Sen's Lectures
General Relativity

Volume 1

ATUL SINGH ARORA

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

Metric and Tensors

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 $d\vec{r} = dr\hat{r} + rd\theta\hat{\theta} + r\sin\theta d\phi\hat{\phi}$ 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 $x^3 = \pm\sqrt{a^2 - \left(x^1\right)^2 - \left(x^2\right)^2}$. 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.

Proof. $g'_{kl}(\vec{x}'_0) = \partial'_k x^i\partial'_l x^j g_{ij}|_{\vec{x}'=\vec{x}'_0\leftrightarrow\vec{x}=\vec{x}_0} = A^i{}_kA^j{}_lg_{ij}(\vec{x}_0) = \left(A^TgA\right)_{kl}$ 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. $S^TS = \mathbb{I}$ and $g = S^Tg_dS$ 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 $g = S^Tg_dS = S^TR^T\eta RS$. We're almost there, let $g'_{kl} = \left(A^TgA\right)_{kl} = \left(A^TS^TR^T\eta RSA\right)_{kl}$ 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. $\square$

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

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).

[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.

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.

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

---

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

[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.

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

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

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

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

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

---

[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.

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

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 $A_{j_1}{}^{j_2}{}_{j_3\dots j_q} := A_{j_1 i_2 j_3 \dots j_q} g^{i_2 j_2}$. □

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

Geodesic, Parallel Transport and the Monodromy Matrix

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 usual special theory of relativity in the flat space. In addition to studying these we will generalise the notion of parallel vectors to the non-Eucledian case.

2.2. The Geodesic Equation

aeu

Part 2

General Relativity

CHAPTER 3

Conventions, Axioms and the Weak Field Limit