Sen's Lectures
General Relativity

Volume 1

ATUL SINGH ARORA

Contents

Foreword

The content is due to Prof. Ashoke Sen while all errors in its presentation here are mine. Strictly then this contains not a word of Ashoke Sen and effectively no thought of mine (reminds you of something?). That was the disclaimer, wait, one more thing: there's nothing useful any further in this section. So with that I am free to express my motivation for writing this. Ashok Sen needs no introduction. I did not get a chance to learn GR at IISER due to my strange selection of electives. This, however, is something I always wanted to learn. Instead of following my usual route of picking up a text book, I thought perhaps watching Ashoke Sen's lectures, delivered at IISc (I think) which some student (Anurag I think) was kind enough to upload online, would be quicker. Watching the lectures was a little challenging because the video quality sometimes distorts the text on the blackboard and one has to make educated guesses. Of course Prof Sen speaks out everything he writes, well almost everything, so one can check the logical consistency with his cues to be sure that everything is correct. That's a good exercise. I have tried making notes and then presenting them in a neat format here as a way to make the subject clear to myself but I suspect this would be of use to others as well. Prof. Sen has also delivered more lectures on many other topics, both advanced and basic, which I might try to cover in my subsequent efforts.

Part 1

Reimannian Geometry

CHAPTER 1

Describing Curved Spaces

1.1. Motivation

Recall that Newton's law for the force of gravity

$$\vec{F} = -\frac{Gm_1m_2}{r^2}\hat{r}$$

has an uncanny resemblance to that of electrostatic attraction between two charged particles. From this perspective generalisation of electrostatics to Maxwell's theory of electrodynamics is similar to the generalisation of Newtonian gravity to Einstein's general theory of gravity (GR). In particular, both electrostatics and Newtonian gravity have the property of instantaneous propagation of effects as they both depend on the relative distance of particles at the instant of time when the force is to be evaluated. In electrodynamics this instantaneity is replaced with propagation of electromagnetic waves and one of the main goals in generalising Newtonian gravity is to find a similar explanation.

Just as vector calculus is the underlying mathematics of Maxwell's electromagnetism, Reimannian Geometry – a generalisation of Euclidean Geometry – is that of GR. Thus one must familiarise oneself with at least the basic concepts of Reimannian Geometry as a prerequisite to learning GR.

1.2. Euclidean Geometry, Metric, Tensors

Consider two points (x^1, x^2, x^3) and $(x^1+dx^1, x^2+dx^2, x^3+dx^3)$. According to Euclidean geometry the distance between these points is given by $ds = \sqrt{(dx^1)^2 + (dx^2)^2 + (dx^3)^2}$. This can be generalised to N-dimensions as

$$ds^2 = \sum_{i=1}^{N}(dx^i)^2.$$

In Remanian geometry the notion of distance between two close-by points is generalised to

$$ds^2 = \sum_{i,j=1}^{N} g_{ij}(\vec{x})dx^i dx^j$$

where $g_{ij}(\vec{x})$ is called the *metric* and is a function of $\left(x^1, x^2 \ldots x^N\right)$. Since $dx^i dx^j = dx^j dx^i$, viz. it is symmetric under the exchange of i and j, it follows that $g_{ij} = g_{ji}$ (see exercise 3). Euclidean geometry can be obtained as the special case where $g_{ij} = \delta_{ij}$, the Kroneker delta, which is defined as

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

It is possible, however, that two different metrics represent the same space. Consider another co-ordinate system which labels the same point $\vec{x}$ as $\vec{x}'$ with components $\left(x'^1, x'^2 \ldots x'^N\right)$. Again consider two points, $\vec{x'}$ and, a point close to it, $\vec{x'} + \vec{dx}'$ where the old $\vec{dx}$ is related to the new $\vec{dx}'$ by

$$dx^i = \sum_k \frac{\partial x^i}{\partial x'^k} dx'^k.$$

Substituting for dx^i in the expression for ds^2 one gets

$$ds^2 = \sum_{k,l} \underbrace{\left(\sum_{i,j} g_{ij} \frac{\partial x^i}{\partial x'^k} \frac{\partial x^j}{\partial x'^l} \right)}_{g'_{kl}} dx'^k dx'^l = \sum_{k,l} g'_{kl} dx'^k dx'^l.$$

Thus a transformation change can cause g to change even though they both describe the same object.

Example 1 (Flat Space). In cartesian coordinates $ds^2 = \left(dx^1\right)^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2$. Using the polar coordinates one can write

$$\begin{aligned} x^1 &= r \sin\theta \cos\phi \\ x^2 &= r \sin\theta \sin\phi \\ x^3 &= r \cos\theta \end{aligned}$$

which can be used to determine $dx^1 = \sin\theta \cos\phi dr + r \cos\theta \cos\phi d\theta - r \sin\theta \sin\phi d\phi$, $dx^2 = \dots$ and $dx^3 = \dots$. When plugged into the expression for ds^2 one obtains $ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2$ (this can also be obtained more easily by writing $d\vec{r} = dr\hat{r} + r d\theta\hat{\theta} + r\sin\theta d\phi\hat{\phi}$ and then using $ds^2 = d\vec{r}.d\vec{r}$ as described in Landau's book). Manifestly then $g_{rr} = 1$, $g_{\theta\theta} = r^2$, $g_{\phi\phi} = r^2 \sin^2\theta$ while in the cartesian case $g_{11} = g_{22} = g_{33} = 1$ even though they both describe the same space.

Example 2 (Surface of Sphere). Consider the surface of a sphere described by $\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 = a^2$. One parametrisation is to use x^3 as a dependent variable given by $x^3 = \pm\sqrt{a^2 - (x^1)^2 - (x^2)^2}$. One can substitute dx^3 in the expression for $ds^2 = \left(dx^1\right)^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2$ to obtain

$$ds^2 = \underbrace{\left(1 + \frac{\left(x^1\right)^2}{a^2 - \left(x^1\right)^2 - \left(x^2\right)^2}\right)}_{g_{11}} \left(dx^1\right)^2 + \underbrace{\left(1 + \frac{\left(x^2\right)^2}{a^2 - \left(x^1\right)^2 - \left(x^2\right)^2}\right)}_{g_{22}} \left(dx^2\right)^2 + \underbrace{\left(\frac{2x^1x^2}{a^2 - \left(x^1\right)^2 - \left(x^2\right)^2}\right)}_{g_{12}} dx^1 dx^2.$$

Alternatively, one could've started with the polar coordinate system and set $r = a$ to obtain $ds^2 = a^2 d\theta^2 + a^2 \sin^2\theta d\phi^2$ which entails $g_{\theta\theta} = a^2$, $g_{\phi\phi} = a^2 \sin^2\theta$. Note that g is diagonal in the polar representation.

Given only the metric how can one conclude whether or not they represent the same space? Our strategy would be to find appropriate linear combinations of the metric and its derivatives which are invariant under coordinate transformations. To proceed we define the following convention.

(1) Index of a coordinate is given by a superscript, e.g. x^i.
(2) Shorthand for derivatives:
$$\partial_i := \frac{\partial}{\partial x^i}.$$
(3) Summation: Any index appearing twice in a formula, once as a subscript and once as a superscript is summed over.
(4) Index of a matrix appears as a subscript.

Under this convention we have

$$ds^2 = \sum_{i,j} g_{ij} dx^i dx^j \to g_{ij} dx^i dx^j$$

and

$$g'_{kl}(\vec{x'}) = \sum_{i,j} g_{ij} \frac{\partial x^i}{\partial x'^k} \frac{\partial x^j}{\partial x'^l} \to g_{ij} \partial'_k x^i \partial'_l x^j.$$

We can now capture this

EXERCISE 3. Consider a symmetric matrix A with elements $A_{ij} = A_{ji}$. If one wishes to evaluate $\sum_{i,j} A_{ij}B_{ij}$ then show that it suffices to assume B is also symmetric.

SOLUTION. Assume $B^{(s)}_{ij} := \frac{B_{ij}+B_{ji}}{2} = B^{(s)}_{ji}$ and $B^{(a)}_{ij} := \frac{B_{ij}-B_{ji}}{2} = -B^{(a)}_{ji}$. Evidently $B_{ij} = B^{(s)}_{ij} + B^{(a)}_{ij}$ so it suffices to show that $\sum_{i,j} A_{ij}B^{(a)}_{ij} = 0$ to prove the claim. This follows from $\sum_{i,j} \frac{1}{2}\left(A_{ij}B^{(a)}_{ij} + A_{ji}B^{(a)}_{ij}\right) = \sum_{i,j} \frac{1}{2}\left(A_{ij}B^{(a)}_{ij} - A_{ji}B^{(a)}_{ji}\right) = 0$.

We capture the generalisation of objects that transform this way by the following definition.

DEFINITION 4 (Tensors). An object that transforms as

$$C'^{i_1\dots i_p}{}_{g_1\dots g_q}(\vec{x}') = \partial_{k_1}x'^{i_1}\dots\partial_{k_p}x'^{i_p}.\partial'_{g_1}x^{l_1}\dots\partial'_{g_q}x^{l_q}C^{k_1\dots k_p}{}_{l_1\dots l_q}(\vec{x})$$

will be termed a rank (p,q) tensor.

DEFINITION 5 (Scalars, Contravariant and Covariant vectors). A rank

- (0,0) tensor is called a scalar
- (1,0) tensor is called a contravariant vector
- (0,1) tensor is called a covariant vector.

Some observations: Note that $A^{i_1\dots i_p}{}_{j_1\dots j_q}B^{k_1\dots k_r}{}_{l_1\dots l_S}$ transforms as a $(p+r, q+s)$ tensor. Note also that for consistency $A^{i_1}{}_{i_1}$ should behave as a scalar. This follows directly from the transformation law and chain rule as

$$\begin{aligned} A'^{i_1}{}_{i_1} &= \underbrace{\partial_{k_1}x'^{i_1}\partial'_{i_1}x^{l_1}}A^{k_1}{}_{l_1} \\ &= \delta^{l_1}_{k_1}A^{k_1}{}_{l_1} = A^{l_1}{}_{l_1}. \end{aligned}$$

1.3. Understanding the Metric: Christophel Symbol, Riemann Tensor

We try to extract all essential information about the space as described by the metric. The main source of redundancy here is the freedom in choosing the coordinate system. Let us proceed systematically with first simplifying the metric itself.

PROPOSITION 6. *Take a point p and assume its coordinate is given by $\vec{x}_0$ in some coordinate system. Then, there exists a coordinate transformation $\vec{x} \to \vec{x}'$ such that $g_{ij}(\vec{x}_0) \to g'_{ij}(\vec{x}'_0) = diag\{\pm, \pm \dots, \pm\}$.*

The proof is straight forward as we shall see. It is interesting to note that every continuous metric[1] will have a fixed number of $+$ and $-$ entries throughout because otherwise there would have to be a jump for a $+$ to turn into a $-$ (which is forbidden by the continuity assumption).

The following expansion will be used several times in this sub-section. Consider a point denoted by $\vec{x}_0$ and $\vec{x}'_0$ by two coordinate systems which are related by $x^i = f^i(\vec{x}')$. Let us write an expression for a point near $\vec{x}_0$ as

$$x^i = f^i(\vec{x}'_0 + (\vec{x}' - \vec{x}'_0)) = x^i_0 + A^i{}_j(x'^j - x'^j_0) + B^i{}_{jk}(x'^j - x'^j_0)(x'^k - x'^k_0) + \dots$$

where we used $x^i_0 = f^i(\vec{x}_0)$ together with the assumption that x_0 and x'_0 represent the same point. Note that $A^i{}_j = \partial'_j f^i(x')|_{\vec{x}'=\vec{x}'_0} = \partial'_j x^i|_{\vec{x}'=\vec{x}'_0}$ and similarly $B^i{}_{jk}$ would consist of two derivatives and so on.

The basic idea will be that instead of specifying f^i directly we would specify f partially by fixing the constants A, B, ... This will become clear in the following proof of the proposition.

[1] one that changes only a little when the coordinate is changed a little, roughly speaking

PROOF. $g'_{kl}(\vec{x}'_0) = \partial'_k x^i \partial'_l x^j g_{ij}|_{\vec{x}'=\vec{x}'_0 \leftrightarrow \vec{x}=\vec{x}_0} = A^i{}_k A^j{}_l g_{ij}(\vec{x}_0) = \left(A^T g A\right)_{kl}$ where in the last step we regard g_{ij} as a matrix with i as row and j as column. To proceed, first note that one can find S s.t. $S^T S = \mathbb{I}$ and $g = S^T g_d S$ where $g_d = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Now define $R = R^T = \mathrm{diag}(\sqrt{|\lambda_1|}, \sqrt{|\lambda_2|}, \dots, \sqrt{|\lambda_n|})$ so that $R^T \eta R = \mathrm{diag}(\lambda_1, \lambda_2 \dots, \lambda_n)$ where $\eta = \mathrm{diag}(\lambda_1/|\lambda_1|, \lambda_2/|\lambda_2|, \dots, \lambda_n/|\lambda_n|)$. Thus I can write $g = S^T g_d S = S^T R^T \eta R S$. We're almost there, let $g'_{kl} = \left(A^T g A\right)_{kl} = \left(A^T S^T R^T \eta R S A\right)_{kl}$ which entails that for $A = (RS)^{-1}$ we'll get $g'_{kl}(\vec{x}'_0) = \eta_{kl}$. This does it then because η is of the $\mathrm{diag}(\pm, \pm, \cdots \pm)$ form. Note that A is not unique because when $\eta = \mathbb{I}$ for all unitaries U, $A = (URS)^{-1}$ will work. $\square$

The set of eigenvalues of g', $\{\pm, \pm, \cdots \pm\}$, as described above are referred to as the *signature* of the metric. When all the entries are $+1$ then the metric is called *Eucledian*. When one entry is -1 (while the others are $+1$) the metric is called *Lorentzian*.

Remarks: First, note that 'locally' then, viz. from g itself (as opposed to its derivatives), there's no more information which can be extracted. Second, this will not distinguish between the surface of a sphere and a plane (both would be Eucledean). Third, the procedure followed above was more than diagonalisation as A is not such that $A^T A = \mathbb{I}$.

We must explore the first derivative of the metric. Consider $K_{ijk} := \partial_i g_{jk}$ so that in the primed coordinate system we would have

$$\begin{aligned}
K'_{i_1 i_2 i_3} &= \partial'_{i_1}(g'_{i_2 i_3}(x')) \\
&= \partial'_{i_1}(\partial'_{i_2} x^{j_2} \partial'_{i_3} x^{j_3} g_{j_2 j_3}) \\
&= \partial'_{i_1}\partial'_{i_2} x^{j_2} \partial'_{i_3} x^{j_3} g_{j_2 j_3} + \partial'_{i_2} x^{j_2} \partial'_{i_1}\partial'_{i_3} x^{j_3} g_{j_2 j_3} + \partial'_{i_2} x^{j_2} \partial'_{i_3} x^{j_3} \partial'_{i_1} x^{j_1} \partial_{j_1} g_{j_2 j_3}
\end{aligned}$$

where the in the last step we used $\partial'_{i_1} = \partial'_{i_1} x^{j_1} \partial_{j_1}$. The last term above can be written as $\partial'_{i_1} x^{j_1} \partial'_{i_2} x^{j_2} \partial'_{i_3} x^{j_3} K_{j_1 j_2 j_3}$ which is how a tensor should transform. However, manifestly there are other terms which means K_{ijk} doesn't transform as a tensor.

One idea is to take some linear combination of K_{ijk} (or even polynomials of it) to construct a tensor. This turns out to be impossible.

PROPOSITION 7. *Any tensor constructed using a linear combination of $K_{ijk} := \partial_i g_{jk}$ or its powers will be the trivial zero tensor.*

PROOF. Assume it can be shown that given a point, there exists a coordinate transformation such that K vanishes at that point. Now if K were to transform as a tensor, then K will be zero at that point in all frames because tensor transformations only involve multiplicative factors. One can repeat this argument for each point thereby establishing that K must be zero. This establishes the result under the assumption stated which we now prove.

To make $K'_{ijk} = 0$ at a given point we need to adjust terms involving two derivatives of x. These are given by $B^i{}_{jk}$ as described in the expansion above. Now K'_{ijk} has two symmetric indices and one independent index which yields $\frac{n.(n+1)}{2}.n$ constraints while $B^i{}_{jk}$ makes for $\frac{n.(n+1)}{2}.n$ adjustable parameters because again j, k are symmetric indices. One can explicitly find the values of $B^i{}_{jk}$ which achieve the said constraint (which is left to the reader). $\square$

What happens if we take two derivatives of the metric? Consider $S_{ijkl} := \partial_i \partial_j g_{kl}$ and now to apply the same argument we count. The number of constraints is $\frac{n(n+1)}{2}\frac{n(n+1)}{2}$ because both i, j and k, l are symmetric. However this time $C^i{}_{jkl}$ would be important and these are $n.\frac{n(n+1)(n+2)}{3!}$ because j, k, l are symmetric this time. The difference between the number of constraints and number of parameters is $\frac{1}{12} n^2 (n-1)$ which means in general one can't make $S'_{ijkl} = 0$ at a given point and thus we can't apply the argument. This means S can be a candidate for constructing a tensor.[2]

Indeed, with $\Gamma^i{}_{jk} := \frac{1}{2} g^{il}(\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk})$ defined to be the *Cristoffel symbol/Connection* we define the *Riemann Tensor* as

$$R^i{}_{jkl} := \partial_l \Gamma^i{}_{jk} - \partial_k \Gamma^i{}_{jl} + \Gamma^m{}_{jk}\Gamma^i{}_{lm} - \Gamma^m{}_{jl}\Gamma^i{}_{km}.$$

While this is intimidating at first note that Γ is not a tensor and can be shown to transform as $\Gamma'^l{}_{mn} = \partial_i x'^l \partial'_m x^j \partial'_n x^k \Gamma^i{}_{jk} + \partial_k x'^l \partial'_m \partial'_n x^k$ (left as an exercise). Using this one can explicitly check that $R^i{}_{jkl}$ indeed transforms as a tensor (exercise).

1.4. Understanding the Reimann Tensor

We define $R_{ijkl} := g_{im} R^m{}_{jkl}$ and use this to state some symmetries of R_{ijkl} and hint at how they might be proved.

Claim 8. $R_{ijkl} = -R_{jikl}$, $R_{ijkl} = -R_{ijlk}$ that is the Reimann Tensor, R_{ijkl}, is anti-symmetric in the indices i, j and also in k, l.

Proof sketch. Consider the symmetric tensor $A_{ijkl} = R_{ijkl} + R_{jikl}$. Note that $R'_{ijkl} = \partial'_i x^m \partial'_j x^n \partial'_k x^p \partial_l x^q R_{mnpq}$ so that

$$\begin{aligned} A'_{ijkl} &= \partial'_i x^m \partial'_j x^n \partial'_k x^p \partial_l x^q R_{mnpq} + \underbrace{\partial'_j x^m \partial'_i x^n \partial'_k x^p \partial_l x^q R_{mnpq}}_{= \partial'_j x^n \partial'_i x^m \partial'_k x^p \partial_l x^q R_{nmpq}} \\ &= \partial'_i x^m \partial'_j x^n \partial'_k x^p \partial_l x^q A_{mnpq} \end{aligned}$$

which shows that indeed A_{ijkl} is a tensor. Recall that at a coordinate system can be found such that the first derivative of the metric vanishes at a point. Using this and the definition of R_{ijkl} and $\Gamma^i{}_{jk}$ one can show that $A_{ijkl} = 0$ at a point. Since the point was arbitrary one concludes (similar to the arguments used above) that A_{ijkl} is a zero tensor which proves the result. ☐

Claim 9. $R_{ijkl} = R_{klij}$

Proof (hint really). This can be proved by looking at the definitions (apparently!). ☐

Claim 10. $R_{i[jkl]} := \frac{1}{6}\left(R_{ijkl} - R_{ijlk} + R_{iljk} - R_{ilkj} + R_{iklj} - R_{ikjl}\right) = 0$

Proof (same as above). ☐

We now define $g^{ij} := g^{-1}{}_{ij}$ where $g^{-1}{}_{ij}$ is the i, j element of the inverse of the matrix g. It follows that g^{ij} is a tensor (exercise) and thus a well defined quantitiy. This is can be used to extract smaller tensors and scalars out of R_{ijkl} which are important enough to bear their own names. $g^{ik} R_{ijkl} = R_{jl}$ is called the *Ricci Tensor*. Note that from the symmetry property $R_{ijkl} = R_{klij}$ it follows that $R_{jl} = R_{lj}$. Note also that $g^{ij} R_{ijkl} = 0$ as it contracts symmetric and anti-symmetric indices. Finally we define $g^{lj} R_{lj} = R$ to be the *Ricci Scalar* or *curvature scalar*.[3] We conclude by examining whether or not after all the hard work we can distinguish between the surface of a sphere and flat space.

[2] The astute reader might complain that we still have some freedom in the As (for e.g. the Unital freedom pointed out for the Eucledian case). This ignorance is justified because As are made of first derivatives of the metric. These first derivatives appear multiplicatively in the tensor transformation. It is easy to see that if a tensor were zero at a point it will remain so regardless of the As. Thus changes in As can't influence the counting.

[3] Note that for scalars one has $R'(x') = R(x)$. Note also that $A_i := \partial_i R$ is a tensor as $A'_i = \partial'_i R' = \partial'_i x^j \partial_j R = \partial'_i x^j A_j$.

EXAMPLE 11. Consider the following cases:

Case A: A plane described by the metric $ds^2 = \left(dx^1\right)^2 + \left(dx^2\right)^2 = dr^2 + r^2 d\theta^2$ which entails $g_{ij} = \delta_{ij}$ while for the polar case $g_{rr} = 1$ and $g_{\theta\theta} = r^2$.

Case B: Surface of a sphere described by the metric $ds^2 = \left(d\theta^2 + \sin^2\theta d\phi^2\right) a^2$ which entails g is diagonal with $g_{\theta\theta} = a^2$, $g_{\phi\phi} = \sin^2\theta a^2$.

Question: Can one tell (without using any prior knowledge of how the metric was constructed) if the two cases describe the same space?

For case A it follows immediately, for the cartesian coordinates, that $R_{ijkl} = 0$ because $g_{ij} = \text{const}$ and so ${\Gamma^i}_{jk} = \partial_l {\Gamma^i}_{jk} = 0$. Since R_{ijkl} is a tensor, it must also vanish in all coordinate systems and in particular in the polar coordinate system. If $R_{ijkl} \neq 0$ for case B then we know the two cases described different *manifolds*[4]. This indeed turns out to be the case and is left to the reader to verify.

Question: Does scaling leave the manifolds invariant?

By this we mean stretching the manifold, i.e. $x^i \to \lambda x^i$. For case A it is intuitively obvious that stretching an infinite plane would leave it unchanged. This is also consistent with the fact that R_{ijkl} stays zero for both situations. However, for case B it is not as obvious because intuitively the surface of a very large balloon would almost seem flat compared to that of a small balloon. We consider the Ricci scalar to answer the question: $R \overset{\text{claim}}{=} \text{const.} a^{-2}$. Now since stretching changes the radius, a, of the sphere it follows that R will get changed. Since R is a scalar (and hence unchanged by any coordinate transformation) we are forced to conclude that stretching changes the manifold.[5]

In general it is hard to compare metrics to check if they describe the same manifold but these tools give us some handle on them.

1.5. Covariant Derivatives

Taking the derivative of a tensor doesn't produce a tensor in general. Consider $B_{ij} := \partial_i A_j$ where A_j is a tensor so that

$$\begin{aligned} B'_{ij} &= \partial'_i A'_j = \partial'_i(\partial'_j x^l A_l) \\ &= \partial'_i \partial'_j x^l A_l + \underbrace{\partial'_j x^l \partial'_i A_l}_{= \partial'_j x^l \partial'_i x^k \partial_k A_l} \\ &= \partial'_i \partial'_j x^l A_l + \underbrace{\partial'_i x^k \partial'_j x^l B_{kl}}_{\text{how a tensor should transform}} . \end{aligned}$$

[4]We haven't defined the word manifold precisely yet. It suffices for our purposes to consider an n-dimensional manifold to be constituted of points that need n coordinates to be specified.

[5]If you are confused about whether $x^i \to \lambda x^i$ is itself a coordinate transformation then the notation has you confused. Think of a ruler. Stretching the ruler means that you take the 1cm mark and stretch it to the 2cm mark (and so on), say. Contrast this to taking a ruler and changing its 1cm mark to a 2cm (and so on). The coordinate transformation just changes the labels of the points. Stretching changes the locations of points.

It is obviously convenient to have a notion of derivatives that preserves the tensor structure and reduces to the usual notion when the space is flat. Consider the variant $C_{ij} := D_i A_j := \partial_i A_j - \Gamma^k_{ij} A_k$ which would transform as

$$\begin{aligned}
C'_{ij} &= \partial'_i A'_j - {\Gamma'^k}_{ij} A'_k \\
&= \underbrace{\partial'_i \partial'_j x^l A_l}_{\text{cancels with the last term}} \quad + \partial'_i x^k \partial'_j x^l \partial_k A_l - \left(\partial_m x'^k\right) \partial'_i x^n \partial'_j x^p {\Gamma^m}_{np} \left(\partial'_k x^l\right) A_l - \left(\partial_m x'^k\right) \partial'_i \partial'_j x^m \left(\partial'_k x^l\right) A_l. \\
&= \partial'_i x^k \partial'_j x^l \partial_k A_l - \partial'_i x^n \partial'_j x^p {\Gamma^l}_{np} A_l \\
&= \partial'_i x^m \partial'_j x^p (\partial_m A_p - {\Gamma^k}_{mp} A_k) = \partial'_i x^m \partial'_j x^p C_{mp}
\end{aligned}$$

which is indeed how a tensor should transform. We leave it as an exercise to check that $D_i A^j := \partial_i A^j + {\Gamma^j}_{ik} A^k$ also transforms as a tensor (notice the difference in the sign). With these two building blocks it is not hard to generalise to an arbitrary tensor as

$$\begin{aligned}
D_k {A^{i_1 i_2 \dots i_p}}_{j_1 j_2 \dots j_q} =& \partial_k {A^{i_1 i_2 \dots p_p}}_{j_1 j_2 \dots j_q} \\
&+ \left({\Gamma^{i_1}}_{kl} {A^{l i_2 \dots i_p}}_{j_1 j_2 \dots j_q} + {\Gamma^{i_2}}_{kl} {A^{i_1 l i_3 \dots i_p}}_{j_1 j_2 \dots j_q} + \dots \right) \\
&- \left({\Gamma^l}_{k j_1} {A^{i_1 i_2 \dots i_p}}_{l j_2 \dots j_q} + {\Gamma^l}_{k j_2} {A^{i_1 i_2 \dots i_p}}_{j_1 l j_3 \dots j_q} + \dots \right)
\end{aligned}$$

so that we are guaranteed that if ${A^{i_1 i_2 \dots i_p}}_{j_1 j_2 \dots j_q}$ is a tensor then so is $D_k {A^{i_1 i_2 \dots i_p}}_{j_1 j_2 \dots j_q}$. It is also easy to see that the usual product rule holds as

$$D_k \left({A^{\cdots}}_{\dots} {B^{\cdots}}_{\dots} \right) = D_k \left({A^{\cdots}}_{\dots} \right) {B^{\cdots}}_{\dots} + {A^{\cdots}}_{\dots} D_k \left({B^{\cdots}}_{\dots} \right).$$

It would be useful to see how the covariant derivative acts on the metric and the delta function.[6]

Claim 12. $D_i({\delta^i}_j) = 0$.

Proof. $D_i({\delta^j}_j) = \cancel{\partial_i {\delta^j}_k} - {\Gamma^l}_{ik} {\delta^j}_l + {\Gamma^j}_{il} {\delta^l}_k = -{\Gamma^j}_{ik} + {\Gamma^j}_{ik} = 0.$ □

Claim 13. $D_i g_{jk} = 0$ and similarly $D_i g^{jk} = 0$.

Proof. We already proved that first derivatives of the metric can't form a non-trivial tensor. We also saw that D_i of a tensor is a tensor. Hence $D_i g_{jk}$ can only be the trivial tensor, i.e. $D_i g_{jk} = 0$. To prove the second statement we use $0 = D_i({\delta^j}_l) = D_i(g^{jk} g_{kl}) = D_i(g^{jk}) g_{kl} + \cancel{g^{jk} D_i(g_{kl})}$. □

To see the usefulness of these results consider the following.

Claim 14. Contracting before or after taking a covariant derivative are equivalent, viz. $D_k({A^{i_1 \dots i_p}}_{i_1 j_2 \dots j_q}) = D_k({A^{i_1 \dots i_p}}_{j_1 \dots j_q}) {\delta^{j_1}}_{i_1}$ for a tensor ${A^{i_1 \dots i_p}}_{j_1 \dots j_2}$.

Proof sketch. Consider the tensor ${A^{i_1 \dots i_p}}_{j_1 \dots j_q} {\delta^m}_n$ so that $D_k \left({A^{i_1 \dots i_p}}_{j_1 \dots j_q} {\delta^m}_n \right) = D_k \left({A^{i_1 \dots i_p}}_{j_1 \dots j_q} {\delta^m}_n \right) {\delta^m}_n$ using the product rule and the fact that $D_k({\delta^m}_n) = 0$ as proved above. Now set $m = j_1$ and $n = i_1$ to get the required result.

Similarly one can extend this to contractions with the metric. It is helpful to extend our convention to include the following.
(5) Raising an index: To raise a lower index i to an upper index j we multiply with g^{ij}.
(6) Lower an index: To lower an upper index i to a lower index j we multiply with g_{ij}.
E.g. $A_{j_1 j_2 j_3 \dots j_q}$ one defines ${A_{j_1}}^{j_2}{}_{j_3 \dots j_q} := A_{j_1 i_2 j_3 \dots j_q} g^{i_2 j_2}$. □

[6] It is nice to see that while ${\delta^i}_j$ is a tensor, δ_{ij} is not and g_{ij} is!

CLAIM 15. Raising/lowering before or after taking a covariant derivative are equivalent, viz. $D_k(A^{i_1 \dots i_p j_1}{}_{j_2 \dots j_q}) = D_k(A^{i_1 \dots i_p}{}_{k_1 j_2 \dots j_q}) g^{j_1 k_1}$.

PROOF HINT. Same argument as above using $D_k(g^{j_1 k_1}) = 0$ instead. □

1.6. Summary

The following are worth keeping in mind.

- Given a point, one can always find a frame in which the first derivative of the metric is zero (at that point).
- Transformation of the Christoffel Symbol $\Gamma'^i{}_{jk} =$ (TODO).
- The fact that the Reimann Tensor is indeed a tensor.
- One has to use the covariant derivative to preserve the tensor structure and how it has properties similar to that of the usual derivative.

With this we conclude the chapter. One more chapter to go and then we can start general relativity.

CHAPTER 2

Shortest Paths and Parallel Lines

2.1. Motivation

Einstein's general theory of relativity uses the notion of geodesic – the shortest length path between two points – in its description. It reduces to the special theory of relativity in the special case of flat space(time). In addition to studying geodesics we will generalise the notion of parallel vectors to the non-Eucledian case. We will connect this to a scheme for detecting the curvature of the underlying space. (I myself don't know where this would be used in GR but we will see.)

2.2. The Geodesic Equation

Let us first find a suitable way of representing a curve in an n dimensional space.

(1) We could define $n-1$ coordinates in terms of the n^{th} coordinate.

This is how we usually describe a line for instance. The difficulty with this method is that it doesn't preserve the symmetry of the coordinates – one of the coordinates is given a special treatment.

(2) We define $n-1$ equations in n variables.

This is symmetric in its treatment of the coordinates but it is implicit.

(3) Parametric description: Define the curve in terms of f as $x^i = f^i(u)$ where u is a parameter which changes monotonically as you move along the curve.

This is both explicit and preserves the symmetry. We will use this scheme for describing our curves. The drawback of this scheme is that different parametrizations may describe different curves but we will handle this appropriately when needed.

We use the variational approach to derive the differential equation for a geodesic. Consider a curve described by $x^i = f^i(u)$ where $u = 0$ corresponds to the starting point and $u = 1$ corresponds to the ending point. Let us pretend that this curve already represents the shortest path between the two extreme points. Now consider some arbitrary nearby curve starting and ending at the same point as given by $x^i = f^i(u) + \delta f^i(u)$ where $\delta f^i(0) = 0 = \delta f^i(1)$. We can evaluate the length of a given curve as

$$L = \int_{\text{curve}} ds = \int_{\text{curve}} du.\frac{ds}{du} = \int_0^1 du \left.\sqrt{g_{ij}\frac{dx^i}{du}\frac{dx^j}{du}}\right|_{x^i = f^i(u)}.$$

Now the change in length between this curve and the arbitrary nearby curve must vanish to first order in δx^k as we assumed that f^i already represents the geodesic where the small parameter $\delta x^k := f^i - \left(f^i - \delta f^i\right) = \delta f^i$. This difference will be given by

$$\delta L = \int_0^1 du \frac{1}{2}\left(g_{mn}\frac{dx^m}{du}\frac{dx^n}{du}\right)^{-1/2}\left\{\partial_k g_{ij}\delta x^k \frac{dx^i}{du}\frac{dx^j}{du} + 2g_{ij}\frac{d(\delta x^i)}{du}\frac{dx^j}{du} + \cancel{g_{ij}\frac{dx^i}{du}\frac{d(\delta x^j)}{du}}\right\}$$

where the last two terms are the same which explains the factor of two. The goal here is to obtain an expression of the form $\int_0^1 du(\dots)\delta x^k + (\dots)\mathcal{O}(\delta x^2)\dots$ so that we can demand the first bracket to be zero to obtain a differential equation for the geodesic[1]. At this point it helps to use what is known as an *affine parameter* – the parameter u is such that equal u spans equal distance along the curve. Thus an affine parameter must satisfy

$$\frac{d}{du}\left[\left(g_{ij}\frac{dx^i}{du}\frac{dx^j}{du}\right)^{1/2}\right] = 0.$$

We will now perform an integration by parts of the second term in the sum (in the expression for δL) to get $\int_0^1 du\,(\ldots)\,\delta x^i\left[\frac{d}{du}\left(2g_{ij}\frac{dx^j}{du}\right)\right] + \cancel{\left[(\ldots)\,\delta x^i\,(\ldots)\right]_{u=0}^{u=1}}$ which yields[2]

$$\delta L = \int_0^1 du\frac{1}{2}\left(g_{mn}\frac{dx^m}{du}\frac{dx^n}{du}\right)^{-1/2}\delta x^k \underbrace{\left\{\partial_k g_{ij}\frac{dx^i}{du}\frac{dx^j}{du} - 2\frac{dg_{kj}}{du}\frac{dx^j}{du} - 2g_{kj}\frac{d^2x^j}{du^2}\right\}}_{\text{must vanish}}.$$

For δL to be zero to first order in δx^k, manifestly then, the term indicated must vanish. Multiplying this term by $-\frac{1}{2}g^{lk}$, using the chain rule for the derivative of the metric and equating it to zero one obtains

$$\frac{d^2x^l}{du^2} + g^{lk}\frac{dx^j}{du}\partial_m g_{kj}\frac{dx^m}{du} - \frac{1}{2}g^{lk}g_{ij}\frac{dx^i}{du}\frac{dx^j}{du} = 0.$$

From the definition of ${\Gamma^l}_{ij}$ and the symmetry in ij one obtains (small exercise) the final form of the *geodesic equation* as

$$\boxed{\frac{d^2x^l}{du^2} + {\Gamma^l}_{ij}\frac{dx^i}{du}\frac{dx^j}{du} = 0.}$$

Some consistency checks: Note that (exercise) $\frac{d}{du}\left(g_{mn}\frac{dx^m}{du}\frac{dx^n}{du}\right) = 0$ follows from the geodesic equation. Note also (again left as an exercise) that changing the coordinate system leaves the form of the geodesic equation unchanged. This is expected as we did not make any assumptions about the coordinate system in the derivation but we did assume the parameter to be affine.

Let us find the geodesic for flat space. In this case $g_{ij} = \delta_{ij}$ which entails ${\Gamma^i}_{jk} = 0$. Consequently the geodesic equation becomes

$$\frac{d^2x^l}{du^2} = 0 \implies x^l = a^l u + b^l$$

where a^l describes the direction and b^l the location of a straight line. Now choose a point $\vec{x}_0$ and a coordinate system x' such that ${\Gamma'^i}_{jk} = 0$ at $\vec{x}_0$. Then

$$\frac{d^2x'^l}{du^2} = \mathcal{O}(u-u_0) \implies x'^l = a'^l u + b'^l + \mathcal{O}((u-u_0)^3)$$

after integration where note that $\mathcal{O}((u-u_0)^2)$ like terms are absent.

2.3. Tangent Vectors and Tangent Space

Consider a point $\vec{x}_0$ and a geodesic passing through it. To arrive at the notion of tangent vectors we ask the following: How many parameters are needed to describe all possible geodesics that pass through this point? To answer this recall the geodesic equation $\frac{d^2x^l}{du^2} + {\Gamma^l}_{ij}\frac{dx^i}{du}\frac{dx^j}{du} = 0$ and note that these constitute n (where $l, i, j \in \{1, 2, \ldots n\}$) second order differential equations which means we must specify $2n$ constants. We already enforce $x^l|_{u=u_0} = x_0^l$ which is a set of n conditions, that is, the geodesic must pass through the said point. Now we set the

[1]Atul: admittedly we only extremised the length and therefore it might even be the longest length curve but since that in most cases the longest curve would have infinite length I think the result here would apply to shortest length curves as we start with assuming a finite length curve.

[2]note the affine parameter assumption allows us to neglect the effect of the coefficient multiplied by du which lets us cancel the surface term

direction to completely specify the curve by requiring n more conditions:

$$\left.\frac{dx^l}{du}\right|_{\vec{x}=\vec{x}_0} = n^l$$

where we define n^l to be the *tangent vectors* to the geodesic at $\vec{x}$. Note that we only need the direction to be specified by n^l as u can always be scaled. If we enforce n^l exactly then the scaling (of the affine parameter) is also specified. We can now define the *tangent space* at $\vec{x}_0$ to be the space containing all vectors n^l through the point $\vec{x}_0$. In this terminology, any vector in the tangent space specifies a geodesic through the given point. It is worth noting that two different ns might correspond to the same geodesic, e.g. the great circle on a sphere is one curve globally but locally will correspond to two tangent vectors (TODO: add figure). We close this section by relating tangent vectors for the same geodesic written in different coordinate systems. In the x coordinate system if it is given by $n^l = dx^l/du\big|_{\vec{x}_0}$ and in the x' coordinate system it is given by

$$n'^l = \left.\frac{dx'^l}{du}\right|_{\vec{x}'_0} = \partial_m x'^l \left.\frac{dx^m}{du}\right|_{\vec{x}_0} = \partial_m x'^l n^m,$$

that is, n^l transforms as a tensor.

2.4. Parallel Transport

We are now in a position to ask what it would mean for two tangent vectors n^l at $\vec{x}_1$ and m^l at $\vec{x}_2$ to be parallel? Let us try the usual notion of parallel vectors from flat space where $\vec{n} \parallel \vec{m}$ if $n^l = \lambda m^l$ for some λ. Now if we look at their relation in a new coordinate system we would get $n'^l = \partial_k x'^l\big|_{\vec{x}'_1} n^k$ while $m'^l = \partial_k x'^l\big|_{\vec{x}'_2} m^k$. Manifestly then two parallel tangent vectors may cease to be parallel in the new coordinate system the tangent vectors according to our naïve definition. This seems to suggest that there's no unambiguous way of determining if two tangents are parallel. We fix this demanding the said condition in a coordinate system which is "flat" to first order (where the first derivative of the metric disappears). Since we can make the space "flat" only locally let us start by considering two close by points.

First choose a coordinate system x' so that $\Gamma'^i{}_{jk}(\vec{x}') = 0$ (or equivalently $\partial'_i g'_{jk} = 0$) and now consider two points $\vec{x'}_0$ and $\vec{x'}_0 + \delta\vec{x}'_0$. We say that the tangents at these points, n^l and m^l, are parallel if $n^l = \lambda m^l + \mathcal{O}(\delta x'^2)$. We also say that $\vec{m}$ is a *parallel transport* of $\vec{n}$ if $\vec{m}' = \vec{n}' + \mathcal{O}(\delta x'^2)$[3]. Recall that in making $\Gamma'^i{}_{jk} = 0$ we had only used the freedom in Bs (the As anyway didn't affect Γ as we saw) but we still had the freedom in specifying Cs. These will define different coordinate systems each of which would have Γ zero. To ensure invariance under such coordinate transformations we had to add $\mathcal{O}(\delta x'^2)$ terms in our definition of parallel vectors.

So far our definition suffers from two issues. First, it is only defined for infinitesimally close points and second, it requires use to pick a specific coordinate system. We remedy this by determining the condition for parallel transport in an arbitrary frame and use the differential equation so obtained to parallel transport a given vector to finitely spaced points. Consider a curve between two points $\vec{x}_{(0)}$ and $\vec{x}_{(1)}$ parametrised by u which starts at 0 and ends at 1. We wish to start with a tangent vector $n^i(0)$ at point $\vec{x}_{(0)}$ and parallel transport it to $\vec{x}_{(1)}$. Assume that $n^i(u)$, implicitly defined, is known and we need to evaluate $n^i(u+\delta u)$. According to our definition of parallel transport we must first choose a coordinate system x' such that $\partial'_i g'_{jk} = 0$ at $\vec{x}' = \vec{x}'(u)$ and then we can write $n'^i(u+\delta u) = n'^i(u) + \mathcal{O}(\delta u^2)$. Let us translate this condition back to the x coordinate system. Recall

[3] where we dropped the λ for n and m can be assumed to have the same normalisation

$n^i(u) = \frac{\partial x^i}{\partial x'^k}\Big|_{\vec{x}'(u)} n'^k(u)$ and

$$
\begin{aligned}
n^i(u+\delta u) &= \frac{\partial x^i}{\partial x'^k}\bigg|_{\vec{x}'(u+\delta u)} n'^k(u+\delta u) \\
&= \frac{\partial x^i}{\partial x'^k}\bigg|_{\vec{x}'(u+\delta u)} \left(n'^k(u) + \mathcal{O}(\delta u^2)\right) \qquad \text{(using the transport equation)} \\
&= \left(\frac{\partial x^i}{\partial x'^k}\bigg|_{\vec{x}'(u)} + \frac{\partial^2 x^i}{\partial x'^l \partial x'^k}\delta x'^l + \dots \right) \left(n'^k(u) + \mathcal{O}(\delta u^2)\right) \\
&= \left(\frac{\partial x^i}{\partial x'^k}\bigg|_{\vec{x}'(u)} + \frac{\partial^2 x^i}{\partial x'^l \partial x'^k}\frac{dx'^l}{du}\delta u + \dots \right) \left(n'^k(u) + \mathcal{O}(\delta u^2)\right)
\end{aligned}
$$

which yields

$$\frac{dn^i}{du} = \frac{\partial^2 x^i}{\partial x'^l \partial x'^k}\frac{dx'^l}{du} n'^k = \frac{\partial^2 x^i}{\partial x'^l \partial x'^k}\frac{\partial x'^l}{\partial x^m}\frac{\partial x'^k}{\partial x^j}\frac{dx^m}{du} n^j$$

where in the last step we used the chain rule twice. This is better but it still depends on the primed coordinates. To remove this dependence recall the transformation of the Christofel symbol was given by

$$\Gamma'^j{}_{lk} = \frac{\partial x'^j}{\partial x^\mu}\frac{\partial x^\nu}{\partial x'^l}\frac{\partial x^\rho}{\partial x'^k}\Gamma^\mu{}_{\nu\rho} + \frac{\partial x'^j}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x'^l \partial x'^k} = 0$$

which upon multiplication with $\partial x^i / \partial x'^j$ yields

$$\cancel{\delta^i_\mu}\frac{\partial^2 x^{\cancel{\mu}\, i}}{\partial x'^l \partial x'^k} = -\cancel{\delta^i_\mu}\frac{\partial x^\nu}{\partial x'^l}\frac{\partial x^\rho}{\partial x'^k}\Gamma^{\cancel{\mu}\, i}{}_{\nu\rho}.$$

The term involving second derivatives is the one we would like to replace in our initial equation. This trading of second derivatives for a Christofel symbol and first derivatives is useful because the first derivatives become delta functions upon multiplication with tensor like transformation factors. Using this we get

$$
\begin{aligned}
\frac{dn^i}{du} &= -\frac{\partial x^\nu}{\partial x'^l}\frac{\partial x^\rho}{\partial x'^k}\Gamma^i{}_{\nu\rho}\frac{\partial x'^l}{\partial x^m}\frac{\partial x'^k}{\partial x^j}\frac{dx^m}{du}n^j \\
&= -\Gamma^i{}_{\nu\rho}\delta^\nu_m \delta^\rho_j \frac{dx^m}{du} n^j
\end{aligned}
$$

which yields the final *parallel transport equation* given by

$$\boxed{\frac{dn^i}{du} + \Gamma^i{}_{jk} n^j \frac{dx^k}{du} = 0}$$

where we changed dummy indices and used the fact that $\Gamma^i{}_{jk} = \Gamma^i{}_{kj}$.

Note that in our derivation we did not make any assumptions about the curve. In particular it needn't be a geodesic. We can also show that the equation is parametrisation independent as follows. Let $u = f(v)$ and note that

$$\frac{dn^i}{du} = \frac{dv}{du}\frac{dn^i}{dv}, \; \frac{dx^k}{du} = \frac{dv}{du}\frac{dx^k}{du} \implies \frac{dv}{du}\left[\frac{dn^i}{du} + \Gamma^i{}_{jk}n^j\frac{dx^k}{du}\right] = 0.$$

This is unlike the geodesic equation which was derived for a specific parametrisation.

Note also that this equation is coordinate invariant because the coordinate system x was chosen arbitrarily. This can also be seen explicitly by choosing another coordinate system x'' and using $n^i = n''^k \partial''_k x^i$, the transformation of the Christofel symbol and $dx^i/du = \partial''_r x^i dx''^r/du$ to get

$$\frac{\partial x^i}{\partial x''^r}\left(\frac{dn''^r}{du} + \Gamma''^r{}_{jk}n''^j\frac{dx''^k}{du}\right) = 0$$

which we leave as an exercise.

It can also be shown that the norm of a tangent vector is preserved under parallel transport, viz. $d\left(g_{ij}n^i n^j\right)/du = 0$. One way to proceed is to evaluate this explicitly as dn^i/du is known and use the fact that $dg_{ij}/du = \partial_k g_{ij} dx^k/du$. The neater method is to note that a scalar is invariant under coordinate transformations and therefore use a co-ordinate frame where the first derivative of the metric vanishes. Further the transport equation yields $dn'/du = 0$ which completes the proof. This is not surprising because we derived all our results from the requirement that $n'(u+\delta u) = n'(u) + \mathcal{O}(\delta u^2)$ where the norm is explicitly preserved to first order.

Observe that we started with choosing any vector from the tangent space at a given point and parallel transported it along an arbitrary curve to another point. Now instead consider the geodesic equation

$$\frac{d^2x^i}{du^2} + \Gamma^i{}_{jk}\frac{dx^k}{du}\frac{dx^j}{du} = 0$$

and use the definition for a tangent to the curve $n^i = dx^i/du$ to get

$$\frac{dn^i}{du} + \Gamma^i{}_{jk}n^j\frac{dx^k}{du} = 0$$

which is the parallel transport equation! This means that the geodesic is a curve such that its tangent vectors are parallel transported from any given tangent on the curve.

2.5. The Monodromy Matrix

In this section we will try to understand what happens to a vector if it is parallel transported along a closed curve and how it might be an interesting object that captures curvature of the space. Imagine a curve c starting at $\vec{x}_{(1)}$ and ending at $\vec{x}_{(2)}$. Also consider a vector $\vec{n}_{(1)}$ in the tangent space at $\vec{x}_{(1)}$ is transported to the end of the curve c and is denoted by $\vec{n}_{(2)}$. We can find a relation between these by integrating the parallel transport equation along the curve. Let $n^i_{(2)} = M^i{}_j n^j_{(1)}$.

Claim 16. $M^i{}_j$ is independent of $\vec{n}_{(1)}$ and $\vec{n}_{(2)}$ (where the symbols are as defined above).

Proof sketch. To see this, recall that the parallel transport equation is linear in $\vec{n}$. For simplicity imagine that the tangent space is two dimensional (at both ends of the curve). Let $\vec{n}_{(1)}, \vec{n}'_{(1)}$ and $\vec{n}_{(2)}, \vec{n}'_{(2)}$ span the space at $\vec{x}_{(1)}, \vec{x}_{(2)}$ respectively and further assume $n'^i_{(2)} = M^i{}_j n'^j_{(1)}$. Any vector in the first tangent space can be written as $\vec{m}_{(1)} = \alpha\vec{n}_{(1)} + \beta\vec{n}'_{(1)}$ which must transform under $M^i{}_j$ as $M^i{}_j m^j = \alpha M^i{}_j n^j_{(1)} + \beta M^i{}_j n'^j_{(1)}$ by linearity of the transport equation. Finally we get $\vec{m}_{(2)} = \alpha\vec{n}_{(2)} + \beta\vec{n}'_{(2)}$ which shows that once the transformation of the basis is known the transformation of every other vector is fixed and thus $M^i{}_j$ can't depend on the vector itself. □

The dependence therefore would be restricted to $M^i{}_j = M^i{}_j(\vec{x}_{(1)}, \vec{x}_{(2)}, c)$. If $-c$ represents the curve c starting from the other end, what can be said about $M^i{}_j(\vec{x}_{(1)}, \vec{x}_{(2)}, -c)$? First note the following.

Claim 17. The parallel transport equation is reversible.

Proof sketch. This follows from independence of parametrisation. Use $v = 1-u$ so that $u = 1 \leftrightarrow v = 0$ and $u = 0 \leftrightarrow v = 1$ and $du \leftrightarrow -dv$. This means $\frac{dn^i}{du} + \Gamma^i{}_{jk}n^j\frac{dx^k}{du} = 0 \leftrightarrow \frac{dn^i}{dv} + \Gamma^i{}_{jk}n^j\frac{dx^k}{dv} = 0$ so we essentially solve the same equation with the boundary conditions reversed. (Aside: Note that even for Newton's law this holds; in $d^2x/dt^2 = F(x)$ send $dt \to -dt$ and you get the same equation but if the force term depended on dx/dt (like friction for instance) then this doesn't work.) □

Let the reversed solution be represented by $\left(M(\vec{x}_{(1)}, \vec{x}_{(2)}, c)^{-1}\right)^i{}_j$ so that $n^i_{(1)} = \left(M(\vec{x}_{(1)}, \vec{x}_{(2)}, c)^{-1}\right)^i{}_j n^j_{(2)}$. Also by definition of M we have $n^i_{(1)} = M(\vec{x}_{(2)}, \vec{x}_{(1)}, -c)^i{}_j n^j_{(2)}$. Since the two equations hold for arbitrary tangent vectors we conclude

$$M(\vec{x}_{(2)}, \vec{x}_{(1)}, -c) = M^{-1}(\vec{x}_{(1)}, \vec{x}_{(2)}, c).$$

If we had $dn^i/du + \Gamma^i{}_{jk} n' \left| dx^j/du \right| = 0$ then the equation would not be reversible and the aforesaid conclusion wrong.

Now consider two curves c_1 and c_2 joining the points $x_{(1)}$ and $x_{(2)}$. In general $M(\vec{x}_{(1)}, \vec{x}_{(2)}, c_1) \neq M(\vec{x}_{(1)}, \vec{x}_{(2)}, c_2)$ but what happens if we consider the closed curve $c_1 - c_2$? We can write this transformation as first going from x_1 to x_2 along c_1 then returning from x_2 to x_1 along $-c_2$. We write this in the reversed order in the matrix (because it operates on objects on the right conventionally) as below.

$$\begin{aligned} M(\vec{x}_{(1)}, \vec{x}_{(1)}, c_1 - c_2) &= M(\vec{x}_{(2)}, \vec{x}_{(1)}, -c_2) M(\vec{x}_{(1)}, \vec{x}_{(2)}, c_1) \\ &= M(\vec{x}_{(1)}, \vec{x}_{(2)}, c_2)^{-1} M(\vec{x}_{(1)}, \vec{x}_{(2)}, c_1) \\ &\neq \mathbb{I} \text{ (in general)} \end{aligned}$$

where $\mathbb{I}$ is the identity matrix corresponding to δ^i_j. This object, the[4] parallel transport operation of a vector around a closed loop c given by $M(\vec{x}, \vec{x}, c)$ is called the *Monodromy matrix* around c. It has an interesting coordinate invariant property.

Claim 18. Consider coordinate systems x and x'. If $M = \mathbb{I}$ then $M' = \mathbb{I}$.

Proof. Let $\tilde{n}^i = M(\vec{x}, \vec{x}, c)^i{}_j n^j$. Translate to another coordinate system to get

$$\begin{aligned} \tilde{n}'^i &= \frac{\partial x'^i}{\partial x^p} \tilde{n}^p \\ &= \frac{\partial x'^i}{\partial x^p} M^p{}_q n^q \\ &= \underbrace{\frac{\partial x'^i}{\partial x^p} M^p{}_q \frac{\partial x^q}{\partial x'^j}}_{M'^i{}_j} n'^j \end{aligned}$$

where if $M^p{}_q = \delta^p{}_q$ then $M'^p{}_q$ is also $\delta^p{}_q$. □

Note that this proof would fail for $M(\vec{x}_1, \vec{x}_2, c)$ because the the two other factors in the bracketed expression would not become a kronecker delta as they would have been evaluated at $\vec{x}_1$ and $\vec{x}_2$. We now wish to show that $R_{ijkl} = 0$ entails that $M = \mathbb{I}$ under some assumptions about connectedness of the underlying space. The reverse statement also holds as we shall see soon.

Proposition 19. *The Reimann Tensor $R_{ijkl} = 0$ entails that $M = \mathbb{I}$ (if the space has no holes [TODO: make this precise]).*

Proof sketch for small curves to $\mathcal{O}(\epsilon^3)$. Consider a small curve of length $\mathcal{O}(\epsilon)$ starting and ending at $x_{(0)}$ with n_1 an initial tangent vector (a vector in the tangent space at $x_{(0)}$) and n_2 its parallel transport along the curve terminating at $x_{(0)}$ again. We use the convention that u varies between 0 and ϵ. This is to ensure that $dx^i/du \sim 1$ that is we can't assume it to be arbitrarily small. If we had taken u to range from 0 to 1 then $\delta x/\delta u \sim \mathcal{O}(\epsilon)$ as illustrated in figure 2.5.1. (Issue: I admit this is not as clear as I would have liked it to be but if you imagine making equal slices on u and on x to define δu and δx respectively, and consider the limit of infinite slices then the aforesaid result should follow neatly. Whether there is another sequence which might yield a different limit I will not consider here for things should be well behaved and the limit should be unique). Recall that

[4] matrix corresponding to the

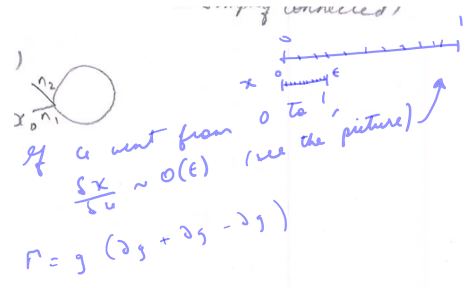


FIGURE 2.5.1. Illustratively justifying why $0 \leq u \leq \epsilon$

$$\frac{dn^i}{du} = -\Gamma^i{}_{jk} n^j(u) \frac{dx^j}{du}.$$

The strategy would be to keep terms of order 1 and then integrate to find n^i. The result would be correct to order $\mathcal{O}(\epsilon)$. Then we use this n^i and integrate to get n^i to $\mathcal{O}(\epsilon^2)$. We already know that $n^j_{(0)}$ is of $\mathcal{O}(1)$ as $dx^i/du = \mathcal{O}(1)$. There's no reason to assume $\Gamma^i{}_{jk}$ is small at x_0 so we simply integrate the following expression

$$\frac{dn^i}{du} = -\Gamma^i{}_{jk}(x_{(0)}) n^j_{(0)} \frac{dx^k}{du} + \mathcal{O}(\epsilon)$$

to get

$$n^i(u) = n^i_{(0)} - \Gamma^i{}_{jk}(x_0) n^j_{(0)} (x^j(u) - x^j_{(0)}) + \mathcal{O}(\epsilon^2)$$

as recall $x^j = \mathcal{O}(\epsilon)$. Note also that we integrate to some arbitrary u not to $u = \epsilon$. We will now integrate the same expression for dn^i/du to obtain n^i correct to $\mathcal{O}(\epsilon^3)$ but set $u = \epsilon$. Plugging in and retaining first order in ϵ we have

$$\begin{aligned} \frac{dn^i}{du} &= -\left\{ \Gamma^i{}_{jk}(x_{(0)}) + \partial_l \Gamma^i{}_{jk}(x_{(0)}) \underbrace{(x^l - x^l_{(0)})}_{\mathcal{O}(\epsilon)} \right\} \left\{ n^i_{(0)} - \Gamma^i{}_{jk}(x_0) n^j_{(0)} \underbrace{(x^j(u) - x^j_{(0)})}_{\mathcal{O}(\epsilon)} \right\} \frac{dx^k}{du} \\ &= -\Gamma^i{}_{jk} n^j_{(0)} \frac{dx^k}{du} - \Gamma^i{}_{jk} \Gamma^j{}_{pq} n^p_{(0)} (x^q - x^q_{(0)}) \frac{dx^k}{du} + \partial_q \Gamma^i{}_{pk} (x^q - x^q_{(0)}) n^p_{(0)} \frac{dx^k}{du} + \mathcal{O}(\epsilon^2) \end{aligned}$$

where we suppressed the $x_{(0)}$ while writing Γ in the second line for brevity and changed dummy indices for aesthetics. The first term would integrate to zero as x^k starts at x_0 at $u = 0$ and ends at x_0 at $u = \epsilon$. We will replace dx^k with $d(x^k - x^k_{(0)})$ and write the last term first to obtain

$$n^i(\epsilon) = n^i_{(0)} + \left(\partial_q \Gamma^i{}_{pk} - \Gamma^i{}_{jk} \Gamma^j{}_{pq} \right) n^p_{(0)} \oint_0^\epsilon (x^q - x^q_{(0)}) \frac{d(x^k - x^k_{(0)})}{du} du + \mathcal{O}(\epsilon^3).$$

Note that if the boundary term is zero we can shift the derivative with a minus sign, viz. for $y = x - x_{(0)}$ $\int \left(y^q dy^k + dy^q y^k \right) = \int d(y^q y^k) = 0 \implies \int y^q dy^k = -\int y^k dy^q$. This means the integral is anti-symmetric in the indices q and k. Thus the symmetric part of the multiplying tensor will not contribute to the final expression. We therefore keep only the anti-symmetric part (in q and k) of it to obtain

$$\begin{aligned} n^i(\epsilon) &= n^i_{(0)} + \frac{1}{2} \left[\partial_q \Gamma^i{}_{pk} - \partial_k \Gamma^i{}_{pq} - \Gamma^i{}_{jk} \Gamma^j{}_{pq} + \Gamma^i{}_{jq} \Gamma^j{}_{pk} \right] n^p_{(0)} \oint_0^\epsilon (x^q - x^q_{(0)}) \frac{d(x^k - x^k_{(0)})}{du} \\ &= n^i_{(0)} + \frac{1}{2} R^i{}_{pkq}|_{\vec{x}_{(0)}} \oint (x^q - x^q_{(0)}) \frac{d(x^k - x^k_{(0)})}{du}. \end{aligned}$$

See what we did there? We got the Reimann Tensor into the equation. Now since $R^i{}_{pqk}|_{\vec{x}_{(0)}} = 0$ we have

$$n^i(\epsilon) = n^i_{(0)} + \mathcal{O}(\epsilon^3)$$

which means $M^i{}_j(\vec{x}_0, \vec{x}_0, c) = \delta^i{}_j + \mathcal{O}(\epsilon^3)$. $\square$

That was already quite satisfying but of course valid only for small loops and still not exact. We will actually use that as a lemma in the full proof. To be able to proceed we will need another small result. With reference to figure 2.5.2a we want to know how the monodromy matrix evaluated at $\vec{x}_1$ around a loop and that evaluated at $\vec{x}_2$ around the same loop are related. Note that because the parallel transport was an integral we can break it as

$$M(\vec{x}_1, \vec{x}_1, c_1 + c_2) = M(x_2, x_1, c_2)M(x_1, x_2, c_1)$$

where we need to use the reversed order as it is an operator. Note that using $M(x_2, x_2, c_2+c_1) = M(x_1, x_2, c_1)M(x_2, x_1, c_2)$ and $M(x_2, x_1, c_2)^{-1} = M(x_1, x_2, -c_2)$ we can write the second term of the aforesaid equation as

$$M(\vec{x}_1, \vec{x}_1, c_1 + c_2) = M(x_2, x_1, c_2)M(x_2, x_2, c_2 + c_1)M(x_1, x_2, -c_2)$$

which means that monodromy matrices evaluated at two different points in a loop are related by conjugation (roughly speaking A and B are conjugates if $A = SBS^{-1}$ for some S). This amounts to the following claim.

Claim 20. Consider two points $\vec{x}_1$ and $\vec{x}_2$. Consider two distinct paths, c_1 from $\vec{x}_1$ to $\vec{x}_2$ and c_2 that continues from $\vec{x}_2$ and goes back to $\vec{x}_1$. The monodromy evaluated around this loop starting from $\vec{x}_1$ and $\vec{x}_2$ are related as

$$M(x_1, x_1, c_1 + c_2) = M(x_2, x_1, c_2)M(x_2, x_2, c_2 + c_1)M^{-1}(x_2, x_1, c_2)$$

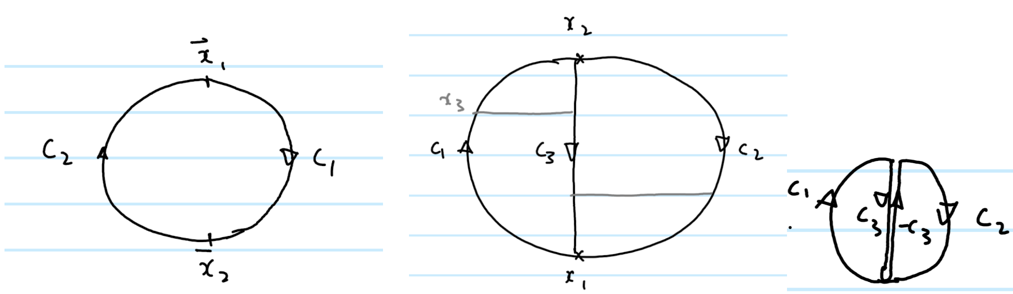
where $M(x_i, x_f, c)$ is the parallel transport transformation matrix for a vector from $\vec{x}_i$ to $\vec{x}_f$ along the path c.

In particular if $M(x_2, x_2, c_2 + c_1) = \mathbb{I}$ then so is $M(x_1, x_1, c_1 + c_2)$, viz. if the monodrome is $\mathbb{I}$ then the point from which you start doesn't matter. We can generalise a little. If $A = SBS^{-1}$ and we know $B = \mathbb{I} + \mathcal{O}(\epsilon)$ then we also know $A = \mathbb{I} + S\mathcal{O}(\epsilon)S^{-1} = \mathbb{I} + \mathcal{O}(\epsilon)$, viz. if B is close to $\mathbb{I}$ then so is A.

Remark 21. (Using the notation from the claim) If $M(x_2, x_2, c_1 + c_2) = \mathbb{I} + \mathcal{O}(\epsilon)$ then $M(x_1, x_1, c_1 + c_2) = \mathbb{I} + \mathcal{O}(\epsilon)$.

With these ingredients in place we are now ready to give the complete proof of the proposition. The basic

Proof. The basic idea $\square$



(A) Visual aid to establishing a relation between monodromy matrices evaluated at $\vec{x}_1$ and $\vec{x}_2$.

(B) Systematically breaking a larger loop into smaller constituents.

Figure 2.5.2. Illustrations to facilitate the understanding of results about Monodromy matrices.

Part 2

General Relativity

CHAPTER 3

Conventions, Axioms and the Weak Field Limit