Introduction to General Relativity 2025
Lesson 4: General Relativity and Differential Geometry

Mario C Díaz

1 Gravity and Geometry

We already discussed at the end of the First lecture the key ingredients that Einstein considered to develop a General Theory of Relativity. He realized that he needed a geometrical formulation of a new theory capable of including Gravitation and containing the special theory as well.

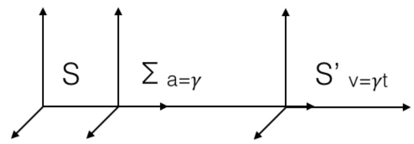
One of them was:

1. The need to include the gravitational field within the special theory of relativity. But it was clear to him that a major role in the formulation was to be played by

2. The equivalence principle (i.e. the interchangeability or indistinguishability of the gravitational force acting on a body with the acceleration experimented by a body attached to a frame experimenting constant acceleration).

The first question leads to the second one: "Is it conceivable that the principle of relativity also holds for systems that are accelerated relative to each other?" One of the first startling results is what it is called the **gravitational redshift**. I presented an heuristic discussion based on the bending of Iight rays. The actual formulation by Einstein was more elaborate and I will discuss it as presented in the book by Pais (*Subtle is the Lord*) following the treatment by Einstein in the article published in 1907 in Jahrbuch der Radioaktivität und Elektronik:

Let's consider two coordinate systems $S(x, y, z, t)$ and $\Sigma(\xi, \eta, \zeta, \tau)$ which are coincident at one time and have $v = 0$. Synchronize two similar network of clocks at both reference systems starting at $t = \tau = 0$. Now system Σ starts moving with constant acceleration γ. Let's introduce now a third reference system $S'(x', y', z', t')$ which relative to S moves with uniform velocity in the x direction and in such a way that for a fixed instant of time t, $x' = \xi, y' = \eta, z' = \zeta$. And at that time $v = \gamma t$. We now go one step further and synchronize at that time all clock in S' with those in Σ.



With all this:

1. Consider an interval δ so small after the coincidence of S' and Σ that we can neglect all effects $O(\delta^2)$.

This means that we still can use the times of the clocks in the S' (a Lorentz frame) to describe the rates of the Σ clocks. We can then use what we have learned from Special relativity because we're considering small light paths. S and S' are inertial frames and the changes in Σ are very small yet, so we can still identify the measurements in S' with those in Σ up to higher order effects.

2. How do clocks in two distinct space points run relative to each other?

At $t = \tau = 0$, the two clocks were synchronous with each other and they remain for a such interval of time that way. But S' is not synchronous to them, because of the Lorentz transformations, so then they do not remain synchronous to each other, i.e. their difference is not constant.

Let's single out one clock in Σ and take τ and setting it to $= t$. We can now define simultaneity of events 1 and 2 in Σ as

$$t_1 - \frac{vx_1}{c^2} = t_2 - \frac{vx_2}{c^2}$$

where $v = \gamma t = \gamma\tau$.

Let now 1 corresponds to the origin of Σ and 2 to a point $(\xi, 0, 0)$ where the clock reading is σ. And we make one last approximation: the time τ between S' and Σ coincidence is taken such $O(\tau^2)$ effects are negligible. Then $x_2 - x_1 = x'_2 - x'_1 = \xi$, $t_1 \equiv \tau$ and $t_2 \equiv \sigma$, so that the previous equation becomes:

$$t_1 - t_2 = \frac{v(x_1 - x_2)}{c^2} \to \sigma = \tau\left(1 + \frac{\gamma\xi}{c^2}\right)$$

where of course $v = \gamma\tau$.

But if the time difference depends on the acceleration, this means that consistent with our knowledge of physics any quantity strongly dependent on a measurement related to time, like a frequency, will change (actually be reduced), i.e. be shifted in frequencies to the red.

From this Einstein said:

”there are clocks which are available at different locations with distinct gravitational potential...these are the generators of spectral lines. It follows ...that light coming from the solar surface has a longer wavelength than the light generated from the same material on earth...”

(Historically the gravitational weakening of light from massive stars was predicted by John Michell in 1783 and Pierre-Simon Laplace in 1796. They were thinking, a la Newton, of light as particles and they speculated with fields strong enough that light could not escape them).

2 The Principle of Equivalence

Is there a difference between inertial mass and gravitational mass?

We will assume it's not, based historically on Galileo's famous ”thought” experiment. But we need to explore the meaning of each of these masses:
From $\vec{F} = m_i \vec{a}$ when the force is gravity we have:

$$\frac{G m_g M_{source}}{r^3} \hat{\mathbf{r}} = m_i \vec{a}$$

Compare with the Coulomb force for a particle of also mass m_i and charge q in the field of another charge Q:

$$k_e \frac{qQ}{r^3} \hat{\mathbf{r}} = m_i \vec{a}$$

Galileo's experiment showed that $m_i = m_g$ while it is very clear that m_i has little to do with the ”electrical mass” q.

Galileo's discovery can be thought as meaning that there is something intrinsically related to gravity in the mass-energy.

2.1 Einstein's first formulation of the Equivalence Principle

Let the frame S be at rest and let it carry a homogeneous gravitational field in the negative z direction. Σ is a field-free frame that moves with constant acceleration relative to S in the positive z direction. In both systems Newton's mechanical laws are equivalent. ”One can speak as little of the *absolute acceleration* of the reference frame as one can of the *absolute velocity* in the ordinary -special- relativity theory” (italics from Einstein). From this“according to this theory, the equal fall of all bodies in a gravitational field is self-evident”. With gravity we cannot attempt to describe the trajectory followed by particles unaffected by its action. We cannot define an inertial frame at rest on Earth. Not even a photon could help to characterize it. All will be affected. But there is a frame where all particles keep uniform velocity: this is a frame in fall under the action of gravity itself. Einstein will develop a general theory of relativity by taking these frames as inertial.

2.2 Deflection of light

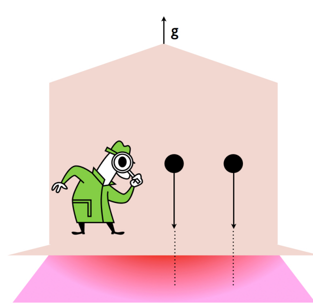
Einstein also realized that gravity would produce on light rays another effect, additional to the redshift.

If the light ray is in any direction that is not completely perpendicular or perfectly aligned with the acceleration, then the light beam would experience an additional effect. The light beam is effectively like a sum of two beams, one in the direction of the acceleration and the other in the perpendicular direction. Because it is made of both, the component in the direction of the acceleration will experience a redshift, but the perpendicular part will need to adjust itself to this change effectively feeling a pull in the direction of the acceleration. The net effect is that the accelerated system will be tucking the light ray in its direction.

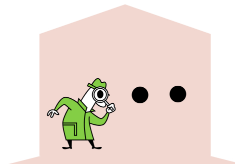
The equivalence principle states that such a thought experiment is indistinguishable from a gravitational field: Einstein concluded that a gravitational field would bend light rays passing near it.

3 Why curvature?

After discovering the gravitational redshift Einstein realized that a flat geometry realization for the physical space of a real mechanics could not be consistent. He made the comment that just looking at a uniformly rotating system of reference there will be observers who will not ”see” the ratio of the circumference of motion to the diameter being π anymore. And we have seen that accelerated systems could be indistinguishable from a uniform gravitational field.The picture below is showing a rocket with $a = g$.

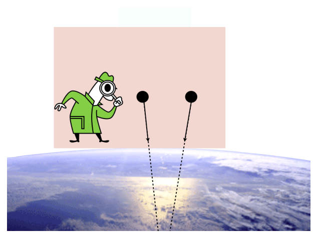


And in this other one $(a = 0)$the objects remain at rest with the observer.

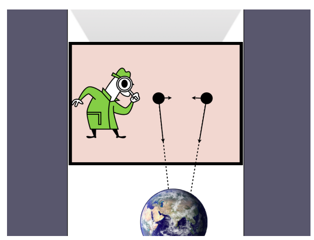


No acceleration in the rocket

The objects (on a lift now standing on earth) fall on **converging** paths.



But if we drop the lift down a shaft.



The particles are getting closer together and this is how we know there is a gravitational field. This is a central idea leading to a theory of relativity that needs geometry in its formulation: If we call *geodesics* the paths followed by test particles, how these paths diverge or converge, i.e. how they deviate from parallelism, is what indicates the presence of a gravitational field.
But before we go into curvature we need to have another look at tensors.

4 Vectors and one forms revisited

In the exercises using polar and spherical coordinates we worked on the transformations using matrices and vectors. For example a coordinate transformation from a system x, y to a system ξ, η.

$$\begin{pmatrix} \partial/\partial\xi \\ \partial/\partial\eta \end{pmatrix} = \begin{pmatrix} \partial\xi/\partial x & \partial\xi/\partial y \\ \partial\eta/\partial x & \partial\eta/\partial y \end{pmatrix} \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \end{pmatrix} \tag{1}$$

We have accordingly defined a vector as something whose components transforms according to:

$$V^{\alpha'} = \Lambda^{\alpha'}{}_{\beta} V^{\beta}, \tag{2}$$

There is a more modern way of defining a vector. If we consider a scalar field ϕ, given coordinates (ξ, η) it is always possible to form the derivatives $\partial\phi/\partial\xi$ and $\partial\phi/\partial\eta$.

And we define the one-form $\tilde{d}\phi$ as the object whose components are;

$$\tilde{d}\phi \rightarrow (\partial\phi/\partial\xi, \partial\phi/\partial\eta) \tag{3}$$

Recalling the chain rule for derivation:

$$\frac{\partial\phi}{\partial\xi} = \frac{\partial x}{\partial\xi}\frac{\partial\phi}{\partial x} + \frac{\partial y}{\partial\xi}\frac{\partial\phi}{\partial y} \tag{4}$$

And of course there is a similar one for $\partial\phi/\partial\eta$ But in matrix notation you have to write as acting on row-vectors:

$$(\partial\phi/\partial\xi, \partial\phi/\partial\eta) = \begin{pmatrix} \partial\phi/\partial x & \partial\phi/\partial y \end{pmatrix} \begin{pmatrix} \partial x/\partial\xi & \partial x/\partial\eta \\ \partial y/\partial\xi & \partial y/\partial\eta \end{pmatrix} \tag{5}$$

and:

$$(\Lambda^{\alpha}{}_{\beta'}) = \begin{pmatrix} \partial x/\partial\xi & \partial x/\partial\eta \\ \partial y/\partial\xi & \partial y/\partial\eta \end{pmatrix} \tag{6}$$

We have to use row vectors the way we did and not column vectors at the right of the matrix because when the matrix is not symmetric, it will not work! Got to be careful as to how to use the matrix representation! We see that $\tilde{d}\phi$ in the primed coordinates transforms:

$$(\tilde{d}\phi)_{\beta'} = \Lambda^{\alpha}{}_{\beta'}(\tilde{d}\phi)_{\alpha} \tag{7}$$

Notice that

$$\Lambda^{\alpha'}{}_{\beta} = \begin{pmatrix} \partial\xi/\partial x & \partial\xi/\partial y \\ \partial\eta/\partial x & \partial\eta/\partial y \end{pmatrix} \tag{8}$$

When doing $\Lambda^{\alpha'}{}_{\gamma}\Lambda^{\gamma}{}_{\beta'}$ we get:

$$\begin{pmatrix} \partial\xi/\partial x & \partial\xi/\partial y \\ \partial\eta/\partial x & \partial\eta/\partial y \end{pmatrix} \begin{pmatrix} \partial x/\partial\xi & \partial x/\partial\eta \\ \partial y/\partial\xi & \partial y/\partial\eta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{9}$$

5 Differential Geometry

Curve: it is a mapping of an interval of the real line (i.e. a set of numbers) into a path in the plane. This number is called a parameter. Usually called s.

$$\xi = f(s), \;\; \eta = g(s), \;\; a \leq s \leq b \tag{10}$$

The above equation defines a curve in the plane. If we change s to s'.

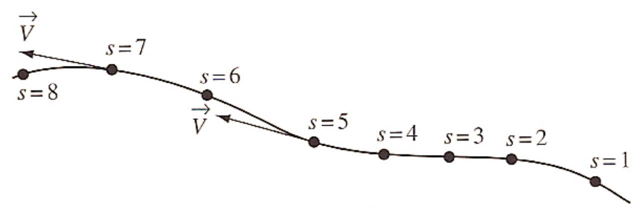
$$\xi = f'(s'),\ \ \eta = g'(s'),\ \ a' \le s' \le b' \tag{11}$$

where f' and g' are new functions and where $a' = s'(a)$ and $b' = s'(b)$. For mathematicians this is a new curve. Although its image, the real thing, remains the same. We can calculate, given a scalar field how much does it change along the curve. This derivative will be given by $d\phi/ds$.

We can write this as $d\phi/ds = \langle \tilde{d\phi}, \vec{V}\rangle$ where $\vec{V}$ is the vector with components are $(d\xi/ds, d\eta/ds)$.

$$\begin{pmatrix} d\xi/ds \\ d\eta/ds \end{pmatrix} = \begin{pmatrix} \partial\xi/\partial x & \partial\xi/\partial y \\ \partial\eta/\partial x & \partial\eta/\partial y \end{pmatrix} \begin{pmatrix} dx/ds \\ dy/ds \end{pmatrix} \tag{12}$$

The modern view of a vector then is that it is a geometrical object which transforms with certain properties and it is tangent to a given curve.



Similarly it can be said that it is the result of the function that evaluated on a one-form $\tilde{d}\phi$ gives $d\phi/ds$.

6 A summary of polar coordinates

$$\vec{e}_r = \cos\theta \vec{e}_x + \sin\theta \vec{e}_y, \tag{13}$$

and similarly:

$$\vec{e}_\theta = -r\sin\theta \vec{e}_x + r\cos\theta \vec{e}_y, \tag{14}$$

The basis one-forms are:

$$\tilde{d}\theta = -\frac{1}{r}\sin\theta \tilde{d}x + \frac{1}{r}\cos\theta \tilde{d}y \tag{15}$$

and

$$\tilde{d}r = \cos\theta \tilde{d}x + \sin\theta \tilde{d}y \tag{16}$$

Notice that this basis does not have unit length.

6.1 Metric tensor

What are the components of the metric $\mathbf{g}_{\alpha\beta}$ in polar coordinates? We know that in Cartesian is: $\mathbf{g}_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$$g_{\alpha'\beta'} = \mathbf{g}(\vec{e}_{\alpha'}, \vec{e}_{\beta'}) = \vec{e}_{\alpha'} \cdot \vec{e}_{\beta'} \tag{17}$$

And using eqs (15) and (16):

$$g_{rr} = \vec{e}_r \cdot \vec{e}_r = 1, \quad g_{\theta\theta} = \vec{e}_\theta \cdot \vec{e}_\theta = r^2, \quad g_{r\theta} = \vec{e}_r \cdot \vec{e}_\theta = 0 \tag{18}$$

$$\mathbf{g}_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \tag{19}$$

And we display the result most of the time:

$$\vec{dl}\vec{dl} = ds^2 = |dr\vec{e}_r + d\theta\vec{e}_\theta|^2 = dr^2 + r^2 d\theta^2 \tag{20}$$

The inverse is:

$$\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix} \tag{21}$$

$$g^{rr} = 1, \quad g^{\theta\theta} = 1/r^2, \quad g^{r\theta} = 0 \tag{22}$$

Example

If ϕ is a scalar field and $\tilde{d}\phi$ is the gradient, then the vector $\vec{d}\phi$ has components:

$$(\vec{d}\phi)^\alpha = g^{\alpha\beta}\phi_{,\beta}, \tag{23}$$

$$(\vec{d}\phi)^r = g^{r\beta}\phi_{,\beta} = g^{rr}\phi_{,r} + g^{r\theta}\phi_{,\theta} = \partial\phi/\partial r \tag{24}$$

$$(\vec{d}\phi)^\theta = g^{\theta r}\phi_{,r} + g^{\theta\theta}\phi_{,\theta} = \frac{1}{r^2}\partial\phi/\partial\theta \tag{25}$$

6.2 Tensor calculus

Notice that the polar basis are not constant like the Cartesian basis vectors.

Consider the vector $\vec{e}_x$. In polar coordinates it has components $(\cos\theta, -r^{-1}\sin\theta)$. The reason is that they are components of a non constant basis (local basis).

Derivatives of basis vectors

$$\frac{\partial}{\partial r}\vec{e}_r = \frac{\partial}{\partial r}(\cos\theta\vec{e}_x + \sin\theta\vec{e}_y) = 0 \tag{26}$$

$$\begin{aligned}\frac{\partial}{\partial \theta}\vec{e}_r &= \frac{\partial}{\partial \theta}(\cos\theta\vec{e}_x + \sin\theta\vec{e}_y) &(27)\\ &= -\sin\theta\vec{e}_x + \cos\theta\vec{e}_y = \frac{1}{r}\vec{e}_\theta &(28)\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\partial}{\partial r}\vec{e}_\theta &= \frac{\partial}{\partial r}(-r\sin\theta\vec{e}_x + r\cos\theta\vec{e}_y) &(29)\\ &= -\sin\theta\vec{e}_x + \cos\theta\vec{e}_y = \frac{1}{r}\vec{e}_\theta &(30)\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial \theta}\vec{e}_\theta &= -r\cos\theta\vec{e}_x - r\sin\theta\vec{e}_y &(31)\\ &= -r\vec{e}_r, &(32)\end{aligned}$$

6.3 Derivatives of general vectors

A general vector $\vec{V}$ has components (V^r, V^θ) on the polar basis. Its derivative is,

$$\frac{\partial\vec{V}}{\partial r} = \frac{\partial}{\partial r}(V^r\vec{e}_r + V^\theta\vec{e}_\theta) = \tag{33}$$

$$\frac{\partial V^r}{\partial r}\vec{e}_r + V^r\frac{\partial\vec{e}_r}{\partial r} + \frac{\partial V^\theta}{\partial r}\vec{e}_\theta + V^\theta\frac{\partial\vec{e}_\theta}{\partial r} \tag{34}$$

In component notation:

$$\frac{\partial\vec{V}}{\partial r} = \frac{\partial V^\alpha}{\partial r}\vec{e}_\alpha + V^\alpha\frac{\partial\vec{e}_\alpha}{\partial r} \tag{35}$$

And the generalized form is:

$$\frac{\partial\vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\vec{e}_\alpha + V^\alpha\frac{\partial\vec{e}_\alpha}{\partial x^\beta} \tag{36}$$

6.4 The Christoffel symbols

Since $\partial\vec{e}_\alpha/\partial x^\beta$ is a vector it can be decomposed as a linear combination of basis vectors.

$$\frac{\partial\vec{e}_\alpha}{\partial x^\beta} = \Gamma^\mu{}_{\alpha\beta}\vec{e}_\mu \tag{37}$$

The interpretation is:

1. α gives the basis vector being differentiated;

2. β gives the coordinate respect to which it is being differentiated.

3. μ denotes de component resulting in this array.

$$\frac{\partial}{\partial r}\vec{e}_r = 0 \Rightarrow \Gamma^{\mu}{}_{rr} = 0 \quad for\ \ all\ \ \mu, \tag{38}$$

$$\frac{\partial}{\partial r}\vec{e}_r = \frac{1}{r}\vec{e}_\theta \Rightarrow \Gamma^{r}{}_{r\theta} = 0 \quad , \Gamma^{\theta}{}_{r\theta} = \frac{1}{r} \tag{39}$$

$$\frac{\partial}{\partial r}\vec{e}_\theta = \frac{1}{r}\vec{e}_\theta \Rightarrow \Gamma^{r}{}_{\theta r} = 0 \quad , \Gamma^{\theta}{}_{\theta r} = \frac{1}{r} \tag{40}$$

$$\frac{\partial}{\partial \theta}\vec{e}_\theta = -r\vec{e}_\theta \Rightarrow \Gamma^{r}{}_{\theta\theta} = -r \quad , \Gamma^{\theta}{}_{\theta\theta} = 0 \tag{41}$$

(37) can be cast in the language of the Christoffel symbols, using (38).

$$\frac{\partial \vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\vec{e}_\alpha + V^\alpha \Gamma^{\mu}{}_{\alpha\beta}\vec{e}_\mu \tag{42}$$

Relabeling to factor out $\vec{e}_\mu$

$$\frac{\partial \vec{V}}{\partial x^\beta} = (\frac{\partial V^\alpha}{\partial x^\beta} + V^\mu \Gamma^{\alpha}{}_{\mu\beta})\vec{e}_\alpha \tag{43}$$

Then if $\frac{\partial V^\alpha}{\partial x^\beta} = V^\alpha_{,\beta}$ the derivative defined in (47) will be denoted, the covariant derivative:

$$V^{\alpha}{}_{;\beta} = V^{\alpha}{}_{,\beta} + V^\mu \Gamma^{\alpha}{}_{\mu\beta} \tag{44}$$

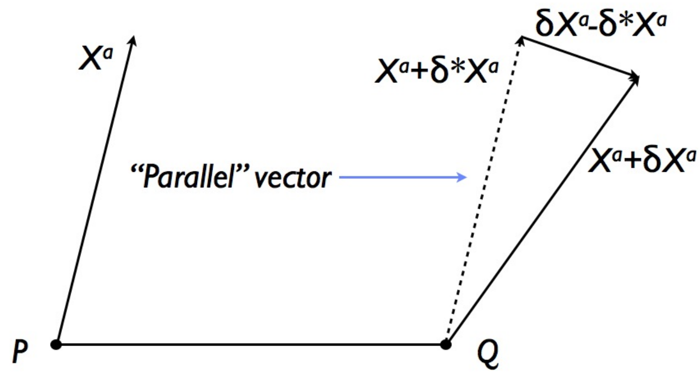
Notice that now $\frac{\partial V^\alpha}{\partial x^\beta}$ is a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor, called the covariant derivative of $\vec{V}$

$$(\nabla \vec{V})^{\alpha}{}_{\beta} = V^{\alpha}{}_{;\beta} \tag{45}$$

On a cartesian basis the components are just $V^{\alpha}{}_{\beta}$.

What about a scalar? A scalar does not depend on the basis vectors so its covariant derivative is the regular one.

6.5 What is the meaning of covariant differentiation?



The parallel vector $X^a + \delta X^a$ at Q

The vector field $X^a(x)$ has coordinates x^a at P. and coordinates $x^a + \delta x^a$ at another point Q near P. Then

$$X^a(x + \delta x) = X^a(x) + \delta x^a \partial_b X^a \tag{46}$$

We can denote the second term by δX^a

$$\delta X^a(x) = \delta x^b X^a{}_{,b} = X^a(x + \delta x) - X^a(x), \tag{47}$$

which is not tensorial in nature because it evaluates the difference between tensors at two different points.
We will try to define a "tensorial" derivative by introducing a vector at Q which in "some sense" is parallel to X^a.

Since $x^a + \delta x^a$ is close to x^a we can assume that the vector only differs from $X^a(x)$ by a small amount, which we denote $\delta^* X^a(x)$.

And we will make the difference $\delta X^a(x) - \delta^* X^a(x)$ a tensor:

$$X^a(x) + \delta X^a(x) - [X^a(x) + \delta^* X^a(x)] = \delta X^a(x) - \delta^* X^a(x) \tag{48}$$

If we want $\delta^* X^a(x)]$ to vanish wherever $X^a(x)$ or δx^a does it, we assume it to be linear in both $X^a(x)$ and δx^a.

$$\delta^* X^a(x)] = -\Gamma^a{}_{bc}(x) X^b(x) \delta x^c \tag{49}$$

Now we define the covariant derivative:

$$\nabla_c X^a = \lim_{\delta x^c \to 0} \frac{1}{\delta x^c} \{X^a(x + \delta x) - [X^a(x) + \delta^* X^a(x)]\} \tag{50}$$

Which using (47) and (50) becomes:

$$\nabla_c X^a = X^a{}_{,c} + \lim_{\delta x^c \to 0} \frac{1}{\delta x^c} \Gamma^a{}_{bd}(x) X^b(x) \delta x^d \tag{51}$$

And using that

$$\lim_{\delta x^c \to 0} \frac{\delta x^c}{\delta x^d} = \frac{\partial x^c}{\partial x^d} = \delta x^c{}_d$$

we finally get

$$X^a{}_{;c} = X^a{}_{,c} + \Gamma^a{}_{bd}(x) X^b(x) \delta x^d{}_c = X^a{}_{,c} + \Gamma^a{}_{bc}(x) X^b(x) \tag{52}$$

6.6 Divergence and Laplacian

Divergence is a scalar.

Just from the definitions:

$$V^\alpha{}_{,\alpha} \equiv V^{\beta'}{}_{;\beta'} \tag{53}$$

It is easy to see how this is implemented working in polar coordinates:

$$V^\alpha{}_{;\alpha} = V^\alpha{}_\alpha + \Gamma^\alpha{}_{\mu\alpha} V^\mu \tag{54}$$

$$\Gamma^\alpha{}_{r\alpha} = \Gamma^r{}_{rr} + \Gamma^\theta{}_{r\theta} = 1/r \tag{55}$$

$$\Gamma^\alpha{}_{\theta\alpha} = \Gamma^r{}_{\theta r} + \Gamma^\theta{}_{\theta\theta} = 0 \tag{56}$$

And then we have:

$$V^\alpha{}_{;\alpha} = V^r{}_{,r} + V^\theta{}_{,\theta} + \frac{1}{r} V^r \tag{57}$$

$$= \frac{1}{r}\frac{\partial}{\partial r}(r V^r) + \frac{1}{r}\frac{\partial}{\partial \theta} V^\theta \tag{58}$$

If we apply this to the gradient of a scalar:

$$\nabla \cdot \nabla \phi \equiv \nabla^2 \phi = \frac{1}{r}\frac{\partial}{\partial r}(r \frac{\partial \phi}{\partial r}) + \frac{1}{r^2}\frac{\partial^2 \phi}{\partial \theta^2} \tag{59}$$

we get the Laplacian in polar coordinates.

6.7 Derivatives for all tensors

One forms

ϕ is a scalar, so $\tilde{d}\phi = \nabla\phi$. i.e. if we look at the formula (13) -which is the same one as (5.19) on page 121 of Schutz book- we have:

$$d\phi/ds = \langle d\tilde{\phi}, \vec{V}\rangle = \nabla\phi \cdot \vec{V} = \phi_{,\alpha}V^{\alpha} \tag{60}$$

where s is a parameter used to define a given curve, ϕ is a scalar field and $\nabla\phi$ is the gradient of the scalar field and $\vec{V}$ is a vector tangent to the curve.

The question now is what is the covariant derivative of a one-form $\tilde{d}\phi$, i.e. what is $\nabla_c X_a$

We can do the following trick: let's calculate the covariant derivative of $\phi_{,\alpha}V^{\alpha}$ and applying the Leibniz rule:

$$(\phi_{,\alpha}V^{\alpha})_{;\beta} = (\phi_{,\alpha}V^{\alpha})_{,\beta} = \phi_{\alpha,\beta}V^{\alpha} + \phi_{,\alpha}V^{\alpha}{}_{,\beta} \tag{61}$$

Now we use our result from (48) solving for $V^{\alpha}_{,\beta}$ and get

$$(\phi_{,\alpha}V^{\alpha})_{,\beta} = \phi_{\alpha,\beta}V^{\alpha} + \phi_{,\alpha}(V^{\alpha}{}_{;\beta} - \Gamma^{\alpha}{}_{\mu\beta}V^{\mu}) \tag{62}$$

Rearranging the indices in the last term after doing the factor multiplication we get:

$$(\phi_{,\alpha}V^{\alpha})_{;\beta} = (\phi_{\alpha,\beta} - \phi_{,\mu}\Gamma^{\mu}{}_{\alpha\beta})V^{\alpha} + \phi_{,\alpha}V^{\alpha}{}_{;\beta} \tag{63}$$

We can now identify in the parenthesis of the right hand side the covariant derivative of the $\nabla\phi = \phi_{,\alpha}$:

$$\phi_{\alpha;\beta} = \phi_{\alpha,\beta} - \phi_{,\mu}\Gamma^{\mu}{}_{\alpha\beta} \tag{64}$$

Notice the reversal in the utilization of the sign in front of the Christoffel symbol as opposed to the covariant derivative of a vector field.

So for a general $\begin{pmatrix} m \\ n \end{pmatrix}$ tensor the covariant derivative will be given by:

$$\nabla_{\alpha}T^{\beta\ldots}_{\delta\ldots} = T^{\beta\ldots}_{\delta\ldots}{}_{,\alpha} + \Gamma^{\beta}{}_{\omega\alpha}T^{\omega\ldots}_{\delta\ldots} + \cdots - \Gamma^{\gamma}{}_{\delta\alpha}T^{\beta\ldots}_{\gamma\ldots} - \cdots \tag{65}$$

6.8 What about the metric?

How is the metric fitting in all this?

Remember that to get a one-form from a given vector we need a metric. There is an easy way to introduce the metric if we remember that we are requiring that tensorial equations transform a covariant index to a contravariant one with the metric, i.e.:

$$V_{\alpha;\beta} = g_{\alpha\mu}V^{\mu}{}_{;\beta} \tag{66}$$

We can calculate the Christoffel symbols from the metric, the fact that it is a rank $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor and some properties that are specific to the "manifolds" we will be studying. These properties are the following:

1. $\Gamma^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\nu\mu}$

2. $\nabla_\rho g_{\mu\nu} = 0$

The first one indicates that the torsion tensor, which is defined below, is zero.

$$T^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\mu\nu} - \Gamma^{\lambda}{}_{\nu\mu} = 0$$

.

This means the metric is torsion free. The second one is called **metric compatibility**. The connection is metric compatible if the covariant derivative of the metric with respect to that connection is everywhere zero.

There are a few of results which can obtained from the above ones which are nice:

1. $\nabla_\lambda \epsilon_{\mu\nu\rho\sigma} = 0$

2. $\nabla_\rho g^{\mu\nu} = 0$

3. $g_{\mu\lambda} \nabla_\rho V^\lambda = \nabla_\rho V^\mu$

Now using these properties:

$$\nabla_\rho g_{\mu\nu} = g_{\mu\nu,\rho} - \Gamma^{\lambda}{}_{\rho\mu} g_{\lambda\nu} - \Gamma^{\lambda}{}_{\rho\nu} g_{\mu\lambda} = 0 \quad (67)$$
$$\nabla_\mu g_{\nu\rho} = g_{\nu\rho,\mu} - \Gamma^{\lambda}{}_{\mu\nu} g_{\lambda\rho} - \Gamma^{\lambda}{}_{\mu\rho} g_{\nu\lambda} = 0 \quad (68)$$
$$\nabla_\nu g_{\rho\mu} = g_{\rho\mu,\nu} - \Gamma^{\lambda}{}_{\nu\rho} g_{\lambda\mu} - \Gamma^{\lambda}{}_{\nu\mu} g_{\rho\lambda} = 0 \quad (69)$$

If we now do (69)+(70) -(68), and use the symmetry of the connection we get:

$$g_{\mu\nu,\rho} - g_{\nu\rho,\mu} - g_{\rho\mu,\nu} + 2\Gamma^{\lambda}{}_{\mu\nu} g_{\lambda\rho} = 0 \quad (70)$$

Multiplying by $g^{\sigma\rho}$ we solve for the connection:

$$\Gamma^{\sigma}{}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} (g_{\nu\rho,\mu} + g_{\rho\mu,\nu} - g_{\mu\nu,\rho}) \quad (71)$$

Please store (71) in your permanent memory!

6.9 The nature of the Christoffel symbols

The Christoffel symbols are defined by (55) and by (74).

$$X^a{}_{;c} = X^a{}_{,c} + \Gamma^a{}_{bc}(x)X^b(x) \tag{72}$$

If we ask that $X^a{}_{;c}$ transforms as a tensor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ under a coordinate transformation, we get:

$$\Gamma'^a{}_{bc} = \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^d}\frac{\partial x^f}{\partial x'^c}\Gamma^d{}_{ef} - \frac{\partial x^d}{\partial x'^b}\frac{\partial x^e}{\partial x'^c}\frac{\partial^2 x'^a}{\partial x^d \partial x^e} \tag{73}$$

This shows a transformation law which is linear but inhomogeneous. A quantity which transforms like $\Gamma^a{}_{bc}$ is called an affinity or **affine connection** or in the gobbledygook of relativists, just **connection** .

6.10 Non coordinate basis

What is a non-coordinate basis? Given a coordinate system X^α we want to use $\{\partial/\partial x^\alpha\}$ as a basis of vectors. Any linearly independent set of vector fields could be used, but not all of them are derivable from coordinate systems. Although $\partial/\partial x^\alpha$ and $\partial/\partial x^\beta$ may commute, two arbitrary vectors $\vec{V} = d/d\lambda$ and $\vec{W} = d/d\eta$ may not: Let's use that $d/d\lambda = (dx^\alpha/d\lambda)\partial/\partial x^\alpha$

$$\frac{d}{d\lambda}\frac{d}{d\mu} - \frac{d}{d\mu}\frac{d}{d\lambda} = V^\alpha W^\beta{}_{,\alpha}\frac{\partial}{\partial x^\beta} - W^\beta V^\alpha{}_{,\beta}\frac{\partial}{\partial x^\alpha} \tag{74}$$

which can eventually be written:

$$\frac{d}{d\lambda}\frac{d}{d\mu} - \frac{d}{d\mu}\frac{d}{d\lambda} = \left(V^\alpha W^\beta{}_{,\alpha} - W^\alpha V^\beta{}_{,\alpha}\right)\frac{\partial}{\partial x^\beta} \tag{75}$$

6.11 Lie bracket

Notice that there is a double sum in α and β. So the commutator:

$$[\vec{V},\vec{W}] = \left[\frac{d}{d\lambda},\frac{d}{d\mu}\right] \equiv \frac{d}{d\lambda}\frac{d}{d\mu} - \frac{d}{d\mu}\frac{d}{d\lambda} \tag{76}$$

is a vector field whose components don't vanish in general. This vector field is called a Lie Bracket.

This means that if these are two elements of a basis and their Lie bracket does not vanish, they will not be expressible as derivatives with respect to any coordinates. A basis like this is called a non-coordinate basis. The existence of these integrability condition (or commutativity of its derivatives) is key to the existence of a coordinate basis associated to the vector fields. Example: just compare polar coordinates $(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta})$ which is a coordinate basis with an orthonormal basis derived from it (i.e. one where the basis vectors are such that the determinant of the transformation matrix is also 1 and not r).