A quick Look through the mathematical prerqusites for General Theory of Relativity

Biplab Mahato
K Philip Thomas

Contents

0.1 Introduction

The general theory of relativity is a theory, which if one wants to see and understand in its full glory, requires one to have a lot of mathematical prerequisites in the fields of calculus, analysis and differential geometry. In this report, we have tried to present a small section of formal mathematical constructions which we learnt over the summer, to understand the Einstein equations. We have defined mathematical objects and tried to give an intuition about the object whenever possible. Few proofs are given, and a few of the proofs are sketched. The report is more of a mathematical nature and the physics part is not given much elaboration. A basic knowledge of multivariable calculus, topology and linear algebra is assumed. All the sections until the last two are purely of mathematical nature.

0.2 Manifolds

Roughly speaking, a manifold is a topological space that "locally" looks like $\mathbb{R}^n$ for some fixed n, where n is called the dimension of the manifold. One could say that the space is locally "flat". From experience, we know that a sphere should intuitively be a manifold of dimension 2. This notion is formalized as follows:

Definition 0.2.1. A topological space M is called a d-dimensional manifold if $\forall p \in M$, there exist a bijective map x from some open set $U \subset M$ containing p to open set $V \subset \mathbb{R}^d$ such that x and x^{-1} are continuous(Such a map x is called a *homeomorphism* between U and V).

We call such pairs (U, x) as *charts* and x as a *chart map*. We see that by using x you can see the space locally in $\mathbb{R}^d$, i.e acts as a local coordinate system on the manifold. An immediate consequence is that you have a collection of these charts, say $\mathcal{A}$ so that for any point you have some chart which maps a neighbourhood of it to $\mathbb{R}^d$. We say that two charts (U, x) and (V, y) to be C^k-compatible if $(x \circ y^{-1})$ and $(y \circ x^{-1})$ (These functions are called *chart transition maps*) are C^k functions on their respective domains, where a function being C^k means it is differentiable k times. Note that $\mathcal{A}$ as above consists of compatible C^0 charts.This motivates the definition for a smooth manifold, meaning the 'local maps' are smooth.

Definition 0.2.2. An smooth atlas $\mathcal{A}$ on M is a collection of charts (U_α, x_α)(not necessarily countable) such that the chart transition maps are smooth and $\bigcup_\alpha U_\alpha = M$

Definition 0.2.3. A d-dimensional manifold M along with a smooth atlas is called a smooth manifold of dimension d.

One usually adds the further condition that an atlas be maximal (i.e contains all compatible charts) so that two equivalent spaces (upto a diffeomorphism) equipped with different atlases are not counted as different manifolds. Also every atlas is contained in a unique maximal atlas. Example: Of course, $\mathbb{R}^n$ is a smooth manifold of dimension n.
This imposition of smoothness to the chart maps help us generalize differentiability of functions between manifolds in the following way:

Definition 0.2.4. A function $f : M \to N$ is smooth at $p \in M$ if there exist chart (U, x) and (V, y) such that $p \in U$, $f(p) \in V$ and $(y \circ f \circ x^{-1}) : x(U) \to y(V)$ is smooth.

It is easy to see that if for one chart the above condition holds, then compatibility of charts imply that it holds for all charts. Hence it is a well-defined notion. We say f is smooth if it is smooth at every point in M.

0.3 Multilinear algebra

We shall deal with vector spaces over the field $\mathbb{R}$. Let V be a vector space.

Definition 0.3.1. The dual space of V, $V^* :=$ set of all linear maps from V to $\mathbb{R}$ along with '+','·' defined as below

$$
\begin{aligned}
(F + G)(v) &:= F(v) + G(v) \\
(c.F)(v) &:= cF(v)
\end{aligned}
$$

where $c \in \mathbb{R}$, $F, G \in V^*$, so that $(V^*, +, .)$ is a vector space.

For finite dimensional vector space V, let (e_i) be a bases, then we have $e^i : e^i(e_j) = \delta^i_j$ to form a basis of V*. Elements of V^* are called co-vectors.

Definition 0.3.2. A (p, q)-tensor T is a multilinear map $T : V^{*p} \times V^q \to \mathbb{R}$, i.e. T is linear in each of its arguments.

For example:
Every covector is a (0,1) tensor.
If T is a $(1, 1)$-tensor, it takes in a dual vector and a vector to output a scalar. Given an endomorphism G of V (i.e. a linear transformation from V to V), one has the $(1, 1)$ tensor T such that $T(w^*, v) = w^*(G(v))$. One important example is that of the inner product, which is a symmetric , non-degenerate $(0, 2)$ tensor, usually depicted by g(where non-degeneracy means $g(v_1, v_2) = 0$ for any $v_2 \implies v_1 = 0$). The Minkowskian metric η is an important example from special relativity.
The set $T^p_q V$ of (p, q)-tensors can be given a vector space structure by pointwise addition and scaling at each evaluation. It is called the (p, q)-tensor space over V.

Definition 0.3.3. Given (p, q) , (r, s) tensors T and Q respectively, one can define a $(p + r, q + s)$ tensor $T \otimes Q$:

$$
T \otimes Q(w_1, .., w_p, w_{p+1}, ..w_{p+r}, v_1, .., v_{q+s}) = T(w_1, .., w_p, v_1, ..v_q).Q(w_{p+1}, ..w_{p+r}, v_{q+1}, .., v_{q+s})
$$

Thus tensor product lets us construct new tensors from other tensors. One other way of getting new tensor is by contracting a tensor. We will come to this later. Suppose V has basis e_i, $i \leq n$ .Then we have a basis for the (p, q) tensor space over V , given by $e_{j_1} \otimes \cdots \otimes e_{j_p} \otimes e^{i_1} \otimes \cdots \otimes e^{i_q}$ where $j_1, \ldots, j_p, i_1, \ldots, i_q \leq n$, where the component of the (p, q) tensor T is given by $T(e^{j_1}, \ldots, e_{i_q})$ with respect to the basis element $e_{j_1} \otimes \cdots \otimes e_{j_p} \otimes e^{i_1} \otimes \cdots \otimes e^{i_q}$. Denoted by

$$
\begin{aligned}
T &= T(e^{j_1}, \ldots, e_{i_q}) e_{j_1} \otimes \cdots \otimes e_{j_p} \otimes e^{i_1} \otimes \cdots \otimes e^{i_q} \\
&= T^{j_1 \ldots j_p}_{i_1 \ldots i_q} e_{j_1} \otimes \cdots \otimes e_{j_p} \otimes e^{i_1} \otimes \cdots \otimes e^{i_q}
\end{aligned}
$$

Remark. We shall be following the Einstein Summation convention so that whenever an index is there as a subscript and superscript in the same term, we mean the sum over the index.

Given a (p,q)-tensor T, p and $q > 0$, we have a $(p-1, q-1)$-tensor Q by plugging in a basis element e_i of V and its corresponding dual e^i of V* in T and taking the sum, i.e, $Q = \sum_i T(., \ldots, e^i, \ldots, e_i, \ldots)$ It can be checked easily that Q defined in this way is again a multilinear map and hence a valid tensor, moreover the components of this tensor are:

$$Q^{j_1 \ldots j_{p-1}}_{i_1 \ldots i_{q-1}} = \sum_{j_k} T^{j_1 .. j_k .. j_p}_{i_1 ., j_k, .. i_q}$$

Change of basis: First let us consider how vectors and co-vectors behave under basis transformations. Let $\mathcal{B} = (e_i)$ and $\mathcal{B}' = (e'_j)$ be two bases of V. Let $e_i = A^j_i e'_j$, Then if vector $v = v^j e_j = v'^j e'_j$ (note that upon change of basis vector itself does not change but its components do change as follows), Then

$$
\begin{aligned}
v &= & v^i A^j_i e'_j = v'^j e'_j \\
&\implies & v^i A^j_i = v'^j \\
& & v^i = v'^j B^i_j
\end{aligned}
$$

where B is the inverse of A i.e. $A^j_i B^k_j = \delta^k_i$. Similarly if w covector, then $w_i = w(e_i) = w(A^j_i e'_j) = A^j_i w(e'_j) = A^j_i w'_j$. Therefore vector components transform with the inverse transformation of the basis elements, while covector components transform in the same way as the basis elements. Hence they are also called *contravariant* and *covariant* respectively. If you have a (p,q) tensor T, then the tensor components transform as

$$T^{j_1, \ldots, j_p}_{i_1, .., i_q} = A^{i'_1}_{i_1} .. A^{i'_q}_{i_q} B^{j_1}_{j'_1} .. B^{j_p}_{j'_p} T^{j'_1, \ldots, j'_p}_{i'_1, .., i'_q}$$

We first note that if you have the components of a tensor and you define operations on these components, then the operations are well defined as operations on tensors, if the result is independent of the basis. So, if you have objects which obeys transformation laws of a tensor in a change of basis, then the object is a well defined tensor.

Using the metric tensor to change the tensor type:

A vector space with an inner product g has a few additional properties. One can, using g construct a natural isomorphism f between V and V^* (finite dimensional). Define: $f : V \to V^*$ such that $(f(v_1))(v_2) = g(v_1, v_2)$ for vectors v_1, v_2. Let $f(v) = v^* = g(v, .)$ (It is a linear map to $\mathbb{R}$).

A simple evaluation shows that f is linear transformation from V to V^*. The fact that it is one-one follows from the non-degeneracy of g, and rank-nullity theorem shows that it is surjective. Hence f is a vector space isomorphism and for each vector v , you have a dual vector corresponding to it through the natural isomorphism induced by g.

Given a (p,q)-tensor T, one can use this isomorphism f to actually change the tensor type. You can compose the ith input of T with f (or f^{-1}) so that it takes one vector (or covector) in place of a covector(or vector). Let the components of g be g_{ab} with respect to $e^a \otimes e^b$. Then,

$$
\begin{aligned}
(f(v))(w) = g(v, w) &= & g_{ij} e^i \otimes e^j (v, w) = g_{ij} e^i(v).e^j(w) \\
&\implies & f(v) = (g_{ij} e^i(v)) e^j
\end{aligned}
$$

Also, one has the inverse metric tensor g^{-1} defined so that $g^{-1} = g^{ij} e_i \otimes e_j : g^{ij} g_{jk} = \delta^i_k$. This inverse metric tensor can be used to compute f^{-1}: we have

$$f(g^{-1}(e^i, .)) = g^{ij} f(e_j) = g^{ij} g_{jk} e^k = e^i$$

Therefore we have , by linearity of that

$$f(g^{-1}(v , .)) = v \implies g^{-1}(v , .) = f^{-1}(v)$$

Also,

$$
\begin{aligned}
g(f^{-1}(e^j), f^{-1}(e^k)) &= & f(f^{-1}(e^j))(f^{-1}(e^k)) = g^{-1}(e^j, e^k) \\
&\implies & g(v, w) = g^{-1}(f(v), f(w))
\end{aligned}
$$

A simple calculation shows that g^{-1} is an innerproduct on V^*.

0.4 Tangent Spaces and Tangent Bundle

Let us take a look at the example S^2, the unit sphere in $\mathbb{R}^3$ given a smooth manifold structure by using the stereographic projection maps as chart maps. In this case, we can visualize S^2 as sitting in $\mathbb{R}^3$, and have a notion of tangent vectors to the surface at every point of S^2,i.e. the vectors lying in the tangent plane at a given point. But given an abstract manifold, can we get a notion of tangent vectors to the manifold at a point, without actually picturing it in some $\mathbb{R}^n$?(This way of picturing a manifold in some $\mathbb{R}^n$ is achieved by a map which is called an *embedding* and a theorem proved by Whitney shows that any manifold can be embedded in $\mathbb{R}^n$ for some n)
This is in fact the case and is achieved by the following construction:

Definition 0.4.1. Let M be a smooth manifold. Define $C^\infty(M)$ to be the set of all smooth maps from M to $\mathbb{R}$. $C^\infty(M)$ equipped with

$$+ \; : \; C^\infty(M) \times C^\infty(M) \to C^\infty(M)$$
$$\cdot \; : \; \mathbb{R} \times C^\infty(M) \to C^\infty(M)$$

pointwise gives a vector space structure on $C^\infty(M)$

Definition 0.4.2. Let $\gamma : \mathbb{R} \to M$ be a smooth curve through a point $p \in M$. W.l.o.g $p = \gamma(0)$.Then the directional derivative operator at p along γ is the linear map:

$$v_{\gamma,p} : C^\infty(M) \to \mathbb{R}$$
$$v_{\gamma,p}(f) = (f \circ \gamma)'(0)$$

$v_{\gamma,p}$ is called a tangent vector to the curve γ at p.

Definition 0.4.3. The tangent vector space T_pM for $p \in M$ is the collection of $v_{\gamma,p}$ such that γ is a smooth curve via point p.

Note that given some scalar field $f \in C^\infty(M)$, $v_{\gamma,p}(f)$ gives directional derivative of f along γ at p. One again equips T_pM with pointwise addition and scaling(over $\mathbb{R}$) to get a vector space structure. To see that $+$ is closed, given $v_{\gamma,p}$ and $v_{\delta,p}$ one chooses a chart (U,x) such that $p \in U$ and define a curve $\sigma = x^{-1} \circ (x \circ \gamma + x \circ \delta - x(p))$ on $\gamma^{-1}(U) \cap \delta^{-1}(U)$ and extend sigma(the rest doesn't affect the tangent vector at all). An evaluation immediately gives $v_{\sigma,p} = v_{\gamma,p} + v_{\delta,p}$ But for manipulation purposes, one would have to find a basis for the tangent vector space, so that it is enough to manipulate the components. Also how do we draw information about this vector space from the base manifold? One natural way of finding a basis is to use the charts. This motivates the following:

Theorem. $dim(T_pM) = dim(M)$ (where the former is the dimension of a vector space and the latter dimension of a manifold)

Proof. We will construct a basis for T_pM using a chart (U,x) containing p. Suppose $dim(M) = n$, then $x = (x^1, x^2, \ldots, x^n)$ where $x^i : M \to \mathbb{R}$. Consider γ_i to be the curve along x^i (i.e $x \circ \gamma_i$ changes only in the ith direction) i.e $x^b \circ \gamma_a = \delta_a^b \lambda + x^b(p)$. Then put $e_i = X_{\gamma_i,p}$.
Now let $g \in C^\infty(M)$. Then

$$\begin{aligned}
e_i(g) = (g \circ \gamma_i)'(0) &= \left(g \circ x^{-1} \circ x \circ \gamma_i\right)'(0) \\
&= \partial_b \left(g \circ x^{-1}\right)(x(\gamma_i(0))) \cdot \left(x^b \circ \gamma_i\right)'(0) \\
&= \partial_b \left(g \circ x^{-1}\right)(x(p)) \cdot \delta_i^b \\
&= \partial_i \left(g \circ x^{-1}\right)(x(p))
\end{aligned}$$

which is the ith partial derivative of $(g \circ x^{-1})$ at $x(p)$.
Notation: $\left(\frac{\partial g}{\partial x^i}\right)_p = \partial_i \left(g \circ x^{-1}\right)(x(p))$
Then $e_i = \left(\frac{\partial}{\partial x^i}\right)_p$.
To show that (e_i) generates T_pM: Consider $X \in T_pM$ then $X = X_{\mu,p}$ for some smooth curve μ. Put $X^a = (x^a \circ \mu)'(0) \implies X = X_{\mu,p} = X^i e_i$
Plugging x^i in the argument of e_j one can easily show the linear independence. $\qquad \square$

Therefore we can write $\forall X \in T_p M, X = X^a \frac{\partial}{\partial x^a}$. We shall now construct the tangent bundle on the manifold which consists of all the tangent vectors.

Definition 0.4.4. Let M be a smooth manifold of dimension n. Then the tangent bundle, as a set is : $TM := \sqcup T_p M$, where $\sqcup$ is disjoint union of sets.

Definition 0.4.5. The bundle projection $\pi : TM \to M : \pi(v) = p$ where $v \in T_p M$.

But (TM, π, M) as a set bundle is not useful. We need a smooth manifold structure on TM to talk about 'vector fields'(which, vaguely is a function which attaches a tangent vector at every point on M). We shall impose the additional structure on TM using the bundle map π. Make TM into a topological space by giving the initial topology on TM using π and topology on M ,i.e. the open sets in TM are the preimages of open sets of M under π. Note that this construction makes π a continuous map. But this topology is far from *Hausdorff* as any 2 points in the same fibre can't be separated. So, we add open sets to this topology in the following way. Consider a point $X \in TM$. Then $X \in T_p M$ for some $p \in M$. Then we have a chart (U, x) containing p. Consider $preim_\pi(U) = V$ which is open in TM and contains X. Then we will define $\xi : V \to \mathbb{R}^{2n} : \xi(Y) = (x^1(\pi(Y)), x^2(\pi(Y)), \ldots, x^n(\pi(Y)), Y^1, \ldots, Y^n)$ where Y^i is the component of Y along $\frac{\partial}{\partial x^i}$. Now we add open sets to the topology of TM such that $\xi^{-1}(O)$ is open in $\pi^{-1}(U)$, where O is open in $x(U) \times \mathbb{R}^n$. Thus we have chart (V, ξ) containing X. It is easy to check that such maps are C^∞ compatible. Thus the topology induced on TM via this method makes TM a topological manifold. Doing this for every point in TM, we construct the smooth atlas on TM to make it into a smooth manifold.

Definition 0.4.6. Let M be a smooth manifold, TM is its tangent bundle, π the bundle map. A vector field X is a smooth map $X : M \to TM : \pi \circ X = id_M$ where id_M is the identity map on M(such a map is called a *section* of the bundle (M, π, TM)).

Note that 'smoothness' of the vector field X makes sense because we gave a smooth manifold structure to TM. Similarly, one can construct the cotangent bundle $T^* M$ consisting of the covectors and give a smooth manifold structure using charts constructed from the chart map and the dual basis components, and covector fields can then be defined as a section to the cotangent bundle.The dual basis of cotangent vector space using chart map x is usually denoted by (dx^i) where $dx^i(\frac{\partial}{\partial x^j}) = \delta^i_j$.
The set of all vector field over M is denoted by $\Gamma(TM)$ and the covector fields as $\Gamma(T^* M)$

Definition 0.4.7. A (r, s)-tensor field T over a smooth manifold M is a $C^\infty(M)$ multilinear map $T : (\Gamma(T^* M))^r \times (\Gamma(TM))^s \to C^\infty(M)$.

Vector fields in terms of components:

The set $\Gamma(TM)$ defined above can be given a $C^\infty(M)$ module structure by pointwise addition and scaling. For manipulation purposes, it would be very convenient to have a basis for the module of vector fields. But this is not always possible, as a module, in general does not have a basis.
As an example, consider $M = S^2$. Then $\Gamma(TM)$ doesn't have a basis. This is a corollary of the theorem that there exists no everywhere non-vanishing vector field on S^2 (*Hairy Ball Theorem*). Since the tangent space has dimension 2, we need atleast 2 tangent vectors in the tangent spaces at all points for any tangent vector to be written as a linear combination of the tangent vectors given by the basis vector fields. The theorem then necessitates that there should be at least 3 vector fields, say e_1, e_2, e_3 to generate any vector field (as one can consider the point where one of them vanishes, and hence has only atmost 2 non-zero tangent vectors). But then you lose linear independence as you have a non-zero $\lambda^i \in C^\infty(M)$ for $i = 1, 2, 3$ such that $\lambda^1 e_1 + \lambda^2 e_2 + \lambda^3 e_3 = 0$.
But one does have a basis for the vector fields 'locally'. This is the basis induced by the charts. If (U, x) is a chart, then on U , one can write a vector field X (the restriction of X to U, to be precise) as : $X = X^i e_i$ where $e_i(p) = \left(\frac{\partial}{\partial x^i}\right)_p$. Similarly for co-vector field $\omega = \omega_i e^i$ where $e^i(p) = dx^i(p)$. Hence one can locally write down a tensor field in the component form using the tensor product of elements of these bases.

One can now define the action of a vector field X on a $C^\infty(M)$ map f as follows:

$$(X(f))(p) = (X(p))(f)$$

i.e., at each point $p \in M, X(f)$ takes the value of the directional derivative of f along the tangent vector $X(p)$ at p. But tangent vectors act as derivations on $C^\infty(M)$ to $\mathbb{R}$(i.e it is linear over $\mathbb{R}$ and obeys the Leibnitz Rule : if $Y \in T_p M$, then $Y(fg) = Y(f)g(p) - f(p)Y(g)$).Hence the vector fields act as derivations from $C^\infty(M)$ to $C^\infty(M)$ through the action as defined above.

0.5 Connections

We have seen that a smooth vector field X can be used to take directional derivative of a function $f \in C^\infty(M)$

$$\nabla_X f := Xf$$

This notion of derivative of functions can be generalised even for tensors as follows.

Definition 0.5.1. A connection ∇ on a smooth manifold M is a map that takes a vector field X and a (p,q)-tensor field T and sends them to a (p,q)-tensor field $\nabla_X T$ s.t.

$$
\begin{aligned}
\nabla_X f &= Xf \\
\nabla_X (T + S) &= \nabla_X T + \nabla_X S \\
\nabla_{fX+gY} T &= f\nabla_X T + g\nabla_X T \\
\nabla_X (T(\omega, Y)) &= (\nabla_X T)(\omega, Y) + T(\nabla_X \omega, Y) + T(\omega, \nabla_X Y)
\end{aligned}
$$

where the last equation is only for (1,1) tensor field but can be generalised for any tensor.

Lets look how the above definition applies to a vector field Y in a chart (U, x).

$$
\begin{aligned}
\nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} Y^j \frac{\partial}{\partial x^j} \\
&= X^i \nabla_{\frac{\partial}{\partial x^i}} (Y^j \frac{\partial}{\partial x^j}) \\
&= X^i \left((\nabla_{\frac{\partial}{\partial x^i}} Y^j)(\frac{\partial}{\partial x^j}) + Y^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right) \\
&= X^i \left((\nabla_{\frac{\partial}{\partial x^i}} Y^j)(\frac{\partial}{\partial x^j}) + Y^j \Gamma^k_{ji} \frac{\partial}{\partial x^k} \right)
\end{aligned}
$$

Here in the last step we assumed that the vector field $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$ has components Γ^k_{ji}. In component this looks

$$(\nabla_X Y)^k = X^i \frac{\partial}{\partial x^i}(Y^k) + \Gamma^k_{ij} Y^i X^j$$

Specifying Γ^k_{ij} for each index (total $dim(M)^3$ functions) in a neighbourhood U in M suffices to fix the action of ∇ on any vector field. Fortunately those $dim(M)^3$ functions are enough to fix its action on any tensor field. For example for some (0,1)-tensor field(co-vector field) ω and (1,2)-tensor field T we have

$$
\begin{aligned}
(\nabla_X \omega)_k &= X(\omega_k) - \Gamma^m_{kj} \omega_j X^m \\
(\nabla_X T)^i_{jk} &= X(T^i_{jk}) + \Gamma^i_{sm} T^s_{jk} X^m - \Gamma^s_{jm} T^i_{sk} X^m - \Gamma^s_{mk} T^i_{js} X^m
\end{aligned}
$$

Remark. If all the functions Γ^k_{ij} are zero then the manifold is called Euclidean space. Here the connections or co-variant derivatives become the usual differentiation in $\mathbb{R}^n$.

Remark. Choice of Γ^k_{ij} are not independent if metric g_{ij} is given in the manifold M. Then the metric tensor itself induce the $dim(M)^3$ functions as

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{mi} \left(\frac{\partial g_{mj}}{\partial x^k} + \frac{\partial g_{mk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^m} \right)$$

To see the change in the componenents upon a change in co-ordinate assume two charts (U, x) and (V, y). In their region of overlap

$$\Gamma^i_{jk(y)} = dy^i \left(\nabla_{\frac{\partial}{\partial y^k}} \frac{\partial}{\partial y^j} \right) = \frac{\partial y^i}{\partial x^q} dx^q \left(\nabla_{\frac{\partial x^p}{\partial y^k} \frac{\partial}{\partial x^p}} \frac{\partial x^s}{\partial y^j} \frac{\partial}{\partial x^s} \right)$$

which upon further simplfication gives

$$\Gamma^i_{jk(y)} = \frac{\partial y^i}{\partial x^q} \frac{\partial x^s}{\partial y^j} \frac{\partial x^p}{\partial y^k} \Gamma^q_{sp(x)} + \frac{\partial y^i}{\partial x^q} \frac{\partial^2 x^q}{\partial y^k \partial y^j} \tag{1}$$

As can be readily inferred from the above equation, the Christoffel Symbols Γ^i_{jk} do not transform like a tensor. Infact transformation behaviour of this sort allows covariant derivative of a tensor to transform like a tensor.

Remark. One can always choose a chart such that Γ^i_{jk} become particularly simple. In fact one can show that for any point $p \in M$, choice of so called *normal co-ordinate* vanishes all the Γ's at that point. However this is not possible globally.

0.6 Parallel Transport

Introduction of a connection in a manifold allows one to "compare" vectors (tensors) at different points. In a curved space the result of parallely transporting a vector from a point to another one depends on the path taken as clearly described by the figure below. And one can naively think that comparing difference in result of such parallel transport will shed some light on the curvature of the space.
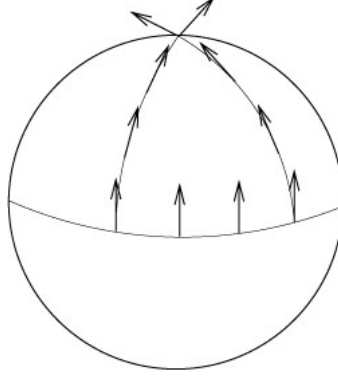

Figure 1: In a sphere start with a vector on the equator, pointing along a line of constant longitude. Parallelly transport it up to the north pole along a line along longitude in the obvious way. Then take the original vector, parallel transport it along the equator by an angle θ, and then move it up to the north pole as before. It is clear that the vector, parallel transported along two paths, arrived at the same destination with two different values (rotated by θ).

To formulate these notions one has to define what is meant by *parallel transport*.

Definition 0.6.1. A vector field X on M is said to be parallely transported along a smooth curve $\gamma : \mathbb{R} \to M$ if $\nabla_{v_\gamma} X = 0$.

Definition 0.6.2. A curve $\gamma : \mathbb{R} \to M$ is called autoparallely transported if $\nabla_{v_\gamma} v_\gamma = 0$.

Remark. v_γ is not exactly a vector field, it is only defined on the curve γ. So to make sense of the definition one has to analytically continue v_γ throughout M and then apply the above definition. However the continuation might not be unique. Fortunately one can show that value of the vectors outside the curve is irrelevent for the above definition hence the definition makes sense.

Now lets look how the equation looks in a chart (U, x).

$$\nabla_{v_\gamma} v_\gamma = \nabla_{\dot{\gamma}^m \frac{\partial}{\partial x^m}} \dot{\gamma}^n \frac{\partial}{\partial x^n} = \dot{\gamma}^m \frac{\partial \dot{\gamma}^n}{\partial x^m} \frac{\partial}{\partial x^n} + \dot{\gamma}^m \dot{\gamma}^n \Gamma^q_{mn} \frac{\partial}{\partial x^q} = \ddot{\gamma}^q \frac{\partial}{\partial x^q} + \dot{\gamma}^m \dot{\gamma}^n \Gamma^q_{mn} \frac{\partial}{\partial x^q}$$

hence we get for an autoparallely transported curve

$$\ddot{\gamma}^q + \Gamma^q_{mn} \dot{\gamma}^m \dot{\gamma}^n = 0$$

The equation is famously known as *geodesic equation*. A particle without any external field in a manifold will follow the path described by this equation.
For euclidean space $U = \mathbb{R}^2$ $x = id_{\mathbb{R}^2}$ and $\Gamma^q_{mn} = 0$. Then the geodesic equation becomes

$$\ddot{\gamma}^q = 0 \implies \gamma^q = a^q \lambda + b^q$$

which is exactly the equation of straight line.

0.7 Torsion

Definition 0.7.1. The torsion of a connection ∇ is the $(1, 2)$-tensor field

$$T(\omega, X, Y) := \omega \left(\nabla_X Y - \nabla_Y X - [X, Y] \right)$$

where $[X, Y]$ is a vector field defined such a way that $[X, Y]f := X(Yf) - Y(Xf)$.

In a chart,

$$T^i_{jk} := T\left(dx^i, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right)$$
$$= dx^i\left(\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k} - \nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^j} - \left[\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right]\right)$$
$$= \Gamma^i_{kj} - \Gamma^i_{jk} - 0$$

The last term vanishes as partial derivatives commute with each other.

Definition 0.7.2. A smooth manifold M with a connection is called torsion-free if $T = 0$.

In terms of Christoffel symbols this translates to the fact that Γ^i_{jk} is symmetric under two lower indices i.e. $\Gamma^i_{kj} = \Gamma^i_{jk}$. The assumption that a manifold is torsion-free is always taken in General Relativity. Geometrical significance of this assumption is not clearly understood.

0.8 Curvature

Definition 0.8.1. Riemann curvature of a connection is the $(1,3)$-tensor field

$$R(\omega, Z, X, Y) := \omega\left(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z\right)$$

To get the idea why this measures the curvature of space note that the covariant derivative of a tensor in a certain direction measures how much the tensor changes relative to what it would have been if it had been parallel transported (since the covariant derivative of a tensor in a direction along which it is parallel transported is zero). The commutator of two covariant derivatives, then, measures the difference between parallel transporting the tensor first one way and then the other, versus the opposite ordering as shown by the picture below. Components of Riemann tensor-field in a chart has the following form

$$R^a_{bcd} = \partial_c\Gamma^a_{bd} - \partial_d\Gamma^a_{bc} + \Gamma^e_{bd}\Gamma^a_{ec} + \Gamma^e_{bc}\Gamma^a_{ed}$$

where $\partial_c := \frac{\partial}{\partial x^c}$. From the above equation it can be easily seen that R is antisymmetric in its last two entry. And one can also show by computation that

$$[\nabla_\mu, \nabla_\nu]X^\rho = R^\rho_{\sigma\mu\nu}X^\sigma - T^\lambda_{\mu\nu}\nabla_\lambda X^\rho$$

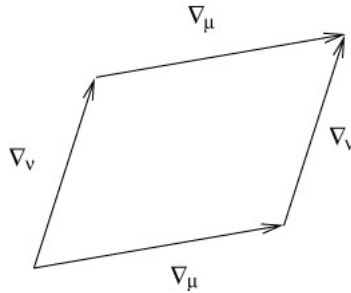where $\nabla_\mu = \nabla_{\frac{\partial}{\partial x^\mu}}$.

Figure 2: Geometric interpretation of Riemann tensor

Riemann tensor is extremely important in general relativity. Here is some properties of the components R^a_{bcd}.

$$R^a_{bcd} + R^a_{cdb} + R^a_{dbc} = 0$$
$$\partial_e R^a_{bcd} + \partial_d R^a_{bec} + \partial_c R^a_{bde} = 0 \tag{2}$$

where the last identity is termed as Bianchi identity.
Riemann tensor is the key to measure curvature in a differentiable manifold. It essentially captures the geometry of the manifold. Not to say, it plays a central role in General relativity, though in GR

one only needs a restricted knowledge of the Riemann tensor, namely its contraction, the Ricci tensor. Contraction of 1st and 3rd indices of Riemann tensor produces the Ricci tensor $R_{bd} := R^a_{bad}$. We are using same notation for both Riemann tensor as well as Ricci tensor but this should not confuse the reader as the number of indices are different. Note that Ricci tensor is symmetric in its two indices i.e. $R_{ab} = R_{ba}$. Also one can define a scalar called Ricci scalar by contracting Ricci tensor with the metric tensor $R = g^{ab}R_{ab}$. This scalar also captures the curvature of space but not quite as accurate as Riemann tensor. But neverthless it is very important in GR. There is one more important tensor left to be defined. Einstien tensor G_{ab} which is defined to be component-wise as

$$G_{ab} = R_{ab} - \tfrac{1}{2}g_{ab}R$$

This tensor has one very interesting property that it's divergence is zero i.e. $\nabla_{\frac{\partial}{\partial x^b}} G_{ab} = 0$. This can be proven using Bianchi identity. This fact will be used later.

Above paragraph consists of various example of tensors and scalar which measure the curvature at a point in a differential manifold. Those are enough to talk about geometry of curved space as far as General Relativity is concerned. Now that we have developed how to talk about the curvature of space (spacetime) next order of business will be to understand what changes the curvature and a self-consisting equation relating the two to determine the dynamics of a particle in spacetime.

0.9 Energy Momentum Tensor

The most important tensor in physics, especially in astrophysics, is the energy-momentum tensor. As the name suggests, this tensor keeps track of energy and momentum densities in space-time. Historically, this tensor is written as T. One can expect this tensor to be dependent on the system of interest. In mechanics we encode all the information of the system in a single scalar called Lagrangian $\mathcal{L}$. Components of energy momentum tensor T_{ab} can be derived from the Lagrangian as follows:

$$T^{ab} = \frac{\partial \mathcal{L}}{\partial(\partial_a \phi)}\partial^b \phi - g^{ab}\mathcal{L}$$

Here are few examples of energy-momentum tensors

- *Isolated Particle*: $T^{ab}(\mathbf{x},t) = \dfrac{mv^a(t)v^b(t)}{\sqrt{1-(v/c)^2}}\delta(\mathbf{x}-\mathbf{x}_p(t))$ where $\mathbf{x}_p(t)$ is the trajectory of the particle.

- *Perfect fluid in equlibrium*: $T^{ab} = \left(\rho + \frac{p}{c^2}\right)u^a u^b + pg^{ab}$ where ρ and p are respectively the density and pressure of the fluid.

- *Electromagnetic Energy-momentum tensor*: $T^{ab} = \frac{1}{\mu_0}\left(F^{ac}g_{cd}F^{bd} - \frac{1}{4}g^{ab}F_{cd}F^{cd}\right)$ where F^{ab} is the electromagnetic field tensor.

The components of energy-momentum tensor can be interpreted as follows: T^{ab} *is the flux of ath component of the momentum vector across a surface with constant x^b co-ordinate*. In relativity, momentum refers to the four-momentum. In this case time-time component is the density of relativistic mass i.e. the energy density divided by c^2. Flux of relativistic mass across x^i surface is represented in the component T^{0i}. T^{ii} (not summed) represents the normal stress i.e. pressure whereas T^{ik} represent shear stress where i, k take values 1,2,3.

Energy momentum tensor in General relativity is symmetric in its components i.e. $T^{ab} = T^{ba}$. Also, from energy and momentum conservation we have the following relation:

$$\nabla_{\frac{\partial}{\partial x^b}} T^{ab} = 0$$

or equivalently $\nabla_{\frac{\partial}{\partial x^b}} T_{ab} = 0$.

0.10 Einstein Equation

We have developed how to describe curvature in space(space-time) and defined energy momentum tensor which represents the mass distribution, the source of curvature. Einstein equation relates these two

quantities to get a field equation. Lets try to get to the equation by some informal reasoning and analogy and of-course,a little bit of hand-waving.

We seek an equation much like Maxwell equation in electromagnetism, motivated from Newtonian Gravity, namely Poisson equation

$$\nabla^2 \Phi = 4\pi G \rho$$

The right hand side represents the mass density ρ which we are tempted to replace with the energy-momentum tensor. Whereas the left hand side Laplacian of gravitational potential should be replaced by curvature of space as argued by Einstein using *Equivalence Principle*. Curvature is given by Riemann tensor, Ricci tensor and Ricci scalar. Which one should we use? Also the metric tensor g_{ab} encodes local curvature. Riemann tensor option can be ruled out as the right hand side contains only two indices. So some combination of Ricci tensor, Ricci scalar and metric tensor is the most general choice of the left hand side. But note that by energy conservation ,we have that the divergence of the right hand side has to be zero. So, for consistency we need divergence of the left hand side to vanish. Only one combination of R_{ab}, R and g_{ab} is divergenceless, which has been introduced earlier, the Einstein Tensor G_{ab}. Hence the final equation should look like

$$G_{ab} = \kappa T_{ab}$$

or in full glory

$$R_{ab} - \tfrac{1}{2} R g_{ab} = \kappa T_{ab}$$

where $\kappa = \frac{8\pi G}{c^2}$ is a constant which can be fixed by taking limits to Newtonian gravity.

Reference

1. SPACETIME AND GEOMETRY: An introduction to General Relativity; Carroll, Sean ISBN 0-8053-8732-3

2. MANIFOLD THEORY: An introduction for mathematical physicists; Martin, Daniel ISBN 0-13-543877-2

3. Lecture notes on General Relativity S. Carroll: https://ned.ipac.caltech.edu/level5/March01/Carroll3/Carroll3.html

4. Lectures on Geometrical Anatomy of Theoretical Physics by Fredric P. Schuller given at Friedrich-Alexander-Universität Erlangen-Nürnberg.

5. Lectures from the WE-Heraeus International Winter School on Gravity and Light given by Fredric P. Schuller in 2015.