys the rules of vector addition and multiplication with any other vector in that same space.

As well as vector let's give the definition of a vector field. A vector field $\vec{v}(x^\mu)$ is the assignment of a vector to each point in the manifold such that $\vec{v}(x^\mu)$ evaluated at any point P is a vector that exists in the tangent space T_P.

Note that vectors at two different points in a curved space exist in two different tangent spaces, this means that they cannot be directly compared (unlike the corresponding case of flat space)

1.7 Basis vectors

In order to do calculations, it is often useful to introduce a set of "basis vectors" $\vec{e}_a$ at each point P, if these basis vectors are linearly independent, then we can use them to write any vector $\vec{v}$ in the following way

$$\vec{v} = v^a \vec{e}_a$$

where a has the value of $0, 1, 2, 3$. Different tangent spaces have different basis vectors, whereas in the same space you can assign the same basis vectors. Where v^a are the contravariant components of the vector $\vec{v}$ in the basis $\vec{e}_a$. In a 4D manifold require 4 basis vectors to form a complete set, one of which will have a timelike direction.

A particularly useful choice of basis vectors is the coordinate basis defined by

$$\vec{e}_\mu \equiv \lim_{\delta x^\mu \to 0} \frac{\delta \vec{s}}{\delta x^\mu} = \frac{\partial \vec{s}}{\partial x^\mu}$$

where $\partial \vec{s}$ is the vector displacement between p and a nearby point q, whose coordinate separation from p is δx^μ. If we were to choose the curve that connects p and q to vary in only one coordinate then we see that $\vec{e}_\mu$ is the tangent vector to that curve. This isn't the only choice of basis, however it does make certain calculations easier, but is not necessarily the easiest for calculating observables.

From this definition, we can write the following

$$d\vec{s} = \frac{\partial \vec{s}}{\partial x^\mu} dx^\mu = \vec{e}_\mu dx^\mu$$

If we want to use this to find the magnitude of the displacement of the vector between points p and q, we take the inner product of $d\vec{s}$ with itself, which gives

$$ds^2 = d\vec{s} \cdot d\vec{s} = (\vec{e}_\mu dx^\mu) \cdot (\vec{e}_\nu dx^\nu) = (\vec{e}_\mu \cdot \vec{e}_\nu) dx^\mu dx^\nu$$

If we now promote these basis vectors to a set of basis vector fields then we can define the metric. The metric, which is a function of coordinates is defined as the inner product of basis vector fields at the point x^σ

$$g_{\mu\nu}(x^\sigma) = \vec{e}_\mu(x^\sigma) \cdot \vec{e}_\nu(x^\sigma)$$

Because the inner product is symmetric this means that $g_{\mu\nu} = g_{\nu\mu}$. This expression gives the magnitude squared along the hypothetical curve C that we have been considering, which is the square of the magnitude of the infinitesimal separation between points p and q. Having said that the distance cannot depend on our choice of coordinates. So the general expression for displacement must be given by

$$\boxed{ds^2 = g_{\mu\nu}dx^\mu dx^\nu}$$

for <u>any</u> choice of coordinates and for <u>any</u> curve C. Note that any arbitrary basis vectors can always be written in terms of coordinate vectors in the following way

$$\vec{e}_a = e_a{}^\mu \vec{e}_\mu$$

where $e_a{}^\mu$ are the contravariant components.

Similarly, we could choose to express the coordinate basis vectors in any other system by writing

$$\vec{e}_\mu = e_\mu{}^a \vec{e}_a$$

Note that substituting one of these expression into the other gives

$$e^a{}_\mu e^\mu{}_b = \delta^a{}_b$$

and

$$e^\mu{}_a e^a{}_\nu = \delta^\mu{}_\nu \,.$$

1.8 The Metric

The metric $g_{\mu\nu}$ can be used to calculate the lengths of curves, as well as the angles at which they meet. The first of these properties follows from integrating $|d\vec{s}|$ along the curve C

$$S = \int_C |d\vec{s}| = \int_C \sqrt{|g_{\mu\nu} dx^\mu dx^\nu|}$$

To go further, in order to write this in terms of the parameter λ along the curve C, we will use the tangent vector expressed in a coordinate basis

$$\vec{t} \equiv \frac{ds}{d\lambda} = \frac{dx^\mu}{d\lambda}\vec{e}_\mu = t^\mu \vec{e}_\mu$$

which gives $dx^\mu = t^\mu d\lambda$ and hence

$$\boxed{S = \int_C \sqrt{|g_{\mu\nu} t^\mu t^\nu|} d\lambda}$$

Note that it is possible to choose this parameter λ such that $g_{\mu\nu} t^\mu t^\nu = \pm 1$. The plus or minus depends on whether C is a spacelike or timelike curve. If I've chosen λ such that this is true then λ measures distance along the curve directly as $S = \int_C d\lambda$. If λ satisfies this requirement, then it is known as an <u>affine parameter</u>.

Example: what is the length of the curve with the following tangent vector, between two points labelled by parameter choices $\lambda = 0$ and $\lambda = 1$,

$$\vec{t} = t^\nu \vec{e}_\nu = t^x \vec{x} + t^y \vec{y} = \sin\lambda \vec{x} + \cos\lambda \vec{y} \qquad ?$$

You may assume that the curve exists in flat Euclidean space.

Solution: the equation that we use to measure the length is

$$S = \int_C \sqrt{g_{\mu\nu} t^\mu t^\nu} d\lambda$$

recall that $g_{\mu\nu} = \delta_{\mu\nu}$ in Euclidean space

$$\Rightarrow g_{\mu\nu} t^\mu t^\nu = \delta_{xx} t^x t^x + \delta_{yy} t^y t^y = (t^x)^2 + (t^y)^2 = \sin^2\lambda + \cos^2\lambda = 1$$

so $S = \int_0^1 d\lambda = 1$, in this case.

In GR we take the length of a timelike curve to correspond to "proper time" and the length of a spacelike curve to correspond to "proper distance". These are the measures of time and distance that clocks and measuring tapes would record if they were following C.

To use the metric to infer angles between vectors we note that it acts as the inner product between two vectors in the same tangent space

$$\vec{v} \cdot \vec{w} = (v^\mu \vec{e}_\mu) \cdot (w^\nu \vec{e}_\nu) = v^\mu w^\nu (\vec{e}_\mu \cdot \vec{e}_\nu) = v^\mu w^\nu g_{\mu\nu}$$

As the tangent space has a flat geometry, the angle between the vector $\vec{v}$ and the vector $\vec{w}$ is given by the usual expression

$$\vec{v} \cdot \vec{w} = |\vec{v}||\vec{w}| \cos\theta \quad \Rightarrow \quad \cos\theta = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}||\vec{w}|} = \frac{g_{\mu\nu} v^\mu w^\nu}{\sqrt{g_{\mu\nu} v^\mu v^\nu} \sqrt{g_{\mu\nu} w^\mu w^\nu}}$$

If $\vec{v}$ and $\vec{w}$ are both unit vectors that satisfy such that $|\vec{v}| = |\vec{w}| = 1$ then this expression is simply

$$\boxed{\cos\theta = g_{\mu\nu} v^\nu w^\mu}$$

1.9 Dual basis

For any set of vectors $\vec{e}_a$, we can define a second set by the equation

$$\vec{e}^a \cdot \vec{e}_b = \delta^a{}_b$$

As before, we can expand any arbitrary vector in terms of this new dual basis

$$\vec{v} = v_a \vec{e}^a$$

and if these new vectors are dual to a coordinate basis, $\vec{e}_\mu$, then we can use them to define a new metric

$$g^{\mu\nu} = \vec{e}^\mu \cdot \vec{e}^\nu$$

Strictly speaking, the dual basis vectors live in what's called a dual tangent space, but this distinction is not important here. Let's consider the different ways we can express the inner product, in terms of $\vec{e}_\mu$ and $\vec{e}^\mu$

$$\vec{v} \cdot \vec{w} = (v^\mu \vec{e}_\mu) \cdot (w^\nu \vec{e}_\nu) = (\vec{e}_\mu \cdot \vec{e}_\nu) v^\mu w^\nu = g_{\mu\nu} v^\mu w^\nu$$

$$= (v_\mu \vec{e}^\mu) \cdot (w^\nu \vec{e}_\nu) = (\vec{e}^\mu \cdot \vec{e}_\nu) v_\mu w^\nu = \delta^\mu{}_\nu v_\mu w^\nu$$

$$= (v^\mu \vec{e}_\mu) \cdot (w_\nu \vec{e}^\nu) = (\vec{e}_\mu \cdot \vec{e}^\nu) v^\mu w_\nu = \delta_\mu{}^\nu v^\mu w_\nu$$

$$= (v_\mu \vec{e}^\mu) \cdot (w_\nu \vec{e}^\nu) = (\vec{e}^\mu \cdot \vec{e}^\nu) v_\mu w_\nu = g^{\mu\nu} v_\mu w_\nu$$

These equations show

$$g_{\mu\nu} v^\mu w^\nu = v_\nu w^\nu = v^\nu w_\nu = g^{\mu\nu} v_\mu w_\nu$$

From which we can infer

$$w_\nu = g_{\mu\nu} w^\nu, \qquad w^\mu = g^{\mu\nu} w_\nu$$

as well as $g_{\mu\nu}g^{\nu\rho} = \delta_\mu{}^\rho$. The $g^{\mu\nu}$ and $g_{\mu\nu}$ are therefore inverses of each other, and raise and lower coordinate indices. Finally, we can note that that the $g_{\mu\nu}$ and $g^{\mu\nu}$ also relate $\vec{e}_\mu$ and $\vec{e}^\mu$, as follows:

$$\vec{v} = v_\mu \vec{e}^\mu = g_{\mu\nu} v^\nu \vec{e}^\mu = v^\nu \vec{e}_\nu$$

and

$$\vec{v} = v^\mu \vec{e}_\mu = g^{\mu\nu} v_\nu \vec{e}_\mu = v_\nu \vec{e}^\nu$$

$$\Rightarrow \vec{e}_\nu = g_{\mu\nu} \vec{e}^\mu$$

and

$$\vec{e}^\nu = g^{\mu\nu} \vec{e}_\mu$$

The $g^{\mu\nu}$ and $g_{\mu\nu}$ and therefore also raise and lower indices on the basis vectors themselves.

Example: by considering the different ways it is possible to expand the inner product of $\vec{e}_\mu$ and $\vec{e}^\mu$, we can prove the following

$$g_{\mu\nu} g^{\mu\sigma} = \delta_\nu{}^\sigma$$

Solution: to show this, let's start with

$$\vec{v} \cdot \vec{w} = (v_\mu \vec{e}^\mu) \cdot (w_\nu \vec{e}^\nu) = v_\mu w_\nu (\vec{e}^\mu \cdot \vec{e}^\nu) = v_\mu w_\nu g^{\mu\nu}$$

and

$$\vec{v} \cdot \vec{w} = (v^\mu \vec{e}_\mu) \cdot (w_\nu \vec{e}^\nu) = v^\mu w_\nu (\vec{e}_\mu \cdot \vec{e}^\nu) = \delta_\mu{}^\nu v^\mu w_\nu$$

$$\Rightarrow g^{\mu\nu} v_\mu = \delta_\mu{}^\nu v^\mu = v^\nu$$

also

$$\vec{v} \cdot \vec{w} = (v^\mu \vec{e}_\mu) \cdot (w^v \vec{e}_\nu) = (\vec{e}_\mu \cdot \vec{e}_\nu) v^\mu w^\nu = g_{\mu\nu} v^\mu w^\nu$$

and

$$\vec{v} \cdot \vec{w} = (v_\mu \vec{e}^\mu) \cdot (w^\nu \vec{e}_\nu) = (\vec{e}^\mu \cdot \vec{e}_\nu) v_\mu w^\nu = \delta^\mu{}_\nu v_\mu w^\nu$$

$$\Rightarrow g_{\mu\nu} v^\nu = \delta^\mu{}_\nu v_\mu = v_\nu$$

Now substitute these results into each other to get

$$v^\nu = g^{\mu\nu} v_\mu = g^{\mu\nu} g_{\mu\sigma} v^\sigma = \delta^\nu{}_\sigma v^\sigma$$

$$\Rightarrow \delta^\nu{}_\sigma = g^{\mu\nu} g_{\mu\sigma}$$

1.10 Coordinate transformations

Consider $\vec{e}_\mu$ under a change of coordinates from x^μ to x'^μ

$$\vec{e}_\mu \equiv \frac{\partial \vec{s}}{\partial x^\mu} = \frac{\partial x'^\nu}{\partial x^\mu}\frac{\partial \vec{s}}{\partial x'^\nu} = \frac{\partial x'^\nu}{\partial x^\mu}\vec{e}\,'_\nu$$

This is called a linear transformation. Now to maintain $\vec{e}^\mu \cdot \vec{e}_\nu = \delta^\mu{}_\nu$ we require

$$\vec{e}^\mu = \frac{\partial x^\mu}{\partial x'^\nu}\vec{e}\,'^\nu$$

$$\Rightarrow \vec{e}^\mu \cdot \vec{e}_\nu = \left(\frac{\partial x^\mu}{\partial x'^\rho}\vec{e}\,'^\rho\right) \cdot \left(\frac{\partial x'^\sigma}{\partial x^\nu}\vec{e}\,'_\sigma\right) = \frac{\partial x^\mu}{\partial x'^\rho}\frac{\partial x'^\sigma}{\partial x^\nu}\delta^\rho{}_\sigma$$

$$= \frac{\partial x^\mu}{\partial x'^\rho}\frac{\partial x'^\rho}{\partial x^\nu} = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu{}_\nu$$

This immediately gives the transformation law for $g_{\mu\nu}$ and $g^{\mu\nu}$:

$$g_{\mu\nu} \equiv \vec{e}_\mu \cdot \vec{e}_\nu = \left(\frac{\partial x'^\rho}{\partial x^\mu}\vec{e}\,'_\rho\right) \cdot \left(\frac{\partial x'^\sigma}{\partial x^\nu}\vec{e}\,'_\sigma\right) = \frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial x'^\sigma}{\partial x^\nu}g'_{\rho\sigma}$$

$$g^{\mu\nu} \equiv \vec{e}^\mu \cdot \vec{e}^\nu = \left(\frac{\partial x^\mu}{\partial x'^\rho}\vec{e}\,'^\rho\right) \cdot \left(\frac{\partial x^\nu}{\partial x'^\sigma}\vec{e}\,'^\sigma\right) = \frac{\partial x^\mu}{\partial x'^\rho}\frac{\partial x^\nu}{\partial x'^\sigma}g'^{\rho\sigma}$$

Finally, let's derive the transformation laws for the coordinate components of $\vec{v}$:

$$\vec{e}^\mu \cdot \vec{v} = \vec{e}^\mu \cdot (v^\nu \vec{e}_\nu) = v^\nu(\vec{e}^\mu \cdot \vec{e}_\nu) = v^\nu \delta^\mu{}_\nu = v^\mu$$

$$= \frac{\partial x^\mu}{\partial x'^\rho} \vec{e}'^\rho \cdot \vec{v} = \frac{\partial x^\mu}{\partial x'^\rho} \vec{e}'^\rho \cdot (v'^\sigma \vec{e}'_\sigma) = \frac{\partial x^\mu}{\partial x'^\rho} v'^\sigma \delta^\rho{}_\sigma = \frac{\partial x^\mu}{\partial x'^\rho} v'^\rho$$

Similarly, $\vec{e}_\nu \cdot \vec{v} \quad \Rightarrow \quad v_\nu = \frac{\partial x'^\nu}{\partial x^\mu} v'_\nu$

General rule: a raised index transforms as

$$X^\mu = \frac{\partial x^\mu}{\partial x'^\nu} X'^\nu$$

and a lowered index transforms as

$$X_\nu = \frac{\partial x'^\mu}{\partial x^\mu} X'_\nu$$

Example: explain how considering the inner product $\vec{e}_\mu \cdot \vec{v}$ leads to the transformation law

$$v_\mu = \frac{\partial x'^\nu}{\partial x^\mu} v'_\nu$$

Solution: let's start with

$$\vec{e}_\mu \cdot \vec{v} = \vec{e}_\mu \cdot (v_\nu \vec{e}^\nu) = v_\nu (\vec{e}_\mu \cdot \vec{e}^\nu) = v_\nu \delta_\mu{}^\nu = v_\mu$$

and

$$\vec{e}'_\mu \cdot \vec{v} = \vec{e}'_\mu \cdot (v'_\nu \vec{e}'^\nu) = v'_\nu \delta_\mu{}^\nu = v'_\mu$$

now

$$\vec{e}'_\mu \equiv \frac{\partial \vec{s}}{\partial x'^\mu}, \quad \vec{e}_\mu \equiv \frac{\partial \vec{s}}{\partial x^\mu} = \frac{\partial x'^\nu}{\partial x^\mu}\frac{\partial \vec{s}}{\partial x'^\nu} = \frac{\partial x'^\nu}{\partial x^\mu}\vec{e}'_\nu$$

so

$$\vec{e}_\mu \cdot \vec{v} = \left(\frac{\partial x'^\nu}{\partial x^\mu}\vec{e}'_\nu\right) \cdot \vec{v} = \frac{\partial x'^\nu}{\partial x^\mu}(\vec{e}'_\nu \cdot \vec{v}) = \frac{\partial x'^\nu}{\partial x^\mu}v'_\mu$$

which gives

$$\Rightarrow \vec{e}_\mu \cdot \vec{v} = v_\mu = \frac{\partial x'^\nu}{\partial x^\mu}v'_\nu$$

1.11 Outer product

We can choose to think of the inner product between two vectors as the action of one of those vectors on the other:

$$\vec{u}(\vec{v}) \equiv \vec{u} \cdot \vec{v}$$

Mathematically, $\vec{u}$ is a linear function that maps its argument ($\vec{v}$) to a real number ($\vec{u} \cdot \vec{v}$). Is it possible to construct new objects that are functions of not one but two vectors? The answer is yes, and that we can construct these new objects using a new operation called the outer product $\otimes$, defined such that:

$$(\vec{u} \otimes \vec{v})(\vec{p}, \vec{q}) \equiv \vec{u}(\vec{p})\vec{v}(\vec{q}) = (\vec{u} \cdot \vec{p})(\vec{v} \cdot \vec{q})$$

Note that in general $\vec{u} \otimes \vec{v} \neq \vec{v} \otimes \vec{u}$, and that $\otimes$ should not be confused with the vector product.

Example: show that the object $\vec{h} = \vec{u} \otimes \vec{v}$ satisfies

$$\vec{h}(\alpha\vec{p}, \vec{q}) = \alpha\vec{h}(\vec{p}, \vec{q}) = \vec{h}(\vec{p}, \alpha\vec{q})$$

Solution: to see this let's consider

$$\vec{h}(\alpha, \vec{p}, \vec{q}) = (\vec{u} \otimes \vec{v})(\alpha\vec{p}, \vec{q}) = \vec{u}(\alpha\vec{p})\vec{v}(\vec{q})$$

where

$$\vec{u}(\alpha\vec{p}) = \vec{u} \cdot (\alpha\vec{p}) = (u^\mu \vec{e}_\mu) \cdot (\alpha p^\nu \vec{e}_\nu) = u^\mu \alpha p^\nu (\vec{e}_\mu \cdot \vec{e}_\nu)$$

$$= \alpha(u^\mu \vec{e}_\mu) \cdot (p^\nu \vec{e}_\nu) = \alpha\vec{u} \cdot \vec{p} = \alpha\vec{u}(\vec{p})$$

so

$$\Rightarrow \vec{h}(\alpha\vec{p}, \vec{q}) = \alpha\vec{u}(\vec{p})\vec{v}(\vec{q}) = \alpha(\vec{u} \otimes \vec{v})(\vec{p}, \vec{q}) = \alpha\vec{h}(\vec{p}, \vec{q})$$

as required; the proof that $\vec{h}(\vec{p}, \alpha\vec{q}) = \alpha\vec{h}'(\vec{p}, \vec{q})$ proceeds in the same way.

1.12 Tensors

The object $\vec{h} = \vec{u} \otimes \vec{v}$, that takes two vectors in its two arguments and returns a real number, is an example of a rank-2 tensor. A rank n tensor is an objects that is linear in each of its arguments, and that takes n vectors to return a real number.

For example, if $\vec{t}$ is a rank-3 tensor and $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors, then $\vec{t}(\vec{u}, \vec{v}, \vec{w})$ is a real number and

$$\vec{t}(\vec{u} + \vec{v}, \vec{v}, \vec{w}) = t(\vec{u}, \vec{v}, \vec{w}) + t(\vec{v}, \vec{v}, \vec{w})$$

$$\vec{t}(\alpha\vec{u}, \vec{v}, \vec{w}) = \alpha\vec{t}(\vec{u}, \vec{v}, \vec{w})$$

Tensors are of fundamental importance in GR. They are defined without any reference to coordinates, so that if we use these objects to write equations for physical laws we know we will end up with a set of equations that are valid in all coordinate systems. This is exactly what is required from the principle of covariance.

Just as it was convenient to express vectors in terms of a set of basis vectors, here it is useful to construct a tensor basis so that we can express our tensors. For a rank n tensor the appropriate basis is made from the outer product of n basis vectors:

$$\vec{e}_a \otimes \vec{e}_b \otimes \ldots \otimes \vec{e}_n$$

Example: for the rank 3 tensor $\vec{t}$ we can write

$$\vec{t} = t^{abc}(\vec{e}_a \otimes \vec{e}_b \otimes \vec{e}_c)$$

We could equally well have constructed our tensor basis from dual basis vectors

$$\vec{t} = t_{abc}(\vec{e}^a \otimes \vec{e}^b \otimes \vec{e}^c)$$

or a mixture of basis and dual basis vectors

$$\vec{t} = t^a{}_{bc}(\vec{e}_a \otimes \vec{e}^b \otimes \vec{e}^c)$$

An explicit example of a tensor is the metric tensor g constructed from the metric $g_{\mu\nu}$

$$\vec{g} = g_{\mu\nu}(\vec{e}^\mu \otimes \vec{e}^\nu)$$

$$\Rightarrow g(\vec{u}, \vec{v}) = g_{\mu\nu}(\vec{e}^\nu \otimes \vec{e}^\nu)(\vec{u}, \vec{v}) = g_{\mu\nu}(\vec{e}^\mu \cdot \vec{u})(\vec{e}^\nu \cdot \vec{v})$$

$$= g_{\mu\nu} u^\mu v^\nu = \vec{u} \cdot \vec{v}$$

1.13 Tensor operations

From the definitions in the previous sections we can show that a rank-2 tensor $\vec{t}$ acting on vectors $\vec{u}$ and $\vec{v}$ can be written:

$$\vec{t}(\vec{u},\vec{v}) = t_{ab}u^a v^b = t^{ab}u_a v_b = t_a{}^b u^a v_b = t^a{}_b u_a v^b$$

with similar expressions applying to higher rank tensors. This shows:

• raised and lowered pairs of dummy indices can be exchanged at will

• $g_{\mu\nu}$ and $g^{\mu\nu}$ lower and raise the indices of tensors, just as with vectors.

• rank-n tensors with two indices contracted are the components of rank-$(n-2)$ tensors.

• components of rank-n tensors multiplying the components of rank-m tensors are rank-$(m+n)$ tensors.

• coordinate transformations on raised and lowered indices of tensors transform in the same way as raised and lowered vector indices.

Exercise: Prove the above statements

Example: prove that the metric $g_{\mu\nu}$ and $g^{\mu\nu}$ lower and raise the indices of tensor components, just as they do vector components.

solution: consider a tensor $\vec{t}$. If this is of rank 2 then

$$\vec{t}(\vec{u},\vec{v}) = t_{ab}(\vec{e}^a \otimes \vec{e}^b)(\vec{u},\vec{v}) = t_{ab}(\vec{e}^a \cdot \vec{u})(\vec{e}^b \cdot \vec{v}) = t_{ab}u^a v^b$$

$$= t^{ab}(\vec{e}_a \otimes \vec{e}_b)(\vec{u},\vec{v}) = t^{ab}(\vec{e}_a \cdot \vec{u})(\vec{e}_b \cdot \vec{v}) = t^{ab}u_a v_b$$

$$= t_a{}^b(\vec{e}^a \otimes \vec{e}_b)(\vec{u},\vec{v}) = t_a{}^b(\vec{e}^a \cdot \vec{u})(\vec{e}_b \cdot \vec{v}) = t_a{}^b u^a v_b$$

$$= t^a{}_b(\vec{e}_a \otimes \vec{e}^b)(\vec{u},\vec{v}) = t^a{}_b(\vec{e}_a \cdot \vec{u})(\vec{e}_b \cdot \vec{v}) = t^a{}_b u_a v^b$$

Combining the second and third equation of the above

$$t_{ab}v^b = t_a{}^b v_b = t_a{}^b g_{bc} v^c$$

$$\Rightarrow t_{ab} = t_a{}^c g_{bc}$$

second and fourth

$$t^{ab}v_b = t^a{}_b v^a = t^a{}_b g^{bc} v_c$$

$$\Rightarrow t^{ab} = t^a{}_c g^{bc}$$

first and fourth

$$t^a{}_b u_a = t_{ab}u^a = t_{ab} g^{ac} u_c$$

$$\Rightarrow t^a{}_b = t_{cb} g^{ac}$$

second and third

$$t_a{}^b u^a = t^{ab} u_a = t^{ab} g_{ac} v^c$$

$$\Rightarrow t_a{}^b = t^{cb} g_{ac}$$

1.14 Connection

So far we have only considered objects defined at points in the manifold. If we want to start considering fields (objects taking values at all points), we will also be interested in how they change value from point to point. This requires differentiation.

For scalars this is not a problem, as the rate of change can simply be written as a partial derivative with respect to the coordinates:

$$\frac{\partial \phi}{\partial x^\mu} \equiv \partial_\mu \phi$$

For vectors it is more difficult. Two vectors, at two different points, lie in two different tangent spaces. To find the derivative of a vector we require a connection between these two tangent spaces. Firstly, we need to know how basis vectors in each space are related to each other. This is done by supposing they change as:

$$\nabla_{\vec{e}_a}\vec{e}_b = \Gamma^c{}_{ba}\vec{e}_c$$

The ∇ is known as "the connection", and can be thought of as the change of its argument $\vec{e}_b$ in the direction indicated by the subscript ($\vec{e}_a$). The $\Gamma^c{}_{ba}$ are known as the connection coefficients, and encode how the basis vectors change between tangent spaces. In what follows, we will mostly be interested in changes of coordinate basis, and for this we use the shorthand $\nabla_\mu \equiv \nabla_{\vec{e}_\mu}$.

To be consistent with the derivatives of scalars above we require that

$$\nabla_\mu \phi = \frac{\partial \phi}{\partial x^\mu}$$

Using $\nabla_{\vec{e}_a}(\vec{e}^{\,b} \cdot \vec{e}_c) = \nabla_{\vec{e}_a}(\delta^b{}_c) = 0$, we can also deduce

$$\nabla_{\vec{e}_a}\vec{e}^{\,b} = -\Gamma^b{}_{ca}\vec{e}^{\,c},$$

which tells us how to connect dual basis vectors in different tangent spaces.

Exercise: Prove the above relation by assuming the connection obeys the Leibnitz rule

With a given connection, we can work out how to calculate the change of vectors and tensors as we move around the manifold.

Finally, the connection is assumed to obey the following mathematical relationships:

i) $\nabla_{\vec{X}_1+\vec{X}_2}\vec{Y} = \nabla_{\vec{X}_1}\vec{Y} + \nabla_{\vec{X}_2}\vec{Y}$

ii) $\nabla_{\vec{X}}(\vec{Y}_1 + \vec{Y}_2) = \nabla_{\vec{X}}\vec{Y}_1 + \nabla_{\vec{X}}\vec{Y}_2$

iii) $\nabla_{\phi\vec{X}}\vec{Y} = \phi\nabla_{\vec{X}}\vec{Y}$

iv) $\nabla_{\vec{X}}(\phi\vec{Y}) = \phi\nabla_{\vec{X}}\vec{Y} + (\nabla_{\vec{X}}\phi)\vec{Y}$

for any vectors $\vec{X}$, $\vec{Y}$ and any scalar ϕ

Note: the connection ∇ can be used to connect any two vectors in neighbouring tangent spaces, and does not apply solely to the basis vectors.

1.15 The Levi-Civita Connection

We need to impose restrictions on the connection, in order to make it useable. In GR, these restrictions are chosen to be:

$$\Gamma^{\mu}{}_{\nu\rho} = \Gamma^{\mu}{}_{\rho\nu}$$

and

$$\nabla_{\nu} g = 0$$

The first of these says that the connection is "torsionless", the second says it is "metric compatible". Note that these conditions are imposed in a coordinate basis. A connection that obeys these conditions is known as the Levi-Civita connection and its connection components are called the Christoffel symbols. The two equations above, the Leibnitz rule, and the definition of the connection are sufficient to determine.

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

Exercise: Prove the above

1.16 Covariant derivative

We can use the Levi-Civita connection to define a covariant derivative operator:

$$\nabla_\nu \vec{v} \equiv (D_\mu v^\nu)\vec{e}_\nu$$

Let's calculate an explicit expression for $D_\mu v^\nu$

$$\nabla_\mu \vec{v} = \nabla_\mu (v^\nu \vec{e}_\nu) = (\nabla_\mu v^\nu)\vec{e}_\nu + v^\nu (\nabla_\mu \vec{e}_\nu) = (\partial_\mu v^\nu)\vec{e}_\nu + v^\nu \Gamma^\sigma{}_{\nu\mu}\vec{e}_\sigma = (\partial_\mu v^\nu + \Gamma^\nu{}_{\rho\mu} v^\rho)\vec{e}_\nu$$

$$\Rightarrow \boxed{D_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu{}_{\rho\mu} v^\rho}$$

Direct transformation of coordinates shows that this quantity transforms as the components of a tensor. If we had chosen to write $\vec{v} = v_\nu \vec{e}^\nu$ we would have found

$$\boxed{D_\mu v_\nu = \partial_\mu v_\nu - \Gamma^\rho{}_{\mu\nu} v_\rho}$$

Now let's consider the covariant derivative of a tensor:

$$\nabla_\mu \vec{t} \equiv (D_\mu t^{\nu_1 \ldots \nu_p})\vec{e}_{\nu_1} \otimes \ldots \otimes \vec{e}_{\nu_p} \otimes \vec{e}^{\rho_1} \otimes \ldots \otimes \vec{e}^{\rho_r}$$

Following a similar longer procedure one finds

$$D_\mu t^{\nu_1 \ldots \nu_p}{}_{\rho_1 \ldots \rho_r} = \partial_\mu t^{\nu_1 \ldots \nu_p}{}_{\rho_1 \ldots \rho_r} + t^{\sigma \ldots \nu_p}{}_{\rho_1 \ldots \rho_r} \Gamma^{\nu_1}{}_{\mu\sigma} + \ldots - t^{\nu_1 \ldots \nu_p}{}_{\sigma \ldots \rho_r} \Gamma^\sigma{}_{\mu\rho_1} - \ldots$$

where ... denotes extra similar terms corresponding to each raised/lowered index on $t^{\nu_1 \dots \nu_p}{}_{\rho_1 \dots \rho_r}$. Again, direct transformations of coordinates shows that this quantity transforms as the components of a tensor.

In fact, using $\nabla_{\vec{v}} \vec{t} = v^\mu \nabla_\mu \vec{t}$ it can be seen immediately that $D_\mu t^{\nu_1 \dots \nu_p}{}_{\rho_1 \dots \rho_r}$ transforms as the components of a tensor. This is due to the fact that $\nabla_{\vec{v}} \vec{t}$ is coordinate independent and v^μ, $\vec{e}_\mu$ and e^μ all transform in the known way, then each index of $D_\mu t^{\nu_1 \dots \nu_p}{}_{\rho_1 \dots \rho_r}$ must transform like a tensor, if they didn't then $\nabla_{\vec{v}} \vec{t}$ would change under some change of coordinates, something which cannot happen due to the covariance principle

1.17 Intrinsic derivative

The covariant derivatives discussed so far assume the vector or tensor fields under consideration are fields over the whole manifold. However, sometimes they are only defined along a single curve. In this case we may be interested in the derivative along the curve (the intrinsic derivative).

For a vector:

$$\frac{d\vec{v}}{d\lambda} \equiv \frac{Dv^\mu}{D\lambda} \vec{e}_\mu$$

where λ is the parameter along a curve $x^\mu(\lambda)$. If the tangent vector to the curve is $\vec{u}$, then

$$\frac{d\vec{v}}{d\lambda} = \nabla_{\vec{u}}\vec{v} = u^\mu \nabla_\mu \vec{v} = \frac{dx^\mu}{d\lambda}(\partial_\mu v^\nu + \Gamma^\nu{}_{\mu\rho} v^\rho)\vec{e}_\nu = \left(\frac{dv^\mu}{d\lambda} + \Gamma^\mu{}_{\nu\rho}\frac{dx^\mu}{d\lambda}v^\rho\right)\vec{e}_\mu$$

so

$$\boxed{\frac{Dv^\mu}{D\lambda} = \frac{dv^\mu}{d\lambda} + \Gamma^\mu{}_{\nu\rho}\frac{dx^\nu}{d\lambda}v^\rho}$$

Similarly, if we had written $\vec{v} = v_\mu \vec{e}^\mu$

$$\boxed{\frac{Dv_\mu}{D\lambda} = \frac{dv_\mu}{d\lambda} - \Gamma^\rho{}_{\mu\nu}\frac{dx^\nu}{d\lambda}v_\rho}$$

1.18 Parallel transport

As well as calculating how a pre-specified set of vectors changes along a curve, we can use the intrinsic derivative to propagate a vector that initially exists at one point on a curve to all points on the curve

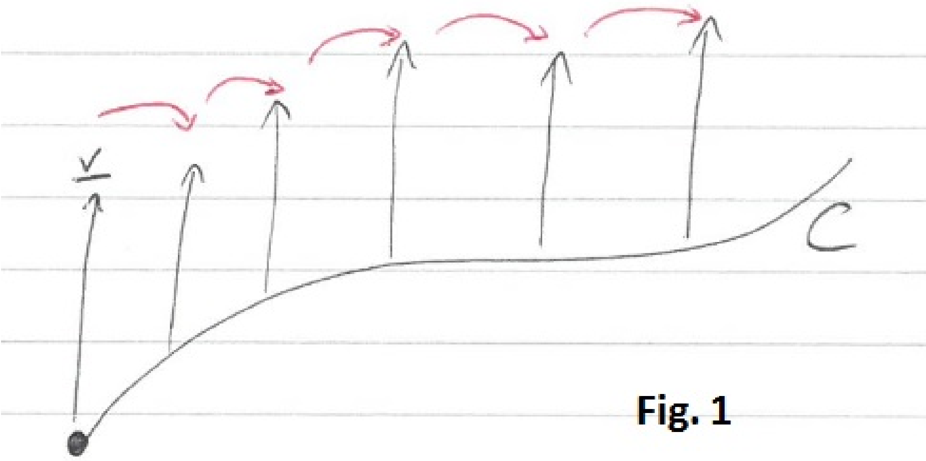

Fig. 1

There are many ways that we could do this, but one particularly useful one is known as parallel transport. We say that $\vec{v}$ has been parallel transported along C if

$$\frac{d\vec{v}}{d\lambda} = 0$$

along C. Re-writing this in components:

$$\frac{dv^\mu}{d\lambda} = -\Gamma^\mu{}_{\nu\rho} v^\nu \frac{dx^\rho}{d\lambda}$$

If we specify v^μ at any point on C, we can use this equation to work out its value at any other. The result of this is a set of vectors that are "parallel" at each point along C. This interpretation should be taken with some care, however, as in general the vectors at different points exist in different tangent spaces.

A case where the interpretation is clear is in flat space. In this case the tangent spaces at different points overlap, and vectors can be compared (and seem to be parallel) at any two points along C.

In a curved space it is not so easy. If we imagine our curved space as existing in a higher-dimensional flat space then the result of parallel transport corresponds to taking a parallel vector, shifting it to a new point on the curve so that it points in the same direction in the higher dimensional space, and then taking the components that lie in the tangent space to the curved space at the new point. This is the closest thing to parallel that exists in a curved space, but is a path-dependent process:

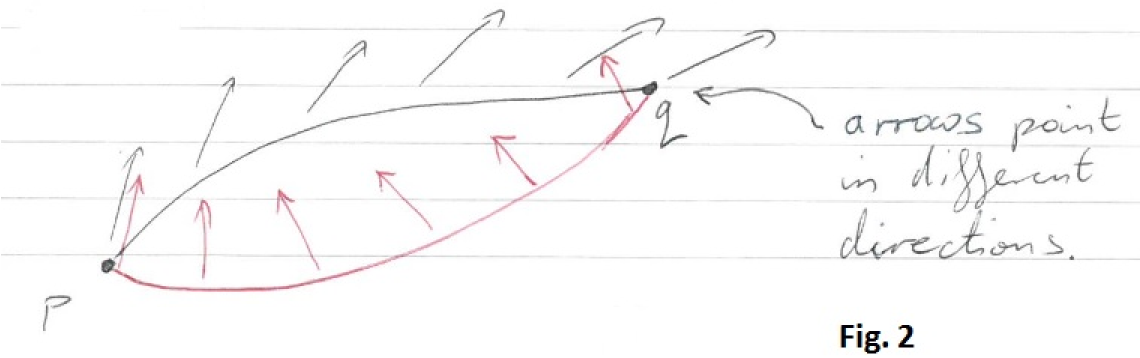

Fig. 2

If the tangent vector to a curve, $\vec{u}$, obeys the parallel transport condition then the curve us said to be "auto-parallel":

$$\frac{d\vec{u}}{d\lambda} \equiv \nabla_{\vec{u}}\vec{u} = 0$$

This is closely related to the idea of a curve being "geodesic", which is true if

$$\ddot{x}^{\mu} + \Gamma^{\mu}{}_{\nu\rho}\dot{x}^{\nu}\dot{x}^{\rho} = 0$$

Exercise: Show that if the connection is the Levi-Civita connection, then a curve that is "auto-parallel" is also a geodesic

Example: Parallel transport in vector $\vec{v}$ along a curve with constant $\theta = \theta_0$ on a 2-sphere with $ds^2 = r^2 d\theta^2 - r^2 \sin^2 d\phi^2$

solution: take ϕ as the parameter along the curve so

$$\frac{dx^\nu}{d\phi} + v^\theta \Gamma + v^\phi \Gamma = 0$$

The only non-zero Christoffel symbols are

$$\Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta \qquad \text{and} \qquad \Gamma^\phi{}_{\theta\phi} = \Gamma^\phi{}_{\phi\theta} = \frac{\cos\theta}{\sin\theta}$$

$$\Rightarrow \frac{dv^\theta}{d\phi} = \sin\theta_0 \cos\theta_0 v^\phi$$

$$\frac{dv^\phi}{d\phi} = -\frac{\cos\theta_0}{\sin\theta_0} v^\theta$$

Differentiate the first of the above equations and sub into the second

$$\Rightarrow \frac{d^2 v^\theta}{d\phi^2} = \sin\theta_0 \cos\theta_0 \frac{dv^\phi}{d\phi} = \sin\theta_0 \cos\theta_0 \Big(-\frac{\cos\theta_0}{\sin\theta_0} \Big) v^\theta = -\cos^2\theta_0 v^\theta$$

$$\Rightarrow v^\theta = A\cos(\alpha\phi) + B\sin(\alpha\phi)$$

similarly

$$v^\phi = C\cos(\alpha\phi) + D\sin(\alpha\theta)$$

1.19 Euler-Lagrange formulation of the geodesic equation

If we consider the following Lagrangian

$$\boxed{L \equiv g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = \left(\frac{ds}{d\lambda}\right)^2}$$

then the geodesic equations are equivalent to the Euler-Lagrange equations

$$\boxed{\frac{\partial L}{\partial x^\mu} - \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) = 0}$$

Here is the proof:

$$\frac{\partial L}{\partial x^\sigma} = \frac{\partial g_{\mu\nu}}{\partial x^\sigma}\dot{x}^\mu \dot{x}^\nu \qquad \text{and} \qquad \frac{\partial L}{\partial \dot{x}^\sigma} = 2g_{\sigma\mu}\dot{x}^\mu$$

$$\Rightarrow \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\sigma} = 2\frac{\partial g_{\sigma\mu}}{\partial x^\rho}\dot{x}^\mu \dot{x}^\rho + 2g_{\mu\sigma}\ddot{x}^\mu = \frac{\partial g_{\sigma\mu}}{\partial x^\mu}\dot{x}^\mu \dot{x}^\rho + \frac{\partial g_{\sigma\rho}}{\partial x^\mu}\dot{x}^\mu \dot{x}^\rho + 2g_{\mu\sigma}\ddot{x}^\mu$$

Combining these equations gives

$$2g_{\sigma\mu}\ddot{x}^\mu = -(g_{\sigma\mu,\rho} + g_{\sigma\rho,\mu} - g_{\mu\rho,\sigma})\dot{x}^\mu \dot{x}^\rho$$

or

$$\ddot{x}^\nu = -\Gamma^\nu{}_{\mu\rho}\dot{x}^\mu \dot{x}^\rho$$

which is the geodesic equation.

Example: Find the geodesic equations for Schwarzschild geometry:

$$ds^2 = -\Big(1 - \frac{2Gm}{r}\Big)dt^2 + \Big(1 - \frac{2Gm}{r}\Big)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

Solution: choose a particle that lies in the plane $\theta = \frac{\pi}{2}$

$$\Rightarrow L = -\Big(1 - \frac{2Gm}{r}\Big)\dot{t}^2 + \frac{\dot{r}^2}{(1 - \frac{2Gm}{r})} + r^2\dot{\phi}^2$$

t-equation:

$$\frac{\partial L}{\partial t} = 0 \qquad \text{and} \qquad \frac{\partial L}{\partial \dot{t}} = -2\Big(1 - \frac{2Gm}{r}\Big)\dot{t}$$

so

$$\frac{\partial L}{\partial t} - \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{t}} = 0 \quad \Rightarrow \quad \frac{d}{d\lambda}\left[\Big(1 - \frac{2Gm}{r}\Big)\dot{t}\right] = 0 \quad \Rightarrow \quad \dot{t} = \frac{A}{\Big(1 - \frac{2Gm}{r}\Big)}$$

ϕ-equation:

$$\frac{\partial L}{\partial \phi} = 0 \qquad \text{and} \qquad \frac{\partial L}{\partial \dot{\phi}} = 2r^2\dot{\phi}$$

so

$$\frac{\partial L}{\partial \phi} - \frac{d}{d\lambda}\frac{\partial}{\partial \dot{\phi}} = 0 \quad \Rightarrow \quad \frac{d}{d\lambda}[r^2\dot{\phi}] = 0 \quad \Rightarrow \quad \dot{\phi} = \frac{B}{r^2}$$

and similarly for the r equation.