

---

---

# THE WE-HERAEUS INTERNATIONAL WINTER SCHOOL ON GRAVITY AND LIGHT

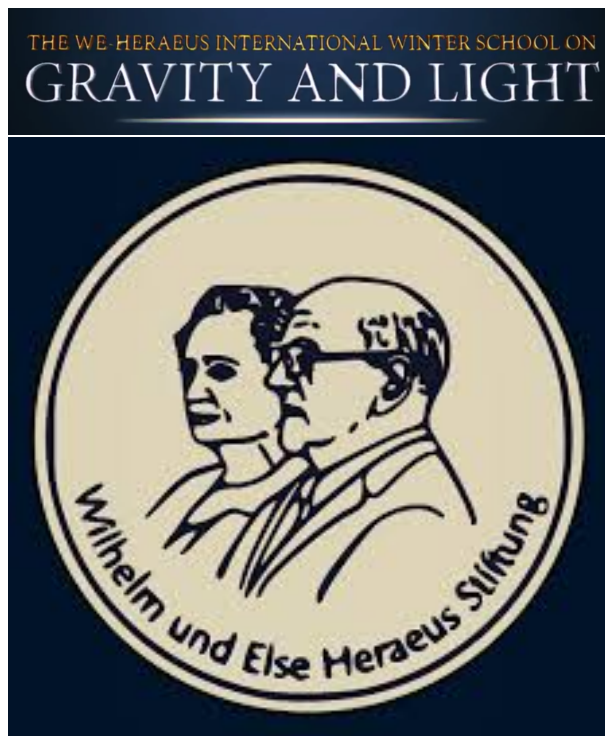
---

---

COURSE DELIVERED IN 2015 BY

DR. FREDERIC P. SCHULLER

*Friedrich-Alexander-Universität Erlangen-Nürnberg,  
Institut für Theoretische Physik III*



NOTES TAKEN BY

RICHIE DADHLEY

*richie@dadhley.com*

## ACKNOWLEDGEMENTS

---

This set of notes is intended to accompany the lecture course on String Theory taught by Dr. Frederic P. Schuller. The videos for the course are available on YouTube via the following link:

<https://www.youtube.com/watch?v=7G4SqIboeig&t=3811s>

I have tried to correct any typos and/or mistakes I think I have noticed over the course. I have also tried to include additional information that I think supports the taught material well, which sometimes has resulted in modifying the order the material was taught. Obviously, any mistakes made because of either of these points are entirely mine and should not reflect on the taught material in any way.

I would like to extend a message of thanks to Dr. Schuller for making this courses available and for his great explanations of the topic.

If you have any comments and/or questions please feel free to contact me via the email provided on the title page.

These notes are currently a work in progress, so for updated notes, as well as a list of other notes/works I have available, visit my blog site

<https://richie291.wixsite.com/theoreticalphysics>

These notes are not endorsed by Dr. Schuller or The WE-Heraeus Institute.

Richie Dadhley

# Contents

<b>0</b>	<b>Introduction</b>	<b>1</b>
<b>1</b>	<b>Topology</b>	<b>3</b>
1.1	Topology . . . . .	3
1.2	Continuous Maps . . . . .	5
1.3	Composition of Continuous Maps . . . . .	7
1.4	Inheriting a Topology . . . . .	8
<b>2</b>	<b>Topological Manifolds</b>	<b>9</b>
2.1	Topological Manifolds . . . . .	9
2.2	Chart Transition Maps . . . . .	11
2.3	Manifold Philosophy . . . . .	12
<b>3</b>	<b>Multilinear Algebra</b>	<b>14</b>
3.1	Vector Spaces . . . . .	14
3.2	Linear Maps . . . . .	16
3.3	Vector Space of Homomorphisms . . . . .	17
3.4	Dual Vector Space . . . . .	18
3.5	Tensors . . . . .	18
3.6	Vectors and Covectors as Tensors . . . . .	20
3.7	Bases . . . . .	21
3.8	Components of Tensors . . . . .	22
<b>4</b>	<b>Differentiable Manifolds</b>	<b>24</b>
4.1	Strategy . . . . .	24
4.2	Compatible Charts . . . . .	25
4.3	Diffeomorphisms . . . . .	27
<b>5</b>	<b>Tangent Spaces</b>	<b>29</b>
5.1	Velocities . . . . .	29
5.2	Tangent Vector Space . . . . .	30
5.3	Components of a Vector w.r.t. a Chart . . . . .	33
5.4	Chart Induced Basis . . . . .	34
5.5	Change of Vector Components Under Change of Chart . . . . .	34
5.6	Cotangent Spaces . . . . .	35

5.7	Change of Components of a Covector Under a Change of Chart . . . . .	36
<b>6</b>	<b>Fields</b>	<b>38</b>
6.1	Bundles, Fibres and Sections . . . . .	38
6.2	Tangent Bundle of Smooth Manifold . . . . .	39
6.3	Vector Fields . . . . .	42
6.4	The $C^\infty(\mathcal{M})$ -Module . . . . .	42
6.5	Tensor Fields . . . . .	45
<b>7</b>	<b>Connections</b>	<b>47</b>
7.1	Directional Derivatives of Tensor Fields . . . . .	48
7.2	New Structure on $(\mathcal{M}, \mathcal{O}, \mathcal{A})$ Required to Fix $\nabla$ . . . . .	49
7.3	Change of $\Gamma$ s Under Change of Chart . . . . .	52
7.4	Normal Coordinates . . . . .	54
<b>8</b>	<b>Parallel Transport &amp; Curvature</b>	<b>55</b>
8.1	Parallelity of Vector Fields . . . . .	56
8.2	Autoparallely Transported Curves . . . . .	57
8.3	Autoparallel Equation . . . . .	58
8.4	Torsion . . . . .	60
8.5	Curvature . . . . .	60
<b>9</b>	<b>Newtonian Spacetime Is Curved</b>	<b>64</b>
9.1	Laplace's Question . . . . .	64
9.2	The Full Wisdom of Newton I . . . . .	65
9.2.1	Tidal Forces . . . . .	67
9.3	The Foundation of the Geometric Formulation of Newton's Axioms . . . . .	68
9.4	Acceleration . . . . .	70
<b>10</b>	<b>Metric Manifolds</b>	<b>72</b>
10.1	Metrics . . . . .	72
10.2	Signature . . . . .	73
10.3	Length Of A Curve . . . . .	75
10.4	Geodesics . . . . .	76
<b>11</b>	<b>Symmetry</b>	<b>80</b>
11.1	Push-Forward Map . . . . .	80
11.2	Pull-back Map . . . . .	82
11.3	Induced Metric . . . . .	83
11.4	Flow of a Complete Vector Field . . . . .	84
11.5	Lie Subalgebras of the Lie Algebra $(\Gamma\mathcal{M}, [\cdot, \cdot])$ of Vector Fields . . . . .	86
11.6	Symmetry . . . . .	87
11.7	Lie Derivatives . . . . .	88
11.7.1	Killing Vector Fields . . . . .	89

<b>12 Integration</b>	<b>90</b>
12.1 Review of Integration on $\mathbb{R}^d$	90
12.2 Integration on One Chart	91
12.3 Volume Forms	92
12.4 Integration On One Chart Domain	94
12.5 Integration Over The Entire Manifold	95
<b>13 Relativistic Spacetime</b>	<b>96</b>
13.1 Time Orientation	96
13.2 Observers	99
13.3 Role Of Lorentz Transformations	102
<b>14 Matter</b>	<b>103</b>
14.1 Point Matter	103
14.1.1 Without External Forces	103
14.1.2 Presence of External Forces	104
14.2 Field Matter	104
14.3 Energy-Momentum Tensor Of Matter Fields	106
<b>15 Einstein Gravity</b>	<b>107</b>
15.1 Hilbert	107
15.2 Variation of $S_H$	108
15.3 Solution Of The $(\nabla_a T)^{ab} = 0$ Issue	110
15.4 Variations of The Field Equations	110
15.4.1 Ricci Scalar Expression	110
15.4.2 Cosmological Constant	111
<b>16 Optical Geometry I</b>	<b>112</b>
16.1 Fermat's Principle	112
<b>17 Optical Geometry II</b>	<b>114</b>
<b>18 Canonical Formulation of GR I</b>	<b>115</b>
<b>19 The Canonical Formulation of GR II</b>	<b>116</b>
<b>20 Cosmology: The Early Epoch</b>	<b>117</b>
20.1 Assumption of Spatial Homogeneity & Isotropy at Large Scale	117
20.2 Einstein Equations	120
20.2.1 Ricci Tensor	120
20.2.2 Matter	121
20.2.3 Reduction of Einstein Equations	122
20.3 Models of Perfect Fluid Matter	122
20.4 Solutions	123

<b>21 Cosmology: The Late Epoch</b>	<b>125</b>
21.1 Density Parameters	125
21.2 Dominant Matter At Various Epochs	126
21.3 A More Realistic Late Universe	126
<b>22 Black Holes</b>	<b>128</b>
22.1 Radial Null Geodesics	129
22.2 Eddington-Finkelstein	131
22.3 Kruskal-Szekeres	133
22.4 Other Types of Black Hole	135
22.4.1 Event Horizons & Infinite Redshift Surfaces	136
22.4.2 Ergoregion	138
22.4.3 Multiple Event Horizons	138
<b>23 Penrose Diagrams</b>	<b>139</b>
23.1 Recipe To Construct A Diagram	139
23.2 Minkowski	142
23.3 Other Spacetimes	144
23.3.1 Schwarzschild Black Hole	144
23.3.2 Kerr Black Hole	144
23.3.3 Reissner-Nordström Black Hole	146
23.3.4 Kerr-Newmann Black Hole	146
23.3.5 Gravitational Collapse	147
23.3.6 Isotropic & Homogeneous Universe	147
<b>24 Perturbation Theory I</b>	<b>149</b>
24.1 Perturbation Of Exact Solutions	149
24.2 The Perturbed Metric	150
24.3 Helmholtz-Hodge	151
24.4 The Price Paid For The Luxury Of Working In A Chart	153
<b>25 Perturbation Theory II</b>	<b>154</b>
25.1 Calculate $\Delta_\xi a$ , $\Delta_\xi c$ , $\Delta_\xi b_\alpha$ & $\Delta_\xi e_{\alpha\beta}$	154
25.2 Scalar Perturbations	156
25.3 Vector Perturbations	157
25.4 Tensor Perturbations	157
<b>26 Tutorials</b>	<b>158</b>
26.1 Topology	158
26.1.1 Topologies On A Simple Set	158
26.1.2 Continuous Functions	158
26.1.3 The Standard Topology On $\mathbb{R}^d$	159
26.2 Topological Manifolds	159
26.2.1 An Atlas From A Real World — the Möbius river	159
26.2.2 A Real World From An Atlas	160
26.2.3 Before The Invention Of The Wheel	160

26.3	Multilinear Algebra . . . . .	162
26.3.1	Vector Spaces . . . . .	162
26.3.2	Indices . . . . .	164
26.3.3	Linear Maps As Tensors . . . . .	166
26.4	Differential Manifolds . . . . .	166
26.4.1	Restricting The Atlas . . . . .	166
26.4.2	Soft Squares on $\mathbb{R} \times \mathbb{R}$ . . . . .	167
26.4.3	Undergraduate Multi-Dimensional Analysis . . . . .	168
26.4.4	Differentiability On A Manifold . . . . .	168
26.5	Tangent Spaces . . . . .	168
26.5.1	Virtuoso Use Of The Symbol $(\frac{\partial}{\partial x^i})_p$ . . . . .	168
26.5.2	Transformation Of Vector Components . . . . .	169
26.5.3	The Gradient . . . . .	171
26.5.4	Is There A Well-Defined Sum Of Curves? . . . . .	171
26.6	Fields . . . . .	173
26.6.1	Vector Fields For Practitioners . . . . .	173
26.6.2	The Cotangent Bundle $T^*\mathcal{M} \xrightarrow{\pi} \mathcal{M}$ . . . . .	174
26.7	Connections . . . . .	175
26.7.1	Practical Rules For How $\nabla$ Acts . . . . .	175
26.7.2	Connection Coefficients . . . . .	175
26.8	Parallel Transport and Curvature . . . . .	177
26.8.1	Where Connection Coefficients Appear . . . . .	177
26.8.2	The Round Sphere . . . . .	178
26.8.3	How Not To Define Parallel Transport . . . . .	178
26.10	Metric Manifolds . . . . .	179
26.10.1	Recognising & Dealing With Different Signatures . . . . .	179
26.10.2	Levi-Civita Connection . . . . .	180
26.10.3	Massaging The Length Functional . . . . .	180
26.10.4	A Practical Way To Quickly Determine Christoffel Symbols . . . . .	182
26.10.5	Properties Of The Riemann-Christoffel Tensor . . . . .	183
26.11	Symmetry . . . . .	184
26.11.1	Pull-Back & Push-Forward . . . . .	184
26.11.2	Lie Derivative — The Pedestrian Way . . . . .	186
26.12	Integration . . . . .	190
26.12.1	Integrals & Volumes . . . . .	190
26.13	Schwarzschild Spacetime . . . . .	191
26.13.1	Geodesics In A Schwarzschild Spacetime . . . . .	191
26.13.2	Gravitational Redshift . . . . .	193
26.14	Relativistic Spacetime, Matter & Gravitation . . . . .	195
26.14.1	Lorentz Force Law . . . . .	195
26.14.2	Which Curvature Can Feature In Einstein's Equations? . . . . .	197
26.20	Cosmology . . . . .	197
26.20.1	Killing's Equation . . . . .	197
26.20.2	Age Of The Universe... . . . .	198
26.23	Diagrams . . . . .	200

26.23.1 Penrose Diagram Of A Radiation-Filled Universe . . . . .	200
26.24 Perturbation Theory . . . . .	203



# 0 | Introduction

This course shall discuss the structure of space and time with an ultimate aim of understanding the theories of gravity and relativistic matter. We will connect these two theories using the famous Einstein equations and show how it is all related to the curvature of spacetime. Collectively this forms a beginning introduction to general relativity.

In order to be able to even start having this discussion, we will have to build up our understanding of the notion of spacetime. It is important we build up a rigorous understanding of what spacetime actually is and not simply just ‘it’s space and time somehow put together’. Instead we shall build up to a point where we can understand the following statement:

Spacetime is a four-dimensional topological manifold with a smooth atlas carrying a torsion-free connection compatible with a Lorentzian metric and a time orientation satisfying the Einstein equations.

Now, it would be surprising if the reader was already completely familiar with exactly what this statement means. The first part of the lecture series is basically devoted to clarifying/defining what all the terms above mean. The disclaimer that comes with this is that the first part of this course will be heavily mathematical and it won’t always be immediately clear why what we’re doing will lead to an understanding of gravity or matter. Dr. Schuller does a great job in trying to keep our minds on track with what we’re doing, however should you get a little lost on how everything will come together, my advice would be trust that it will and just focus on understanding the material fully, as this understanding will be vital later.

As a brief overview of how we will build up to and understanding of this statement, the following table tells us which lectures will tackle which parts of the statement:

Terms	Lecture
Topological	1
4-Dimensional Manifold	2
Smooth atlas	4
Connection	7
Torsion-free	8
Lorentzian Metric	10
Time Orientation	13
Einstein Equations	15

The remainder of the course will then be used to discuss this interplay between matter and gravity and to discuss objects such as black holes.

There are also tutorials provided with the course, and I shall equally type up these and place them at the end of the notes, so as to not interrupt the flow of the notes. It is highly recommended that the reader also go through these tutorials after the corresponding lecture.

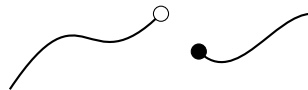
I have also included exercises throughout the notes to give the reader a chance to check they understand what's going on. Some of these are actually answered in the lectures, so can be checked against them. I have only done this if I thought the proof was relatively straight forward. Other exercises are based off comments made by Dr. Schuller while teaching. As well as these some exercises are of my own invention. I encourage the reader to attempt them all, and should they get stuck if they email me I shall try get back with some further hints and/or the solutions.

All of the diagrams in these notes have been drawn by myself in Tikz, and anyone who is interested in using them please feel free to email me. Alternatively, the code for these notes is available on via [my GitHub](#).

# 1 | Topology

## 1.1 Topology

At its coarsest level, spacetime is just a *set*. In other words, spacetime is just a collection of points, known as the *elements* of a set.<sup>1</sup> However, this definition is not enough to talk about even the simplest notions we discuss in classical physics, namely *continuity of maps*. It is important that we are able to talk about and require continuity of maps as the motion of a particle is given by a map, and it is clear we want this map to be continuous. That is, we don't want the position of a particle to all of a sudden 'jump' from one point to another:



So we need to introduce some new structure onto our set that allows us to talk about continuity. There are lots of things we could introduce in order to do this, however we need to be careful; we do not want to start adding additional properties to our set that will later come back to bite us. We therefore want to use the *weakest* structure we can. So what is it? Luckily the answer to this question is already known: it is a so-called *topology*.

**Definition** (Powerset). The **powerset**  $\mathcal{P}(S)$  of a set  $S$  is the set of all subsets of  $S$ .

**Definition** (Topology). Let  $\mathcal{M}$  be a set. A **topology**  $\mathcal{O}$  on  $\mathcal{M}$  is a subset  $\mathcal{O} \subseteq \mathcal{P}(\mathcal{M})$  satisfying:

- (i)  $\emptyset \in \mathcal{O}$  and  $\mathcal{M} \in \mathcal{O}$ .
- (ii) Given  $U, V \in \mathcal{O}$  then  $U \cap V \in \mathcal{O}$ .
- (iii) Let  $A$  be an arbitrary index set. Given  $U_\alpha \in \mathcal{O}$  then  $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{O}$ .

*Remark 1.1.1.* Conditions (ii) and (iii) look deceptively similar, but there is an important difference. Condition (ii) says that a *finite* intersection of elements is still in  $\mathcal{O}$ , whereas (iii) says that an *arbitrary* union of elements is in  $\mathcal{O}$ .

---

<sup>1</sup>I might include a short section on set theory here. This footnote is just to remind me to consider it.

Example 1.1.2. Let  $\mathcal{M} = \{1, 2, 3\}$ . Then we could choose to define

$$\begin{aligned}\mathcal{O}_1 &:= \{\emptyset, \{1, 2, 3\}\} \\ \mathcal{O}_2 &:= \{\emptyset, \{1\}, \{2\}, \{1, 2, 3\}\} \\ \mathcal{O}_3 &:= \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.\end{aligned}$$

The question is, which are topologies on  $\mathcal{M}$ ? A quick check shows that  $\mathcal{O}_1$  and  $\mathcal{O}_3$  are, but  $\mathcal{O}_2$  is *not* as  $\{1\} \cup \{2\} = \{1, 2\} \notin \mathcal{O}_2$ .

So we see, given the set  $c\mathcal{M}$  calculating whether something is a topology is rather easy (although perhaps a bit boring). What is we don't know the form of  $\mathcal{M}$ , can we still define specific topologies? The answer is yes and the example shows us what.

**Definition** (Chaotic Topology). Let  $\mathcal{M}$  be a set. The **chaotic** topology on  $\mathcal{M}$  is defined as

$$\mathcal{O}_{\text{chaotic}} := \{\emptyset, \mathcal{M}\}.$$

**Definition** (Discrete Topology). Let  $\mathcal{M}$  be a set. The **discrete** topology on  $\mathcal{M}$  is defined as

$$\mathcal{O}_{\text{discrete}} := \mathcal{P}(\mathcal{M}).$$

So in Example 1.1.2,  $\mathcal{O}_1$  is the chaotic topology and  $\mathcal{O}_3$  the discrete topology. However, both the chaotic and discrete topology are utterly useless, they are simply the extreme cases of topologies with the least amount and most amount of elements, respectively. However, on  $\mathbb{R}^d = \{(p_1, \dots, p_d) \mid p_i \in \mathbb{R}\}$  there is a very important topology which we shall use throughout these notes.

**Definition** (Standard Topology on  $\mathbb{R}^d$ ). Let  $\mathcal{M} = \mathbb{R}^d$ . The **standard** topology on  $\mathcal{M}$  is defined as

$$\mathcal{O}_s := \{U \in \mathcal{P}(\mathbb{R}^d) \mid \forall p \in U \exists r \in \mathbb{R}^+ : B_r(p) \subseteq U\},$$

where

$$B_r(p) := \left\{ (q_1, \dots, q_d) \in \mathbb{R}^d \mid \sum_{i=1}^d (q_i - p_i)^2 < r^2 \right\}$$

is called a *soft-ball of radius  $r$  about  $p$* , also known as the neighbourhood of  $p$  with radius  $r$ .

### Exercise

Show that the standard topology is in fact a topology, i.e. show it meets the conditions (i), (ii) and (iii).

Remark 1.1.3. To those familiar with vector space structures and normed spaces, you might be tempted to say 'Ah the soft ball is just the Euclidean norm'. However, the definition above does not need a full vector space structure (which a norm does) in order for it to hold. All we require is that we know what  $(q_i - p_i)^2$  means.

**Definition** (Topological Space). Let  $\mathcal{M}$  be a set and  $\mathcal{O}$  be a topology on  $\mathcal{M}$ . We call the double  $(\mathcal{M}, \mathcal{O})$  a **topological space**.

*Remark 1.1.4.* In these notes if we talk about  $\mathbb{R}^d$  being a topological space (for example when saying a map is continuous, see below), if no specific topology is given it will be assumed that we equip it with the standard topology.

*Remark 1.1.5.* In this lecture course we might not always write down the topology and simply call  $\mathcal{M}$  a topological manifold. Obviously if we say that, there is an invisible  $\mathcal{O}$  kicking about, we just want to save some typing. However, we shall try to always be explicit in order to avoid confusion.

Intuitively we can think of the standard topology on  $\mathbb{R}^d$  as shapes that don't include their boundaries and the soft ball as a circle<sup>2</sup> of radius  $r$  that doesn't contain the boundary. With this intuition we easily see the extension of the standard topology to general topologies: they are the sets of *open sets* within that set. In fact it actually works the other way around, we use a topology in order to define what we mean by an open set.

**Definition (Open Set).** Let  $(\mathcal{M}, \mathcal{O})$  be a topological space. We call  $U \subseteq \mathcal{M}$  an **open set** if and only if  $U \in \mathcal{O}$ .

**Definition (Closed Set).** Let  $(\mathcal{M}, \mathcal{O})$  be a topological space. We call a set  $V \subseteq \mathcal{M}$  a **closed set** if and only if  $\mathcal{M} \setminus V \in \mathcal{O}$ , where  $\mathcal{M} \setminus V$  is known as the *complement* of  $V$ .

*Remark 1.1.6.* It is tempting to think that a closed set is simply a set that is not open. However, this is not true. In fact a set can be

- (i) Open and not closed, e.g.  $(0, 1)$  in  $(\mathbb{R}, \mathcal{O}_s)$ ,
- (ii) Not open and closed, e.g.  $[0, 1]$  in  $(\mathbb{R}, \mathcal{O}_s)$ ,
- (iii) Open and closed, e.g.  $\emptyset$  in any topological space,
- (iv) Not open and not closed, e.g.  $[0, 1)$  in  $(\mathbb{R}, \mathcal{O}_s)$ .

### Exercise

Show that the examples given in the above remark are correct.

## 1.2 Continuous Maps

Let's first just recall the terminology/notation for a map. We say that  $f$  is a map from a set  $\mathcal{M}$ , known as the *domain*, to another set  $\mathcal{N}$ , known as the *target*, and we write this as  $f : \mathcal{M} \rightarrow \mathcal{N}$ . We say that the element  $m \in \mathcal{M}$  is mapped to  $n \in \mathcal{N}$ , which we write as  $f : m \mapsto n$ . A map will take *every* element in its domain to *some* element in its target. It is possible that a two different elements of the domain are mapped to the same element of the target and it is not required that every element in the target is hit. Based on this we have the following definitions.

<sup>2</sup>In  $\mathbb{R}^2$  at least. In higher values of  $d$  you just take that dimensional equivalent of a circle, e.g. a ball in 3D.

**Definition** (Injective Map). A map  $f : \mathcal{M} \rightarrow \mathcal{N}$  is said to be **injective** if it is one-to-one. That is

$$f(m_1) = f(m_2) \iff m_1 = m_2 \quad \forall m_1, m_2 \in \mathcal{M}.$$

**Definition** (Surjective Map). A map  $f : \mathcal{M} \rightarrow \mathcal{N}$  is said to be **surjective** if every element of the target is hit. That is

$$\forall n \in \mathcal{N} \quad \exists m \in \mathcal{M} : f(m) = n.$$

**Definition** (Bijective Map). A map  $f : \mathcal{M} \rightarrow \mathcal{N}$  is called **bijective** if it is both injective and surjective.

The answer to whether a map  $f : \mathcal{M} \rightarrow \mathcal{N}$  is *continuous* depends, *by definition*, on the choice of which topologies are chosen on the sets  $\mathcal{M}$  and  $\mathcal{N}$ .

**Definition** (Continuous Map). Let  $(\mathcal{M}, \mathcal{O}_{\mathcal{M}})$  and  $(\mathcal{N}, \mathcal{O}_{\mathcal{N}})$  be topological spaces. A map  $f : \mathcal{M} \rightarrow \mathcal{N}$  is called **continuous** with respect to  $\mathcal{O}_{\mathcal{M}}$  and  $\mathcal{O}_{\mathcal{N}}$  if and only if

$$\forall V \in \mathcal{O}_{\mathcal{N}}, \quad \text{preim}_f(V) \in \mathcal{O}_{\mathcal{M}},$$

where

$$\text{preim}_f(V) := \{m \in \mathcal{M} \mid f(m) \in V\}.$$

That is "the preimages of open sets in  $\mathcal{N}$  are open in  $\mathcal{M}$ ".

*Remark 1.2.1.* Note the preimage of a map  $f$  is *not* the same thing as its inverse. For example, we can not define an inverse for a non-injective map; if two elements in the domain map to the same element in the target, there is no clear way to decide which element you get under the inverse map. However the preimage in the case is the collection of both points. So continuity does not require the map to be injective. Equally note that surjectivity is not required as any element that isn't hit has preimage  $\emptyset \in \mathcal{O}_{\mathcal{M}}$ .

*Remark 1.2.2.* Note if we choose the topology on  $\mathcal{M}$  to be the discrete topology then *every* map  $f : \mathcal{M} \rightarrow \mathcal{N}$  is continuous. This is easily seen because the preimage of any set in  $\mathcal{O}_{\mathcal{N}}$  is either a subset of  $\mathcal{M}$  or the empty set, both of which are in the discrete topology on  $\mathcal{M}$ .

**Definition** (Homeomorphism). Let  $f : \mathcal{M} \rightarrow \mathcal{N}$  be a bijective map. Then the map is said to be a **homeomorphism** if both  $f$  and its inverse  $f^{-1}$  are continuous. They are the *structure-preserving*<sup>3</sup> maps of topology.

*Example 1.2.3.* Let  $\mathcal{M} = \mathcal{N} = \{1, 2\}$  and let  $f : \mathcal{M} \rightarrow \mathcal{N}$  be given by

$$f(1) = 2 \quad \text{and} \quad f(2) = 1.$$

Now define

$$\mathcal{O}_{\mathcal{M}} := \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \quad \text{and} \quad \mathcal{O}_{\mathcal{N}} := \{\emptyset, \{1, 2\}\}.$$

<sup>3</sup>In other words, they are the topological *isomorphisms*.

- (a) To check whether  $f$  is continuous w.r.t.  $\mathcal{O}_{\mathcal{M}}$  and  $\mathcal{O}_{\mathcal{N}}$  we need to check the preimages of open sets in  $\mathcal{N}$  are open sets in  $\mathcal{M}$ .

$$\begin{aligned}\text{preim}_f(\emptyset) &= \emptyset \in \mathcal{O}_{\mathcal{M}}, \\ \text{preim}_f(\{1, 2\}) &= \mathcal{M} \in \mathcal{O}_{\mathcal{M}},\end{aligned}$$

and so  $f$  is continuous.

- (b) How about the inverse map  $f^{-1} : \mathcal{N} \rightarrow \mathcal{M}$ . Well it satisfies

$$f^{-1}(1) = 2 \quad \text{and} \quad f^{-1}(2) = 1,$$

however we have

$$\text{preim}_{f^{-1}}(\{1\}) = \{2\}, \quad \text{and} \quad \text{preim}_{f^{-1}}(\{2\}) = \{1\},$$

neither of which are in  $\mathcal{O}_{\mathcal{N}}$ , and so this map is not continuous.

### 1.3 Composition of Continuous Maps

**Definition** (Composition of Maps). Given two maps  $f : \mathcal{M} \rightarrow \mathcal{N}$  and  $g : \mathcal{N} \rightarrow \mathcal{P}$ , we can define their **composition** as a new map

$$g \circ f : \mathcal{M} \rightarrow \mathcal{P}$$

where  $(g \circ f)(m) := g(f(m))$ .

**Theorem 1.3.1.** *Let  $f : \mathcal{M} \rightarrow \mathcal{N}$  and  $g : \mathcal{N} \rightarrow \mathcal{P}$  be two continuous maps w.r.t. the relevant topologies. Then the composition map  $g \circ f : \mathcal{M} \rightarrow \mathcal{P}$  is also continuous.*

*Proof.* Let  $V \in \mathcal{O}_{\mathcal{P}}$ . Now we have

$$\begin{aligned}\text{preim}_{g \circ f}(V) &:= \{m \in \mathcal{M} \mid (g \circ f)(m) \in V\} \\ &= \{m \in \mathcal{M} \mid f(m) \in \text{preim}_g(V) \in \mathcal{O}_{\mathcal{N}}\} \\ &= \text{preim}_f(\text{preim}_g(V)) \in \mathcal{O}_{\mathcal{M}},\end{aligned}$$

where the second line follows from the continuity of  $g$  and the last line from the continuity of  $f$ . □

#### Exercise

Show that Theorem 1.3.1 extends to the composition of an arbitrary number of continuous maps.

## 1.4 Inheriting a Topology

So we know how to define a topology on a set  $\mathcal{M}$ . We now want to ask the question: given some other set  $S$  is it possible to use the topology on  $\mathcal{M}$  to define one on  $S$ ? The answer is obviously yes<sup>4</sup> and it is known as the *inherited topology*. How you do this obviously depends on the situation. The way that is important for spacetime physics is the following.

**Definition.** Let  $(\mathcal{M}, \mathcal{O}_{\mathcal{M}})$  be a topological space and let  $S \subset \mathcal{M}$ . We define the **subset topology** to be

$$\mathcal{O}|_S := \{U \cap S \mid U \in \mathcal{O}_{\mathcal{M}}\}.$$

*Proof.* We want to show that the subset topology is indeed a topology, i.e. need to check it meets the three conditions.

- (i) We have  $\emptyset = \emptyset \cap S$  and  $\emptyset \in \mathcal{O}_{\mathcal{M}}$  which tells us that  $\emptyset \in \mathcal{O}|_S$ . Equally we have  $S = \mathcal{M} \cap S$  and  $\mathcal{M} \in \mathcal{O}_{\mathcal{M}}$  and so  $S \in \mathcal{O}|_S$ .
- (ii) Let  $A, B \in \mathcal{O}|_S$ , then we know there exists  $\tilde{A}, \tilde{B} \in \mathcal{O}_{\mathcal{M}}$  such that  $A = \tilde{A} \cap S$  and  $B = \tilde{B} \cap S$ . From which we have  $A \cap B = (\tilde{A} \cap S) \cap (\tilde{B} \cap S) = (\tilde{A} \cap \tilde{B}) \cap S$ . Finally using  $\tilde{A} \cap \tilde{B} \in \mathcal{O}_{\mathcal{M}}$  we have  $A \cap B \in \mathcal{O}|_S$ .
- (iii) Let  $U_{\alpha} \in \mathcal{O}|_S$ , which tells us that there exists  $\tilde{U}_{\alpha} \in \mathcal{O}_{\mathcal{M}}$  such that  $U_{\alpha} = \tilde{U}_{\alpha} \cap S$ . As above this tells us that  $\bigcup_{\alpha \in A} U_{\alpha} = \left(\bigcup_{\alpha \in A} \tilde{U}_{\alpha}\right) \cap S$  where  $A$  is an arbitrary index set. Finally using  $\bigcup_{\alpha \in A} \tilde{U}_{\alpha} \in \mathcal{O}_{\mathcal{M}}$  we get  $\bigcup_{\alpha \in A} U_{\alpha} \in \mathcal{O}|_S$ .

□

We might wonder why on earth we would choose to define such a map, the answer is to do with the continuity of maps.

*Claim 1.4.1.* Suppose we have some continuous map  $f : \mathcal{M} \rightarrow \mathcal{N}$  between topological spaces  $(\mathcal{M}, \mathcal{O}_{\mathcal{M}})$  and  $(\mathcal{N}, \mathcal{O}_{\mathcal{N}})$ . Then if we have some subset  $S \subset \mathcal{M}$  which we turn into a topological space with the subset topology  $(S, \mathcal{O}|_S)$ , then we are guaranteed that the restricted map  $f|_S : S \rightarrow \mathcal{N}$  is also continuous w.r.t.  $\mathcal{O}|_S$  and  $\mathcal{O}_{\mathcal{N}}$ .

### Exercise

Prove Claim 1.4.1.

<sup>4</sup>Otherwise the title of this section would seem silly.



## 2 | Topological Manifolds

It's a fact of life<sup>1</sup> that there are so many different topological spaces that mathematicians can't even classify<sup>2</sup> them. In other words, there is no such set of topological notions known such that we can work out whether two spaces are homeomorphic by simply 'ticking' whether two spaces have these notions or not.

For classical<sup>3</sup> spacetime physics, we may focus on topological spaces  $(\mathcal{M}, \mathcal{O}_{\mathcal{M}})$  that can be *charted*, analogously to how the surface of the Earth is charted in an atlas.

### 2.1 Topological Manifolds

**Definition.** Let  $U_p$  denote an open neighbourhood containing the point  $p$  in some topological space. A topological space  $(\mathcal{M}, \mathcal{O})$  is called a  *$d$ -dimensional topological manifold* if

$$\forall p \in \mathcal{M} \exists U_p \in \mathcal{O} : \exists x : U_p \rightarrow x(U_p) \subseteq \mathbb{R}^d$$

such that

- (i)  $x$  is *invertible*:  $x^{-1} : x(U_p) \rightarrow U_p$ ,
- (ii)  $x$  is *continuous*,<sup>4</sup>
- (iii)  $x^{-1}$  is *continuous*.

*Remark 2.1.1.* Note, from the required continuity of  $x$  and its inverse, we see that the image  $x(U_p)$  must be open w.r.t. the standard topology on  $\mathbb{R}^d$ .

*Example 2.1.2.* Let  $\mathcal{M}$  be the surface of a torus. This is a subset of  $\mathbb{R}^3$ , and so we can inherit the subset topology from the standard topology on  $\mathbb{R}^3$ . This is an example of a 2-dimensional topological manifold. We see this by taking any open neighbourhood on the torus, which is just a boundaryless closed shape on the surface, and we map all the points within it to an open set in  $\mathbb{R}^2$  (see diagram below). With some thought/workings one can convince themselves that this map will be injective (and so invertible), the inverse map is surjective (so that the whole torus surface is mapped) and continuous in both directions.

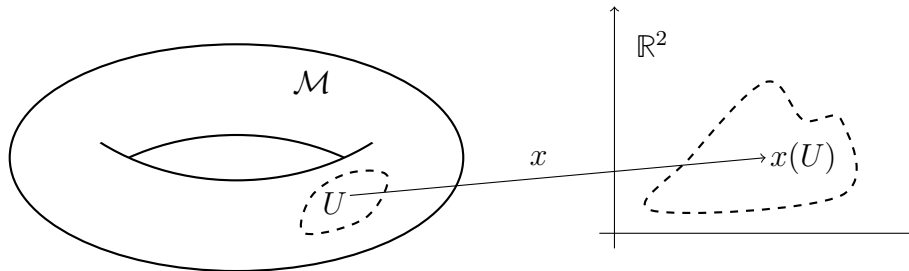
---

<sup>1</sup>Or mathematics to be less dramatic.

<sup>2</sup>In the sense that one can classify all Lie groups.

<sup>3</sup>As in not quantum mechanical. Obviously we are talking about relativistic physics.

<sup>4</sup>We use the standard topology on  $\mathbb{R}^d$ .



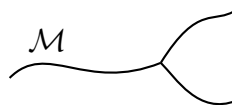
This is exactly the same idea as what one does when charting the surface of the Earth to make road maps and atlases.

Remark 2.1.3. It might be tempting to say that a topological manifold is homeomorphic<sup>5</sup> to  $\mathbb{R}^d$  given the explanation in the previous example. However, this is not true because our map is only surjective to a *subset* of  $\mathbb{R}^d$ , not the whole set. So the correct statement is that a topological manifold is homeomorphic to some particular subset of  $\mathbb{R}^d$ .

It is important to note that the values in the chart (i.e. the coordinates in  $\mathbb{R}^2$  above) bare no physical significance whatsoever. They simply act as a way for us to compare the positions of things in the real world. It is the surface of the torus itself that has the physical significance. To clarify, if the base of the Eiffel tower was at point  $p \in \mathcal{M}$  and we mapped it to the coordinates  $(x_1(p), x_2(p)) = (1, 2)$ , say, the values 1 and 2 do not mean anything *physical*, they simply tell us that *in this chart* the position of the Eiffel tower's base is  $(1, 2)$ . Of course if we picked a different chart (for example consider just rotating our chart by 90 degrees) these coordinate values would change to something new, however the Eiffel tower itself is completely unaffected by this.

Example 2.1.4. Let  $\mathcal{M}$  be a wire loop. We again can imagine this in  $\mathbb{R}^3$  and inherit the subset topology from the standard topology. Following the same idea as the previous example, we see that this is a 1-dimensional topological manifold.

Example 2.1.5. Now consider the following diagram



<sup>5</sup>That is there exists a bijective map that is continuous and so is its inverse.

Again this is clearly a subset of  $\mathbb{R}^3$  (or even  $\mathbb{R}^2$  if you view it as flat on the page) and so we can inherit a topology onto it. However this topological space fails to be a topological manifold because of the splitting point. This point essentially stops us being able to define an invertible, both ways continuous map.

*Terminology.* • The pair  $(U, x)$  is called a **chart** of  $(\mathcal{M}, \mathcal{O})$ .

- The set  $\mathcal{A} = \{(U_{(\alpha)}, x_{(\alpha)}) \mid \alpha \in A\}$ , for some arbitrary index set  $A$ , is called an **atlas** of  $(\mathcal{M}, \mathcal{O})$  if  $\bigcup_{\alpha \in A} U_{(\alpha)} = \mathcal{M}$ .
- $x : U \rightarrow x(U) \subseteq \mathbb{R}^d$  is called a **chart map** defined by  $x(p) = (x^1(p), \dots, x^d(p))$ , where  $x^i(p)$  is the  $i^{\text{th}}$  coordinate of  $p$  w.r.t. the chosen chart  $(U, x)$ .
- $x^i : U \rightarrow \mathbb{R}$  are called the **coordinate maps**.

**Definition** (Maximal Atlas). An atlas that contains every possible chart for a topological manifold is called a **maximal atlas**.

## 2.2 Chart Transition Maps

As the name suggests, a chart transition map is a chart dependent thing and therefore have no physical meaning at all. However, they are incredibly useful (especially for physicists) and so we shall study them.

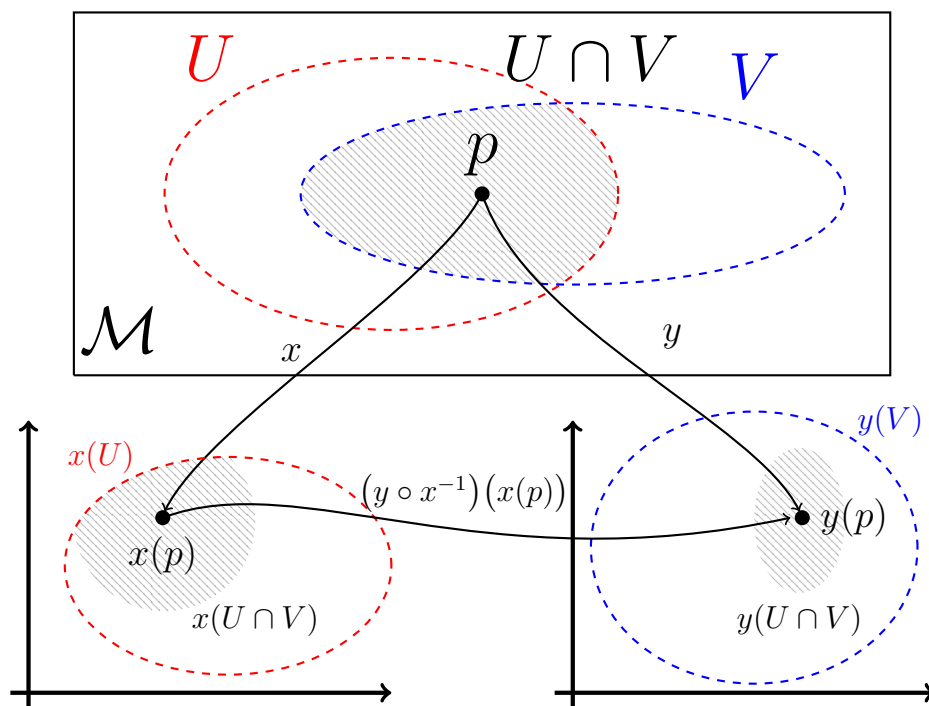
Imagine two charts  $(U, x)$  and  $(V, y)$  for the same topological space  $(\mathcal{M}, \mathcal{O})$  with overlapping regions, i.e.  $U \cap V \neq \emptyset$ . A point in this overlap region can be mapped by both  $x$  and  $y$  to their respective patches of  $\mathbb{R}^d$ . We can go between these two chart representatives of the point using the **chart transition maps**. For example if we want to go from the chart  $(U, x)$  to  $(V, y)$  we use the chart transition map  $(y \circ x^{-1}) : x(U \cap V) \rightarrow y(U \cap V)$  (see Figure 2.1).

We can draw this idea just in terms of maps by the following:

$$\begin{array}{ccc}
 & U \cap V & \\
 x \swarrow & & \searrow y \\
 x(U \cap V) & \xrightarrow{y \circ x^{-1}} & y(U \cap V)
 \end{array}$$

Informally, the chart transition maps contain the information about how to ‘glue together’ the pages of an atlas. That is, given 10 pages of an atlas each of which overlaps with the two others, the chart transition maps tell us what order to put them together to get the geographical order<sup>6</sup> correct.

<sup>6</sup>By which we obviously mean that page 3 follows on from page 2 in the same way that page 2 follows on from page 1.



**Figure 2.1:** Chart representations  $(U, x)$  and  $(V, y)$  with a non-empty overlap. The overlap region (shaded)  $U \cap V$  is mapped by both  $x$  and  $y$  to their respective representations. A chart transition map  $y \circ x^{-1}$  can be used to map the overlap region from one representation into the other. The chart transition map is continuous as it is the composition of two continuous maps.

### 2.3 Manifold Philosophy

Often it is desirable (or indeed the only way) to define properties (e.g. continuity) of real world objects (e.g. the curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$ ) by judging suitable conditions, not on the real world object itself but on a chart representative/image of that real world object. The main advantage of doing this is we can then use undergraduate analysis to study these properties. For example if  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  is the real world trajectory of a particle, we can work out whether the path is continuous by asking whether the composite map  $(x \circ \gamma) : \mathbb{R} \rightarrow \mathbb{R}^d$  is continuous, using the undergraduate notion of continuity of such a map.

We shall see, however, that we must be careful when doing this. Just because a real world object has a certain undergraduate behaviour in *some* chart, it does not mean the real world object has it too. What we will actually require is that we can form an atlas such that in *every* chart the representative of the object has our desired property. We will see next lecture, that this can be thought of as the idea that we want the chart transition maps to also have our desired undergraduate property, and that the property is maintained under the composition of maps. What we're saying here is that the property of the real world object can't depend on how we imagine it drawn on a piece of paper. It is a *chart independent* property.

**Exercise**

Show that the ‘lifted’ notion of undergraduate continuity corresponds to the definition of a continuous map given earlier. That is, if  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  is a path on our manifold, show that if we know all the chart representative maps  $(x \circ \gamma) : \mathbb{R} \rightarrow \mathbb{R}^d$  are undergraduate continuous, we can conclude that the preimage of open sets in  $\mathcal{M}$  under  $\gamma$  are open sets in  $\mathbb{R}$ .

*Hint: Use Theorem 1.3.1 along with the definition of a topological manifold.*

## 3 | Multilinear Algebra

Multilinear algebra, as the name suggests, is just an extension of linear algebra. When studying linear algebra, one invariably studies vector space structures. We wish to emphasise here that we will *not* equip space(time) with a vector space structure. This might seem like a strange thing to say, but in which case ask yourself ‘where is  $5 \times \text{Paris}$ ?’ or ‘where is  $\text{Paris} + \text{Vienna}$ ?’ However, the so-called *tangent spaces*  $T_p\mathcal{M}$  to smooth manifolds<sup>1</sup> will carry a natural vector space structure.

It is beneficially to first study vector spaces abstractly for two reasons

- (i) For the construction of  $T_p\mathcal{M}$ , one need an intermediate vector space  $C^\infty(\mathcal{M})$ , and
- (ii) Tensor techniques are most easily understood in an abstract setting.

### 3.1 Vector Spaces

In order to define a vector space, we first need to make sure we know what a *field* is.<sup>2</sup>

**Definition** (Abelian Group). Let  $K$  be a set and let  $\bullet : K \rightarrow K$ . The double  $(K, \bullet)$  is a **Abelian** (or commutative) group if the following axioms are satisfied

- (i) Commutative;  $a \bullet b = b \bullet a$ ,
- (ii) Associative;  $(a \bullet b) \bullet c = a \bullet (b \bullet c)$ ,
- (iii) Neutral element;  $\exists 0 \in K$  such that  $a \bullet 0 = 0 \bullet a = a$ ,
- (iv) Inverse;  $\exists a^{-1} \in K$  such that  $a \bullet a^{-1} = a^{-1} \bullet a = 0$ .

*Example 3.1.1.* The real numbers equipped with addition form an Abelian group. However, the real numbers do *not* form an Abelian group when equipped with multiplication. This is because the neutral element is clearly  $1 \in \mathbb{R}$ , but  $0 \in \mathbb{R}$  and there is no  $a \in \mathbb{R}$  such that  $a \times 0 = 1$ .<sup>3</sup>

**Definition** (Field). A **field** is a triple  $(\mathbb{F}, +, \cdot)$  where

- $\mathbb{F}$  is a set, and

---

<sup>1</sup>What these terms strictly mean shall be explained in the course.

<sup>2</sup>There is a lot more information on this in Dr. Schuller’s Lectures on Geometric Anatomy of Theoretical Physics.

<sup>3</sup>Infinity is not counted as a well defined element of the reals.

- $+, \cdot : \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F}$  are maps.

They must satisfy the following axioms

- (i)  $(\mathbb{F}, +)$  is an Abelian group.
- (ii)  $(\mathbb{F}^*, \cdot)$  is an Abelian group, where  $\mathbb{F}^* = \mathbb{F} \setminus \{0\}$ .
- (iii) Distributive;  $\forall a, b, c \in \mathbb{F} \ a \cdot (b + c) = a \cdot b + a \cdot c$ .

Remark 3.1.2. If we don't require condition (ii) above, but in its place just require the associativity condition,  $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ , then we get a weaker notion called a *ring*. If we also require the existence of a neutral element  $1 \in \mathbb{F}$  then we get a *unital ring*. Similarly, if we require the commutative condition we get a *commutative ring*. We will use rings later in the course (starting in lecture 6).

**Definition** ( $\mathbb{F}$ -Vector Space). A  **$\mathbb{F}$ -vector space** is the triple  $(V, +, \cdot)$  where

- $V$  is a set,
- $+$  is the addition map,  $+: V \times V \rightarrow V$ , and
- $\cdot$  is the s-multiplication map,  $\cdot : \mathbb{F} \times V \rightarrow V$ ,

satisfying, for all  $v, w, u \in V$  and  $a, b \in \mathbb{F}$

- (i) Commutative w.r.t.  $+$ ;  $v + w = w + v$ ,
- (ii) Associative w.r.t.  $+$ ;  $(v + w) + u = v + (w + u)$ ,
- (iii) There is a neutral element w.r.t.  $+$ ;  $\exists e \in V$  such that  $v + e = v$ ,
- (iv) There is an inverse element w.r.t.  $+$ ;  $\exists \tilde{v} \in V$  such that  $\tilde{v} + v = v + \tilde{v} = e$ .
- (v) Associative w.r.t.  $\cdot$ ;  $a \cdot (b \cdot v) = (a \cdot b) \cdot v$ ,
- (vi) Distributive 1;  $(a + b) \cdot v = a \cdot v + b \cdot v$ ,
- (vii) Distributive 2;  $a \cdot (v + w) = a \cdot v + a \cdot w$ ,
- (viii) Unitary w.r.t.  $\cdot$ ;  $1 \cdot v = v$ .

In these notes we will only really consider  $\mathbb{R}$ -vector spaces, and so most of the definitions that follow will use  $\mathbb{R}$  as the field. Obviously we could extend these definitions to general  $\mathbb{F}$ -vector spaces.

Remark 3.1.3. We should actually be careful in the above definition when we write  $+$  and  $\cdot$ . There are two kinds floating about. One is the  $+/ \cdot$  we are defining for our vector space, and the other is the  $+/ \cdot$  on  $\mathbb{F}$ . For example in (vi) we have  $(a+b) \cdot v = a \cdot v + b \cdot v$ . The black  $+$ s here are the ones defined for our vector space, whereas the red one is the addition on  $\mathbb{F}$ . The same idea goes for condition (v). If we were being really particular, we would give these different names, however we shall just assume that we can work it out given the context (i.e. both  $a$  and  $b$  are real numbers so  $a + b$  is clearly the addition on  $\mathbb{F}$ .)

*Remark 3.1.4.* Just as we can build a  $\mathbb{F}$ -vector space over a field  $(\mathbb{F}, +, \cdot)$ , we can build a so-called  $R$ -module over a ring  $(R, +, \cdot)$ . It is done in exactly the same fashion.

*Terminology.* An element of a vector space is often referred to *informally* as a vector.

We emphasise the word *informally* in the above terminology, for a reason that the next example demonstrates.

*Example 3.1.5.* First define the set

$$P := \left\{ p : (-1, 1) \rightarrow \mathbb{R} \mid p(x) = \sum_{n=1}^N p_n x^n, p_n \in \mathbb{R} \right\}.$$

We could then ask whether  $\square : (-1, 1) \rightarrow \mathbb{R}$  defined by  $\square(x) = x^2$  is a vector? The answer is ‘no, of course it’s not!’ Why? Well because we don’t even have a vector space so can’t have vectors. That is we haven’t defined an addition and s-multiplication for  $P$ .

Now let’s imagine that we define an addition and s-multiplication *pointwise*, meaning

$$+ : (p, q) \mapsto p +_P q$$

is defined via

$$(p +_P q)(x) := p(x) +_{\mathbb{R}} q(x),$$

and similarly for  $\cdot_P : \mathbb{R} \times V \rightarrow \mathbb{R}$ , where the subscripts indicate which  $+$  we’re talking about. If we *now* ask whether  $\square$  is a vector, the answer is ‘well, yes!’

The point of the above example is to demonstrate that you can’t just look at something itself and decide whether it is a vector or not, you need to know whether there is an underlying vector space or not. This might seem like a rather pedantic point to prove, however it is an important point to note as people often ask ‘what *is* a tensor?’ A tensor is an extension of a vector<sup>4</sup> and so they are defined as elements of a tensor space, which in itself is a rather abstract object. This often leads people to being very confused, however once you understand the above point, this confusion should die away.

#### Exercise

Prove that  $(P, +_P, \cdot_P)$  as defined in the above example is in fact a vector space, i.e. show it meets the 8 axioms.

## 3.2 Linear Maps

It is a standard procedure in mathematics that once you introduce a new structure to a object that you consider the structure preserving maps. That is the maps that map two objects with the same types of structures that have the property that the structure on one can be derived from the structure on the other. Such maps are generally known as *isomorphisms* of the relevant structure(s). We have already done this when considering topological spaces: we considered the homeomorphisms between two topological spaces. As you might have pieced together, the structure preserving maps for the vector space structure are known as *linear maps*.

<sup>4</sup>Or perhaps more correctly, a vector is a specific type of tensor.



**Definition** (Linear Maps). Let  $(V, +_V, \cdot_V)$  and  $(W, +_W, \cdot_W)$  be vector spaces. Then a map  $\varphi : V \rightarrow W$  is called **linear** if: for all  $v, \tilde{v} \in V$  and  $\lambda \in \mathbb{R}$ ,

$$(i) \quad \varphi(v +_V \tilde{v}) = \varphi(v) +_W \varphi(\tilde{v}), \text{ and}$$

$$(ii) \quad \varphi(\lambda \cdot_V v) = \lambda \cdot_W \varphi(v).$$

Example 3.2.1. Again let's consider the space  $P$  as defined in Example 3.1.5. Let's consider the map  $\delta : P \rightarrow P$  defined by  $\delta(p) := p'$ , i.e. the differential operator. This is a linear map as

$$\begin{aligned} \delta(p + q) &= (p + q)' = p' + q' = \delta(p) + \delta(q), & \text{and} \\ \delta(\lambda \cdot p) &= (\lambda \cdot p)' = \lambda \cdot p' = \lambda \cdot \delta(p). \end{aligned}$$

*Notation.* We write a linear map  $\varphi : V \rightarrow W$  by putting a tilde on the arrow, i.e.  $\varphi : V \xrightarrow{\sim} W$ .

**Theorem 3.2.2** (Composition of Linear Maps). *Suppose we have the following linear maps  $\varphi : V \xrightarrow{\sim} W$  and  $\psi : W \xrightarrow{\sim} U$ . Then the map  $(\psi \circ \varphi) : V \xrightarrow{\sim} U$  is also linear.*

#### Exercise

Prove Theorem 3.2.2 and show that it holds for arbitrary compositions.

Example 3.2.3. Let  $\delta : P \xrightarrow{\sim} P$  be the same map as before. Now consider the composite map  $\delta \circ \delta$ , the second derivative operator. Then Theorem 3.2.2 along with the previous example tells us that this is also a linear map  $(\delta \circ \delta) : P \xrightarrow{\sim} P$

### 3.3 Vector Space of Homomorphisms

**Definition** (The Vector Space of Homomorphisms). Let  $(V, +, \cdot)$  and  $(W, +, \cdot)$ <sup>5</sup> be vector spaces. Then we can define the *set*

$$\text{Hom}(V, W) := \{\varphi : V \xrightarrow{\sim} W\}.$$

We can turn this into a vector space by defining

$$\begin{aligned} \oplus : \text{Hom}(V, W) \times \text{Hom}(V, W) &\rightarrow \text{Hom}(V, W) \\ (\varphi, \psi) &\mapsto \varphi \oplus \psi, \end{aligned}$$

where  $(\varphi \oplus \psi)(V) = \varphi(V) + \psi(V)$ , and similarly for the s-multiplication  $\odot : \mathbb{R} \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W)$ . The triple  $(\text{Hom}(V, W), \oplus, \odot)$  is the **vector space of homomorphisms**.

Example 3.3.1. Again we use our polynomial set  $P$ . We can turn  $\text{Hom}(P, P)$  into a vector space by defining  $\oplus/\odot$  as above. So we obtain things like

$$5 \odot \delta \oplus (\delta \circ \delta) \in \text{Hom}(P, P),$$

and so a sum of derivative operators of different orders is again a linear map on the set of polynomials.

<sup>5</sup>From now on we shall drop the subscripts on the  $+/\cdot$  and assuming we know which is which based on the context of the equation.

### 3.4 Dual Vector Space

**Definition** (Dual Vector Space). Let  $(V, +, \cdot)$  be a vector space. We define the **dual vector space** (to  $V$ )  $(V^*, \oplus, \odot)$ , where

$$V^* := \text{Hom}(V, \mathbb{R}) := \{\varphi : V \xrightarrow{\sim} \mathbb{R}\},$$

and where  $\oplus/\odot$  are necessarily defined.

*Terminology.* As with the vector, an element  $\varphi \in V^*$  is informally called a *covector*.

We actually need to be even more careful when talking about covectors; as we have defined it a covector is an element of a vector space, but our previous terminology tells us that that's a vector. So in order to call something a covector, we not only need to know that the set it belongs to has an underlying vector space, we also need to know that it is a dual to some other vector space, whose elements we have already called vectors.

*Example 3.4.1.* Consider a map  $I : P \xrightarrow{\sim} \mathbb{R}$ , which tells us that  $I \in P^*$ . We define it by

$$I(p) := \int_0^1 dx p(x).$$

This tells us that the integration operator  $I := \int_0^1 dx$  is a covector.

#### Exercise

Prove that  $I : P \xrightarrow{\sim} \mathbb{R}$  is indeed linear.

**Theorem 3.4.2.** Let  $(V, +, \cdot)$  be a vector space. If it is finite dimensional<sup>6</sup> then the double dual is the vector space itself. That is<sup>7</sup>

$$(V^*)^* = V,$$

when  $\dim V < \infty$ .

*Remark 3.4.3.* When we learn physics lower down in school we actually meet lots of covectors that we, at that time, called vectors. This was obviously done so as not to have to introduce the idea of a covector. However, we just want to point out here that covectors are not some new thing we have never seen before.

### 3.5 Tensors

If we are consider finite dimensional vector spaces, then there is a very natural definition for tensors as multilinear maps.

<sup>6</sup>We shall soon clarify what we mean by the 'dimension' of a vector space.

<sup>7</sup>Really we should use an isomorphic symbol here, but we shall ignore this in these notes.

**Definition (Tensor).** Let  $(V, +, \cdot)$  be a vector space. A  $(r, s)$ -tensor,  $T$ , over  $V$  is a multilinear map

$$T : \underbrace{V^* \times \dots \times V^*}_{r\text{-terms}} \times \underbrace{V \times \dots \times V}_{s\text{-terms}} \xrightarrow{\sim} \mathbb{R}.$$

*Remark 3.5.1.* Others flip the definition of an  $(r, s)$ -tensor, in the sense that  $r$  tells you how many  $V$  terms appear in the above map and  $s$  tells you how many  $V^*$  terms appear there. It is important to make sure you know which convention you are dealing with before moving forward.

*Example 3.5.2.* Let  $T$  be a  $(1, 1)$ -tensor. This means it takes in as its argument one covector and one vector. The multilinearity of  $T$  tells us that: for all  $\varphi, \psi \in V^*$ ,  $v, w \in V$  and  $\lambda \in \mathbb{R}$

$$\begin{aligned} T(\varphi + \psi, v) &= T(\varphi, v) + T(\psi, v), \\ T(\lambda \cdot \varphi, v) &= \lambda T(\varphi, v), \\ T(\varphi, v + w) &= T(\varphi, v) + T(\varphi, w) \\ T(\varphi, \lambda \cdot v) &= \lambda T(\varphi, v). \end{aligned}$$

*Example 3.5.3.* We can give an example of a tensor using our polynomial space. The map  $g : P \times P \xrightarrow{\sim} \mathbb{R}$  defined by

$$g(p, q) = \int_{-1}^1 dx p(x)q(x)$$

is a  $(0, 2)$ -tensor over  $P$ . This is just the inner product on the real numbers. So the inner product is a  $(0, 2)$ -tensor. This example will lead nicely into later when we discuss so-called metrics.

*Terminology.* As defined above, the number  $r$  is often known as the *covariant order* of  $T$  and  $s$  the *contravariant order*. Their sum  $r + s$  is known as the *rank* of  $T$ .

The definition for the tensor we have given is actually only one way that you might see a tensor defined. We shall now give a couple others to make reading other texts easier. Both of these require our vector spaces to be finite dimensional.

**Definition (Tensor (via Tensor Product)).** Let  $(V, +, \cdot)$  be a vector space. A  $(r, s)$ -tensor is defined by

$$T = \underbrace{V \otimes \dots \otimes V}_{r\text{-terms}} \otimes \underbrace{V^* \otimes \dots \otimes V^*}_{s\text{-terms}} \equiv V^{\otimes r} \otimes (V^*)^{\otimes s},$$

where  $\otimes$  is the so-called *tensor product*.

One can give a strict definition of the tensor product, but for our purposes, we can just view it to be such that this definition and the first tensor definition coincide. Note how here the  $r$  and  $s$  are switched, so we have  $r$   $V$  terms and  $s$   $V^*$  terms. Now, because we're assuming our vector space is finite dimensional, Theorem 3.4.2 tells us that  $V = (V^*)^*$ , and so we can think of  $V$  as the set of all linear maps from  $V^*$  to  $\mathbb{R}$ . We therefore just take the tensor product to mean 'we have  $r$  linear maps  $V : V^* \xrightarrow{\sim} \mathbb{R}$  and  $s$  linear maps  $V^* : V \xrightarrow{\sim} \mathbb{R}$ .'

This definition is useful because it shows us easily how to make higher order tensors: simply tensor product it with another tensor. For example if  $T$  is a  $(r, s)$ -tensor and  $S$  is a  $(p, q)$ -tensor then  $T \otimes S$  is a  $(r + p, s + q)$ -tensor.

There is a third common way people like to think/define tensors. It is easiest explained through an example.

*Example 3.5.4.* Let  $(V, +, \cdot)$  be a *finite dimensional* vector space and let  $T$  be a  $(1, 1)$ -tensor. Then  $T$  maps one covector and one vector to a real number. However, if we only feed it the one vector we are left with a linear map from  $V^*$  to  $\mathbb{R}$ . This is, by definition, an element of  $(V^*)^*$ , but because our vector space is finite dimensional, Theorem 3.4.2 tells us that  $(V^*)^* = V$ . We can thus define the map  $\phi_T : V \xrightarrow{\sim} V$  by  $\phi(v) = T(\bullet, v)$ , where  $\bullet$  indicated an empty slot. It is for this reason people often refer to a  $(1, 1)$ -tensor as a linear map that takes a vector to a vector. Similar words are used for higher order tensors.

*Remark 3.5.5.* Personally,<sup>8</sup> I am not a fan of talking about  $(r, s)$ -tensors at all. The reason for this is the notation is highly misleading as it fails to take into account the ordering of the vector spaces in the Cartesian product. To clarify what I mean, the two spaces with corresponding sets

$$\begin{aligned} V \times V^* &:= \{(v, \tilde{v}) \mid v \in V \text{ and } \tilde{v} \in V^*\}, \\ V^* \times V &:= \{(\tilde{v}, v) \mid \tilde{v} \in V^* \text{ and } v \in V\} \end{aligned}$$

are not the same. The addition on the two spaces separately say ‘add two elements entry wise. So the first entries are added together and the second entries are added together.’ Clearly, then, we can not add an element from the former to an element of the later as the ordering of the entries is switched. The only way we could do this addition would be to redefine our notion of addition to account for this.

However, people would call a linear map from either of these spaces a  $(1, 1)$ -tensor, but a tensor can be made into a tensor space (as we will do shortly) and so we should be able to add elements in this set together, given some rule. However, we have just established that there is no consistent rule in order to do this.

In the language of our second definition of a  $(r, s)$ -tensor, this problem is related to the fact that one can’t simply compare  $V \otimes W$  and  $W \otimes V$  for two general spaces. That is, they are completely different spaces and need not be related by any sort of symmetry property.

This is a problem that is very rarely highlighted in textbooks,<sup>9</sup> as the two spaces are clearly isomorphic (all you are doing is switching the order of the entries around) but that I think it is an important point to note.<sup>10</sup> I therefore think it is best to just give the explicit form of a tensor in terms of its tensor product array, as then there can be no confusion. However, this is not what Dr. Schuller does in his course and so I shall not do this in the rest of the notes.

### 3.6 Vectors and Covectors as Tensors

**Corollary 3.6.1.** *A covector  $\varphi \in V^*$ , i.e.  $\varphi : V \xrightarrow{\sim} \mathbb{R}$ , is a  $(0, 1)$ -tensor.*

**Corollary 3.6.2.** *For a finite dimensional vector space, a vector  $v \in V$  is a  $(1, 0)$ -tensor.*

<sup>8</sup>As in me, Richie.

<sup>9</sup>At least the ones I’ve read.

<sup>10</sup>An attitude past on to me by my lecturer at university who pointed this all out.

### 3.7 Bases

So far we have talked about vectors without mentioning any numbers for them at all. That is at no point have we said the vector  $(1, 2) \in \mathbb{R}^2$  or something. In order to even make such a statement, we need to introduce a basis, which tells us what the entries mean, for our vector space. For example,  $(1, 2) \in \mathbb{R}^2$  could mean ‘go along the  $x$ -axis 1 unit and along with  $y$ -axis 2 units’, in which case  $x$  and  $y$  are our choice of basis. Now obviously this is not the only choice of basis for  $\mathbb{R}^2$ , it is simply one of an uncountably many. Often making this choice allows us to progress greatly with the problem at hand, however one must always be conscious of the fact that they made a choice and that *anything* derived using that choice could be completely dependent on that choice, by which we mean that the result could be different had a different choice been made. If one wants to assign the result of our calculation as a property of the vector itself (e.g. its *length*) one needs to show that the result is completely basis independent. It is therefore best to try and avoid bases all together and only use them when absolutely necessary.

As we have already said, using a basis can often greatly simplify a calculation and so, with the above comment in mind, we shall now proceed to studying bases.

**Definition** (Basis for Vector Space). Let  $(V, +, \cdot)$  be a vector space. A subset  $B \subseteq V$  is called a (Hamel-)<sup>11</sup> **basis** if

$$\forall v \in V, \exists! \text{ finite } F = \{f_1, \dots, f_n\} \subset B : \exists! v_1, \dots, v_n \in \mathbb{R} : v = v^1 f_1 + \dots + v^n f_n.$$

This is not the only way to define a (Hamel-)basis and indeed is not the most useful for calculations. Instead we have the following definition.

**Definition** (Basis for Vector Space (linear independence)). Let  $(V, +, \cdot)$  be a vector space. A subset  $B = \{e_1, \dots, e_d\} \subseteq V$  is called a **basis** if

- (i) The basis *spans/generates*  $V$ ; that is any  $v \in V$  can be written as a linear combination of the basis elements, and
- (ii) The basis elements are linearly independent; that is

$$\sum_{i=1}^d \lambda^i e_i = 0 \quad \implies \quad \lambda^i = 0 \quad \forall i \in \{1, \dots, d\}.$$

**Definition** (Dimension of Vector Space). If there exists a basis  $B \subseteq V$  for a vector space  $(V, +, \cdot)$  with finitely many elements, say  $d$  many, then we call  $d$  the **dimension of the vector space**, denoted  $\dim V = d$ .

*Claim 3.7.1.* We claim that the dimension of a vector space is well defined. That is every basis for  $V$  will have  $d$  elements.

Note that the definition above holds for both infinite and finite dimensional vector spaces, as we did not require  $d < \infty$ . However from now on in these notes, we shall always assume we are dealing with a finite dimensional vector space, unless otherwise specified.

<sup>11</sup>As apposed to a Schauder-basis. For more info. see Dr. Schuller’s Quantum Theory course.

*Remark 3.7.2.* Let  $(V, +, \cdot)$  be a vector space. Then having chosen a basis  $\{e_1, \dots, e_n\}$ , we may uniquely associate

$$v \mapsto (v^1, \dots, v^n),$$

called the **components** of  $v$  with respect to the chosen basis, such that  $v = v^1 e_1 + \dots + v^n e_n$ .

Of course given the vector space  $(V, +, \cdot)$  and its dual space  $(V^*, +, \cdot)$ , you can define a basis on each space completely independently of each other. However, there is a very nice, and incredibly helpful, way we can related these bases.

**Definition (Dual Basis).** Let  $(V, +, \cdot)$  be a vector space and let  $\{e_1, \dots, e_n\}$  be a basis on it. We define *the dual basis* of the dual space  $(V^*, +, \cdot)$  as  $\{\epsilon^1, \dots, \epsilon^n\}$  satisfying

$$\epsilon^i(e_j) = \delta_j^i = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

### Exercise

Prove that the constraint above uniquely defines the elements  $\{\epsilon^1, \dots, \epsilon^n\}$  and show that they do indeed form a basis for  $(V^*, +, \cdot)$ .

*Hint: Use the linearity of the elements  $\varphi \in V^*$ .*

*Example 3.7.3.* Let's set  $N = 3$  in our polynomial space (i.e. the highest order is cubed). Then the set  $\{e_0, e_1, e_2, e_3\}$ <sup>12</sup> is a basis for this space if we identify

$$e_0(x) = 1, \quad e_1(x) = x, \quad e_2(x) = x^2, \quad e_3(x) = x^3.$$

It is easy to see that this is a basis because any polynomial of order 3 can be written a linear combination of these terms. The dual basis for the dual space is given by

$$\epsilon^a = \frac{1}{a!} \partial^a \Big|_{x=0},$$

for  $a = 0, 1, 2, 3$ . Direct calculation shows that this satisfies our necessary condition.

## 3.8 Components of Tensors

**Definition (Components of Tensors).** Let  $T$  be a  $(r, s)$ -tensor over a finite dimensional vector space  $(V, +, \cdot)$  and let  $\{e_1, \dots, e_{\dim V}\}$  be a basis for it with corresponding dual basis  $\{\epsilon^1, \dots, \epsilon^{\dim V}\}$ . Then define the  $(r + s)^{\dim V}$  many numbers

$$T^{i_1 \dots i_r}_{j_1 \dots j_s} := T(\epsilon^{i_1}, \dots, \epsilon^{i_r}, e_{j_1}, \dots, e_{j_s}) \in \mathbb{R}$$

for  $i_1, \dots, i_r, j_1, \dots, j_s \in \{1, \dots, \dim V\}$ . These are known as the **components of  $T$**  with respect to the chosen basis.

<sup>12</sup>Note we start the index at 0 because it relates nicely to the order of the term. Obviously it's just an index and we can call it whatever we like so this is fine.

Remark 3.8.1. It is not actually necessary that you take the dual basis above in order to define the components of a tensor. Any basis for the dual space will do, but that is almost never done.

Knowing the components (and corresponding basis) of a tensor, we can reconstruct the entire tensor.

Example 3.8.2. Let  $T$  be a  $(1, 1)$ -tensor. Then  $T^i_j := T(\epsilon^i, e_j)$ , then, using the multilinearity of the tensor, we have

$$T(\varphi, v) = T\left(\sum_{i=1}^{\dim V} \varphi_i \epsilon^i, \sum_{j=1}^{\dim V} v^j e_j\right) = \sum_{i,j=1}^{\dim V} \varphi_i v^j T(\epsilon^i, e_j) = \sum_{i,j=1}^{\dim V} \varphi_i v^j T^i_j.$$

Terminology. We call a raised index a *contravariant* index and a lower index a *covariant* index and we stick with the convention that a vector has components with contravariant indices and basis elements with covariant indices and covectors the other way around. This makes contact with what the naming of the contravariant/covariant order of a tensor given previously.

In order to not always have the summation symbols everywhere Einstein came up with a clever notation, known as **Einstein summation convention**, whereby any repeated indices where one is up and one is down is implicitly summed over. For example

$$v^i e_i = \sum_i v^i e_i.$$

This notation instantly tells us that any time the same index label appears more than twice, we are dealing with something ill-defined. Equally it is not good to have the same index appearing both up or both down. That is we don't want things like

$$T^i_j S^i_i \quad \text{or} \quad v^i w^i.$$

With the Einstein summation convention, one can only add two terms that have the same indices in the same places, so

$$T^{ij}_k + S^{ij}_k = R^{ij}_k$$

is good, but

$$T^{ij}_k + S^i_{jk} = R_{ij}^k$$

is not well defined. In note of Remark 3.5.5, it is also not well defined to write things like

$$T^{ij}_k + S_k^{ij},$$

*unless* some sort of property is given, e.g.  $S_k^{ij} = 3S^{ij}_k$ .

Remark 3.8.3. Note that we can only really use the Einstein summation convention because we are considering multilinear maps. That is we could have

$$\varphi(v^1 e_i) = \varphi\left(\sum_i v^i e_i\right) \quad \text{or} \quad \varphi(v^1 e_i) = \sum_i \varphi(v^i e_i),$$

and it is only because the multilinearity equates the two that we're OK.

## 4 | Differentiable Manifolds

So far we have looked at topological manifolds, which allowed us to talk about the continuity of a curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$ . If we are to (and we will) associate the motion of a particle in spacetime as a curve on a manifold, we want more than just continuity; we want to be able to associate a *velocity* to each point of the curve. Roughly speaking, we think of the velocity of a curve as a tangent vector to the curve which we obtain by differentiating the curve.

The question is, then, is the structure we already have on our  $d$ -dimensional topological manifold  $(\mathcal{M}, \mathcal{O})$  sufficient in order to talk about differentiability or do we need some more structure? The short answer is ‘no, you need more structure.’ In this lecture we will show this, while also working out *what* extra structure we need in order to talk about differentiability of curves.

In fact, we wish to define a notion of differentiable for more than just curves, we wish to define it also for: functions,  $f : \mathcal{M} \rightarrow \mathbb{R}$ , and maps,  $\varphi : \mathcal{M} \rightarrow \mathcal{N}$ .

### 4.1 Strategy

Let’s first consider curves,  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$ . Recall in lecture 2 we said that we can assign properties to manifolds by considering the chart representatives, which were maps  $x \circ \gamma : \mathbb{R} \rightarrow U \subseteq \mathbb{R}^d$ . We already have a notion of differentiability of such curves from undergraduate courses, and so we seek to use this in order to define what we mean for the curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  to be differentiable.

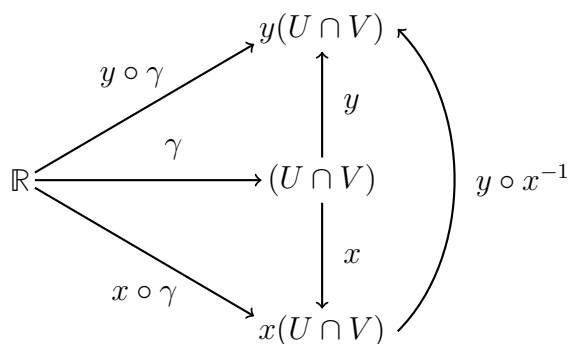
Let’s consider the part of our curve that lies in the chart domain  $\gamma : \mathbb{R} \rightarrow U$ .<sup>1</sup> Once we have worked out differentiability here, we can extend it to a global notion for the whole curve. Recall that if we are going to do this ‘lifting’ of notion to the manifold level, we have to make sure the lifted notion is chart independent, i.e. it doesn’t matter which chart we use, we always get the same result. When we encountered this before we were fine, because we knew that the composition of continuous maps is continuous and so our chart transition maps were continuous. However, the continuity of the chart transition map does *not* guarantee their undergraduate differentiability as there could be a sudden turning point in the curve.

At first sight, our strategy doesn’t work out. However, there is a remedy to this that we hinted at in lecture 2, it is the content of the next section.

---

<sup>1</sup>We should really rename it something like  $\gamma_U$ , but we don’t want to clutter notation too much, so shall just call it  $\gamma$ .





**Figure 4.1:** Two charts  $(U, x)$  and  $(V, y)$  used to represent a physical curve  $\gamma$ . It is assumed that one knows the map  $x \circ \gamma$  is so called ‘undergraduate differentiable’. However, one can not yet conclude whether  $\gamma$  itself is differentiable as the continuity of the chart transition map  $y \circ x^{-1}$  does not guarantee differentiability (there could be a sudden turning point in the curve representation).

## 4.2 Compatible Charts

The problem mentioned above stems from the fact that we took  $(U, x)$  and  $(V, y)$  to be *any* charts for our topological manifold  $(\mathcal{M}, \mathcal{O})$ . To emphasise this, we can say that we took them from the maximal atlas  $\mathcal{A}_{max}$ . If instead we had insisted that our charts come from a smaller atlas that we knew contained no overlapping charts with only continuous (and not differentiable) charts, we could solve our problem. In other words, we ‘tear out’ any pages of our atlas that correspond to chart transition functions that are not differentiable. Now we are not guaranteed that after doing this we are left with an atlas, as we may no longer cover the whole space, but if we do, then we stand a much better chance at defining what we mean by the differentiability of a curve. So we consider a *restricted atlas*.

Note this is a *huge* choice to make. By doing this, anything we want to talk about later on that relies on differentiability of the curve can *only* be discussed in a chart if that chart comes from our restricted atlas.

**Definition** (Compatible Charts). Two charts  $(U, x)$  and  $(V, y)$  of a topological manifold  $(\mathcal{M}, \mathcal{O})$  are called  $\square$ -**compatible**<sup>2</sup> if either

- (a)  $U \cap V = \emptyset$ , or
- (b)  $U \cap V \neq \emptyset$  and the chart transition maps  $(x \circ y^{-1}) : y(U \cap V) \rightarrow x(U \cap V)$  and  $(y \circ x^{-1}) : x(U \cap V) \rightarrow y(U \cap V)$  are ‘undergrad’  $\square$ .<sup>3</sup>

**Definition** (Compatible/Restricted Atlas). An atlas  $\mathcal{A}_{\square}$  is a  $\square$ -**compatible atlas** if all of its charts are  $\square$ -compatible.

**Definition.** A  $\square$ -manifold is the triple  $(\mathcal{M}, \mathcal{O}, \mathcal{A}_{\square})$ .

<sup>2</sup>I tried to get a flower like Dr. Schuller uses, but overleaf was not having it. Apologies for me meagre  $\square$ .

<sup>3</sup>That is they have the property  $\square$  as maps from  $\mathbb{R}^d \rightarrow \mathbb{R}^d$  that we know from undergraduate courses.

Now it might be possible that two separate criteria for ‘tearing our pages’ in order to obtain a  $\square$ -compatible atlas exist. That is there might be more than one atlas that is  $\square$ -compatible. In this case, we have to make a choice about which one to use, but must always remember that we have made this choice, as one of these atlases might allow a different property, say  $\blacksquare$ , to be defined, whereas the other might not. Physically, this is obviously a very important thing to keep in mind, as  $\square$  and  $\blacksquare$  are physical properties of the curve and therefore don’t depend at all on what charts or atlases we choose to use, so we must make sure we pick the ones that match up to the physics.

Before moving on, let’s consider the types of things  $\square$  can be.

$\square$	Undergraduate $\square$
$C^0$	$C^0(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , continuous w.r.t. the standard topology on $\mathbb{R}^d$ .
$C^1$	$C^1(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , once differentiable with continuous result w.r.t. the standard topology on $\mathbb{R}^d$ .
$C^k$	$C^k(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , $k$ -times continuously differentiable.
$D^k$	$D^k(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , $k$ -times differentiable, don’t need to be continuous.
$C^\infty$	$C^\infty(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , infinitely differentiable with continuous result, known as <b>smooth</b> .
$C^\omega$	$C^\omega(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ , analytic functions (can be Taylor expanded).
$\mathbb{C}^\infty$	$\mathbb{C}^\infty(\mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n})$ , $\dim \mathcal{M} = 2n$ for integer $n$ , they satisfy the Cauchy-Riemann equations pairwise. This gives us a <i>complex</i> manifold.

**Theorem 4.2.1** (Whitney Theorem). *Any  $C^k$ -atlas  $\mathcal{A}_{C^k}$  for  $k \geq 1$  for a topological manifold, contains as a subatlas a  $C^\infty$  atlas.*

We shall not prove this theorem, but simply give an motivation for it via the following example.

*Example 4.2.2.* Say we were only interested in curves that were  $C^2$ , so the third derivative was discontinuous. In order to talk about such curves we would need a  $C^2$ -atlas. However, any function that is  $C^3$  is also  $C^2$  and so we would still be able to talk about these curves on a  $C^3$ -atlas as  $C^3 \circ C^2 = C^2$ , roughly speaking. That is, if we insist that our transition functions are  $C^3(\mathbb{R}^d \rightarrow \mathbb{R}^d)$  then we can still obtain a well defined notion for the curve being  $C^2$ . Repeating this again, it follows that we could use a  $C^\infty$ -atlas in order to describe our curves.

Note that the theorem does not say that we can turn a  $C^2$  curve into a  $C^\infty$  curve, only that we can talk about it on both a  $C^2$ -atlas and a  $C^\infty$ -atlas.

This is a very useful theorem for physics, because we now don’t need to worry about ‘how many derivatives should I be worried about ensuring?’ Just make sure it’s at least  $C^1$  and then take the subatlas. Thus, we may, without loss of generality, always consider  $C^\infty$ -manifolds, or **smooth** manifolds (unless we wish to define Taylor expandability or complex differentiability, etc).

### 4.3 Diffeomorphisms

Recall that whenever we introduce a new structure to our objects that it is always worth studying the structure preserving maps. These maps are generally known as **isomorphisms**. For two sets, the isomorphism is a bijection. For topological spaces we saw that the isomorphism is a continuous, bijection whose inverse is also continuous, we called these *homeomorphisms*. Two objects that are related by an isomorphism are said to be **isomorphic**.

**Definition** (Smooth Maps). Let  $(\mathcal{M}, \mathcal{O}_{\mathcal{M}}, \mathcal{A}_{\mathcal{M}})$  and  $(\mathcal{N}, \mathcal{O}_{\mathcal{N}}, \mathcal{A}_{\mathcal{N}})$  be two smooth manifolds of dimension  $m$  and  $n$ , respectively. A map  $\varphi : \mathcal{M} \rightarrow \mathcal{N}$  is said to be  $C^\infty$  (or **smooth**) if the map  $y \circ \varphi \circ x^{-1}$  is undergrad  $C^\infty$  for charts  $(U, x) \in \mathcal{A}_{\mathcal{M}}$  and  $(V, y) \in \mathcal{A}_{\mathcal{N}}$ .

$$\begin{array}{ccc}
 U & \xrightarrow{\varphi} & V \\
 x \downarrow & & \downarrow y \\
 \mathbb{R}^m & \xrightarrow{\quad} & \mathbb{R}^n \\
 & y \circ \varphi \circ x^{-1} &
 \end{array}$$

*Remark 4.3.1.* Note that we just gave the definition above for two charts  $(U, x)$  and  $(V, y)$ . We already know that if it holds for one chart it will hold for all because our manifolds are smooth and therefore switching charts is  $C^\infty$ . That is we have the following diagram:

$$\begin{array}{ccc}
 & \tilde{y} \circ \varphi \circ \tilde{x}^{-1} & \\
 & \mathbb{R}^m \xrightarrow{\quad} \mathbb{R}^n & \\
 \tilde{x} \uparrow & & \uparrow \tilde{y} \\
 C^\infty \curvearrowleft & U \xrightarrow{\varphi} V & \curvearrowright C^\infty \\
 & x \downarrow & \downarrow y \\
 & \mathbb{R}^m \xrightarrow{\quad} \mathbb{R}^n & \\
 & y \circ \varphi \circ x^{-1} &
 \end{array}$$

**Definition** (Diffeomorphism). Let  $(\mathcal{M}, \mathcal{O}_{\mathcal{M}}, \mathcal{A}_{\mathcal{M}})$  and  $(\mathcal{N}, \mathcal{O}_{\mathcal{N}}, \mathcal{A}_{\mathcal{N}})$  be two smooth manifolds. They are isomorphic if there exists a bijection  $\varphi : \mathcal{M} \rightarrow \mathcal{N}$  such that  $\varphi$  and  $\varphi^{-1}$  are  $C^\infty$  maps. Such a map is known as a **diffeomorphism** and the manifolds are said to be **diffeomorphic**.

We can think of diffeomorphisms as relating two surfaces that can be ‘moulded’ into each other without cutting/tearing/folding the surface. For example, the surface of a sphere as a differential manifold is diffeomorphic to the surface of a potato,<sup>4</sup> but it is not diffeomorphic

<sup>4</sup>Provided there’s not sharp edges and/or holes in the potato.

to a doughnut. In other words, at the smooth manifold level, things don't have a shape yet, but are made out of sort of fluidy-substance.<sup>5</sup>

**Theorem 4.3.2.** *The number of  $C^\infty$ -manifolds one can make from a given  $C^0$ -manifold (if any), up to diffeomorphism<sup>6</sup> is given by the following table:*

$\dim \mathcal{M}$	<i>No. <math>C^\infty</math>-manifolds</i>
<i>1</i>	<i>1</i>
<i>2</i>	<i>1</i>
<i>3</i>	<i>1</i>
<i>4</i>	<i>uncountably infinitely many</i>
<i>5</i>	<i>finitely many</i>
<i>6</i>	<i>finitely many</i>
<i>7</i>	<i>finitely many</i>
$\vdots$	$\vdots$

The first three results in the table are the so-called *Moise-Radon* theorems, and the 5,6,7,... results are shown using an area of topology known as *surgery theory*. Unfortunately as physicists, we are most interested in  $\dim \mathcal{M} = 4$  for spacetime. Ahh, Sod's law!

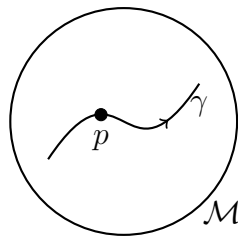
---

<sup>5</sup>This is not a technical term, please do not quote me on it.

<sup>6</sup>That is, any two  $C^\infty$ -manifolds that are diffeomorphic count as the same one.

# 5 | Tangent Spaces

The aim of the lecture is going to be answer the following question: what is the velocity of a curve  $\gamma$  at a point  $p \in \mathcal{M}$ ?



In doing this, we first want to completely forget everything we already know about what we mean by ‘velocity’. We are going to rediscover what it means during this lecture.

## 5.1 Velocities

**Definition** (Scalar Fields). The vector space with set

$$C^\infty(\mathcal{M}) := \{f : \mathcal{M} \rightarrow \mathbb{R} \mid f \text{ is a smooth function}\}$$

equipped with point-wise addition  $(f \oplus g)(p) = f(p) + g(p)$  and s-multiplication  $(\lambda \odot f)(p) = \lambda \cdot f(p)$ , is known as the space of **scalar fields** (or smooth functions) on  $\mathcal{M}$ .

*Example 5.1.1.* An example of a smooth function on  $\mathcal{M}$  is a temperature distribution. To each point in the room (which is  $\mathcal{M}$ ) we associate a real number, the temperature of that point.

*Remark 5.1.2.* We should actually be a little careful with the terminology above. A smooth function is defined for any two manifolds of arbitrary dimension, provided the map is smooth obviously. A scalar field is strictly a map to a one-dimensional manifold, in this case the real numbers  $\mathbb{R}$ . The notation  $C^\infty(\mathcal{M})$  means that we are considering the map to the reals. We would indicate a general smooth function more explicitly as  $C^\infty(\mathcal{M}, \mathcal{N})$  or something.

**Definition.** Consider a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  and a curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  that is at least  $C^1$ . Suppose  $\gamma(\lambda_0) = p \in \mathcal{M}$ . The **velocity** of  $\gamma$  at  $p$  is the *linear map*

$$v_{\gamma,p} : C^\infty(\mathcal{M}) \xrightarrow{\sim} \mathbb{R},$$

defined by

$$v_{\gamma,p}(f) := (f \circ \gamma)'(\lambda_0).$$

The intuition here is that as you run along in the world (i.e. move along  $\gamma$ ) you ask how something (the scalar field) changes in your direction of motion. So you take the directional derivative of the scalar field. We will make this a lot more concrete shortly.

## 5.2 Tangent Vector Space

**Definition** (Tangent Vector Space). For each point  $p \in \mathcal{M}$  we define the vector space, known as the **tangent (vector) space** to  $\mathcal{M}$  at  $p$ , whose set is

$$T_p\mathcal{M} := \{v_{\gamma,p} \mid \gamma \text{ smooth curve through } p\},$$

and whose addition and s-multiplication is given by

$$\begin{aligned} (v_{\gamma,p} \oplus v_{\delta,p})(f) &:= v_{\gamma,p}(f) + v_{\delta,p}(f), \\ (\alpha \odot v_{\gamma,p})(f) &:= \alpha \cdot v_{\gamma,p}(f). \end{aligned}$$

We need to show that the right-hand sides of the last two expressions do indeed lie in  $T_p\mathcal{M}$ . That is, we need to show that

- (i) There exists a  $\tau : \mathbb{R} \rightarrow \mathcal{M}$  such that  $\alpha \odot v_{\gamma,p} = v_{\tau,p}$ , and
- (ii) There exists a  $\sigma : \mathbb{R} \rightarrow \mathcal{M}$  such that  $v_{\gamma,p} \oplus v_{\delta,p} = v_{\sigma,p}$ .

*Proof.* It is clear that both the right-hand side expressions will be elements of  $\text{Hom}(C^\infty(\mathcal{M}), \mathbb{R})$ , but we need to check that they are velocities to some curves through  $p$ . Let's consider them in turn

- (i) Let  $\lambda_0 \in \mathbb{R}$  such that  $\gamma(\lambda_0) = p$ . Construct the curve  $\tau : \mathbb{R} \rightarrow \mathcal{M}$  by

$$\tau(\lambda) := \gamma(\alpha\lambda + \lambda_0) = (\gamma \circ \mu_\alpha)(\lambda),$$

where  $\mu_\alpha : \mathbb{R} \rightarrow \mathbb{R}$  defined by  $\mu_\alpha(\lambda) := \alpha\lambda + \lambda_0$ . We claim this curve satisfies our condition.

First note that  $\tau(0) = \gamma(\lambda_0) = p$  and so it passes through the point, which we need. Then

$$\begin{aligned} v_{\tau,p}(f) &:= (f \circ \tau)'(0) \\ &= (f \circ \gamma \circ \mu_\alpha)'(0) \\ &= \alpha \cdot (f \circ \gamma)'(\lambda_0) \\ &=: \alpha \cdot v_{\gamma,p}(f), \end{aligned}$$

where we have used the multidimensional chain rule to go from the second to third line along with  $\mu_\alpha(0) = \lambda_0$  and  $\mu_\alpha'(0) = \alpha$ . Since this holds for any  $f \in C^\infty(\mathcal{M})$ , we get the result.

- (ii) This is slightly more involved. In order to show it, we shall introduce a chart  $(U, x)$ . However, as we have explained already, it is important that this choice of chart plays no vital role in the result; that is the result must be chart independent, so we will have

to check this at the end. Again let  $\lambda_0 \in \mathbb{R}$  such that  $\gamma(\lambda_0) = p$ . Similarly, let  $\lambda_1 \in \mathbb{R}$  such that  $\delta(\lambda_1) = p$ .

Construct the curve  $\sigma_x : \mathbb{R} \rightarrow \mathcal{M}$ , where the subscript reminds us that we are working in a chart, by

$$\sigma_x(\lambda) := x^{-1}((x \circ \gamma)(\lambda_0 + \lambda) + (x \circ \delta)(\lambda_1 + \lambda) - (x \circ \gamma)(\lambda_0)).$$

Again we claim this curve satisfies our condition. First, we need to check it goes through the point  $p$ , and a quick calculation shows that  $\sigma_x(0) = p$ , so we can proceed. We have

$$\begin{aligned} v_{\sigma_x, p}(f) &:= (f \circ \sigma_x)'(0) \\ &= (f \circ x^{-1} \circ x \circ \sigma_x)'(0) \end{aligned}$$

Now we have  $(f \circ x^{-1}) : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $(x \circ \sigma_x) : \mathbb{R} \rightarrow \mathbb{R}^d$  and so we use the multidimensional chain rule for the derivative. We have<sup>1</sup>

$$\begin{aligned} v_{\sigma_x, p}(f) &:= (x^i \circ \sigma_x)'(0) \cdot \partial_i (f \circ x^{-1}) \Big|_{(x \circ \sigma_x)(0)} \\ &= ((x^i \circ \gamma)'(\lambda_0) + (x^i \circ \delta)'(\lambda_1)) \cdot \partial_i (f \circ x^{-1}) \Big|_{x(p)} \\ &= (f \circ \gamma)'(\lambda_0) + (f \circ \delta)'(\lambda_1) \\ &=: v_{\gamma, p}(f) + v_{\delta, p}(f), \end{aligned}$$

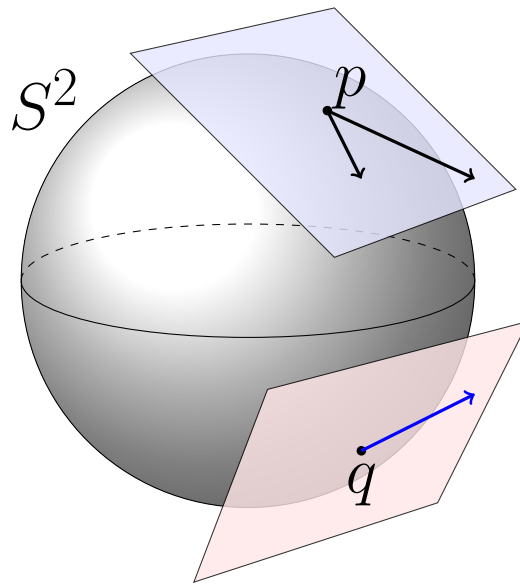
where we used the ‘evaluated at’ notation  $|$  in order to reduce potential confusion, and where to get to the penultimate line we did the multidimensional chain rule in reverse (i.e. we did the steps up to that point backwards but not with  $\gamma$  and  $\delta$ ). The final line make no reference to the chart  $(U, x)$  and so we know we can use *any* chart in our atlas to do this and so the result is chart independent. Finally, again since this holds for a general  $f \in C^\infty(\mathcal{M})$  we get the result. □

*Remark 5.2.1.* Dr. Schuller gives some nice picture descriptions of the above proofs in his lecture (starting about 39:00), these are worth looking at. I have not drawn them here as they will be a reasonable amount of work (especially the (ii) property) in Tikz, and I’m feeling too lazy for that, but they really are worth seeing, so go look at them if you haven’t already!

It is important to note that in all of the above we are always considering the same point  $p \in \mathcal{M}$ . It does not make sense to add two velocities that are the tangents at different points, i.e.  $v_{\gamma, p} \oplus v_{\delta, q}$  only makes sense when  $p = q$ . One way to remember this is to think about the velocities being little arrows in planes tangent to the manifold. For example if  $\mathcal{M} = S^2$ , the 2-sphere,<sup>2</sup> then we have something like Figure 5.1. Thought about this way, it becomes clear why we can’t add velocities that are tangent to different points: they live on completely separate  $\mathbb{R}^2$  planes, so it doesn’t make sense to add them. You might think ‘well can’t we just put the velocity at  $q$  onto the tangent plane at  $p$ ?’ The proper answer to this question comes later, but the short answer is ‘only if we take into consideration the so-called intrinsic curvature of the manifold’.

<sup>1</sup>We use an index  $i$  to denote which element in  $\mathbb{R}^d$  we are considering.  $\partial_i$  obviously means the derivative w.r.t. the  $i^{\text{th}}$  element.

<sup>2</sup>For those unfamiliar, a 2-sphere is what we think of as the surface of a 3d ball. The surface is 2-dimensional and so we call it the 2-sphere.



**Figure 5.1:** *Tangent planes at two points  $p, q \in \mathcal{M} = S^2$ . The arrows are the tangent velocities to curves (not drawn) on the manifold. The two velocities at  $p$  (black arrows) can be added because they live in the same tangent space, but it does not make sense to add one of them to the velocity at  $q$  blue arrow.*

*Remark 5.2.2.* As the above discussion highlights, we often think of the velocities as being little arrows that lie tangent to our curves and ‘point out’ of the manifold. In order to do this, we obviously need to first *embed* our manifold into a higher dimensional space (so we look at the 2-sphere in  $\mathbb{R}^3$ ). However, as soon as we start considering manifolds of dimension  $d \geq 3$  then we have a problem: we need to picture a at least 4-dimensional space to embed in, and I can’t see 4D spaces.<sup>3</sup> Besides that obstacle, when we start talking about the universe, if we embed it into something we then are talking about things that lie outside the universe, which is a rabbit whole we do not want to go down.<sup>4</sup>

Luckily, our formulation of what a velocity is made no reference whatsoever to some higher dimensional embedding space. It was defined *intrinsically* to the manifold itself. This seems promising, but we need to make sure that the two ideas coincide with each other. The answer is that they do and so we can choose how we want to think about our tangent vectors on a case by case basis: taking the embedding idea when we can for some nice intuition, and using the intrinsic definition when we are dealing with things too hard to imagine.

<sup>3</sup>If you can, props!

<sup>4</sup>I’ll take the blue pill, Morpheus.



### 5.3 Components of a Vector w.r.t. a Chart

Let  $(U, x)$  be a chart of a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  and let  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  be a curve that passes through point  $p \in U$  as  $\gamma(0) = p$ . Now we have the calculation

$$\begin{aligned} v_{\gamma,p}(f) &:= (f \circ \gamma)'(0) \\ &= (f \circ x^{-1} \circ x \circ \gamma)'(0) \\ &= (x^i \circ \gamma)'(0) \cdot \partial_i(f \circ x^{-1})|_{x(p)} \end{aligned}$$

The first thing to note, as we touched on before, is that the index on  $\partial_i$  tells us which *entry* to derive by. That is it make no reference whatsoever to  $x$ , but simply says ‘what ever the  $i^{\text{th}}$  entry is, derive by that.’<sup>5</sup> Now this is a lot of writing and so we introduce some new notation in order to simplify it: we define

$$\left( \frac{\partial f}{\partial x^i} \right)_p := \partial_i(f \circ x^{-1})|_{x(p)}, \quad \text{and} \quad \dot{\gamma}_x^i(0) := (x^i \circ \gamma)'(0),$$

where the colours are just used to show that the terms appear on both sides. The first thing we have to point out is that this is *just notation*. The first term looks an awful lot like a partial derivative, however strictly it is something completely different; it is just notation for the right-hand side. Obviously this notation is not done by accident and it will turn out that it will possess all of the properties we’d want from a partial derivative, but that still doesn’t make it one.

Given the above, we can write down the velocity to a curve at point  $p$  in the following form

$$v_{\gamma,p}(f) = \dot{\gamma}_x^i(0) \cdot \left( \frac{\partial}{\partial x^i} \right)_p (f),$$

or, as a *map*, we can write

$$v_{\gamma,p} = \dot{\gamma}_x^i(0) \cdot \left( \frac{\partial}{\partial x^i} \right)_p.$$

**Definition** (Components of a Vector w.r.t. a Chart). We call  $\dot{\gamma}_x^i(0)$  the  $i^{\text{th}}$  **component** of the velocity vector at point  $p \in \mathcal{M}$  w.r.t. the chart  $(U, x)$ .

**Definition** (Basis Elements of  $T_p U$ ). We call  $\left( \frac{\partial}{\partial x^i} \right)_p$  the  $i^{\text{th}}$  **basis element** of  $T_p U$  w.r.t. which the components need to be understood.

Note that in the above, we only have a basis element for  $T_p U$ , not  $T_p \mathcal{M}$  as the chart is only defined for the subset  $U$ .

**Corollary 5.3.1.** *The action of a basis element on the  $j^{\text{th}}$  coordinate function  $x^j$  satisfies*<sup>6</sup>

$$\left( \frac{\partial}{\partial x^i} \right)_p (x^j) = \delta_i^j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

<sup>5</sup>This is analogous to the fact that given  $f : \mathbb{R} \rightarrow \mathbb{R}$  we define  $f' : \mathbb{R} \rightarrow \mathbb{R}$  completely independently of what variable we’re using. So  $f' = \frac{df}{dx}$  is not a general expression, but is a notation choice once we have decided that  $x$  is our variable.

<sup>6</sup>Note we have used the angle bracket here. This makes sense as  $x^j : U \subseteq \mathcal{M} \rightarrow \mathbb{R}$  is  $C^\infty$  (as its a smooth manifold) and  $\left( \frac{\partial}{\partial x^i} \right)_p \in T_p \mathcal{M}$  is a vector. This highlights the benefit of using this notation.

*Proof.* Use the fact that  $x^j \circ x^{-1}$  only gives us the  $j^{\text{th}}$  entry of  $x(p)$ . Obviously, then, if we try to differentiate w.r.t. any of entries we get 0 (as the entry is already 0), but if we differentiate w.r.t. this entry we get 1. This is just  $\delta_i^j$ .  $\square$

#### 5.4 Chart Induced Basis

**Theorem 5.4.1.** *Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be a  $d$ -dimensional smooth manifold. The set*

$$\left\{ \left( \frac{\partial}{\partial x^1} \right)_p, \dots, \left( \frac{\partial}{\partial x^d} \right)_p \right\}$$

*constitutes a basis for the tangent space  $T_p U$ , and it's known as the **chart induced basis**.*

*Proof.* We have already known that they generate  $T_p U$  as any vector in  $T_p U$  can be written in terms of them. All that remains to be shown is that they are linearly independent, that is we require that

$$\sum_{i=1}^d \lambda^i \left( \frac{\partial}{\partial x^i} \right)_p = 0 \quad \implies \quad \lambda^i = 0 \quad \forall i.$$

Consider the action on the  $j^{\text{th}}$  coordinate function,  $x^j$ . We have

$$\sum_{i=1}^d \lambda^i \left( \frac{\partial}{\partial x^i} \right)_p (x^j) = \sum_{i=1}^d \lambda^i \delta_i^j = \lambda^j$$

and so we get the result.  $\square$

**Corollary 5.4.2.** *The dimension of the tangent space is equal to the dimension of the manifold*

$$\dim T_p \mathcal{M} = \dim \mathcal{M}.$$

*Proof.* This just follows from the fact that there are  $d$ -basis elements for  $T_p U$  for all the chart domains and the fact that the  $d$  came from the dimension of  $\mathcal{M}$ .  $\square$

#### 5.5 Change of Vector Components Under Change of Chart

One often comes across statements like ‘a vector transforms as [insert equation] under a change of chart’. However, we know that this statement is not complete as vectors (and also tensors) are abstract objects that are completely independent of the charts. The velocity of the bird is the velocity of the bird. So the only thing we could insert into the statement is ‘they don’t transform’, but this in itself is not a super useful for calculations. A better, and much more useful, statement is ‘the *components*<sup>7</sup> of a vector transforms as [insert equation] under a change of chart’.

Let  $(U, x)$  and  $(V, y)$  be overlapping charts for a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  and  $p \in U \cap V$ . If  $X \in T_p \mathcal{M}$  then we can decompose it in either chart,

$$X = X_{(x)}^i \left( \frac{\partial}{\partial x^i} \right)_p \quad \text{and} \quad X = X_{(y)}^i \left( \frac{\partial}{\partial y^i} \right)_p$$

<sup>7</sup>The components of a vector are simply given in relation to the basis, see Remark 3.7.2.

To study how these relate, consider the following

$$\begin{aligned}
 \left(\frac{\partial}{\partial x^i}\right)_p(f) &:= \partial_i(f \circ x^{-1})|_{x(p)} \\
 &= (f \circ y^{-1} \circ y \circ x^{-1})|_{x(p)} \\
 &= \partial_i(y^j \circ x^{-1})|_{x(p)} \cdot \partial_j(f \circ y^{-1})|_{y(p)} \\
 &= \left(\frac{\partial y^j}{\partial x^i}\right)_p \cdot \left(\frac{\partial f}{\partial y^j}\right)_p \\
 \implies \left(\frac{\partial}{\partial x^i}\right)_p &= \left(\frac{\partial y^j}{\partial x^i}\right)_p \left(\frac{\partial}{\partial y^j}\right)_p
 \end{aligned}$$

Inserting this into the fact that  $X$  can be expressed in either basis, we have

$$\begin{aligned}
 X_{(y)}^j \left(\frac{\partial}{\partial y^j}\right)_p &= X_{(x)}^i \left(\frac{\partial}{\partial x^i}\right)_p \\
 &= X_{(x)}^i \left(\frac{\partial y^j}{\partial x^i}\right)_p \left(\frac{\partial}{\partial y^j}\right)_p \\
 \implies X_{(y)}^j &= X_{(x)}^i \left(\frac{\partial y^j}{\partial x^i}\right)_p,
 \end{aligned}$$

where to get to the last line we have used the fact that  $\left(\frac{\partial}{\partial y^j}\right)_p$  is a basis and so the coefficients must be equal.

It is important that we evaluate the derivative at the point  $p \in \mathcal{M}$  as we did not say that our transformation needed to be linear. Indeed the transformation can be wildly nonlinear (provided the expression still makes sense), but once we evaluate this result at a point we are just left with a number, which is exactly what we want.

*Remark 5.5.1.* In special relativity, one often hears people talking about Minkowski *vector space*, i.e. the vector space whose set is made up of the positions  $x^\mu$ . This goes against what we said at the start of lecture 3: "We wish to emphasise here that we will *not* equip space(time) with a vector space structure." A counter would be 'but the coordinate transformations work!', however the transformations considered in special relativity are not general transformations: we restrict ourselves to linear transformations, which we further restrict to be Lorentz transformations. This seems like a reasonable thing to do, but we should be able to study special relativity in polar coordinates if we want to.<sup>8</sup> We can make such a transformation (Cartesian to polar) and the *velocities* at a point will change via linear maps as described above, but the position space will not transform linearly! In other words, it is an over structuralisation to equip Minkowski space with a vector space structure, as in doing so we must restrict ourselves to Lorentz transformations. This just highlights again that the positions are *not* vectors, it is the *velocities* that are the vectors.

## 5.6 Cotangent Spaces

We have constructed the tangent space as a vector space, but our work from the lecture 3 tells us that we can take the dual to this space.

<sup>8</sup>As, once again, the choice of chart/coordinates has no impact whatsoever on the real world physics.

**Definition** (Cotangent Space). Let  $T_p\mathcal{M}$  be the tangent space to some point  $p \in \mathcal{M}$ . The dual of this space is known as the **cotangent space**

$$T_p^*\mathcal{M} \equiv (T_p\mathcal{M})^* := \{\varphi : T_p\mathcal{M} \xrightarrow{\sim} \mathbb{R}\}.$$

**Definition** (Gradient of  $f$  at  $p$ ). Let  $f \in C^\infty(\mathcal{M})$ . Then we can define the linear map

$$(df)_p : T_p\mathcal{M} \xrightarrow{\sim} \mathbb{R}, \quad (df)_p X := X(f).$$

Clearly this makes  $(df)_p$  an element of the cotangent space. It is known as the **gradient** of  $f$  at point  $p \in \mathcal{M}$ .

*Remark 5.6.1.* Note that we do not need to use a chart in order to define the gradient, as one might think we would from undergraduate classes.

The gradient is a  $(0,1)$ -tensor over the vector space  $T_p\mathcal{M}$  and so we can find its components w.r.t. the chart induced basis using the method discussed previously:

$$((df)_p)_j := (df)_p \left( \left( \frac{\partial}{\partial x^j} \right)_p \right) = \left( \frac{\partial f}{\partial x^j} \right)_p = \partial_j (f \circ x^{-1})|_{x(p)}.$$

**Corollary 5.6.2.** *The chart induced basis for  $T_p^*\mathcal{M}$  is the set*

$$\{(dx^1)_p, \dots, (dx^d)_p\},$$

where  $x^i : U \rightarrow \mathbb{R}$  are the coordinate maps for the chart  $(U, x)$ .

*Proof.* By direct calculation we have

$$(dx^i)_p \left( \left( \frac{\partial}{\partial x^j} \right)_p \right) := \left( \frac{\partial x^i}{\partial x^j} \right)_p = \delta_j^i,$$

which is the dual basis of dual space condition. □

## 5.7 Change of Components of a Covector Under a Change of Chart

Just as with the vector above, the covector itself remains invariant under a change of charts (its a tensor!), but the *components* change under a change of chart. Proceeding analogously to the vector component calculation we get: if  $\omega \in T_p^*\mathcal{M}$  and  $(U, x)$  and  $(V, y)$  are the two charts, then

$$\omega_{(y)j} = \left( \frac{\partial x^i}{\partial y^j} \right)_p \omega_{(x)i}.$$

Note that here the fraction is flipped in comparison to the vector components (i.e. the  $x$  and  $y$  have changed places). This reflects the fact that  $\omega$  is a covector and so its components transform inversely to the components of a vector. This highlights an important point that we have hinted at a few times: we must not think of the gradient as a vector. It is a covector

and its transformation properties prove it. If you need extra convincing, if it was a vector, we would expect it to transform under the chain rule, but that would give us

$$((df)_p)_{(x)i} = \left( \frac{\partial f}{\partial x^i} \right)_p = \left( \frac{\partial y^j}{\partial x^i} \right)_p \left( \frac{\partial f}{\partial y^j} \right)_p = \left( \frac{\partial y^j}{\partial x^i} \right)_p ((df)_p)_{(y)j},$$

which is in contradiction to our result!

#### Exercise

Show that the above transformation property is true.

*Hint:* Write  $\omega = \omega_{(x)i}(dx^i)_p$

There is a general rule to check that your transformation properties are right: look at the left-hand side and look at the placement of indices (plural for the case of higher order tensors) and which basis labelling they correspond to ( $x$  or  $y$ ). Every lower index becomes a denominator index in your fraction and comes with the relevant basis label. You then write the component in the new coordinate (i.e.  $\omega_{(x)i}$ ) and then, as there was no  $x$  or  $i$  on the left-hand side, use the Einstein summation convention to remove it from the right-hand side by placing it in the numerator. When you consider higher order tensors, you just make sure you pair up the correct indices with each other: for example

$$T_{(x)}^{ij} = \left( \frac{\partial x^i}{\partial y^k} \right)_p \left( \frac{\partial x^j}{\partial y^\ell} \right)_p T_{(y)}^{k\ell},$$

where the first indices ( $i$  and  $k$ ) are paired and the second indices ( $j$  and  $\ell$ ) are paired. We will see this rule more generally when we consider the change of components of tensors shortly.

## 6 | Fields

So far we have discussed a single tangent space and vectors lying in it. What we now want to study are vector *fields*, which is essentially a vector for every point on the manifold. We need to give a proper technical way to introduce vector fields, as simply saying ‘imagine a vector at every point’ isn’t good enough (two people might imagine differently). The answer to doing this is known as the *theory of bundles*.

### 6.1 Bundles, Fibres and Sections

**Definition (Bundle).** A (smooth) **bundle** is a triple  $(E, \pi, \mathcal{M})$  where  $E$  and  $\mathcal{M}$  are smooth manifolds known as the *total space* and the *base space*, respectively.  $\pi : E \rightarrow \mathcal{M}$  is a smooth, surjective map, known as the *projection map*.

*Notation.* It is also common to denote a bundle in the following notation  $E \xrightarrow{\pi} \mathcal{M}$ . It is important to know, though, that the bundle is the complete triple and not just the map, as one might think using this notation.

**Definition (Fibre over  $p$ ).** Let  $(E, \pi, \mathcal{M})$  be a bundle. We define the **fibre over**  $p \in \mathcal{M}$  as  $\text{preim}_{\pi}(p)$ .

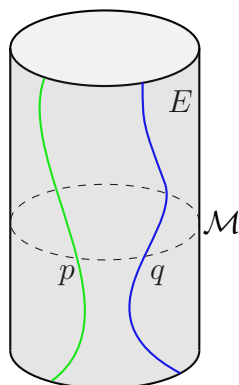
**Definition (Section).** A **section**,  $\sigma$ , of a bundle  $(E, \pi, \mathcal{M})$  is a map  $\sigma : \mathcal{M} \rightarrow E$  such that  $(\pi \circ \sigma) = \mathbb{1}_{\mathcal{M}}$ , the identity on  $\mathcal{M}$ .

As we shall see shortly, sections are the fields over our manifolds. The rough idea is that we make the fibres the tangent spaces to each point and then by taking a section, we pick one vector from each tangent space, giving us a vector field.

*Example 6.1.1.* In quantum mechanics, we are taught to think of the wavefunction as a function. This is technically not true. The wavefunction is a scalar field over the base space, and a scalar field is not a function (despite us maybe thinking it is). More technically, the wavefunction is a section over a  $\mathbb{C}$ -line bundle (that is a bundle whose fibres are the complex line). This is actually an important distinction when one comes to studying quantum mechanics in curved coordinates as the covariant derivative<sup>1</sup> acts in a non-trivial manner on sections.

---

<sup>1</sup>Which we will discuss later.



**Figure 6.1:** Example of a bundle and fibre. The total space,  $E$ , is the surface of the cylinder and the base space,  $\mathcal{M}$ , is the ring. The bundle is the triplet consisting of  $E$ ,  $M$  and a smooth, surjective projection map  $\pi : E \rightarrow M$ . The preimage of the the point  $p$  w.r.t. the projection map  $\pi$  is the green line — that is  $\pi$  maps every point on the green line to  $p$  — known as the fibre over  $p$ . Similarly the blue line is the fibre over  $q$ . The section w.r.t.  $p$ ,  $\sigma_p : M \rightarrow E$ , maps  $p$  to a point within its fibre (a point on the green line). A map  $\tau : M \rightarrow E$  which maps  $p$  to a point in  $q$ 's fibre (the blue line) is not a section, as  $(\pi \circ \tau)(p) = q \neq \mathbb{1}_M(p)$ . The complete section is the set of points formed by taking one point from each fibre.

## 6.2 Tangent Bundle of Smooth Manifold

Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be a smooth manifold. We define the **tangent bundle** as the bundle whose base space is our smooth manifold and whose total space has the set<sup>2</sup>

$$T\mathcal{M} := \dot{\bigcup}_{p \in \mathcal{M}} T_p\mathcal{M},$$

where the dot means ‘disjoint union’. The projection map is given by

$$\pi : X \mapsto p,$$

where  $p$  is the *unique* point such that  $X \in T_p\mathcal{M}$ .

We need to show how to turn the above set  $T\mathcal{M}$  into a smooth manifold (as we need for a bundle), but first two quick comments: it is important that we take the disjoint union above as this allows us to identify each vector with its base point  $p$ . It is because we take the disjoint union that we can say the *unique* point; and the projection is surjective as we took the union over all  $p \in \mathcal{M}$  and so we hit every element in  $\mathcal{M}$ .

We now need to make  $T\mathcal{M}$  into a smooth manifold, in such a way that  $\pi : T\mathcal{M} \rightarrow \mathcal{M}$  is a smooth map. So we need to define a topology on  $T\mathcal{M}$ , the question is ‘how do we do this?’ With a little thought the answer becomes clear: we already have a topology on our base space and if we are going to require  $\pi$  to be smooth, why don’t we just use the coarsest<sup>3</sup> topology

<sup>2</sup>We shall make this into a smooth manifold below.

<sup>3</sup>Recall coarsest means it has the least number of elements such that  $\pi$  is *just* continuous.

on  $T\mathcal{M}$  such that  $\pi$  is continuous (as continuity is needed for smoothness). This topology is known as the **initial topology w.r.t.  $\pi$** . It is defined simply as<sup>4</sup>

$$\mathcal{O}_{T\mathcal{M}} := \{\text{preim}_\pi(U) \mid U \in \mathcal{O}\}$$

So far we have a topological manifold  $(T\mathcal{M}, \mathcal{O}_{T\mathcal{M}})$  and a continuous map  $\pi$  to a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$ . We now need to define a  $C^\infty$ -atlas for  $(T\mathcal{M}, \mathcal{O}_{T\mathcal{M}})$  in such a way that our map becomes smooth. As with the topology, we are going to construct this atlas from  $\mathcal{A}$ .

The question is ‘how?’ Well we know that  $X \in T\mathcal{M}$  is described by two pieces of information: it is a vector and it has a base point. We can easily obtain the coordinates of the base point by simply projecting  $X$  down using  $\pi$  and then using our atlas on  $\mathcal{M}$  to find its coordinates. What about the vector part? Well, we have a chart on  $\mathcal{M}$  and so we can induce a chart on the tangent space and decompose  $X$  as

$$X =: X_{(x)}^i \frac{\partial}{\partial x^i}.$$

It is the components  $X_{(x)}^i$  that we want to use, but we want to get them by just using the chart  $(U, x)$ . The answer is very straight forward: just consider the gradient of the chart maps. That is

$$X_{(x)}^i = (dx^i)_{\pi(X)}(X).$$

So we construct the atlas

$$\mathcal{A}_{T\mathcal{M}} := \{(TU, \xi_x) \mid (U, x) \in \mathcal{A}\},$$

where

$$\xi_x : TU \rightarrow \mathbb{R}^{2 \cdot \dim \mathcal{M}},$$

given by

$$\xi_x(X) = \underbrace{\left( (x^1 \circ \pi)(X), \dots, (x^d \circ \pi)(X) \right)}_{(U,x)\text{-coordinate of } \pi(X)}, \underbrace{\left( (dx^1)_{\pi(X)}(X), \dots, (dx^d)_{\pi(X)}(X) \right)}_{\text{Vector components w.r.t. } (U,x)}$$

We also need the inverse map:

$$\xi_x^{-1} : \mathbb{R}^{2 \cdot \dim \mathcal{M}} \rightarrow TU.$$

With a bit of thought it is clear that it must satisfy

$$\xi_x^{-1}(\alpha^1, \dots, \alpha^d, \beta_1, \dots, \beta^d) := \beta^i \left( \frac{\partial}{\partial x^i} \right)_{x^{-1}(\alpha^1, \dots, \alpha^d)} = \beta^i \left( \frac{\partial}{\partial x^i} \right)_{\pi(X)}.$$

Now we need to check that these maps are smooth (as we need a smooth atlas). Consider

---

<sup>4</sup>In the tutorial we show that  $\mathcal{O}_{T^*\mathcal{M}}$  is a topology for the cotangent bundle. An analogous proof can be inserted here to show that  $\mathcal{O}_{T\mathcal{M}}$  is also a topology.



another chart  $(V, y)$  with  $V \cap U \neq \emptyset$ , we have<sup>5</sup>

$$\begin{aligned}
 (\xi_y \circ \xi_x^{-1})(\alpha^1, \dots, \alpha^d, \beta^1, \dots, \beta^d) &:= \xi_y \left( \beta^m \left( \frac{\partial}{\partial x^m} \right)_{\pi(X)} \right) \\
 &= \left( \dots, (y^i \circ \pi) \left( \beta^i \left( \frac{\partial}{\partial x^i} \right)_{\pi(X)} \right), \dots, \dots, (dy^i)_{\pi(X)} \left[ \left( \beta^m \left( \frac{\partial}{\partial x^m} \right)_{\pi(X)} \right) \right], \dots \right) \\
 &= \left( \dots, (y^i \circ x^{-1})(\alpha^1, \dots, \alpha^d), \dots, \dots, \beta^m \left( \frac{\partial y^i}{\partial x^m} \right)_{\pi(X)}, \dots \right) \\
 &= \left( \dots, (y^i \circ x^{-1})(\alpha^1, \dots, \alpha^d), \dots, \dots, \beta^m \partial_m (y^i \circ x^{-1})|_{(\alpha^1, \dots, \alpha^d)}, \dots \right),
 \end{aligned}$$

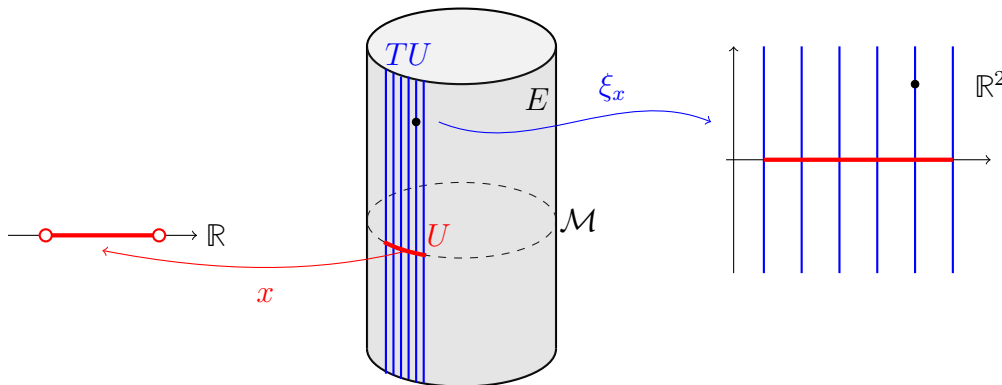
where to go to the third line we have used the fact that  $\pi(X) = p = x^{-1}(\alpha^1, \dots, \alpha^d)$  and to get to the last line we have used the definition for the derivative fraction along with  $(x \circ \pi)(X) = x(p) = (\alpha^1, \dots, \alpha^d)$ . Now  $(y^i \circ x^{-1})$  is smooth because  $\mathcal{A}$  is smooth and so the above result is smooth. We therefore have a smooth atlas  $\mathcal{A}_{T\mathcal{M}}$ .

**Definition** (Tangent Bundle). The triple  $(T\mathcal{M}, \pi, \mathcal{M})$  is a bundle, known as the **tangent bundle**.

This all seems rather abstract and complicated, but the following example shows it's actually rather natural and intuitive.

*Example 6.2.1.* Let  $\mathcal{M} = S^1$  (a circle) and let the fibres just run straight up and down. The further up/down the fibre one goes, the greater the value of the vector, with going downwards corresponding to placing a minus sign in front of the vector.

$U$  here is a small part of the circle and is mapped by  $x$  to an open interval in the real line.  $TU$  is the set of fibres that run through  $U$ . These are mapped via  $\xi_x$  to  $\mathbb{R}^2$  in the following way. Consider a point on one of the fibres, call it  $X$ . The horizontal axis value in the  $\mathbb{R}^2$  chart is given by the value the base point  $p = \pi(X) \in \mathcal{M}$  takes in the  $\mathbb{R}$  chart, as mapped by  $x$ . The vertical value in the  $\mathbb{R}^2$  chart is just given by the size of the vector (as it is only one-dimensional so the component is the size) and is plotted accordingly. That is, the vertical axis is 'length of vector', again with the negative axis corresponding to a vector that is lower down on the fibre than the base point.



<sup>5</sup>Sorry this doesn't look very nice. It's lots of brackets and indices!

### 6.3 Vector Fields

We just put in a lot of work to check/prove that the set  $T\mathcal{M}$  can be made into a smooth manifold and so we have a bundle. It is reasonable to wonder why we did such a crazy calculation. The answer is that it allows the next definition.

**Definition** (Smooth Vector Field). A **smooth vector field** is a *smooth section* on the tangent bundle. That is

$$\chi : \mathcal{M} \rightarrow T\mathcal{M}, \quad \pi \circ \chi = \mathbb{1}_{\mathcal{M}}.$$

*Remark 6.3.1.* Note we have used the Greek letter  $\chi$  here to denote a smooth vector field. We do this to make the distinction between a vector  $X \in T_p\mathcal{M}$  and a smooth vector field  $\chi$ . We will continue to use Greek letters for fields (in general, so we will also use Greek letters for covector fields and tensor fields) in this lecture.<sup>6</sup> As we shall see shortly, we also introduce a new notation for the action of smooth vector fields. We shall point these out as we introduce them.

The smooth part of the above definition is what all the work was for. Intuitively when we think of a vector field (a vector at each point) we see a smooth vector field, i.e. one where the vectors appear to naturally flow from one to another, rather than just pointing randomly at each point. Smooth vector fields will obviously play a vital role in general relativity: the velocity of a particle is a smooth vector field over the manifold that is the worldline of the particle. If this vector field was not smooth, it would correspond to the particle's velocity all of a sudden changing, which we know is not physical.

### 6.4 The $C^\infty(\mathcal{M})$ -Module

So far we have a definition for a smooth vector field, but we have no way of adding them together or scaling them in any way. This is something we clearly want to be able to do, and so we want to try and make it into a vector space over some field. The addition is straight forward, just add the vectors in the tangent spaces together and take the result to be the new vector at that point. What about scaling? We don't want to limit ourselves to only being able to scale the smooth vector field uniformly, i.e. by the same amount at all  $p \in \mathcal{M}$ . So we need something that is defined all over  $\mathcal{M}$  but that can take different values at each point. This is just a scalar field.

So we want to try and turn the set of smooth sections over the tangent bundle into a  $C^\infty(\mathcal{M})$ -vector space. There is a problem, though. Recall the definition

$$C^\infty(\mathcal{M}) := \{f : \mathcal{M} \rightarrow \mathbb{R} \mid f \text{ is a smooth function}\}.$$

It is possible that a non-vanishing<sup>7</sup> element of  $C^\infty(\mathcal{M})$  can vanish at some points, i.e. there are points  $p \in \mathcal{M}$  that are mapped to  $0 \in \mathbb{R}$ . We cannot turn  $C^\infty(\mathcal{M})$  into a field, then, as we don't have an inverse under multiplication for every element (we can't invert the points that vanish!). The best we can do, then, is to turn it into a *ring*. We clearly have a neutral element – the elements that just maps all points  $p \in \mathcal{M}$  to  $1 \in \mathbb{R}$  – and we can define

<sup>6</sup>We will change our minds next lecture!

<sup>7</sup>That is does not map every point  $p \in \mathcal{M}$  to  $0 \in \mathbb{R}$ .

the commutativity pointwise, using the fact that  $(\mathbb{R}, +)$  is commutative. We therefore get a *commutative, unital ring*. If we build on top of this, we get a *module*.

**Definition** (The  $C^\infty(\mathcal{M})$ -Module,  $\Gamma T\mathcal{M}$ ). The triple  $(\Gamma T\mathcal{M}, \oplus, \odot)$  is a  $C^\infty(\mathcal{M})$ -module where

$$\Gamma T\mathcal{M} := \{\chi : \mathcal{M} \rightarrow T\mathcal{M} \mid \text{smooth section}\},$$

and

$$\begin{aligned} (\chi \oplus \tilde{\chi})\langle f \rangle &:= \chi\langle f \rangle + \tilde{\chi}\langle f \rangle, \\ (g \odot \chi)\langle f \rangle &= g \cdot \chi\langle f \rangle, \end{aligned}$$

where  $+/\cdot$  are the addition/multiplication on  $C^\infty(\mathcal{M})$ .

This is the first point where we have introduced a new notation for the action of a field. Recall we have been denoting the action of a vector (at a point) on a  $C^\infty(\mathcal{M})$  function via standard brackets,  $X(f)$ . In order to distinguish this from the action of a smooth vector field on  $f$ , we use angled brackets for the latter  $\chi\langle f \rangle$ .<sup>8</sup> This might seem like a just a notational problem, however it actually encapsulates an important point: a smooth vector field is a map  $\chi : \mathcal{M} \rightarrow T\mathcal{M}$ , so how does it act on a scalar field? The answer is obviously through the vectors that make up  $\chi$ :

$$\chi\langle f \rangle|_p := (\chi(p))(f).$$

That is, we first evaluate  $\chi(p)$ , which gives us *the*  $X \in T_p\mathcal{M}$ , and then we let this act on the scalar field, giving a real number. We do this for every point  $p \in \mathcal{M}$  and so get a map that associates to each point a real number, this is a scalar field. In other words we can think of smooth vector fields as maps

$$\chi : C^\infty(\mathcal{M}) \xrightarrow{\sim} C^\infty(\mathcal{M}).$$

It is because of this that we take the addition/multiplication on the right-hand side of the expressions in the definition to be those defined on  $C^\infty(\mathcal{M})$ .

#### Exercise

Show that the map  $\chi : C^\infty(\mathcal{M}) \xrightarrow{\sim} C^\infty(\mathcal{M})$  is  $\mathbb{R}$ -linear. That is, for  $f, g \in C^\infty(\mathcal{M})$  and  $\lambda \in \mathbb{R}$

$$\begin{aligned} \chi\langle f + g \rangle &= \chi\langle f \rangle + \chi\langle g \rangle, \\ \chi\langle \lambda \cdot f \rangle &= \lambda \cdot \chi\langle f \rangle. \end{aligned}$$

Also show that it obeys

$$\chi\langle f \bullet g \rangle = f \bullet \chi\langle g \rangle + \chi\langle f \rangle \bullet g,$$

where  $\bullet : C^\infty(\mathcal{M}) \times C^\infty(\mathcal{M}) \rightarrow C^\infty(\mathcal{M})$  is the multiplication on the ring. This property is known as the **Leibniz rule**.<sup>a</sup>

<sup>a</sup>Note for partial differential equations it is the familiar product rule.

<sup>8</sup>This particular choice of notation is used as it is the one I learned while at University.

Now there is an important fact in set theory<sup>9</sup> that every vector space has a basis. However, this incredibly useful fact does not apply to modules. That is, in general, we *cannot* simply take the subset

$$\{\chi_1, \dots, \chi_d\} \subseteq \Gamma T\mathcal{M}$$

such that any other  $\chi \in \Gamma T\mathcal{M}$  can be expressed as a linear combination of this subset

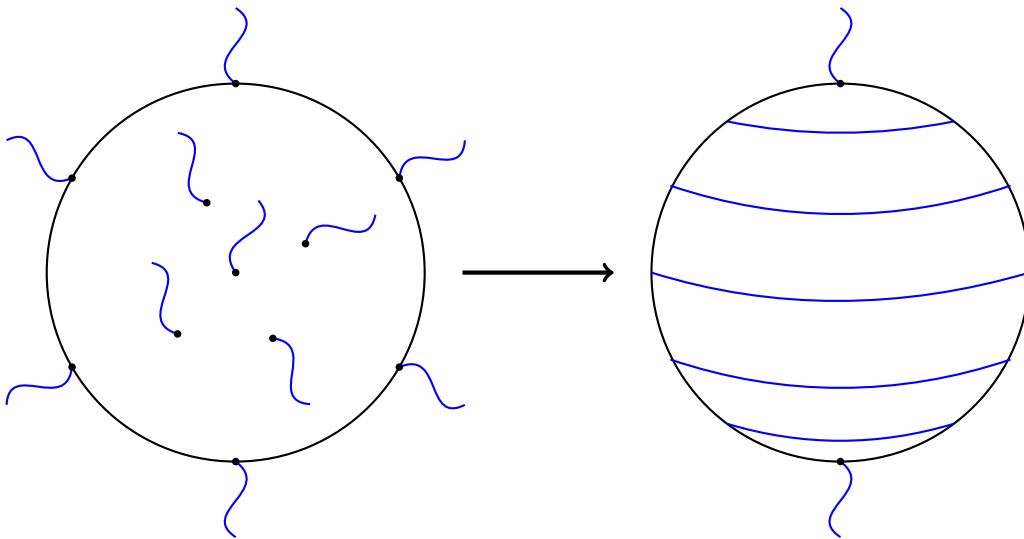
$$\chi = f^i \odot \chi_i.$$

Of course we can do this *locally* by simply decomposing our vector fields locally, but we cannot do it *globally*.

Example 6.4.1. Consider a ball with a smooth vector field over it. If we imagine this smooth vector field as hairs sticking out of the ball, the idea of having a globally defined nowhere vanishing smooth vector field, would be to ‘comb’ the hairs flat to the surface. That is, we want all of the vectors to lie in the tangent spaces and not ‘stick straight out’.

However, in order to do this, we would have to remove some of the hairs: for example in the diagram drawn below, the hair at the top and bottom would have to ‘vanish’ if we wanted the ball to be smooth.

The fact that the vector field is not defined globally means that it can not possibly be a basis element. Of course you could have another vector field that went ‘top-to-bottom’ on the sphere that was defined at the North and South poles, but that would not allow you to define *any* vector field at those points — how would you write a vector that pointed East from the North pole?



This failure to define a global, nowhere vanishing, smooth vector field is related to the fact that we can't chart the space using only a single chart. That is, the minimal atlas for a sphere contains two charts. For example, if we used polar spherical coordinates on the surface to chart the sphere, the North and South poles will not be charted – what's the longitude value at these points?

<sup>9</sup>Strictly speaking we need to use ZFC set theory, because we need the axiom of choice. For more information see Dr. Schuller's Lectures on the Geometric Anatomy of Theoretical Physics.

We can repeat everything we did in order to define the tangent bundle but instead starting from the cotangent space  $(T^*\mathcal{M}, +, \cdot)$ . In doing this we get the cotangent bundle and smooth covector fields. Finally we get the  $C^\infty(\mathcal{M})$ -module  $(\Gamma T^*\mathcal{M}, \oplus, \odot)$  where

$$\Gamma T^*\mathcal{M} := \{\alpha : \mathcal{M} \rightarrow T^*\mathcal{M} \mid \text{smooth section}\}.$$

Example 6.4.2. Recall we had the gradient at a point  $(df)_p : T_p\mathcal{M} \xrightarrow{\sim} \mathbb{R}$ . We now want to extend this to be over the whole manifold. We therefore define

$$df : \Gamma T\mathcal{M} \xrightarrow{\sim} C^\infty(\mathcal{M})$$

by<sup>10</sup>

$$df : \chi := \chi \langle f \rangle.$$

The linearity here is actually  $C^\infty$ -linear.<sup>11</sup> This is different to smooth vector fields, which are only  $\mathbb{R}$ -linear.

### Exercise

Show that the map  $df : \Gamma T\mathcal{M} \xrightarrow{\sim} C^\infty(\mathcal{M})$  is indeed  $C^\infty$ -linear. That is for all  $\chi, \Upsilon \in \Gamma T\mathcal{M}$  and  $g \in C^\infty(\mathcal{M})$ ,

$$\begin{aligned} df : (\chi \oplus \Upsilon) &= (df : \chi) + (df : \Upsilon) \\ df : (g \odot \chi) &= g \cdot (df : \chi). \end{aligned}$$

## 6.5 Tensor Fields

We have the smooth vector fields and the smooth covector fields. We can, therefore, now construct smooth *tensor* fields.

**Definition** (Smooth Tensor Field). A **smooth  $(r, s)$ -tensor field** is a  $C^\infty(\mathcal{M})$  multilinear map

$$T : \underbrace{\Gamma T^*\mathcal{M} \times \dots \times \Gamma T^*\mathcal{M}}_{r\text{-terms}} \times \underbrace{\Gamma T\mathcal{M} \times \dots \times \Gamma T\mathcal{M}}_{s\text{-terms}} \xrightarrow{\sim} C^\infty(\mathcal{M}),$$

or in the other notation

$$T := \underbrace{\Gamma T\mathcal{M} \otimes \dots \otimes \Gamma T\mathcal{M}}_{r\text{-terms}} \times \underbrace{\Gamma T^*\mathcal{M} \otimes \dots \otimes \Gamma T^*\mathcal{M}}_{s\text{-terms}}.$$

Remark 6.5.1. Note in the second notation, the  $\otimes$  now means a map to  $C^\infty(\mathcal{M})$  not just  $\mathbb{R}$ , as it did when we first introduced it. This is the downfall of this notation: people use the same tensor product symbol for all kinds of different things that look similar.<sup>12</sup>

<sup>10</sup>We have used the notation given to me by my lecturer, namely we denote the action of  $df$  on  $X$  by a colon and the action of a vector field on a scalar field via angled brackets.

<sup>11</sup>This is often called  $f$ -linear, for obvious reasons.

<sup>12</sup>For examples, see chapter 14 of my notes from Dr. Schuller's Quantum Theory course.

The two definitions above don't seem to quite match, though. We have seen that  $\alpha \in \Gamma T^* \mathcal{M}$  can map a  $X \in \Gamma T \mathcal{M}$  to a  $C^\infty(\mathcal{M})$  function, but the opposite isn't true — vector fields map scalar fields to scalar fields — so how does the tensor product definition work? The answer is simply that we interpret it as the following example highlights.

Example 6.5.2. Let  $T$  be a smooth  $(1, 1)$ -tensor field given by  $T = X \otimes \alpha$ . Its action as a map is given by

$$T(\beta, Y) = (X \otimes \alpha)(\beta, Y) := (\beta : X) \otimes (\alpha : Y) = (\beta : X) \bullet (\alpha : Y),$$

for  $\beta \in \Gamma T^* \mathcal{M}$ ,  $Y \in \Gamma T \mathcal{M}$  and where we have used the fact that the tensor product of two scalar fields is just their multiplication,  $\bullet$ .

Remark 6.5.3. From this point on wards we shall simply say vector/covector/tensor field when we mean a *smooth* field. This is just to lighten the amount of words.

## 7 | Connections

So far everything that we have introduced has been something we have to introduce by hand, e.g. we provide a topology on our set. As we will see later in the course, Einstein's equations will actually give us a connection<sup>1</sup> for our manifold, and so it is the physics that provides this structure. Nevertheless, we shall continue on wards in a mathematical sense and define connections this way.

*Remark 7.0.1.* Really what we are interested in are so-called covariant derivatives, which are technically slightly different to connections. However, this difference will not manifest here and so we shall use both terms interchangeably.

*Notation.* We shall now undo the notation about labelling vector fields by Greek letters and simply use  $X, Y, Z$  for vector fields. This is done because we will only consider vector fields from this point on wards. If we do use a vector at some stage, it will be clear as we will use the notation for the action of a vector (that is regular brackets) whereas we will continue to use the angular bracket notation for vector fields.

So far we have seen that a vector field  $X$  can be used to provide a directional derivative  $X\langle f \rangle$  of a function  $f \in C^\infty(\mathcal{M})$ . To remind ourselves that we are dealing with directional derivatives, we shall introduce a new notation

$$\nabla_X f := X\langle f \rangle.$$

This seem like a massive notational overkill: we have three equivalent expressions,

$$\nabla_X f = X\langle f \rangle = df : X.$$

However, although the evaluations are equal, the three objects are actually different as maps. That is

$$\begin{aligned} X &: C^\infty(\mathcal{M}) \xrightarrow{\sim} C^\infty(\mathcal{M}), \\ df &: \Gamma T\mathcal{M} \xrightarrow{\sim} C^\infty(\mathcal{M}). \end{aligned}$$

What about  $\nabla_X$ , that as a map

$$\nabla_X : C^\infty(\mathcal{M}) \xrightarrow{\sim} C^\infty(\mathcal{M}),$$

appears to be exactly the same as  $X$ . This is true, but it turns out that we can actually extend the definition of  $\nabla_X$  to be a map from a  $(r, s)$ -tensor field to a  $(r, s)$ -tensor field, which  $X$  cannot.

---

<sup>1</sup>It is actually a bit of a longer route, via so-called metrics, but we will see all of this.

## 7.1 Directional Derivatives of Tensor Fields

We formulate a wish list of the properties which the  $\nabla_X$  acting on a tensor field should have. This wish list might not give a unique form for  $\nabla_X$  and there may be many such objects that satisfy our wish list conditions. It will be important for us to work out how to parameterise these structures so that we can pick the best one for the circumstance we are considering.

**Definition** (Connection/Covariant Derivative). A (linear) **connection**  $\nabla$  on a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  is a map that takes a pair consisting of a vector (field)  $X$  and a  $(r, s)$ -tensor field  $T$  and sends them to a  $(p, q)$ -tensor (field)  $\nabla_X T$ , satisfying:<sup>2</sup> for all  $f \in C^\infty(\mathcal{M})$  and  $(r, s)$ -tensor fields  $T, S$

- (i) Action on scalars;  $\nabla_X f := X\langle f \rangle$ ,
- (ii) +-linearity in the tensor fields;  $\nabla_X(T + S) = \nabla_X T + \nabla_X S$ ,
- (iii) Leibniz; e.g. if  $T$  being a  $(1,1)$  tensor field, and  $\omega \in \Gamma T^* \mathcal{M}, Y \in \Gamma T \mathcal{M}$ ,

$$\nabla_X(T(\omega, Y)) = (\nabla_X T)(\omega, Y) + T(\nabla_X \omega, Y) + T(\omega, \nabla_X Y).$$

This is extended naturally to higher order tensors, and

- (iv)  $f$ -linearity in the vector field;  $\nabla_{f \cdot X + Y} T = f \nabla_X T + \nabla_Y T$ .

*Remark 7.1.1.* The bracketed (field) in the above definition is because it is possible to only feed  $\nabla$  a vector (not a vector field) and get out just a tensor defined at the same point as the vector. It is important thought that we always feed in a tensor field. This is just the extension of the fact that  $X(f) \in \mathbb{R}$  whereas  $X\langle f \rangle \in C^\infty(\mathcal{M})$ .

What the above remark actually highlights is what the covariant derivative does. Recall that the derivative of something corresponds to ‘comparing how it changes as you go along’. If we want to take some form of derivative of a tensor field, then, we obviously require it to be defined at more than one point (so that we have two values to compare). This is why it must be a field. The lower entry, though, simply tells us the *direction* that we wish to take this derivative, and so we can consider just a single vector. So the covariant derivative asks the question ‘how does  $T$  vary as you move along  $X$ ?’ If we use just a vector, the result we get tells us how  $T$  changes along  $X$  *at that point*, and so our result is just defined at that point. If we use a vector field, though, we get how  $T$  varies along the vector field  $X$  and so our result is a field.

We will see later a different derivative structure, the Lie derivative, that requires knowing *both*  $T$  and  $X$  in a neighbourhood and so does not work for  $X$  being a vector.

---

<sup>2</sup>We shall assume it is a vector field for the notation used in these conditions. For just a vector just replace the angular brackets with regular ones and replace  $f$  in condition (iv) with  $\lambda \in \mathbb{R}$ .



## Exercise

Condition (iii) is also given in a different form. It is

$$\nabla_X(S \otimes R) = (\nabla_X S) \otimes R + S \otimes (\nabla_X R).$$

This makes the name Leibniz<sup>a</sup> seem more reasonable. Prove that the expression above can be derived from condition (iii).

*Hint: Let  $T = W \otimes y$  for  $W \in \Gamma T\mathcal{M}$  and  $y \in \Gamma^* T\mathcal{M}$  and then use the result*

$$\nabla_X(\omega : W) = (\nabla_X \omega) : W + \omega : (\nabla_X W),$$

*which you get from applying condition (iii) to a covector field.*

<sup>a</sup>Recall Leibniz basically means an extension of the product rule.

## Exercise

Show that conditions (ii)-(iv) are satisfied for a  $(0, 0)$ -tensor field, i.e. for a  $f \in C^\infty(\mathcal{M})$ .

*Hint: Use the fact that  $f \otimes g = f \bullet g$ , where  $\bullet$  is the multiplication on the ring. Note, if you have done the other exercises you are basically done!*

*Remark 7.1.2.* We have shown/argued that  $\nabla_X$  is the extension of the action of  $X$ , so its natural to ask the question ‘what is  $\nabla$  itself?’ The answer is simply that it is the extension of  $d$ . We see this straight away from  $\nabla_X f = X\langle f \rangle = df : X$ .

**Definition** (Affine Manifold). We say that a **manifold with connection**, or **affine manifold**, is the quadruple of structures  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$ .

7.2 New Structure on  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  Required to Fix  $\nabla$ 

The question we want to answer is whether this is unique or whether different  $\nabla$ s will give the same result. In other words, how much freedom do we have in choosing  $\nabla$ ?<sup>3</sup>

Consider the action of a vector field  $X$  on another vector field  $Y$ . In order to do the calculation, we also introduce a chart  $(U, x)$ .

$$\begin{aligned} \nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} \left( Y^j \frac{\partial}{\partial x^j} \right) \\ &= X^i \nabla_{\frac{\partial}{\partial x^i}} \left( Y^j \frac{\partial}{\partial x^j} \right) \\ &= X^i \left( \nabla_{\frac{\partial}{\partial x^i}} Y^j \right) \frac{\partial}{\partial x^j} + X^i Y^j \left( \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right) \\ &= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} + X^i Y^j \left( \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right) \\ &= X\langle Y^j \rangle \frac{\partial}{\partial x^j} + X^i Y^j \left( \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right), \end{aligned}$$

<sup>3</sup>There is a slightly more generic, nice discussion of this given in Wald’s book, Section 3.1 (pages 32-34).

where we have used the axioms for the connection along the way. Now, what do we do with the last term? Well its the covariant derivative of a vector field, and so we know that it must be a vector field. We can then expand this in the chart induced basis and give it coefficients, which we call  $\Gamma^m$ . These coefficients will actually also need a covariant  $i$  and  $j$  index in order for the summation convention to not be broken. So we have

$$\left( \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right) = \Gamma_{(x)ji}^m \frac{\partial}{\partial x^m}.$$

These coefficients are called the **connection coefficient functions** of  $\nabla$  *w.r.t.* the chart  $(U, x)$ .<sup>4</sup> Note we wrote the  $j$  index before the  $i$  index above, this is important to note as we don't, as yet, have any symmetry condition  $\Gamma_{(x)ji}^m = \Gamma_{(x)ij}^m$ .

We can write these via the following definition.

**Definition** (Connection Coefficient Functions). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$  be a affine manifold and let  $(U, x) \in \mathcal{A}$ . The **connection coefficient functions**, henceforth just 'the  $\Gamma$ s', are the  $(\dim \mathcal{M})^3$  many functions

$$\begin{aligned} \Gamma_{(x)jk}^i &: U \rightarrow \mathbb{R} \\ p &\mapsto \left[ dx^i : \left( \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} \right) \right]_p. \end{aligned}$$

*Remark 7.2.1.* We can see<sup>5</sup> how our two expressions for the  $\Gamma$ s are equivalent by imagining 'inverting' the action of the  $dx^i$  so that it becomes a  $\frac{\partial}{\partial x^i}$  on the left-hand side.

Now, plugging the  $\Gamma$ s into the expression for  $\nabla_X Y$  and relabelling the indices, we can express the right-hand side as  $(\dots)^i \frac{\partial}{\partial x^i}$ , and so extract the components of the derivative. We get

$$(\nabla_X Y)^i = X \langle Y^i \rangle + X^j Y^k \Gamma_{(x)kj}^i.$$

So the answer to our question of how much freedom is left is the  $\Gamma$ s; that is we can tell you exactly which  $\nabla$  we are using by telling you the  $\Gamma$ s. Clearly this only holds on the chart domain  $U \subseteq \mathcal{M}$  as that's where the  $\Gamma$ s are defined. Now you might say 'hold up we only know that this will give us the covariant derivative of a vector field, what about the covariant derivative of different tensors? We will surely need more and more terms to find them!' Luckily the answer is that we don't and it suffices to just know that  $\Gamma$ s. In order to see this, consider the following.

If we wanted to work out the action on a covector basis element  $dx^i$ , we could do a similar thing to above and expand the result in the basis. That is

$$\nabla_{\frac{\partial}{\partial x^i}} dx^j = \Theta_{(x)ki}^j dx^k,$$

where the  $\Theta$ s are defined by this expression. We want to show that we can actually express

<sup>4</sup>Note the subscript  $(x)$ , there to remind us that it is defined with respect to the chart.

<sup>5</sup>Granted rather hand waveling as presented here.

these  $\Theta$ s in terms of the  $\Gamma$ s. In order to do that, consider the following:

$$\begin{aligned}\nabla_{\frac{\partial}{\partial x^i}} \left( dx^j : \frac{\partial}{\partial x^k} \right) &= \left( \nabla_{\frac{\partial}{\partial x^i}} dx^j \right) : \frac{\partial}{\partial x^k} + dx^j : \left( \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^k} \right) \\ \nabla_{\frac{\partial}{\partial x^i}} \delta_k^j &= \Theta_{(x)\ell i}^j dx^\ell : \frac{\partial}{\partial x^k} + dx^j : \left( \Gamma_{(x)ki}^\ell \frac{\partial}{\partial x^\ell} \right) \\ 0 &= \Theta_{(x)\ell i}^j \delta_k^\ell + \Gamma_{(x)ki}^\ell \delta_\ell^j \\ \Theta_{(x)ki}^j &= -\Gamma_{(x)ki}^j,\end{aligned}$$

and so by giving the  $\Gamma$ s we can also tell you the action of the covariant derivative on a covector field.

We have the following mnemonic: ‘when it acts on a vector field, you get a plus sign, when it acts on a covector you have a minus sign.’ Summarising, we have

$$\begin{aligned}(\nabla_X Y)^i &= X \langle Y^i \rangle + \Gamma_{(x)jk}^i X^k Y^j, \\ (\nabla_X \omega)_i &= X \langle \omega_i \rangle - \Gamma_{(x)ik}^j X^k \omega_j.\end{aligned}$$

Note the placement of all the indices, it is very important to know the method of which index corresponds to which term (i.e. the  $X$  or  $Y$  or  $\omega$ ). An easy way to do this is to think summation convention and then to know that on the second term the index on whatever you’re differentiating changes. Then you just remember that the second lower index on the  $\Gamma$ s always corresponds to the index of the  $X$ .

So what about higher order tensors? The answer is obviously just to use the Leibniz rule. For example, for a  $(1, 2)$ -tensor field  $T$  we have

$$(\nabla_X T)^i_{jk} = X \langle T^i_{jk} \rangle + \Gamma_{(x)m\ell}^i X^\ell T^m_{jk} - \Gamma_{(x)j\ell}^m X^\ell T^i_{mk} - \Gamma_{(x)k\ell}^m X^\ell T^i_{jm}.$$

Each term is the contribution from one of the indices on the left-hand side. You consider that index formula and you leave the remaining two untouched.

#### Exercise

Show that the above result is indeed obtained by the Leibniz formula.

*Hint: Let  $T = Y \otimes \omega \otimes \gamma$  for  $Y \in \Gamma T\mathcal{M}$  and  $\omega, \gamma \in \Gamma T^*\mathcal{M}$ .*

*Remark 7.2.2.* We can use the  $\Gamma$ s to define what we mean by a Euclidean space. Let  $\mathcal{M} = \mathbb{R}^3$  be equipped with the standard topology  $\mathcal{O}_{st}$  and a smooth atlas  $\mathcal{A}$ . We define the Euclidean space to be this smooth manifold equipped with a connection such that it is possible to find a chart  $(U, x) \in \mathcal{A}$  such that

$$\Gamma_{(x)jk}^i = 0,$$

for all  $i, j, k \in \{1, \dots, \dim \mathcal{M}\}$ . Note we say ‘it is possible to find a chart’ such that this happens. As we will see, just because the  $\Gamma$ s vanish in one chart does not mean they will vanish in another (that is, they are not tensors!). We will also extend this notion of a Euclidean space to define the spacetime extension known as Minkowski spacetime, which is an intrinsically flat spacetime. We get a hint here about what covariant derivatives do: they detect curvature.

*Notation.* (i) From now on, unless the context requires (e.g. considering change of charts), we shall drop the  $(x)$  subscript on the  $\Gamma$ 's in order to lighten notation.

(ii) Again unless the context requires, we shall also use the notation

$$\nabla_i := \nabla_{\frac{\partial}{\partial x^i}}.$$

**Definition** (Divergence of Vector Field). Let  $X$  be a vector field on a smooth affine manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$ . The **divergence** of  $X$  is the function

$$\operatorname{div} X := (\nabla_i X)^i.$$

*Claim 7.2.3.* The above definition is chart independent.

### 7.3 Change of $\Gamma$ s Under Change of Chart

So far we have defined the  $\Gamma$ s on  $U \subseteq \mathcal{M}$ , we obviously want to extend this to be a global definition on all of  $\mathcal{M}$ . We do this by considering overlapping charts and require compatibility.

Assume we have a affine manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$  and consider two charts  $(U, x)$  and  $(V, y)$  with  $U \cap V \neq \emptyset$ . We want to relate the  $\Gamma$ s in these charts.

$$\begin{aligned} \Gamma_{(y)jk}^i &:= dy^i : \left( \nabla_{\frac{\partial}{\partial y^j}} \frac{\partial}{\partial y^k} \right) \\ &= \frac{\partial y^i}{\partial x^q} dx^q : \left( \nabla_{\frac{\partial}{\partial y^j}} \frac{\partial x^s}{\partial x^p} \frac{\partial}{\partial y^k} \frac{\partial}{\partial x^s} \right) \\ &= \frac{\partial y^i}{\partial x^q} dx^q : \left( \frac{\partial x^p}{\partial y^j} \left[ \frac{\partial}{\partial x^p} \left\langle \frac{\partial x^s}{\partial y^k} \right\rangle \frac{\partial}{\partial x^s} + \frac{\partial x^s}{\partial y^k} \left( \nabla_{\frac{\partial}{\partial x^p}} \frac{\partial}{\partial x^s} \right) \right] \right) \\ &= \frac{\partial y^i}{\partial x^q} dx^q : \left( \frac{\partial x^p}{\partial y^j} \left[ \frac{\partial}{\partial x^p} \left\langle \frac{\partial x^s}{\partial y^k} \right\rangle \frac{\partial}{\partial x^s} + \frac{\partial x^s}{\partial y^k} \Gamma_{(x)sp}^m \frac{\partial}{\partial x^m} \right] \right) \\ &= \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^j} \left( \frac{\partial}{\partial x^p} \left\langle \frac{\partial x^s}{\partial y^k} \right\rangle \delta_s^q + \frac{\partial x^s}{\partial y^k} \Gamma_{(x)sp}^m \delta_m^q \right) \\ &= \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^j} \frac{\partial}{\partial x^p} \left\langle \frac{\partial x^q}{\partial y^k} \right\rangle + \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^j} \frac{\partial x^s}{\partial y^k} \Gamma_{(x)sp}^q \\ &= \frac{\partial y^i}{\partial x^q} \frac{\partial}{\partial y^j} \left\langle \frac{\partial x^q}{\partial y^k} \right\rangle + \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^j} \frac{\partial x^s}{\partial y^k} \Gamma_{(x)sp}^q \\ &= \frac{\partial y^i}{\partial x^q} \frac{\partial^2 x^q}{\partial y^j \partial y^k} + \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^j} \frac{\partial x^s}{\partial y^k} \Gamma_{(x)sp}^q, \end{aligned}$$

where to get to the penultimate line we have used the change of chart rule, that is<sup>6</sup>

$$\frac{\partial x^p}{\partial y^j} \frac{\partial}{\partial x^p} = \frac{\partial}{\partial y^j},$$

<sup>6</sup>This is an important step as we need both the derivatives to be w.r.t. the same chart label  $(y)$  in order for us to be able to use Schwartz's rule for switching the differentiation order.

and where we have introduced the notation<sup>7</sup>

$$\frac{\partial^2 x^q}{\partial y^j \partial y^k} := \frac{\partial}{\partial y^j} \left\langle \frac{\partial x^q}{\partial y^k} \right\rangle.$$

Note that if the expression was simply

$$\Gamma_{(y)jk}^i = \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^j} \frac{\partial x^s}{\partial y^k} \Gamma_{(x)sp}^q,$$

we would say ‘ah this is a (1,2)-tensor component transformation!’ This is not the only term though and so we see that the  $\Gamma$ s *are not tensors!* This second term actually has another, very important, implication: because there is no  $\Gamma_{(x)}$  term present in it, just because the  $\Gamma_{(x)}$ s vanish, it does not mean that they vanish for another chart *for the same manifold*. That is, by simply a nonlinear transformation we can introduce  $\Gamma$ s into our system. This is what the comment in Remark 7.2.2 was on about, we can only talk about the existence of a chart such that the  $\Gamma$ s vanish as they will not vanish on all charts.

*Remark 7.3.1.* Note for linear transformations (also known as *affine maps*) the  $\Gamma$ s behave like the components of a (1,2)-tensor as the second derivative will vanish. This is one of the reasons that people choose to restrict themselves to linear transformation in position in special relativity.

The condition above is our compatibility condition for the overlapping regions in order to define the  $\Gamma$ s globally. Since this is our chart compatibility condition, we can only generally make the  $\Gamma$ s vanish *locally*, i.e. within one chart.<sup>8</sup>

*Remark 7.3.2.* Technically speaking it is the symmetric part (which we denote with regular parentheses around the symmetric indices) of the  $\Gamma$ s that are not the components of a tensor. The antisymmetric part  $\Gamma_{(y)[jk]}^i$ , which means

$$\Gamma_{(y)jk}^i = -\Gamma_{(y)kj}^i,$$

are the components of a (1,2)-tensor. We see this simply from the fact that

$$\frac{\partial^2 x^q}{\partial y^j \partial y^k} = \frac{\partial^2 x^q}{\partial y^k \partial y^j}.$$

So if you have a non-vanishing antisymmetric part to your  $\Gamma$ s you can not use a chart transformation to remove it. It turns out that the antisymmetric part of the  $\Gamma$ s vanish when we have a so-called *torsion free*<sup>9</sup> system. So if we restricted ourselves to torsion free charts, we could then use a chart transformation to obtain *locally* vanishing  $\Gamma$ s.

<sup>7</sup>Obviously this notation is just that for partial derivatives, but recall that our fractions  $\frac{\partial f}{\partial x^i}$  don't mean partial derivative, it means the expression we defined before.

<sup>8</sup>Some manifolds, like Minkowski spacetime, can be covered with a single chart and so we can obtain *globally* vanishing  $\Gamma$ s.

<sup>9</sup>We shall discuss this briefly later.

### 7.4 Normal Coordinates

Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$  be an arbitrary affine manifold and let  $p \in \mathcal{M}$ . Then one can construct a chart  $(U, x) \in \mathcal{A}$  with  $p \in U$  such that<sup>10</sup>

$$\Gamma_{(x)(jk)}^i(p) = 0.$$

This says that we can make the  $\Gamma$ s vanish *at the point*  $p \in \mathcal{M}$ , *not* that we can necessarily make them vanish in some neighbourhood of  $p$ .<sup>11</sup>

*Proof.* Let  $(V, y) \in \mathcal{A}$  be any chart with  $p \in V$ . Thus, in general, the  $\Gamma_{(y)(jk)}^i \neq 0$ . Then consider a new chart  $(U, x)$  to which one transits by virtue of

$$(x \circ y^{-1})(\alpha^1, \dots, \alpha^d) := \alpha^i - \frac{1}{2} \alpha^j \alpha^k \Gamma_{(y)(jk)}^i(p).$$

Then

$$\begin{aligned} \left( \frac{\partial x^i}{\partial y^j} \right)_p &:= \partial_j (x^i \circ y^{-1})|_{(\alpha^1, \dots, \alpha^p)} \\ &= \delta_j^i - \alpha^m \Gamma_{(y)(jm)}^i(p) \\ \implies \left( \frac{\partial^2 x^i}{\partial y^k \partial y^j} \right)_p &= -\Gamma_{(y)(jk)}^i(p). \end{aligned}$$

Now we can choose, w.l.o.g., the chart  $(V, y)$  such that  $y(p) = (0, \dots, 0)$ , then we have

$$\Gamma_{(x)jk}^i(p) = \Gamma_{(y)jk}^i(p) - \Gamma_{(y)(jk)}^i(p) = \Gamma_{(y)[jk]}^i(p),$$

and so we only have an antisymmetric contribution, therefore the symmetric part vanishes.  $\square$

Terminology. The chart  $(U, x)$  is called a **normal coordinate chart** of  $\nabla$  at  $p \in \mathcal{M}$ .

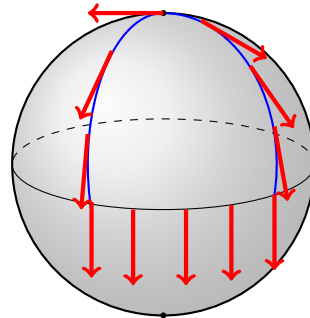
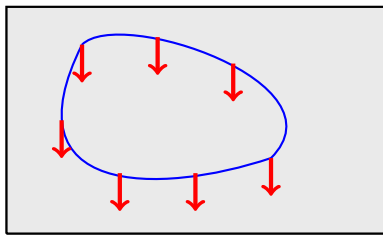
<sup>10</sup>Again, the parentheses denote the symmetric part:  $\Gamma_{(x)(jk)}^i = \frac{1}{2}(\Gamma_{(x)jk}^i - \Gamma_{(x)kj}^i)$ .

<sup>11</sup>This means that we can not set derivative of the  $\Gamma$ s to zero generally.

## 8 | Parallel Transport & Curvature

Consider the following experiment: stand on a surface and stick your arm out directly in-front of you. Make a mental note at where your arm is pointing. Now walk around the room, but do it in a fashion such that you are *not* allowed to rotate your body or move the position of your arm relative to your chest. So if you want to move to your left you continue to face forward with your arm pointing forward and simply step left. Walk around the room in this fashion for however long you like and then finally return to your initial position. Now compare where your arm is pointing to where it was pointing previously. Provided you did follow the instructions, if you are on a *flat* surface your arm will be pointing in exactly the same direction as it was at the start. If you were on a curved surface, it is possible that your arm is now pointing in a different direction.

To see why the latter is true, let the surface be the surface of the earth. Imagine you start at the North Pole. You then walk<sup>1</sup> directly forwards until you reach the equator. Now *side step* to your right for a quarter turn around the equator. Finally walk backwards until you reach the North Pole again. Your arm will now be pointing at a 90 degree angle (to the right) of how it was initially.



Mathematically, what we are talking about is the directional derivative of a vector field. On the plane the vector field does not change no matter what path you take, and so the instructions of how to walk about are simply

$$\nabla_{v_\gamma} X = 0$$

where  $\gamma$  is the path you take and  $X$  is the vector field made by your arms.

The instructions on the sphere are the same, but the result is different. This gives us our first hint that the covariant derivative somehow encodes the (intrinsic) curvature of the

---

<sup>1</sup>You can walk on water for this experiment and there are no buildings or mountains etc in your way.

surface. From here we can convince ourselves that the connection is what gives our manifold ‘shape’. That is both the sphere and the potato have  $(S^2, \mathcal{O}, \mathcal{A})$  as topological manifolds but they have different curvature and so have different connections,  $\nabla_{\text{sphere}}$  and  $\nabla_{\text{potato}}$ . The aim of this lecture is to make this more precise.

### 8.1 Parallely of Vector Fields

In this lecture we shall assume that a connection has already been chosen for our manifold and so we are dealing with a smooth affine manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$ .

**Definition** (Parallely Transported). A vector field  $X$  on  $\mathcal{M}$  is said to be **parallely transported** along a smooth curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  if

$$\nabla_{v_\gamma} X = 0.$$

*Remark 8.1.1.* Note at this point it is important that we don’t need the lower slot in the covariant derivative to be a vector field over all of  $\mathcal{M}$ , as  $v_\gamma$  is only a vector field over the image of the curve.

We also have a slightly weaker condition.

**Definition** (Parallel). A vector field  $X$  in  $\mathcal{M}$  is said to be **parallel** along a curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  if

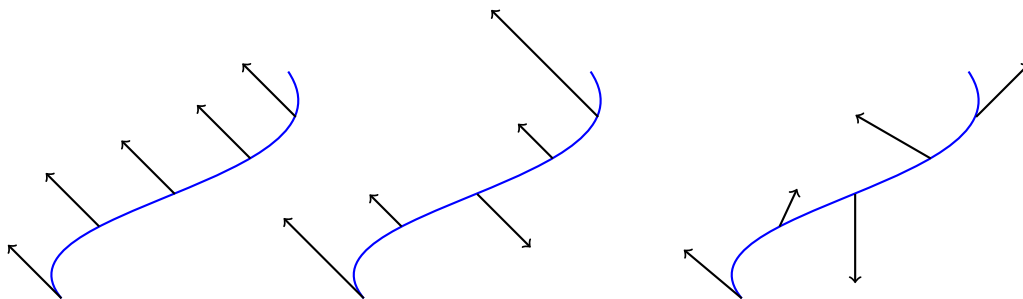
$$\nabla_{v_\gamma} X = \mu \cdot X,$$

for  $\mu : \mathbb{R} \rightarrow \mathbb{R}$  a smooth function. Written pointwise, that is

$$(\nabla_{v_{\gamma, \gamma(\lambda)}} X)_{\gamma(\lambda)} = \mu(\lambda) \cdot X_{\gamma(\lambda)}.$$

Note any parallely transported vector field is parallel – simply choose  $\mu(\lambda) = 0$  for all  $\lambda$ .

*Example 8.1.2.* Let our smooth affine manifold be the Euclidean plane  $(\mathbb{R}^2, \mathcal{O}, \mathcal{A}, \nabla_E)$ . The left drawing below is a parallely transported vector field, the middle drawing is a parallel vector field and the right drawing is not even parallel.



In the middle drawing it is important that the vector field vanishes in-between the points when it points ‘up’ vs. ‘down’, as  $\mu : \mathbb{R} \rightarrow \mathbb{R}$  is smooth.

*Remark 8.1.3.* It is tempting to look at the example above and think of the length of the vector field being constant for a parallely transported vector field whereas the length is allowed to change for a parallel vector field. Although this is intuitively very good, we as of yet have no notion of how to measure a length and so it doesn’t make sense for us to talk about the length staying the same/changing. It is just the connection that gives us the above drawings.



## 8.2 Autoparallely Transported Curves

As the name suggests, an *autoparallely* transported curve is one that is parallelly transported along itself. What we mean by this is to take the starting point of the curve and look at its tangent vector and then tell the curve to follow that direction. You then repeat this for every point along the curve. To use our person-with-their-arm-out analogy, it would be the idea of ‘follow where your arm is pointing’.

This gives us a great intuitive insight: we are travelling along the *straightest* curve between two points. Note we say straightest and not shortest, as we still don’t have a notion of length yet. Note also that the straightest line might not actually look straight when ‘viewed from above’. That is, if we embed the manifold into a higher dimensional one and then look just at the curve, it might look curved. For example, on the sphere a straight line traces out a portion of a circle around the sphere. This does not look straight in the (Euclidean) embedding, however *on the surface* it is the straightest line.

Let’s write this more formally.

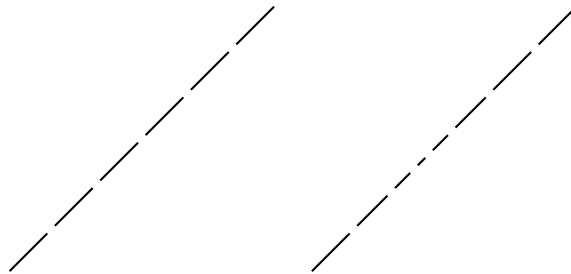
**Definition** (Autoparallely Transported). A smooth curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  is called **autoparallely transported** if

$$\nabla_{v_\gamma} v_\gamma = 0.$$

**Definition.** A smooth curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  is called an **autoparallel** if

$$\nabla_{v_\gamma} v_\gamma = \mu \cdot v_\gamma.$$

*Example 8.2.1.* Again consider the Euclidean plane  $(\mathbb{R}^2, \mathcal{O}, \mathcal{A}, \nabla_E)$ . If we represent equal parameter changes by dashes in our drawings we have the following drawings, where the left is a autoparallely transported curve and the right is just a autoparallel.



*Remark 8.2.2.* The autoparallely transported curve in the above example is what we might think of as a "uniform straight curve", and the autoparallel just a straight curve. This gives us our next nice insight. Recall that Newton’s first law talks about a moving body that experiences no forces moves along a uniform straight path. We see, then, that what Newton’s first law says is that these bodies are autoparallely transported. So we could do such an experiment and use the result to work backwards and determine what the connection is. That is, Newton’s first axiom is a measurement prescription for your geometry.

*Terminology.* People also refer to autoparallely transported vector fields as simply *autoparallels*. As we have seen this actually means a curve where we only require the right-hand side be proportional point-by-point to  $v_\gamma$ . Despite this, in these lectures we shall also adopt this terminology and (unless the case specifically requires it) simply refer to autoparallels, when we really mean autoparallely transported.

### 8.3 Autoparallel Equation

Consider an autoparallel  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  and consider the portion of the curve that lies in  $U \subseteq \mathcal{M}$  where  $(U, x) \in \mathcal{A}$ . We would like to express the condition  $\nabla_{v_\gamma} v_\gamma = 0$  in terms of chart representatives of the objects. The left-hand side is a vector field (along  $\gamma$ ) and so we can express  $v_\gamma$  in the chart as

$$v_{\gamma, \gamma(\lambda)} = \dot{\gamma}_{(x)}^m(\lambda) \cdot \left( \frac{\partial}{\partial x^m} \right)_{\gamma(\lambda)}.$$

So we have (suppressing  $(x)$  for notational convenience)

$$\begin{aligned} \nabla_{v_\gamma} v_\gamma &= \nabla_{\dot{\gamma}^m \cdot \left( \frac{\partial}{\partial x^m} \right)} \left[ \dot{\gamma}^n \cdot \left( \frac{\partial}{\partial x^n} \right) \right] \\ &= \dot{\gamma}^m \frac{\partial \dot{\gamma}^n}{\partial x^m} \frac{\partial}{\partial x^n} + \dot{\gamma}^m \dot{\gamma}^n \Gamma_{nm}^q \frac{\partial}{\partial x^q}. \end{aligned}$$

Now, all the indices are summed over and so we are free to relabel  $n \rightarrow q$  in the first term, and using  $\ddot{\gamma}^n := \dot{\gamma}^m \frac{\partial \dot{\gamma}^n}{\partial x^m}$  gives us

$$\nabla_{v_\gamma} v_\gamma = (\ddot{\gamma}^q + \dot{\gamma}^m \dot{\gamma}^n \Gamma_{nm}^q) \frac{\partial}{\partial x^q}.$$

Now we know the basis elements are linearly independent and so the autoparallely transported condition must be true for each component and so we have (reinserting all the  $(x)$ s and the  $(\lambda)$ s)

$$\ddot{\gamma}_{(x)}^i(\lambda) + \Gamma_{(x)jk}^i|_{\gamma(\lambda)} \dot{\gamma}_{(x)}^k(\lambda) \dot{\gamma}_{(x)}^j(\lambda) = 0,$$

which is the chart expression that the curve  $\gamma$  be autoparallely transported. This is a really important equation for physics, as we shall begin to see next lecture.

*Remark 8.3.1.* We know that the complete autoparallel equation transforms like a vector (as it comes from  $\nabla_{v_\gamma} v_\gamma$ , which is a vector). However we have already seen that the  $\Gamma$ s are not tensors and so do not transform nicely. We see, then, the  $\ddot{\gamma}$  must also not be a tensor itself, but must transform in such a way as to cancel the bad parts from the  $\Gamma$ s. This is an important fact to note, as one is often tempted to call  $\ddot{\gamma}$  the *acceleration* along  $\gamma$ , but it is not (as acceleration is a vector). In fact the acceleration is the complete autoparallel equation. This is actually a very nice result as it tells us that the condition for a straight line is that the acceleration along the line vanishes! It is only in a flat space, in a chart where we take the  $\Gamma$ s to all vanish that we recover  $a = \ddot{\gamma}$ . For emphasis, we also write this in the following definition. We shall also return to acceleration at the end of this lecture.

**Definition** (Acceleration). Let  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  be a smooth curve on an affine manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$ , and let  $v_\gamma$  be the velocity field along  $\gamma$ . Then the **acceleration** field along  $\gamma$  is given by

$$a_\gamma := \nabla_{v_\gamma} v_\gamma.$$

*Example 8.3.2.* Consider the Euclidean plane  $(\mathbb{R}^2, \mathcal{O}, \mathcal{A}, \nabla_E)$  and the chart  $(U, x) = (\mathbb{R}^2, \mathbb{1}_{\mathbb{R}^2})$  so that  $\Gamma_{(x)jk}^i = 0$  for all  $i, j, k = 1, 2$ . Then our autoparallel equation simply reads

$$\ddot{\gamma}_{(x)}^i(\lambda) = 0 \quad \implies \quad \dot{\gamma}_{(x)}^i(\lambda) = a^i \lambda + b^i,$$

where  $a^i, b^i \in \mathbb{R}$ . This is just what we normally think of as the equation for a straight line. Note, however, this is only valid in this chart. If we transformed to polar coordinates the  $\Gamma$ s wouldn't vanish and so the expression for  $\gamma$  would be different.

*Example 8.3.3.* Now consider the so-called round sphere<sup>2</sup>  $(S^2, \mathcal{O}, \mathcal{A}, \nabla_{\text{round}})$  and the chart  $(U, x)$  with  $x(p) = (\theta, \varphi)$ , where  $\theta \in (0, \pi)$  and  $\varphi \in (0, 2\pi)$ , which are the usual spherical coordinates.<sup>3</sup>

We define  $\nabla_{\text{round}}$  to be such that

$$\Gamma_{(x)22}^1|_{x^{-1}(\theta, \varphi)} = -\sin \theta \cos \theta, \quad \Gamma_{(x)12}^2|_{x^{-1}(\theta, \varphi)} = \Gamma_{(x)21}^2|_{x^{-1}(\theta, \varphi)} = \cot \theta,$$

and all other  $\Gamma$ s vanishing. If we now introduce the (sloppy) notation

$$x^1(p) = \theta(p), \quad \text{and} \quad x^2(p) = \varphi(p),$$

then the autoparallel equation tells us

$$\begin{aligned} \ddot{\theta} + \Gamma_{(x)22}^1 \dot{\varphi} \dot{\varphi} &= \ddot{\theta} - \sin(\theta) \cos(\theta) \dot{\varphi} \dot{\varphi} = 0 \\ \ddot{\varphi} + 2\Gamma_{(x)12}^2 \dot{\varphi} \dot{\theta} &= \ddot{\varphi} + 2 \cot(\theta) \dot{\varphi} \dot{\theta} = 0. \end{aligned}$$

Now look at solutions to these equations. One solution is

$$\theta(\lambda) = \frac{\pi}{2} \quad \text{and} \quad \varphi(\lambda) = \omega \cdot \lambda + \varphi_0,$$

for  $\omega, \varphi_0 \in \mathbb{R}$ , which is checked by direct substitution. These equations correspond to just going around the equator of the sphere. at a constant speed.

You can show that any curve that goes right round the sphere (e.g. North pole to South pole and back) will satisfy these equations. So we see that the straightest curves (that is the curves that satisfy the autoparallel equation) on the round sphere are just curves that go all the way around. This is why this choice of  $\Gamma$ s corresponds to the round sphere; we think of a round sphere as one whose straight lines behave like this.

*Remark 8.3.4.* Technically the last example is slightly wrong. This is because the chart domain  $U$  does not cover all of the round sphere but must necessarily miss off two antipodal points (e.g. North and South pole) and a straight line connecting them (e.g. a line of longitude). However, the results of the exercise are still clear.

#### Exercise

Show that the statement in the above remark is true: that  $U$  must miss out two antipodal points and a straight line connecting them.

<sup>2</sup>That is just perfect sphere, but here 'round' tells us to use the connection that gives this and not, say, the one for a potato.

<sup>3</sup>I.e.  $\theta$  is the angle from the  $z$ -axis and  $\varphi$  the angle from the  $x$ -axis. Note the  $x, y, z$ -axes are actually a coordinate system in themselves.

## 8.4 Torsion

Question: Can one use  $\nabla$  to define tensors on  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$ ?

Answer: Yes.

**Definition** (Torsion). The **torsion** of a connection  $\nabla$  is the  $(1, 2)$ -tensor field

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

where  $[\cdot, \cdot] : \Gamma T\mathcal{M} \times \Gamma T\mathcal{M} \rightarrow \Gamma T\mathcal{M}$  is the commutator<sup>4</sup> on  $\Gamma T\mathcal{M}$  given by

$$[X, Y]\langle f \rangle := X\langle Y\langle f \rangle \rangle - Y\langle X\langle f \rangle \rangle.$$

### Exercise

Prove that  $T$  is  $C^\infty$  linear in each entry, which we require if  $T$  is to be a tensor.

**Definition** (Torsion Free Connection). A affine manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$  is called **torsion free** if the torsion tensor  $T$  vanishes everywhere. One can also say that the connection is torsion free. This is often just written as

$$\nabla_X Y - \nabla_Y X = [X, Y]$$

for all  $X, Y \in \Gamma T\mathcal{M}$ .

### Exercise

Show that a torsion free manifold is one such that the  $\Gamma$ s are purely symmetric. That is show  $\Gamma^i_{[ab]} := \frac{1}{2}(\Gamma^i_{ab} - \Gamma^i_{ba}) = 0$ .

*Hint: Calculate  $T^i_{ab} = T(dx^i, \frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b})$ .*

*Remark 8.4.1.* The above exercise is exactly the result we discussed when we first introduced the  $\Gamma$ s are talked about only being able to remove the symmetric part by chart transformation.

People have tried to attach physical significance to torsion (e.g. Schrödinger's "Spacetime Structure") but in the standard theory of general relativity we do not and so from this point on-wards in the lectures<sup>5</sup> we shall only use torsion free connections.

## 8.5 Curvature

There is another, more important, tensor that we can define using our connection.

**Definition** (Riemann Curvature). The **Riemann curvature** of a connection  $\nabla$  is the  $(1, 3)$ -tensor field

$$\text{Riem}(\omega, Z, X, Y) := \omega : (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z).$$

<sup>4</sup>In fact turn this is a Lie bracket by restricting to  $\mathbb{R}$ -linearity instead of  $C^\infty$ -linearity, and define the Lie algebra of vector fields. We will do this in Lecture 11.

<sup>5</sup>Not in the tutorials, though.

Note the order of the entries, its  $(Z, X, Y)$  not  $(X, Y, Z)$ , this is just a convention that makes the right-hand side look neater.

**Definition** (Ricci Curvature). Let  $\text{Riem}$  be the Riemann curvature tensor of a connection  $\nabla$ . We define the **Ricci curvature** tensor as the  $(0, 2)$ -tensor field

$$\text{Ric}(X, Y) := \text{Riem}(e^a, Y, X, Z_a),$$

where  $e^a : Z_b = \delta_b^a$ .

In terms of components<sup>6</sup> the Ricci curvature tensor is given by

$$\text{Ric}_{ab} := \text{Riem}^c{}_{acb}.$$

*Notation.* We have define the Riemann curvature tensor with the symbol  $\text{Riem}$  and the Ricci curvature with the symbol  $\text{Ric}$ . In the literature one often sees just  $R$  used for either. This is done because one is often looking at the components and so you can easily work out which you are dealing with based on that. However, as we shall see, there is a third object called the Ricci scalar (which we can't define until we have defined metrics) which we denote  $R$ . Seeing as it is a scalar, it has no indices and so just appears as  $R$ . It is in order to avoid any potential confusion that we have decided to use  $\text{Riem}$  and  $\text{Ric}$  for these notes.

#### Exercise

Show that  $\text{Riem}$  is  $C^\infty$  linear in all its entries.

#### Exercise

Show that  $\text{Riem}$  is antisymmetric in its final two entries. That is

$$\text{Riem}(\omega, Z, X, Y) = -\text{Riem}(\omega, Z, Y, X).$$

Use this second result to show that  $\text{Riem}$  has  $d^3(d-1)/2$  independent components.

*Hint: The second two parts are done in the tutorial video.*

*Notation.* When there is no confusion about which basis<sup>7</sup> is being used we shall use the short hand notation

$$\nabla_a := \nabla_{\frac{\partial}{\partial x^a}}.$$

In light of this, we shall also use the short hand

$$\partial_a := \frac{\partial}{\partial x^a}.$$

The latter is subtle as we need to remember that the right-hand side is defined in terms of partial derivatives which are written  $\partial_i$ .

<sup>6</sup>See the tutorial for the components of  $\text{Riem}$ .

<sup>7</sup>That is when we're only dealing with one basis. If there is more than one (e.g. a change of basis calculation) it is vital to keep track of which indices are for which basis.

An algebraic relevance of Riem is the following. We have the result

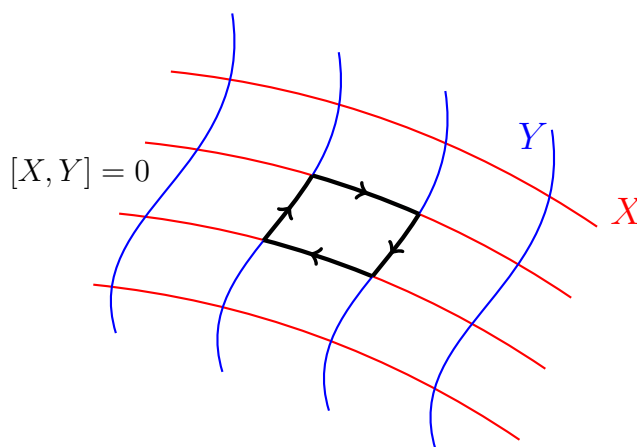
$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \text{Riem}(\cdot, Z, X, Y) + \nabla_{[X, Y]} Z.$$

If we consider a chart  $(U, x)$  and let  $X = \partial_a$  and  $Y = \partial_b$ , this becomes

$$(\nabla_a \nabla_b Z)^m - (\nabla_b \nabla_a Z)^m = \text{Riem}^m{}_{nab} Z^n,$$

where we have used  $[\partial_a, \partial_b] = 0$ .

Claim 8.5.1. The Lie bracket  $[X, Y]$  answers the question of "how well the vector fields  $X$  and  $Y$  can be coordinate vector fields". That is it tells us that if we lay  $X$  and  $Y$  on top of each other, do they form a grid? Pictorially, it asks "does the black shape close?" If  $[X, Y] = 0$  then the answer is yes.



From this claim we can get a nice geometrical idea for the Riemann curvature. The left-hand side  $(\nabla_a \nabla_b Z - \nabla_b \nabla_a Z)$  takes the vector  $Z$  from the bottom corner of the black shape and around in the direction of the arrows drawn (note the minus sign means we go down the  $X$  and left on the  $Y$ ). If Riem vanishes, the result is that the transported  $Z$  and the initial  $Z$  coincide, and therefore we haven't travelled through curvature (recall that parallel transport on a curved surface is path dependent). Whereas if Riem does not vanish then the transported  $Z$  is not the same as the initial  $Z$  and so we must have gone through curvature. Therefore the Riemann curvature tensor encodes information about the curvature of the manifold (hence the name!).

Remark 8.5.2. Note we have used a chart in order to obtain the above result and so we might be worried that Riem vanishes in one chart but not in another (e.g. Cartesian to polar). The answer is obviously that this can't happen because it is a tensor and so if it vanishes in one chart it must vanish in all charts.

**Lemma 8.5.3.** *The Riemann tensor satisfies the differential **Bianchi identity**,*

$$(\nabla_A \text{Riem})(\omega, Z, B, C) + (\nabla_B \text{Riem})(\omega, Z, C, A) + (\nabla_C \text{Riem})(\omega, Z, A, B) = 0,$$

where  $\nabla$  is torsion-free. In component form this reads

$$\nabla_c R^w{}_{zab} + \nabla_a R^w{}_{zbc} + \nabla_b R^w{}_{zca} = 0$$

**Exercise**

Prove that the Bianchi identity holds.

*Hint (from tutorial): Start by rewriting the first term only by repeated use of the Leibniz rule and one-time employment of the definition of the Riemann tensor. From this result, generate the second and third terms by mere cyclic substitution of the appropriate vectors. The rest is systematic and disciplined elimination of terms.*

# 9 | Newtonian Spacetime Is Curved

The title to this lecture sounds shocking: isn't Newtonian spacetime flat? The answer is 'yes in the standard formulation it is.' What this lecture aims to do is to express Newtonian spacetime in a new way such that gravity manifests itself as curvature. It is important to note that this is **not** general relativity, it is simply Newtonian spacetime.

Our argument is going to revolve around showing that gravity must not be considered as a force but instead it must be considered to be encoded in a curvature of the spacetime.

Recall Newton's first two laws:

- (I) A body on which **no** force acts moves uniformly along a straight line.
- (II) Deviation of a body's motion from such straight motion is effected by a force, reduced by a factor of the body's reciprocal mass.

The first thing we note is that, if read as a prescription of what a body does, the first axiom is merely a specific case of the second one (i.e. just let the force vanish in Newton II). We therefore need to read the first axiom in a different manner: you assume that a particle is not experiencing any forces and you use these particles to experimentally check what a straight line is. The first axiom is a measurement prescription for geometry.

The second important point we need to note is that, if we view gravity as a force, the first axiom is only useful if we consider a universe in which a single particle lives. That is, gravity universally acts on all massive objects and so if we have two massive particles in our universe (which our Universe clearly does<sup>1</sup>) they must both experience a force, and so Newton (I) becomes useless... Unless we stop thinking of gravity as a force.

*Remark 9.0.1.* You might think that we're being a bit pedantic here and just say 'oh ok, but we can just use Newton II and go on our merry way!' The problem with that is that Newton II talks about the deviation from a straight line, and without Newton I we don't know what a straight line is.<sup>2</sup>

## 9.1 Laplace's Question

Laplace asked the following question:

*"Can gravity be encoded in the curvature of space, such that its effects show if particles under the influence of (no other) force are postulated to move along straight lines in this curved space?"*

---

<sup>1</sup>Otherwise there is no one else to read these notes, and I have wasted some time.

<sup>2</sup>Checkmate.



The answer to this question is, unfortunately for Laplace, a resounding "no".

*Proof.* Let's consider the 'gravity as a force' point of view. We have

$$m\ddot{x}^\alpha(t) = F^\alpha(x(t)),$$

for  $\alpha = 1, 2, 3$ , and Poisson's equation for  $F^\alpha = mf^\alpha$

$$-\partial_\alpha f^\alpha = 4\pi G\rho.$$

Substituting in  $F^\alpha = mf^\alpha$  we see that we can cancel out the  $m$ s to get a relationship between the acceleration and the force that is independent of mass. This is an experimentally verified fact (see [video at 16:50-20:00](#) if you aren't familiar with such an experiment), and is given the name 'weak equivalence principle'.

So Laplace's question becomes is

$$\ddot{x}^\alpha(t) - f^\alpha(x(t)) = 0$$

of the form of a autoparallel equation. That is is it of the form

$$\ddot{x}^\alpha(t) + \dot{x}^\beta(t)\dot{x}^\gamma(t)\Gamma^\alpha_{\beta\gamma}(x(t)) = 0?$$

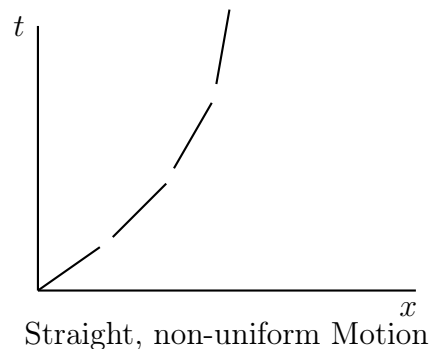
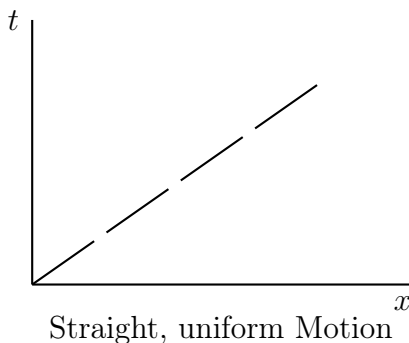
The answer is no, because  $f^\alpha$  is only a function of  $x(t)$  and no its derivatives, but the second term in the autoparallel equation contains derivatives. Along with this, the  $\Gamma$ s are only dependent on  $x(t)$  and so can't cancel out this velocity dependence and so it is just not possible to equate the two expressions.

So we cannot find  $\Gamma$ s such that Newton's equation takes the form of an autoparallel, and since the  $\Gamma$ s are what determine the connection, which we have seen is related to the curvature, we cannot encode the effect of gravity as a curvature in this way.  $\square$

## 9.2 The Full Wisdom of Newton I

We have just shown that the answer to Laplace's question was no, so why did we bother to talk about it? The answer is that it highlights its flaw and then allows us to see how to change it in order to get something correct. The problem was that Laplace didn't read Newton I careful enough. Newton I does not just talk about motion but about *uniform* motion.

Uniform motion involves understanding how something moves in *time* as well as space. Uniform motion is plotted as a straight line on a space-time graph, whereas straight, but not uniform, motion is given by a curve.



In a spacetime picture, then, straight, uniform motion in space is simply just straight motion. So our idea is to alter Laplace's question to be "... curvature of spacetime, ...", and then repeat the process. Again note that here we are talking about Newtonian spacetime, this is not general relativity!

For motion in space we had the particles motion given by  $x : \mathbb{R} \rightarrow \mathbb{R}^3$ . We need to convert this into the particles *worldline*, which we get from the map  $X : \mathbb{R} \rightarrow \mathbb{R}^4$  given by

$$X(t) = (t, x^1(t), x^2(t), x^3(t)) := (X^0(t), X^1(t), X^2(t), X^3(t)).$$

We haven't done anything new, we have simply just turned the parameterisation of the curve (the 'time') into a coordinate and considered the spacetime picture.

Claim 9.2.1. By doing the above, the answer to the modified Laplace's question is "yes".

*Proof.* Assume that (note it is the little  $x$  here)

$$\ddot{x}^\alpha(t) = f^\alpha(x(t))$$

for  $\alpha = 1, 2, 3$ , still holds. We now have the trivial result

$$\dot{X}^0(t) = 1, \quad \implies \quad \ddot{X}^0(t) = 0.$$

We can rewrite the Newton equation in terms of the big  $X$  as<sup>3</sup>

$$\ddot{X}^\alpha(t) = f^\alpha(X(t)),$$

for  $\alpha = 1, 2, 3$ . Now, we can multiply by  $\dot{X}^0(t)$  because its equal to 1, and so we have

$$\ddot{X}^\alpha(t) - f^\alpha(X(t))\dot{X}^0(t)\dot{X}^0(t) = 0.$$

Now combing this with the  $\ddot{X}^0(t) = 0$  equation, we see that we have a autoparallel equation

$$\ddot{X}^a + \Gamma^a_{bc}\dot{X}^b\dot{X}^c = 0$$

where  $a, b, c = 0, 1, 2, 3$ . This is seen by choosing all of the  $\Gamma$ s to vanish apart from

$$\Gamma^\alpha_{00} = -f^\alpha \quad \forall \alpha = 1, 2, 3.$$

Now this could just be a coordinate-choice artefact, and so could be transformed away. In terns out that this is not the case, and you can show it by calculating the Riemann curvature tensor components. The only non-vanishing ones are

$$\text{Riem}^\alpha_{\ 0\beta 0} = -\partial_\beta \langle f^\alpha \rangle.$$

As this is a tensor, if it is non-vanishing in one chart it must be non-vanishing in all charts.  $\square$

<sup>3</sup>Note technically  $f^\alpha(X(t))$  is a new function, but we just define it to be such that it ignores the first entry.

*Remark 9.2.2.* Given the Riemann tensor at the end of the proof above, we can actually work out the Ricci tensor, given by setting  $\alpha = \beta$ ,

$$\text{Ric}_{00} = -\partial_a \langle f \rangle,$$

which, using the Poisson equation gives

$$\text{Ric}_{00} = 4\pi G\rho.$$

This is actually one of the so-called Einstein equations

$$\text{Ric}_{00} = 8\pi GT_{00},$$

where  $T_{00} = \rho/2$ .  $T$  is known as the *energy-momentum tensor*, we shall meet this in much more detail later on.

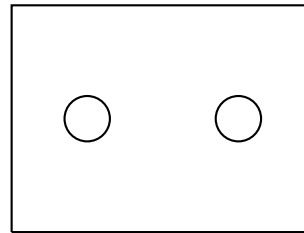
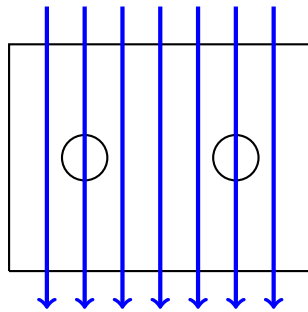
*Remark 9.2.3.* Note the fact that the only non-vanishing  $\Gamma$ s have the lower indices both ‘time’ (i.e. they are 0), it tells us that the curvature is taking place in *spacetime*, not just in space. That is the Riemann tensor vanishes for all spatial indices  $\text{Riem}^\alpha_{\beta\gamma\delta} = 0$  for all  $\alpha, \beta, \gamma, \delta = 1, 2, 3$ .

### 9.2.1 Tidal Forces

The result above about not being able to transform away the curvature result is known as tidal forces. The basic idea is that you can only transform away gravitational fields locally. In other words, the only way you can transform away a gravitational field globally is if it is uniform.

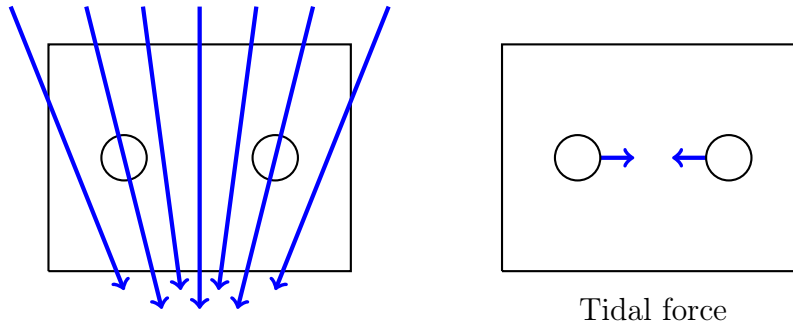
To see why this is the case, imagine being inside a box in space with two balls. now imagine the box is in a gravitational field, and so is in free fall towards some massive object. We shall ignore the gravitational fields generated by our body and by the balls themselves. If the gravitational field is uniform across the box, everything experiences the same pull and so falls exactly the same. That is, if we put the balls out at our sides, they would appear to just float there, and if there was no windows on our box to see things moving past us, we actually wouldn’t even know we were in a gravitational field. Obviously someone sat stationary (w.r.t the massive object) outside the box would see the balls moving down and so would say they are in a gravitational field.

What is going on here is that we have transformed ourselves to a frame of reference (which for this remark is just a chart) which falls with the balls and so we have ‘removed’ the effects of gravity via such a change of chart.



Gravitational effect transformed away

Now imagine we do the same thing, but the gravitational field is not uniform, but comes radially from some spherical object. Again everything still falls at the same rate, but now the ball to our left will be pulled slightly to the right and the ball to our right will be pulled slightly left. To us inside the box, then, the balls slowly move towards each other. This is not an effect that we can remove by going to another frame of reference, and so represents something physical. This is ‘real’ gravity.



The inability to remove this effect by a change of chart is what we refer to as a tidal force.<sup>4</sup> From this we see that when we feel gravity pulling us, it’s actually the inhomogeneous nature of the gravity we feel; it pulls our feet harder than it pulls our head and it pushes our arms towards each other.

### 9.3 The Foundation of the Geometric Formulation of Newton’s Axioms

So far we have managed to change our thinking of gravity as a force into thinking of it as being part of a curvature of spacetime. This is done so that Newton’s first axiom, which now reads "the worldline of a body on which no force acts is a straight line in spacetime", can be taken a measurement prescription for what a straight line is. The problem is, we have had indices flying about everywhere and so have been committing the crime of relying on charts!

We are now going to rederive our result without making reference to a chart at all. We are doing this afresh, and so should not use the results we just obtained (e.g.  $\Gamma^{\alpha}_{00} = -f^{\alpha}$ ). In order to do this, we need to introduce a few definitions.

**Definition** (Newtonian Spacetime). A **Newtonian spacetime** is a quintuple of structures  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla, t)$  where  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  is a 4-dimensional smooth manifold and  $t : \mathcal{M} \rightarrow \mathbb{R}$  is a smooth function called the **absolute time**, which satisfies:

- (i)  $(dt)_p \neq 0$  for all  $p \in \mathcal{M}$  — there is a concept of *absolute space* (defined below),
- (ii)  $\nabla dt = 0$  everywhere — absolute time flows uniformly,
- (iii)  $\nabla$  is torsion free.

**Definition** (Absolute Space). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla, t)$  be a Newtonian spacetime. **Absolute space** at time  $\tau$  is the set

$$S_{\tau} := \{p \in \mathcal{M} \mid t(p) = \tau\}.$$

<sup>4</sup>The name derives from the fact that it is due to this that the moon creates tides in the oceans/seas.

It follows that

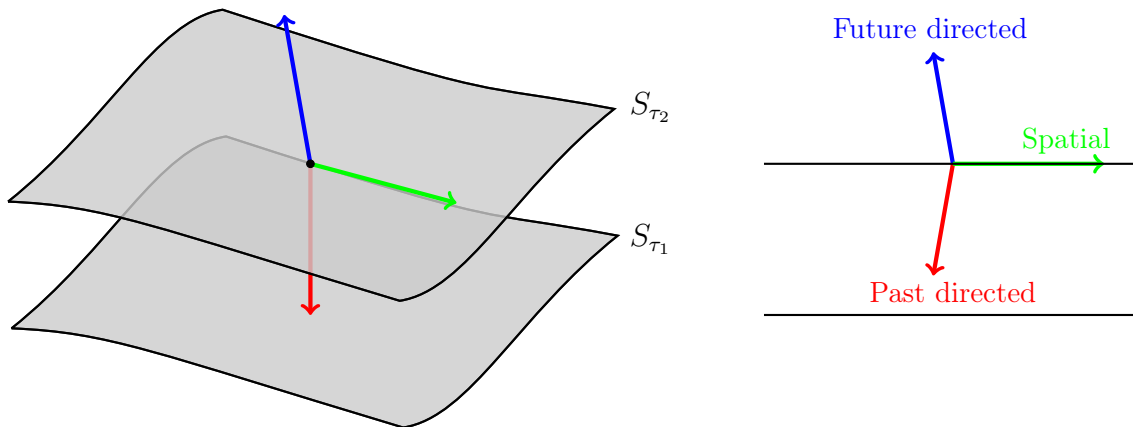
$$\mathcal{M} = \dot{\bigcup}_{\tau} S_{\tau}.$$

Condition (i) in the definition of Newtonian spacetime is what gives us the disjoint union in the definition of absolute time. That is, condition (i) says the surfaces of absolute space at different times must not meet, as if they did the gradient of  $t$  would vanish. Note it is only once we introduce the absolute time function that we can think of splitting spacetime into space and time, before that it was just a 4-dimensional manifold.

**Definition** (Future Directed / Spatial / Past Directed). A vector  $X \in T_p\mathcal{M}$  is called

- (i) **Future directed** if  $dt : X > 0$ ,
- (ii) **Spatial** if  $dt : X = 0$ , and
- (iii) **Past directed** if  $dt : X < 0$ .

We see the above definition nicely pictorially. Let  $\tau_2 > \tau_1$ , then we have the following picture.



We can now reword Newton's laws as

- (I) The worldline of a particle under the influence of no force (gravity is not one here) is a future directed autoparallel. That is  $\nabla_{v_\gamma} v_\gamma = 0$  and  $dt : v_\gamma > 0$  everywhere.
- (II) The acceleration along a worldline is

$$a_\gamma := \nabla_{v_\gamma} v_\gamma = \frac{F}{m},$$

where the force,  $F$ , is a spatial vector field,  $dt : F = 0$ , and where  $m$  is the mass of the particle.

## 9.4 Acceleration

*Convention.* Restrict attention to atlases  $\mathcal{A}_{\text{stratified}}$  where the chart  $(U, x)$  have the property that  $x^0 = t|_U$ . That is the first chart map coincides with the absolute time function. This convention, along with condition (ii) in the definition of Newtonian spacetime gives us

$$0 = (\nabla_a dx^0)_b = -\Gamma^0_{ba}$$

for  $a, b = 0, 1, 2, 3$ . So in a stratified atlas all the  $\Gamma$ s with an upper 0 index vanish.

Let's now evaluate Newton II in a stratified atlas. Let  $X(\lambda)$  denote the particle's worldline, then we have

$$\nabla_{v_X} v_X = \frac{F}{m}.$$

We have

$$(X^\alpha)'' + \Gamma^\alpha_{\gamma\delta}(X^\gamma)'(X^\delta)' + 2\Gamma^\alpha_{0\gamma}(X^\gamma)'(X^0)' + \Gamma^\alpha_{00}(X^0)'(X^0)' = \frac{F^\alpha}{m},$$

for  $\alpha = 1, 2, 3$ , where we have used the fact that Newtonian spacetime is torsion free and so the  $\Gamma$ s are symmetric in the lower indices.

Now, using the fact that  $F$  is a spatial vector field (so  $F^0 = 0$ ) we also have

$$\begin{aligned} (X^0)'' + \Gamma^0_{ab}(X^a)'(X^b)' &= 0 \\ (X^0)'' &= 0 \\ \implies X^0(\lambda) &= a\lambda + b \\ (t \circ X)(\lambda) &= a\lambda + b, \end{aligned}$$

for  $a, b \in \mathbb{R}$ . This gives us the idea that we can reparameterise our curve in terms of the absolute time, and we get

$$\frac{d}{d\lambda} \longrightarrow a \frac{d}{dt}.$$

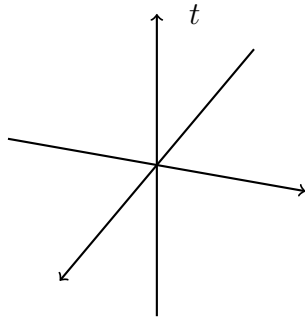
Subbing this into the expression for the spatial components to give

$$\ddot{X}^\alpha + \Gamma^\alpha_{\gamma\delta}\dot{X}^\gamma\dot{X}^\delta + 2\Gamma^\alpha_{0\gamma}\dot{X}^0\dot{X}^\gamma + \Gamma^\alpha_{00}\dot{X}^0\dot{X}^0 = \frac{F^\alpha}{a^2m}.$$

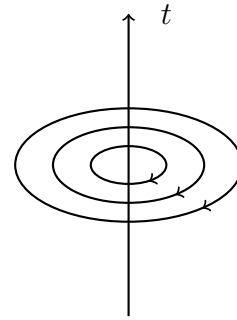
Now recalling Remark 8.3.1, we see that it is the *entire* left-hand side that is the ( $\alpha$  component of the) acceleration, *not* just  $\ddot{X}^\alpha$ . This is a really profound result and it explains a lot of the stuff you hear about lower down in education.

First we note that the  $\Gamma^\alpha_{00}$  term is non-zero in the presence of gravity, it is  $-f^\alpha$ . So let's assume there is no gravity so this term vanishes. Now, there exists a chart such that all the  $\Gamma$ s vanish and we are simply left with  $\ddot{X}^\alpha = F^\alpha/a^2m$ , which is our usual result. However if we simply just choose another chart,  $\Gamma$ s will start to appear! Obviously physically nothing has changed, but it appears that looking at the problem in different ways introduces new 'accelerations' (quotation marks because we know they aren't real accelerations, only their sum is). These are charts in spacetime, not just space and so we need to make sure we account for this.

The  $\Gamma^\alpha_{\gamma\delta}$  terms arise if we simply choose another coordinate system, e.g. instead of considering Cartesian coordinates we could use polar coordinates for the spatial part and leave time unchanged.



All  $\Gamma$ s vanish



$\Gamma^{\alpha}_{\gamma\delta}$  terms present

The  $\Gamma^{\alpha}_{0\gamma}$  terms arise when your charted spatial slices ‘move’ in time. For example if the chart is made up of lots of spatial slices that rotate about the time axis, which we’ll call a rotating chart. It is important to note that it is only the chart that rotates, it is not that the actual, real world, spatial slices are rotating. In this case both the  $\Gamma^{\alpha}_{00}$  and  $\Gamma^{\alpha}_{0\gamma}$  terms appear and they represent the so-called *centrifugal* and *Coriolis* pseudo-accelerations. The ‘pseudo’ tells us that something is not quite right about them being accelerations, and now we understand why: they are not accelerations in themselves but only the sum is an acceleration.

# 10 | Metric Manifolds

We establish a structure on a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  that allows us to assign vectors in each tangent space a length and<sup>1</sup> an angle between vectors in the same tangent space. Such a structure in each tangent space is a *inner product*, and the complete structure over all tangent spaces is what we call a *metric* (i.e. it is a inner produce field).

From this structure, one can then define a notion of length of a curve. Then we can look at shortest (and longest) curves, which are known as *geodesics*. We will develop this completely independently of the notion of straight curves, i.e. of the covariant derivative, but then shall insist, for obvious reasons, at the end that the two coincide. In doing this we will define what we mean by so-called *metric compatible* connections.

## 10.1 Metrics

**Definition** (Metric). A **metric**  $g$  on a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  is a  $(0, 2)$ -tensor field satisfying:

- (i) It's symmetric;  $g(X, Y) = g(Y, X)$ , for all  $X, Y \in \Gamma T\mathcal{M}$ , and
- (ii) Non-degeneracy; the map  $\flat : \Gamma T\mathcal{M} \rightarrow \Gamma T^*\mathcal{M}$ , given by

$$\flat(X) : Y := g(X, Y),$$

is a  $C^\infty$ -isomorphism, i.e. it is invertible and smooth in both directions.

**Definition** (Inverse Metric). The **inverse metric**,  $g^{-1} : \Gamma T^*\mathcal{M} \times \Gamma T^*\mathcal{M} \xrightarrow{\sim} C^\infty(\mathcal{M})$ , w.r.t. a metric  $g$  is the symmetric,  $(2, 0)$ -tensor field defined by

$$g^{-1}(\omega, \sigma) := \omega : \flat^{-1}(\sigma).$$

*Remark 10.1.1.* One needs to be careful when referring to  $g^{-1}$  as an inverse. It is not an inverse in the sense of a map, but in the matrix sense. That is the map inverse of  $g : \Gamma T\mathcal{M} \times \Gamma T\mathcal{M} \xrightarrow{\sim} C^\infty(\mathcal{M})$  would be a map from  $C^\infty(\mathcal{M})$  to  $\Gamma T\mathcal{M} \times \Gamma T\mathcal{M}$ , which  $g^{-1}$  is not. If we denote<sup>2</sup> the components of  $g^{-1}$  simply as  $g^{ab}$  (so no  $-1$  in it), then what we mean by inverse is that the following holds:

$$g^{ac} g_{cb} = \delta_b^a.$$

<sup>1</sup>If we were only looking to define a length, we would just define a norm on our manifold. For more information on norms see Dr. Schuller's Lectures on Quantum Theory.

<sup>2</sup>And we will from this point on-wards.



*Remark 10.1.2.* It is very common for people to talk about ‘raising/lowering’ indices using the metric/inverse metric. What they mean is the idea that  $\flat(X)$  is a covector and so has a covariant index. However, we’re lazy and so we don’t want to have to keep writing the  $\flat$  bit and so we just write

$$X_a := g_{am} X^m,$$

and similarly for a covector made into a vector via  $\flat^{-1}$ . The clear problem is that, unless specified, we don’t know whether  $T_a$  are the components of a covector defined completely independently of the metric or whether they are the ‘lowered’ components of a vector, and so is dependent on the metric. In these notes we shall never suppress the metric and so shall not talk about ‘raised/lowered’ indices.

*Example 10.1.3.* Consider the smooth manifold  $(S^2, \mathcal{O}, \mathcal{A})$  with the chart  $(U, x)$  corresponding to spherical coordinates  $(\theta, \varphi)$ .<sup>3</sup> Define the metric as<sup>4</sup>

$$g_{ij}(x^{-1}(\theta, \varphi)) := \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{pmatrix}_{ij}.$$

This gives us the *round sphere of radius  $R$* . Note, just as with the connection, defining a metric allows us to give the manifold shape. Note also, though, that, unlike the round sphere obtained using the connection, we can talk about the *size* of the round sphere obtained from the metric.

## 10.2 Signature

Recall the eigenvalue equation

$$Av = \lambda \cdot v,$$

where  $v$  is an eigenvector. If we want to express this in terms of components, it is clear that  $A$  must be a  $(1, 1)$ -tensor, otherwise we break Einstein summation convention. That is

$$A^a{}_m v^m = \lambda \cdot v^a$$

is the only valid index placement. If we represent  $A$  as a matrix, it is a well known result of linear algebra that we can bring it to the form

$$A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d),$$

where  $d$  is the dimension of the vector space.

We want to have a similar thing for tensors of different ranks. For  $(0, 2)$ -tensors, in particular the metric, we have the *signature* of  $g$  which has only  $+1$ ,  $-1$ , and  $0$  on the diagonal. One should be careful, though, as these are not simply eigenvalues for  $g$ ; the first reason being we just argued that eigenvalues only make sense for  $(1, 1)$ -tensors, and besides that it turns

<sup>3</sup>Again this chart does not cover the whole manifold, but requires that we remove two antipodal points and a line of longitude. Note also that we have not equipped our manifold with a connection and so it doesn’t actually have a shape yet!

<sup>4</sup>The notation here just means that we collect the 4 components of  $g$  into a matrix where  $i$  tells us the row and  $j$  tells us the column. For more information on this see, for example, section 1.5 of *Manifolds, Tensors, and Forms: An Introduction for Mathematicians and Physicists* by Paul Renteln.

out that we can *always* transform to a chart such that the signature is *only*  $+1, -1$  and  $0$ s, which you can not do for eigenvalues in general.

Actually what we want to define as the signature is the double  $(p, q)$  where  $p$  is the number of  $+1$ s and  $q$  is the number of  $-1$ s. This is the definition we shall use in these notes, but it is important to be aware that others refer to different, but related, things as the signature.<sup>5</sup>

Claim 10.2.1. The signature is independent of the choice of chart. That is the values of  $p$  and  $q$  do not depend on what basis you use in order to write down the matrix components.

Notation. We shall use the standard notation of  $+s$  and  $-s$  as a  $d$ -tuple to indicate the signature. For example if  $d = 3$ ,  $p = 2$  and  $q = 1$  we write  $(+, +, -)$ . The position in the tuple corresponds to the corresponding metric components. So for our example  $g_{11} = 1$ ,  $g_{22} = 1$  and  $g_{33} = -1$  in this basis.

Remark 10.2.2. What we are really interested in is the *relative* sign between the components of the metric, and so we could easily have switched  $p \leftrightarrow q$  in the definition and proceeded from there, i.e. our example in the above notation would become  $(-, -, +)$ . It does not matter which we pick, as long as we are consistent. Our choice of signature is given by the following two definitions.

**Definition (Riemannian Metric).** A metric is called **Riemannian** if its signature is  $(+, +, \dots, +)$ .

A metric with any other signature (apart from  $(-, -, \dots, -)$  of course) is called *pseudo-Riemannian*. Of particular importance in general relativity is the following case.

**Definition (Lorentzian Metric).** A metric is called **Lorentzian** if its signature is  $(+, -, \dots, -)$ .

Remark 10.2.3. The convention given for the Riemannian metric is almost always the one used, however for Lorentzian metrics it's about a 50/50 split between people who use  $(+, -, \dots, -)$  and people who use  $(-, +, \dots, +)$ . We will use the one given in the definition.<sup>6</sup>

#### Exercise

Show that the metric non-degeneracy condition for a Riemannian metric is equivalent to the non-degeneracy of an inner product. That is

$$g(X, Y) = 0, \quad \forall Y \in \Gamma TM \quad \iff \quad X = 0.$$

*Hint: Think about what it means for a matrix to be invertible and then decompose  $X$  and  $Y$  in a basis.*

<sup>5</sup>For example some people call the single number  $(-1)^q$  the signature. This just tells you whether there is an even or odd number of  $-1$ s. This convention is only used when you are considering the case when there are no  $0$ s, which is true if the tensor is non-degenerate, e.g. the metric.

<sup>6</sup>Personally I prefer the second one, as I prefer to think of spatial lengths as positive (this statement will make sense shortly), however I shall stick with Dr. Schuller's for consistency with the videos. This is just a footnote as a warning that I might (but hopefully won't) use the wrong convention in a calculation later.

### 10.3 Length Of A Curve

Let  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  be a smooth curve, then we know its velocity  $v_{\gamma, \gamma(\lambda)}$  at each  $\gamma(\lambda) \in \mathcal{M}$ . On a topological manifold this is as far as we can go, but on a metric manifold we have the following.

**Definition** (Speed of a Curve). On a Riemannian metric manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$ , the **speed** of a curve at  $\gamma(\lambda)$  is the number

$$s(\lambda) := \sqrt{g(v_\gamma, v_\gamma)} \Big|_{\gamma(\lambda)}$$

*Remark 10.3.1.* Although we might expect the velocity components  $v^a$  to have units  $LT^{-1}$ , this is not true; they have units  $T^{-1}$ . The apparent ‘loss’ of the distance comes from the fact that the components are chart dependent objects and the distance in a chart is physically meaningless and so we cannot attach physical units to it. The *speed*, however, does have units  $LT^{-1}$ , which tells us that the metric components must have units  $L^2$ .

**Definition.** Let  $\gamma : (0, 1) \rightarrow \mathcal{M}$ <sup>7</sup> be a smooth curve. Then the **length** of  $\gamma$  is the number<sup>8</sup>

$$L[\gamma] := \int_0^1 d\lambda s(\lambda).$$

What we have just seen is that the velocity is actually the fundamental object and from it we derive the speed and from that we get the length of a curve. This is entirely opposite to what we learn lower down in school!

*Example 10.3.2.* Reconsider the round sphere of radius  $R$ . Consider its equator,

$$\begin{aligned} \theta(\lambda) &:= (x^1 \circ \gamma)(\lambda) = \frac{\pi}{2} \\ \varphi(\lambda) &:= (x^2 \circ \gamma)(\lambda) = 2\pi\lambda^3. \end{aligned}$$

The length of this curve is

$$L[\gamma] = \int_0^1 d\lambda \sqrt{g_{ij}(x^{-1}(\theta(\lambda), \varphi(\lambda))) (x^1 \circ \gamma)'(\lambda) (x^2 \circ \gamma)'(\lambda)}.$$

Using

$$g_{ij} = \text{diag}(R^2, R^2 \sin^2 \theta), \quad \theta'(\lambda) = 0, \quad \text{and} \quad \varphi'(\lambda) = 6\pi\lambda^2,$$

we have

$$\begin{aligned} L[\gamma] &= \int_0^1 d\lambda \sqrt{R^2 \sin^2(\theta(\lambda)) 36\pi^2 \lambda^4} \\ &= 6\pi R \int_0^1 \sin(\pi/2) \lambda^2 \\ &= 6\pi R \cdot \frac{1}{3} \\ &= 2\pi R. \end{aligned}$$

<sup>7</sup>We are free to choose the domain of  $\gamma$  to be  $(0, 1)$  by simply rescaling/shifting  $\lambda$  accordingly.

<sup>8</sup>We have used square brackets around  $\gamma$  below because it is a function. This tells us that  $L$  is a so-called *functional*. Anyone unfamiliar with this terminology is referred to a course on Lagrangian mechanics.

Note that although we used a seemingly funny parameterisation (i.e.  $\lambda^3$  not just  $\lambda$ ) in the above example, the answer still came out as we would like. This is obviously because nowhere in the definitions above did we talk about how we parameterise the curve. Physically it makes sense that the length of the curve is independent of it: the length of your walk does not depend on how quickly you do it, or if you even do it at a constant speed (provided you don't turn around and walk backwards on yourself at any point). This can be written nicely as the following theorem.

**Theorem 10.3.3.** *Let  $\gamma : (0, 1) \rightarrow \mathcal{M}$  be a smooth curve and let  $\sigma : (0, 1) \rightarrow (0, 1)$  be smooth, bijective and increasing, then  $L[\gamma] = L[\gamma \circ \sigma]$ .*

## 10.4 Geodesics

**Definition** (Geodesic). A curve  $\gamma : (0, 1) \rightarrow \mathcal{M}$  is called a **geodesic** on a Riemannian manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  if it is a *stationary*<sup>9</sup> curve w.r.t. the length functional  $L$ .

**Theorem 10.4.1.** *The curve  $\gamma : (0, 1) \rightarrow \mathcal{M}$  is a geodesic if and only if it satisfies the Euler Lagrange equations for the Lagrangian*

$$\begin{aligned} \mathcal{L} : T\mathcal{M} &\rightarrow \mathbb{R} \\ X &\mapsto \sqrt{g(X, X)}. \end{aligned}$$

In a chart, this is

$$\mathcal{L}(\gamma^i, \dot{\gamma}^i) = \sqrt{g_{ij}(\gamma(\lambda))\dot{\gamma}^i(\lambda)\dot{\gamma}^j(\lambda)}.$$

Finding the Euler Lagrange equations proceeds as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m} &= \frac{1}{\sqrt{\dots}} g_{mj}(\gamma(\lambda)) \dot{\gamma}^j(\lambda) \\ \therefore \frac{d}{d\lambda} \left( \frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m} \right) &= \frac{d}{d\lambda} \left( \frac{1}{\sqrt{\dots}} \right) g_{mj}(\gamma(\lambda)) \dot{\gamma}^j(\lambda) + \frac{1}{\sqrt{\dots}} \left( g_{mj}(\gamma(\lambda)) \ddot{\gamma}^j(\lambda) + \dot{\gamma}^s(\lambda) [\partial_s g_{mj}(\gamma(\lambda))] \dot{\gamma}^j(\lambda) \right). \end{aligned}$$

Now we're stuck with the ugly task of trying to work out  $\frac{d}{d\lambda} \left( \frac{1}{\sqrt{\dots}} \right)$ . However, we have already demonstrated that the length of the curve is independent on how we choose to our parameter. We are free, therefore, to choose it to be something convenient, and we simply take it to be such that  $g(\dot{\gamma}, \dot{\gamma}) = 1$ , that is the speed it one along the whole curve. We then just have

$$\frac{d}{d\lambda} \left( \frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m} \right) = g_{mj}(\gamma(\lambda)) \ddot{\gamma}^j(\lambda) + \dot{\gamma}^s(\lambda) \left( \partial_s g_{mj}(\gamma(\lambda)) \right) \dot{\gamma}^j(\lambda).$$

We also need to find

$$\frac{\partial \mathcal{L}}{\partial \gamma^m} = \frac{1}{2} \left( \partial_m g_{ij}(\gamma(\lambda)) \right) \dot{\gamma}^i(\lambda) \dot{\gamma}^j(\lambda),$$

where we have already imposed our parameter choice condition. So our Euler Lagrange equations are (dropping the  $(\lambda)$ s for notational brevity)

$$g_{mj} \ddot{\gamma}^j + (\partial_i g_{mj}) \dot{\gamma}^i \dot{\gamma}^j - \frac{1}{2} (\partial_m g_{ij}) \dot{\gamma}^i \dot{\gamma}^j = 0.$$

<sup>9</sup>In the sense of Lagrangians in classical mechanics.

Multiplying both sides by the inverse metric  $g^{mq}$  and using the condition  $g^{mq}g_{mj} = \delta_j^q$ , we have

$$\ddot{\gamma}^q + g^{qm} \left( \partial_i g_{mj} - \frac{1}{2} \partial_m g_{ij} \right) \dot{\gamma}^{(i} \dot{\gamma}^{j)} = 0,$$

where the brackets on the last two indices indicate the symmetry  $\dot{\gamma}^i \dot{\gamma}^j = \dot{\gamma}^j \dot{\gamma}^i$ . Using this symmetry we can double the first term by switching  $i \leftrightarrow j$ , giving us

$$\ddot{\gamma}^q + \frac{1}{2} g^{qm} (\partial_i g_{mj} + \partial_j g_{mi} - \partial_m g_{ij}) \dot{\gamma}^{(i} \dot{\gamma}^{j)} = 0.$$

This is the **geodesic equation** for the components of  $\gamma$  in a chart. We can write this in the form of an autoparallel<sup>10</sup> equation by introducing the following definition.

**Definition** (Christoffel Symbols). Given a metric  $g$  and a chart  $(U, x)$ , we define the **Christoffel symbols** (or **Levi-Civita connection coefficients**) as

$${}^{LC}\Gamma_{ij}^q(\gamma(\lambda)) := \frac{1}{2} g^{qm} (\partial_i g_{mj} + \partial_j g_{mi} - \partial_m g_{ij}),$$

where the components of the metric (and its inverse, obviously) are taken in the chart given.

*Notation.* We can lighten the notation slightly by defining

$$g_{ij,m} := \partial_m g_{ij},$$

and similarly for any other tensor rank. That is, we simply denote a partial derivative in a chart by a comma and the index follows it is the derivative entry. This notation is very useful, as it can be used along side the semi-colon notation for the covariant derivative

$$T_{jk;i} := (\nabla_i T)_{jk}$$

We shall adopt this notation in these notes, however the reader is warned that Dr. Schuller does not use this notation, and so to make sure they can transition between the two when comparing these notes to the lectures.

This process, specifically the point at which we say that the  ${}^{LC}\Gamma$ s come from a connection,  ${}^{LC}\nabla$ , identifies the shortest<sup>11</sup> curves (geodesics) with straight curves (autoparallels). This is clearly a physically very reasonable, and correct, thing to do. It is important to note, though, that up until this point, geodesics and autoparallels are completely separate entities.

Note by making this identification, we obtain the connection *from* the metric. That is, we do not need to provide both a metric and a connection, but by simply providing a metric we can obtain a unique connection such that the shortest curves and the straight curves coincide. This sounds like a chart dependent thing, and therefore not a good thing to do. However the following theorem puts our minds to rest on this point, letting us know everything is OK.

**Theorem 10.4.2.** *Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g, \nabla)$  be a topological manifold equipped with both a metric and a connection. If*

<sup>10</sup>I.e. in the form  $\ddot{\gamma}^q + \Gamma_{ij}^q \dot{\gamma}^i \dot{\gamma}^j$ .

<sup>11</sup>Again strictly speaking they're are just maximal curves, so it's also true for longest curves and curves corresponding to points of inflection.

(i)  $\nabla$  is torsion free, and

(ii)  $\nabla g = 0$ , known as metric compatibility,

then we can conclude  $\nabla =^{LC} \nabla$ .

*Proof.* See tutorials. □

#### Exercise

Show that the metric compatibility condition allows us to ‘move the metric in and out of the covariant derivative’. That is,

$$g \cdot \nabla T = \nabla g \cdot T.$$

Finally for this lecture, let’s introduce some definitions. As we see, all of them are directly related to the metric.

**Definition** (Riemann Christoffel Curvature). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  be a metric manifold. The components of the **Riemann Christoffel curvature** are defined by

$$\text{Riem}_{abcd} := g_{am} \text{Riem}^m{}_{bcd},$$

where  $\text{Riem}^m{}_{bcd}$  are the Riemann tensor components obtained from the Levi-Civita connection  $^{LC}\nabla$ .

**Definition** (Ricci Scalar). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  be a metric manifold and let  $\text{Riem}$  be the Riemann tensor obtained from the Levi-Civita connection. We then define the **Ricci scalar** as

$$R = g^{ab} \text{Ric}_{ab},$$

where  $\text{Ric}_{ab} := \text{Riem}^c{}_{acb}$  are the components of the Ricci curvature tensor.

**Definition** (Einstein Curvature). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  be a metric manifold and let  $\text{Riem}$  be the Riemann tensor obtained from the Levi-Civita connection. We define the components of the **Einstein curvature** as

$$G_{ab} := \text{Ric}_{ab} - \frac{1}{2} g_{ab} R,$$

where  $\text{Ric}$  and  $R$  are the Ricci curvature and Ricci scalar, respectively.

It is important to note that these quantities are not only related to the metric through its direct appearance in the expressions, but also through the fact that, in order to define the Riemann curvature tensor we need a connection and for all of them we used the Levi-Civita connection, a metric dependent object. For this latter reason, the Ricci curvature tensor (defined previously) is also a metric dependent object.

As all of the names above suggest, we have just established a link between the curvature of the spacetime and the metric structure.<sup>12</sup> This is the first major step into understanding the main principles of general relativity: that matter generates curvature on the spacetime.

<sup>12</sup>In fact we made this identification the moment we insisted geodesics and autoparallels coincided, as we then established a link between the metric and the covariant derivative, which we’ve seen encodes curvature.

*Remark 10.4.3.* The above definitions can, of course, all be expressed in a chart free manner as they are tensors, however the notation can be a bit confusing and it's much easier to see in component form, hence why we have defined it this way.

#### Exercise

Show that  $\text{Riem}_{abcd}$  have the correct transformation behaviour the components of a  $(0, 4)$ -tensor field. From this it follows analogously that the remaining definitions are indeed tensors.

# 11 | Symmetry

We have the intuitive feeling that the round sphere of radius  $R$ ,  $(S^2, \mathcal{O}, \mathcal{A}, g^{\text{round}})$ , has *rotational symmetry*, while the potato  $(S^2, \mathcal{O}, \mathcal{A}, g^{\text{potato}})$  does not.

Prior to this course we have been taught<sup>1</sup> to think of symmetries as a group of maps which map the object to itself, while preserving all of the structures of the object. For example, 3-dimensional rotational symmetry is often given by the  $SO(3)$  group. In teaching this, we make use of the inner product available to us. The method we're about to describe here is actually subtly different. As we have just seen above, it is through the introduction of the metric that we get symmetries. That is, the symmetries are *not* something else we provide as well as providing the metric, they come as a consequence of *which* metric we provide. Now it's reasonable to think 'well the metric provides an inner product in each tangent space, so we could make a connection to the previously taught idea?' This is where the subtle nature comes in. What the metric is a tensor *field*, and so tells us how to *distribute* these inner products over all the tangent spaces. So the symmetry appears not to come from the inner products themselves but somehow from their distribution over the manifold.

So we want to answer the question 'how do we describe the symmetries of a metric?' This is not just a matter of academic interest, but actually is very important when it comes to studying the physical solutions. For example, the only way to solve Einstein's equations is to provide some symmetry conditions for the spacetime (i.e. the Universe).

## 11.1 Push-Forward Map

**Definition** (Push-Forward Map). Let  $\phi : \mathcal{M} \rightarrow \mathcal{N}$  be a smooth map between two smooth manifolds. Then we define the **push-forward map**  $\phi_* : T\mathcal{M} \rightarrow T\mathcal{N}$  by

$$\phi_*(X)\langle f \rangle := X\langle f \circ \phi \rangle,$$

where  $f \in C^\infty(\mathcal{N})$ , i.e.  $f : \mathcal{N} \rightarrow \mathbb{R}$ .

Diagrammatically, the maps in the above definition are related by the following diagram.

$$\begin{array}{ccccc} T\mathcal{M} & \xrightarrow{\phi_*} & T\mathcal{N} & & \\ \pi_{\mathcal{M}} \downarrow & & \downarrow \pi_{\mathcal{N}} & & \\ \mathcal{M} & \xrightarrow{\phi} & \mathcal{N} & \xrightarrow{f} & \mathbb{R} \end{array}$$

<sup>1</sup>Well I was and am assuming the reader was too.



**Corollary 11.1.1.** Recall that the fibres of the tangent bundle are just the tangent spaces to that point, i.e.  $\text{preim}_{\pi_{\mathcal{M}}} p = T_p \mathcal{M}$ . It follows, then, that

$$\phi_*(T_p \mathcal{M}) \subseteq T_{\phi(p)} \mathcal{N}.$$

That is, the image of the  $p$ -fibres on  $\mathcal{M}$  are at least contained within the  $\phi(p)$ -fibres on  $\mathcal{N}$ .

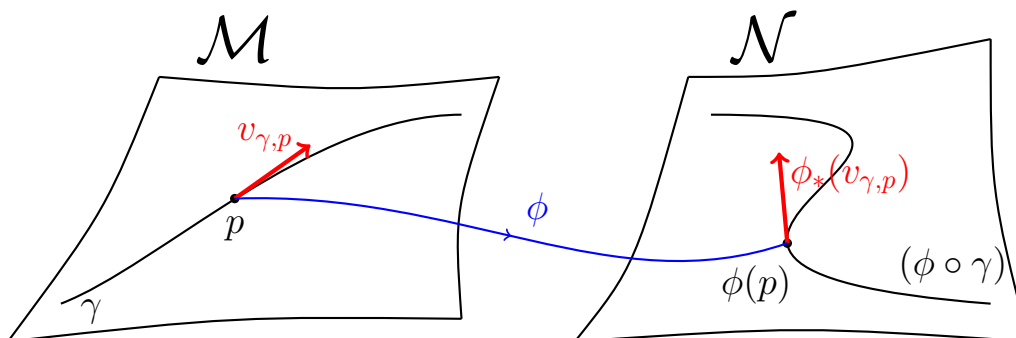
There is a mnemonic to remember what the push forward does: "vectors are pushed forward".

It is worth looking at the components of the push-forward map in the *two* charts  $(U, x) \in \mathcal{A}_{\mathcal{M}}$  and  $(V, y) \in \mathcal{A}_{\mathcal{N}}$ . We have, for  $p \in \mathcal{M}$

$$\phi_*^a := dy^a : \phi_* \left( \left( \frac{\partial}{\partial x^i} \right)_p \right) = \phi_* \left( \left( \frac{\partial}{\partial x^i} \right)_p \right) \langle y^a \rangle = \frac{\partial (y \circ \phi)^a}{\partial x^i} =: \frac{\partial \hat{\phi}^a}{\partial x^i},$$

where  $a \in \{1, \dots, \dim \mathcal{N}\}$  and  $i \in \{1, \dots, \dim \mathcal{M}\}$ . Note that  $\hat{\phi} := (y \circ \phi)$  is a map  $\hat{\phi} : U \rightarrow \mathbb{R}^{\dim \mathcal{N}}$ .

The following figure gives a nice pictorial description of the push-forward map.



**Figure 11.1:** Given two smooth manifolds and a smooth map  $\phi : \mathcal{M} \rightarrow \mathcal{N}$ , the push forward,  $\phi_*$ , maps tangent vector,  $v_{\gamma,p}$  of curve  $\gamma$  at point  $p \in \mathcal{M}$  to from the corresponding tangent vector,  $\phi_*(v_{\gamma,p})$ , of curve  $(\phi \circ \gamma)$  at point  $\phi(p) \in \mathcal{N}$ .

**Corollary 11.1.2.** Looking at Figure 11.1, we see that  $\phi_* : v_{\gamma,p} \mapsto v_{(\phi \circ \gamma), \phi(p)}$ .

*Proof.* Let  $f \in C^\infty(\mathcal{N})$  and let  $p \in \mathcal{M}$  be such that  $\gamma(\lambda_0) = p$ . Then

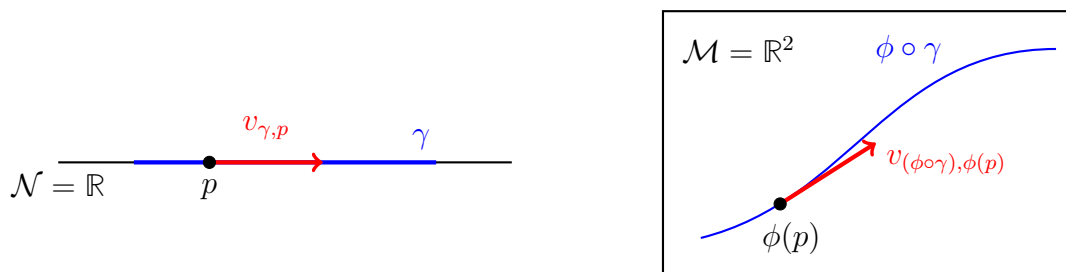
$$\begin{aligned} \phi_*(v_{\gamma,p}) &:= v_{\gamma,p}(f \circ \phi) \\ &= ((f \circ \phi) \circ \gamma)'(\lambda_0) \\ &= (f \circ (\phi \circ \gamma))'(\lambda_0) \\ &= v_{(\phi \circ \gamma), (\phi \circ \gamma)(\lambda_0)} \\ &= v_{(\phi \circ \gamma), \phi(p)}. \end{aligned}$$

□

*Example 11.1.3.* An important/interesting example of use of the push-forward is when  $\phi$  is an *embedding* map<sup>2</sup> from a  $d$ -dimensional manifold to a  $(d + 1)$ -dimensional manifold.

For obvious pictorial reasons, let  $d = 1$ .<sup>3</sup> If  $\gamma : (0, 1) \rightarrow \mathcal{M}$  is a curve in this 1-dimensional manifold, then  $v_{\gamma,p}$  is an element of the 1-dimensional tangent space  $T_p\mathcal{M}$ . Let  $\phi : \mathcal{M} \hookrightarrow \mathcal{N}$  be an embedding of  $\mathcal{M}$  into  $\mathcal{N}$ , where  $\dim \mathcal{N} = 2$ . Then the velocity  $v_{(\phi \circ \gamma), \phi(p)}$  is an element of the 2-dimensional tangent space  $T_{\phi(p)}\mathcal{N}$ . This allows us to make a connection between the *intrinsic* vector  $v_{\gamma,p}$  and the *extrinsic* vector  $v_{(\phi \circ \gamma), \phi(p)}$ .

As an analogy, consider an ant walking along a wire laid down on a table. The vector  $v_{\gamma,p}$  would be what the ant (who is oblivious to the higher dimensional space) itself says its velocity is, whereas  $v_{(\phi \circ \gamma), \phi(p)}$  is what we (who have a birds eye view of the table) would say the ant's velocity is.



## 11.2 Pull-back Map

**Definition** (Pull-back). Let  $\phi : \mathcal{M} \rightarrow \mathcal{N}$  be a smooth map between two smooth manifolds. Then we define the **pull-back map** as  $\phi^* : T^*\mathcal{N} \rightarrow T^*\mathcal{M}$  via

$$\phi^*(\omega) : X := \omega : \phi_*(X).$$

for  $\omega \in T^*\mathcal{N}$  and  $X \in T\mathcal{M}$ .

Again let's look at the components with respect to the two charts  $(U, x) \in \mathcal{A}_{\mathcal{M}}$  and  $(v, y) \in \mathcal{A}_{\mathcal{N}}$ .

$$\phi^*{}^a{}_i := \phi^*((dy^a)_{\phi(p)}) : \left( \frac{\partial}{\partial x^i} \right)_p = (dy^a)_{\phi(p)} : \phi_* \left( \left( \frac{\partial}{\partial x^i} \right)_p \right) =: \phi_*^a{}_i,$$

so the components of the pull-back and the components of the push-forward are the same!

Just as we showed that the push-forward of a velocity to a curve was the velocity of the mapped curve, the pull-back of the gradient of some function is the gradient of a function that is mapped to the other function. That is

$$\phi^*(df) = d(f \circ \phi).$$

This result can be obtained in a similar manner to the push-forward calculation (see tutorial), or it follows immediately from the following proposition and definition.

<sup>2</sup>It is important we use an embedding here and not just an *immersion*, which can have self-intersections. If we had self intersections we would not have a unique tangent vector to the mapped curve. For more details on embeddings and immersions, see section 3.6 of Renteln's *Manifolds, Tensors and Forms* textbook.

<sup>3</sup>We could also use  $d = 2$ , but that will be significantly harder for me to draw in Tikz, so  $d = 1$  it is.

**Proposition 11.2.1.** *The pull-back map and the map  $d$  commute. That is*

$$\phi^*(d\bullet) = d(\phi^*\bullet).$$

**Definition.** Let  $\phi : \mathcal{M} \rightarrow \mathcal{N}$  be a smooth map between two smooth manifolds. Then the pull back of  $f \in C^\infty(\mathcal{N})$  is given by

$$\phi^*(f) := f \circ \phi.$$

The mnemonic phrase here is "covectors are pulled back."

### 11.3 Induced Metric

There is an important application for the pull-back. Again consider  $\phi : \mathcal{M} \hookrightarrow \mathcal{N}$  as an embedding with  $\dim \mathcal{M} < \dim \mathcal{N}$ . Now let the smooth manifold with  $\mathcal{N}$  be equipped with a metric,  $g$ . We now want to ask whether we can use this metric to define one on the manifold with  $\mathcal{M}$ , which we shall call the *induced metric*,  $g_{\mathcal{M}}$ . The metric is a  $(0,2)$ -tensor field, and so can be pulled-back. The question we want to answer is "But how do we define such a metric?"

The way we want this to work is the following. We want to work out the length of a path,  $\gamma$ , between two points on  $\mathcal{M}$  using  $g_{\mathcal{M}}$ . We take the value to be the length of the mapped path,  $\gamma \circ \phi$ , obtained using  $g$ .

Now obviously there is more than one way to embed the space. Each one of these embeddings gives a potentially different length, and so defines a different metric (and shape) for  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$ . To use the examples referred to frequently in these notes, the smooth manifold  $(S^2, \mathcal{O}, \mathcal{A})$  can be either a round sphere of radius  $R$  or a potato. We can decide which it is by defining an embedding  $\phi : S^2 \hookrightarrow \mathbb{R}^3$  such that the induced metric gives the correct shape. This is what our eyes do when differentiating a football<sup>4</sup> from a potato; they look at the lengths between points using our 3D Euclidean metric and conclude that the induced metric is that of a football (or potato).

We can write this mathematically as the following definition.

**Definition** (Induced Metric). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  and  $(\mathcal{N}, \mathcal{O}, \mathcal{A})$  be a smooth manifolds, with  $|\mathcal{M}| \leq |\mathcal{N}|$ .<sup>5</sup> and let  $\phi : \mathcal{M} \hookrightarrow \mathcal{N}$  be an embedding. Now equip  $(\mathcal{N}, \mathcal{O}, \mathcal{A})$  with a metric  $g$ . We define the **induced metric** on  $\mathcal{M}$  as the pull back  $g_{\mathcal{M}} := \phi^*g$ , which satisfies<sup>6</sup>

$$g_{\mathcal{M}}(X, Y) := g(\phi_*(X), \phi_*(Y)),$$

for all  $X, Y \in \Gamma T\mathcal{M}$ .

The above condition in the definition can be written in components as

$$(g_{\mathcal{M}})_{ij} = g_{ab} \frac{\partial \hat{\phi}^a}{\partial x^i} \frac{\partial \hat{\phi}^b}{\partial x^j},$$

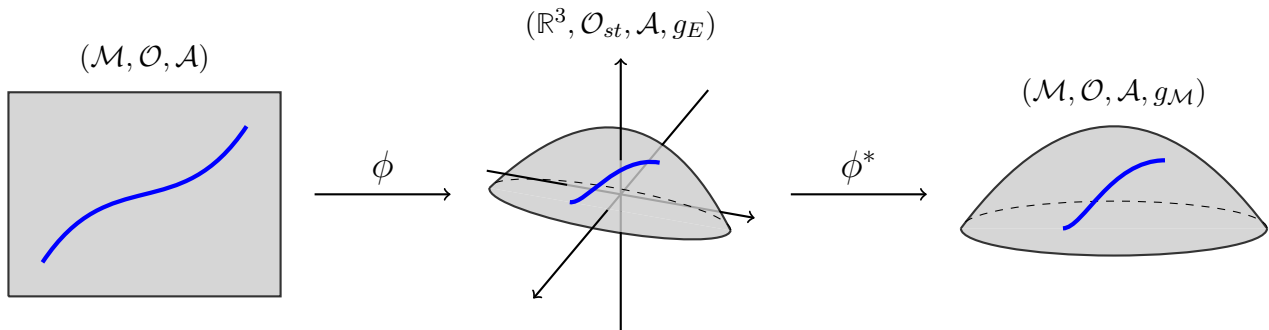
where  $\hat{\phi} = (y \circ \phi)$ , as in the calculation for the components of the push-forward.

<sup>4</sup>That is 'Soccer' to some.

<sup>5</sup>The vertical lines indicate the so-called cardinality of the set, i.e. how many elements are in it.

<sup>6</sup>The push-forward of a vector field is simply defined point wise, i.e. push-forward each vector and make a vector field.

*Example 11.3.1.* Pictorially we can see the above ideas via the following drawings. Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be some 2-dimensional smooth manifold and let  $(\mathcal{N}, \mathcal{O}, \mathcal{A}, g) = (\mathbb{R}^3, \mathcal{O}_{st}, \mathcal{A}, g_E)$ , the Euclidean 3-space. We could define an embedding  $\phi : \mathcal{M} \hookrightarrow \mathbb{R}^3$  such that  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  looks dome shaped w.r.t. the metric  $g_E$ . We can then pull this metric back onto the  $\mathcal{M}$  manifold itself, giving the induced metric space  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g_{\mathcal{M}})$ .



It is important to note that we only have a dome shape both in the embedding and as the induced metric because we are considering the embedding space to be the Euclidean 3-space. That is, when we draw the diagrams on the far right, we are seeing it as being embedded in Euclidean 3-space. This is the comment made about what our eyes do to give differentiate between footballs and potatoes.

## 11.4 Flow of a Complete Vector Field

**Definition** (Integral Curve). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be a smooth manifold and let  $\gamma : (a, b) \rightarrow \mathcal{M}$  be a smooth curve with  $(a, b) \subseteq \mathbb{R}$ . If we have a vector field  $X \in \Gamma T\mathcal{M}$ , then  $\gamma$  is said to be an *integral curve* of  $X$  if

$$v_{\gamma, \gamma(\lambda)} = X_{\gamma(\lambda)}.$$

That is, the tangent vectors to the curve reproduce the vector field constrained to the curve.

*Example 11.4.1.* An example of a integral curve would be that corresponding to a paper ship floating down a river. The vector field  $X$  would be the velocity field of the water molecules and the curve  $\gamma$  would be the trajectory of the ship.

**Definition** (Complete Vector Field). A vector field  $X \in \Gamma T\mathcal{M}$  is called **complete** if all integral curves have domain  $\mathbb{R}$  (i.e.  $(a, b) = \mathbb{R}$ ).

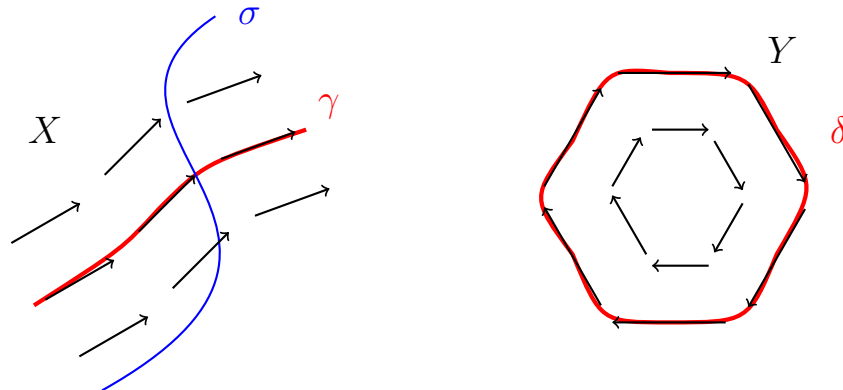
It is tempting to think that this is always possible because you can just reparameterise  $\gamma$  such that  $(a, b) = \mathbb{R}$ , right? Well it's true that you can do this, but in doing so you change the absolute value/length of the tangent vectors and then they no longer coincide with the vector field vectors. So the choice of parameterisation if chosen by the absolute values of the vectors in the vector field.

Following from the above point, note that for a vector field to be complete it is important that we don't remove points from the domain of the vector field. If we did this, the integral curve through that point would then have finite length and so we would not be able to extend

the interval  $(a, b)$  to the whole of  $\mathbb{R}$  without breaking the integral curve nature. This is a really important point because it leads the way to a proper understanding of singularity<sup>7</sup> theorems.

This result is contained within the following theorem.

**Theorem 11.4.2.** *A compactly<sup>8</sup> supported, smooth vector field is complete.*



**Figure 11.2:** *Left:  $\gamma$  is an integral curve of the smooth vector field  $X$  as its tangent vectors at all points reproduce the vector field at those points.  $\sigma$  is not an integral curve as the tangent vectors do not coincide with the vector field vectors at that point. Right: Example of a complete vector field,  $Y$ . The integral curves,  $\delta$ , are closed and therefore have domain  $\mathbb{R}$ . If we were to remove one point in the space, we would no longer have a complete vector field as one of the integral curves would then have finite length.*

**Definition** (Flow of a Complete Vector Field). The **flow of a complete vector field**  $X \in \Gamma T\mathcal{M}$  is a one-parameter family

$$h^X : \mathbb{R} \times \mathcal{M} \rightarrow \mathcal{M} \\ (\lambda, p) \mapsto \gamma_p(\lambda),$$

where  $\gamma_p : \mathbb{R} \rightarrow \mathcal{M}$  is the integral curve of  $X$  with  $\gamma(0) = p$ .

We can use the above definition to introduce a new map by simply taking a fixed value for  $\lambda$ . That is, for fixed  $\lambda \in \mathbb{R}$  we have the smooth map

$$h_\lambda^X : \mathcal{M} \rightarrow \mathcal{M},$$

which takes every point in  $\mathcal{M}$  and moves it a parameter distance  $\lambda$  along the integral curve through that point.

<sup>7</sup>A singularity can be thought of as a point that is removed from the spacetime because, for example, the curvature blows up there.

<sup>8</sup>A topological space is said to be compact if every open cover has a finite subcover. For more details see, e.g., Renteln's *Manifolds, Tensors, and Forms* textbook.

### 11.5 Lie Subalgebras of the Lie Algebra $(\Gamma T\mathcal{M}, [\cdot, \cdot])$ of Vector Fields

**Definition** (Lie Algebra). A **Lie algebra** is a vector space<sup>9</sup>  $\mathfrak{g}$  equipped with a bilinear operation  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ , known as the **Lie bracket**, that also satisfies

- (i) Antisymmetry:  $[x, y] = -[y, x]$ ,
- (ii) The *Jacobi identity*:  $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$

**Definition** (Structure Constants). Let  $(\mathfrak{g}, [\cdot, \cdot])$  be a Lie algebra. We define the **structure constants** of the Lie algebra,  $C^k_{ij} \in \mathbb{F}$ , via

$$[x_i, x_j] = C^k_{ij} x_k$$

for  $x_i \in \mathfrak{g}$  and  $i, j, k \in \{1, \dots, \dim \mathfrak{g}\}$ .

Recall in lecture 8 we defined the commutator of two vector fields as

$$[X, Y]\langle f \rangle = X\langle Y\langle f \rangle \rangle - Y\langle X\langle f \rangle \rangle.$$

We want to make this into a Lie bracket, however we have to address a problem. As it stands we are considering the  $C^\infty$ -module  $(\Gamma T\mathcal{M}, \oplus, \odot)$ , but our commutator does not obey  $C^\infty$ -bilinearity. That is

$$[f \odot X, Y] \neq f \odot [X, Y].$$

However, it does obey  $\mathbb{R}$ -bilinearity.

**Proposition 11.5.1.** *If we therefore restrict ourselves to the  $\mathbb{R}$ -vector space  $(\Gamma T\mathcal{M}, +, \cdot)$  then the commutator becomes a Lie bracket.*

*Notation.* We will denote the Lie algebra of vector fields as just  $(\Gamma T\mathcal{M}, [\cdot, \cdot])$ , but it is important to remember that we are considering the restricted to case of  $\cdot : \mathbb{R} \times \Gamma T\mathcal{M} \rightarrow \Gamma T\mathcal{M}$ , i.e.  $\mathbb{R}$ -vector space.

#### Exercise

- (a) Show the above inequality,  $[f \odot X, Y] \neq f \odot [X, Y]$ .
- (b) Prove Proposition 11.5.1.

**Definition** (Lie Subalgebra). Let  $(\mathfrak{g}, [\cdot, \cdot])$  be a Lie algebra. A vector subspace  $\mathfrak{a} \subseteq \mathfrak{g}$  is called a **Lie subalgebra** if it is closed under the Lie bracket. That is,  $[x, y] \in \mathfrak{a}$  for all  $x, y \in \mathfrak{a}$ .

By restricting to  $\mathbb{R}$ -linearity we get an infinite dimensional vector space. This just comes from the fact that we can only scale the basis vector fields by the same amount at each point (as opposed to with  $C^\infty$ -linearity), and so we need an infinite number of them to have a complete basis. However, we can just restrict ourselves to a subalgebra  $(\text{span}_{\mathbb{R}}\{X_1, \dots, X_s\}, [\cdot, \cdot])$  of finite dimension.

<sup>9</sup>In fact you only need a module over a commutative ring.

Example 11.5.2. An example of such a Lie subalgebra on  $(S^2, \mathcal{O}, \mathcal{A})$  is<sup>10</sup>

$$\mathfrak{so}(3) := (\text{span}_{\mathbb{R}}\{X_1, X_2, X_3\}, [\cdot, \cdot]),$$

where

$$[X_1, X_2] = X_3, \quad [X_3, X_1] = X_2 \quad \text{and} \quad [X_2, X_3] = X_1.$$

This is the 3-dimensional rotation Lie algebra, and finds important use in quantum mechanics.<sup>11</sup>

Remark 11.5.3. In the tutorials we show that

$$\begin{aligned} X_1(p) &= -\sin(\varphi(p)) \frac{\partial}{\partial \theta} - \cot(\theta(p)) \cos(\varphi(p)) \frac{\partial}{\partial \varphi}, \\ X_2(p) &= \cos(\varphi(p)) \frac{\partial}{\partial \theta} - \cot(\theta(p)) \sin(\varphi(p)) \frac{\partial}{\partial \varphi}, \\ X_3(p) &= \frac{\partial}{\partial \varphi}, \end{aligned}$$

is of the form above, justifying why it is often called the 3-dimensional rotation algebra.

Note that we can make no reference to a metric at any point here, and so any  $\{X_1, X_2, X_3\}$  that satisfies the above will hold on both the round sphere of radius  $R$  and on the potato.

## 11.6 Symmetry

**Definition** (Symmetry of a Metric). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  be a metric manifold, and let  $\{X_1, \dots, X_s\} \subset \Gamma T\mathcal{M}$ . Define  $L := \text{span}_{\mathbb{R}}\{X_1, \dots, X_s\}$ , then the  $s$ -dimensional Lie subalgebra  $(L, [\cdot, \cdot])$  is said to be a **symmetry** of a metric tensor field  $g$ , if for all  $X \in L$

$$g((h_\lambda^X)_*(A), (h_\lambda^X)_*(B)) = g(A, B),$$

for  $A, B \in T_p\mathcal{M}$  and  $(h_\lambda^X)_*$  is the push-forward of the flow of  $X$ . We can write this alternatively as

$$(h_\lambda^X)^*g = g,$$

where  $(h_\lambda^X)^*$  is the pull-back associated to the flow of  $X$ .

The first part in the above definition basically says that the angle and projection between  $A$  and  $B$  (which the metric tells you) does not change if you move both  $A$  and  $B$  along the integral curves of  $X$ . For example, for the round sphere of radius  $R$ , if we move  $A$  and  $B$  around the sphere in the ‘ $\theta$ ’-direction then obviously nothing changes.

The second part just says if we move the metric ‘backwards’ along the integral curves, it still looks the same. This is again intuitively clear for a round sphere when we rotate the sphere. It is not true, however, for the potato, because by moving the metric, the shape of the potato moves. This is clearly just the statement that the round sphere is rotationally symmetric, but the potato is not.

<sup>10</sup>We’re actually being a bit clumsy here.  $\mathfrak{so}(3)$  is the Lie algebra of the Lie group  $SO(3)$ , which is a manifold equipped with a group structure.

<sup>11</sup>For more details see Dr. Schuller’s Lectures on Quantum Theory course.

## 11.7 Lie Derivatives

The above test for symmetry is very intuitive but it has the major flaw that you have to do a lot of calculation. We therefore typically don't use that method, but instead use the following one.

It follows from the above that if, for all  $X \in L$ ,

$$\lim_{\lambda \rightarrow 0} \frac{(h_\lambda^X)^* g - g}{\lambda} = 0$$

holds then  $L$  is a symmetry of  $g$ . We actually give the left-hand side its own notation. We define the *Lie derivative* of a metric  $g$ , w.r.t. a vector field  $X$  as

$$\mathcal{L}_X g := \lim_{\lambda \rightarrow 0} \frac{(h_\lambda^X)^* g - g}{\lambda}.$$

The Lie derivative is actually quite a subtle thing to define. The definition we've used above makes contact with the pull back ideas we discussed above and so we can think of it as comparing the 'dragged back'<sup>12</sup> tensor to the tensor as it is. For another explanation of this see [these notes](#).

*Remark 11.7.1.* Alternatively, one can define the Lie derivative using *Cartan's formula*. This is useful when discussing the Lie derivative of differential forms. We shall not discuss this further here, but for the interested reader the formula is  $\mathcal{L}_X := d\iota_X - \iota_X d$ .<sup>13</sup>

Given the above comments, we actually define the Lie derivative in a rather abstract way, but that looks very similar to the definition of the covariant derivative.

**Definition** (Lie Derivative). The *Lie derivative*  $\mathcal{L}$  on a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  sends a pair of a vector field,  $X$ , and a  $(p, q)$ -tensor field,  $T$ , to a  $(p, q)$ -tensor field such that: for  $f \in C^\infty(\mathcal{M})$  and  $Y \in \Gamma T\mathcal{M}$ ,

- (i)  $\mathcal{L}_X = X\langle f \rangle$ ,
- (ii)  $\mathcal{L}_X Y = [X, Y]$ ,
- (iii)  $\mathcal{L}_X(T + S) = \mathcal{L}_X T + \mathcal{L}_X S$ ,
- (iv)  $\mathcal{L}_X(T(\omega, Y)) = (\mathcal{L}_X T)(\omega, Y) + T(\mathcal{L}_X \omega, Y) + T(\omega, \mathcal{L}_X Y)$  and similarly for different rank tensors,
- (v)  $\mathcal{L}_{X+Y} T = \mathcal{L}_X T + \mathcal{L}_Y T$ .

These conditions look very similar to those of the covariant derivative, but with the Lie derivative we don't need to provide any extra structure, i.e. don't need to define any  $\Gamma$ s. You might think that this makes the Lie derivative a more useful derivative, however it comes with its own flaws.

<sup>12</sup>Dragged back as the map  $h$  is an automorphism, so the pull-back just drags the points backwards.

<sup>13</sup> $d$  is the exterior derivative, which we have touched on in these notes, and  $\iota_X$  is the so-called interior derivative w.r.t.  $X$ .



The first thing we notice is that the lower entry for the Lie derivative must be a vector *field*. This is different to the covariant derivative, where we can take just a vector here. This comes from the idea that we need to obtain this flow of  $X$ , and that clearly involves knowing  $X$  in a neighbourhood of the point and so it must be a field. Next we notice that condition (ii) is something not present in the definition of a covariant derivative. It has the drastic effect on condition (v) whereby the Lie derivative is **not**  $C^\infty$ -linear in the lower slot (as the covariant derivative is). This comes simply from

$$\mathcal{L}_f X Y = [fX, Y] = f[X, Y] - Y\langle f \rangle X.$$

#### Exercise

There is another important difference to note. Recall that for the components of the covariant derivative of a tensor, each upper index came with a + sign and each lower index came with a – sign. The opposite is true for the Lie derivative. That is,

$$(\mathcal{L}_X T)^i_j = X^m \frac{\partial T^i_j}{\partial x^m} - \frac{\partial X^i}{\partial x^m} T^m_j + \frac{\partial X^m}{\partial x^j} T^i_m.$$

Show that this result holds.

Using the relation in the above exercise, the condition  $\mathcal{L}_X g = 0$  becomes a very easy thing to solve, and so we obtain a nice way to see if a metric features a symmetry.

#### 11.7.1 Killing Vector Fields

**Definition** (Killing Vector Field). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  be a metric manifold. A vector field  $K \in \Gamma T\mathcal{M}$  is called a **Killing vector field** (or just Killing field) if it is a symmetry of the metric, i.e.  $\mathcal{L}_K g = 0$ , which can equally be written as

$$K\langle g(X, Y) \rangle - g([K, X], Y) - g(X, [K, Y]) = 0.$$

Noether's theorem tells us that there is a link between symmetries and conservation laws, and so we see that Killing vector fields correspond to conservation laws. For example, as we will see later, the vector field which corresponds to temporal translation  $\partial_0$  is a Killing vector field Minkowski spacetime, and gives rise to conservation of energy. Similarly we have Killing vector fields for momentum conservation.

#### Exercise

Show that for the Levi-Civita connection the Killing vector field condition becomes

$$g(\nabla_X K, Y) + g(X, \nabla_Y K) = 0.$$

*Hint: You need to use both the metric compatible and the Torsion free conditions.*

# 12 | Integration

This lecture completes our ‘lift’ of analysis on the charts to analysis at the manifold level. It is the last step in the mathematical foundations before we can move on next lecture to start to discuss general relativity itself.

The aim is to define a notion of integration at the manifold level, i.e. we want to be able to compute  $\int_{\mathcal{M}} f$  where  $f$  is a smooth function on  $\mathcal{M}$ . In order to define this, we need to introduce a mild new structure, known as the *volume form*. We will also need to restrict our atlas, giving us a so-called *orientation*.

## 12.1 Review of Integration on $\mathbb{R}^d$

The simplest case is that of a function  $F : \mathbb{R} \rightarrow \mathbb{R}$ , where we simply have<sup>1</sup>

$$\int_{(a,b)} F := \int_a^b dx F(x),$$

where the right-hand side integral is some known integration operation (e.g. Riemann integrals).

Next we can consider  $F : \mathbb{R}^d \rightarrow \mathbb{R}$ . If we are to do this over a box shaped domain,  $(a,b) \times \dots \times (u,v) \subseteq \mathbb{R}^d$ , the integral is simply

$$\int_{(a,b) \times \dots \times (u,v)} d^d x F(x) := \int_a^b dx^1 \dots \int_u^v dx^d F(x^1, \dots, x^d).$$

We can then extend this to general domains (i.e. not necessarily box shaped)  $G \subseteq \mathbb{R}^d$  by introducing an *indication function*  $\mu_G : \mathbb{R}^d \rightarrow \mathbb{R}$  given by

$$\mu_G(x) = \begin{cases} 1 & \text{if } x \in G \\ 0 & \text{otherwise.} \end{cases}$$

We then define the integral

$$\int_G d^d x F(x) := \int_{-\infty}^{\infty} dx^1 \dots \int_{-\infty}^{\infty} dx^d \mu_G(x) F(x).$$

We now need to ask how this definition changes under a change of variable (which will correspond to a change of chart in the lifted notion).

---

<sup>1</sup>We are assuming that the following exists. We shall assume that the results exist for this whole section.

**Theorem 12.1.1.** Let  $\phi : \text{preim}_\phi(G) \rightarrow G$  denote the change of variable map with  $G, \text{preim}_\phi(G) \subseteq \mathbb{R}^d$ . Then if the integral of  $F : G \rightarrow \mathbb{R}$  is defined as above we have<sup>2</sup>

$$\int_G d^d x F(x) = \int_{\text{preim}_\phi(G)} d^d y |\det(\partial_a \phi^b)(y)| (F \circ \phi)(y),$$

where  $|\det(\partial_a \phi^b)(y)|$  is the **Jacobian** of  $\phi$ .<sup>3</sup>

Example 12.1.2. Consider  $d = 2$  and let  $G = \mathbb{R}^2 \setminus (x, 0)$ , i.e.  $\mathbb{R}^2$  with the  $x$ -axis cut out. Then let

$$\begin{aligned} \phi : \mathbb{R}^+ \times [(0, \pi) \cup (\pi, 2\pi)] &\rightarrow G \\ (r, \varphi) &\mapsto (r \cos \varphi, r \sin \varphi). \end{aligned}$$

We have

$$(\partial_a \phi^b)(r, \varphi) := \begin{pmatrix} \partial_r(r \cos \varphi) & \partial_r(r \sin \varphi) \\ \partial_\varphi(r \cos \varphi) & \partial_\varphi(r \sin \varphi) \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -r \sin \varphi & r \cos \varphi \end{pmatrix},$$

so  $|\det(\partial_a \phi^b)(r, \varphi)| = |r| = r$ . This gives

$$\int_G dx^1 dx^2 F(x^1, x^2) = \int_0^\infty dr \int_0^{2\pi} d\varphi r F(r \cos \varphi, r \sin \varphi).$$

We then say that the volume elements  $dx^1 dx^2$  and  $r dr d\varphi$  correspond to each other in their respective coordinates. Of course what we have just looked at is simply changing from Cartesian to polar coordinates.

## 12.2 Integration on One Chart

Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be a smooth manifold and let  $f \in C^\infty(\mathcal{M})$ . Consider charts  $(U, x), (U, y) \in \mathcal{A}$  with the same domain. Denote  $f \circ x^{-1} =: f_{(x)} : x(U) \rightarrow \mathbb{R}$  and similarly for  $f_{(y)} : y(U) \rightarrow \mathbb{R}$ . We want to define the integral of  $f$  over  $U$  at the manifold level as something along the lines of

$$\int_U f := \int_{x(U)} d^d \alpha f_{(x)}(\alpha),$$

where  $\alpha \in \mathbb{R}^d$  is the coordinate tuple in  $x(U)$ . However, this is not possible as it is not chart independent. This is seen easily by considering the chart transition map  $x \circ y^{-1} : y(U) \rightarrow x(U)$ :

$$\begin{aligned} \int_{x(U)} d^d \alpha f_{(x)}(\alpha) &= \int_{y(U)} d^d \beta |\det(\partial_a (x \circ y^{-1})^b)(\beta)| f_{(y)}(\beta) \\ &= \int_{y(U)} d^d \beta \left| \det \left( \frac{\partial x^b}{\partial y^a} \right)_{y^{-1}(\beta)} \right| f_{(y)}(\beta), \end{aligned}$$

<sup>2</sup>The indices  $a$  and  $b$  in the determinant do not break Einstein summation convention. What is meant here is the determinant of the elements, and we know the determinant is invariant of which chart we use (think about the determinant of a matrix just being the product of the eigenvalues).

<sup>3</sup>Sometimes the Jacobian is defined without the absolute value part, but here we shall use the whole thing.

where we have used  $f_{(x)} \circ x \circ y^{-1} = f_{(y)}$  and our definition of  $\partial_a(x \circ y^{-1})^b$  in terms of the fraction notation. In general this Jacobian will not be unit, and so we don't get  $\int_U f = \int_{y(U)} d^d \beta f_{(y)}(\beta)$ , as is required for chart independence.

The obvious solution to this problem is to try and introduce something to the right-hand side of our definition of  $\int_U f$  that cancels the Jacobian factor we obtain. Such a structure can not be define on just a smooth manifold, and we need to introduce a new structure.

### 12.3 Volume Forms

**Definition** (Volume Form). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be a smooth manifold. We call a  $(0, \dim \mathcal{M})$ -tensor field  $\Omega$  is called a **volume form** if

- (i) it vanishes nowhere;  $\Omega|_p \neq 0$  for all  $p \in \mathcal{M}$ , and
- (ii) it is totally antisymmetric;  $\Omega(\dots, X, \dots, Y, \dots) = -\Omega(\dots, Y, \dots, X, \dots)$  for all entries.

The obvious question to ask is do we need to provide the volume form by hand or can be obtain it from some structure we've already talked about. The answer is, that we can obtain it from a metric on a metric manifold. In order to make this definition we have to introduce the Levi-Civita symbol.

**Definition** (Levi-Civita Symbol). The **Levi-Civita symbol** in  $d$  dimensions is denoted  $\epsilon_{i_1 \dots i_d}$  and is defined via the following two properties:

- (i)  $\epsilon_{123 \dots d} = 1$ , and
- (ii) total antisymmetry, i.e. it flips sign when any two indices are exchanged

$$\epsilon_{i_1 \dots i_j \dots i_k \dots i_d} = -\epsilon_{i_1 \dots i_k \dots i_j \dots i_d}$$

*Remark 12.3.1.* Note condition (ii) for the Levi-Civita symbol tells us that if any two indices repeat the symbol vanishes. It also tells us that permutations of the indices leave the value unchanged. For example for  $d = 4$ ,  $\epsilon_{1123} = 0$  and  $\epsilon_{1234} = \epsilon_{2341}$ .

**Definition** (Oriented Atlas). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  be a manifold. Then the subatlas  $\mathcal{A}^\uparrow \subseteq \mathcal{A}$  is called the (positive) **oriented atlas** if

$$\det \left( \frac{\partial y^m}{\partial x^i} \right) > 0,$$

for all overlapping charts  $(U, x), (V, y) \in \mathcal{A}^\uparrow$ .

We can similarly define  $\mathcal{A}^\downarrow$  to be such that the determinant is negative.

*Remark 12.3.2.* It is important to note that you can not always define a oriented atlas. That is, it's not necessarily true that the charts  $(U_i, x_i)$  that satisfy the determinant condition will cover all of  $\mathcal{M}$ , and so do not form an atlas. Such manifolds are known as *non-orientable* manifolds.

Claim 12.3.3. Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}^\uparrow, g)$  be an oriented metric manifold. We can define the components of the volume form in some chart  $(U, x) \in \mathcal{A}^\uparrow$  as the following

$$\Omega_{(x)i_1 \dots i_d} := \sqrt{\det(g_{(x)ij})} \epsilon_{i_1 \dots i_d}.$$

*Proof.* It is clear that the two conditions for the volume form are satisfied as  $\det(g) \neq 0$  for Riemannian/pseudo-Riemannian metrics (i.e. there are no 0s in the signature), and the Levi-Civita symbol is totally antisymmetric. So we just need to show that the result is well defined, i.e. we need to show the components transform like those of a  $(0, d)$ -tensor field. We have<sup>4</sup>

$$\begin{aligned} \Omega_{(x)i_1 \dots i_d} &= \sqrt{\det\left(g_{(y)mn} \frac{\partial y^m}{\partial x^i} \frac{\partial y^n}{\partial x^j}\right)} \epsilon_{i_1 \dots i_d} \\ &= \sqrt{\det(g_{(y)mn})} \left| \det\left(\frac{\partial y^m}{\partial x^i}\right) \right| \epsilon_{i_1 \dots i_d} \\ &= \sqrt{\det(g_{(y)mn})} \operatorname{sgn}\left(\det\left(\frac{\partial y^m}{\partial x^i}\right)\right) \det\left(\frac{\partial y^m}{\partial x^i}\right) \epsilon_{i_1 \dots i_d}. \end{aligned}$$

Now we use the result

$$\det\left(\frac{\partial y^m}{\partial x^i}\right) \epsilon_{i_1 \dots i_d} = \left(\frac{\partial y^{m_1}}{\partial x^{i_1}} \dots \frac{\partial y^{m_d}}{\partial x^{i_d}}\right) \epsilon_{m_1 \dots m_d}.$$

So we have

$$\begin{aligned} \Omega_{(x)i_1 \dots i_d} &= \operatorname{sgn}\left(\det\left(\frac{\partial y^m}{\partial x^i}\right)\right) \cdot \left[\left(\frac{\partial y^{m_1}}{\partial x^{i_1}} \dots \frac{\partial y^{m_d}}{\partial x^{i_d}}\right) \sqrt{\det(g_{(y)mn})} \epsilon_{m_1 \dots m_d}\right] \\ &= \operatorname{sgn}\left(\det\left(\frac{\partial y^m}{\partial x^i}\right)\right) \cdot \left[\left(\frac{\partial y^{m_1}}{\partial x^{i_1}} \dots \frac{\partial y^{m_d}}{\partial x^{i_d}}\right) \Omega_{(y)m_1 \dots m_d}\right], \end{aligned}$$

which would be the correct transformation property if we didn't have the  $\operatorname{sgn}$  term. So we see that if we restrict our atlas such that  $\det\left(\frac{\partial y^m}{\partial x^i}\right) > 0$ , as stated in the claim, then we simply get our desired transformation property.  $\square$

Remark 12.3.4. The above definition of a volume form is quite a tedious way to define a volume form, as you would have to check all the chart transition maps and ensure your manifold is orientable etc. There is a much nicer way to define a volume form (using the pull-back map), however in order to introduce it here, we would need to introduce the idea of a differential form (which is where the volume form derives the latter part of its name). The interested reader is directed to appendix C of Renteln's *Manifolds, Tensors and Forms* textbook.

---

<sup>4</sup>The index notation here can be a bit confusing, but the main thing to keep an eye on is where determinants appear, because then the indices simply tell us about the positions in matrices and so we can seemingly break summation convention.

**Definition** (Scalar Density). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}^\uparrow)$  be an oriented smooth manifold. We define the **scalar density** in chart  $(U, x)$  as

$$\omega_{(x)} := \Omega_{i_1 \dots i_d} \epsilon^{i_1 \dots i_d},$$

where  $\epsilon^{i_1 \dots i_d}$  is defined the same as  $\epsilon_{i_1 \dots i_d}$ .

It follows from the calculations above that scalar densities on metric manifolds satisfy<sup>5</sup>

$$\omega_{(y)} = \det \left( \frac{\partial x}{\partial y} \right) \omega_{(x)}.$$

## 12.4 Integration On One Chart Domain

**Definition** (Integration on a Chart Domain). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}^\uparrow)$  be an oriented smooth manifold and let  $(U, x) \in \mathcal{A}^\uparrow$ . We define the integral of  $f \in C^\infty(\mathcal{M})$  on the domain  $U$  via

$$\int_U f := \int_{x(U)} d^d \alpha f_{(x)}(\alpha) \cdot \omega_{(x)}(x^{-1}(\alpha)),$$

where  $\omega_{(x)}$  is the scalar density corresponding to the volume form  $\Omega$ .

*Claim 12.4.1.* The above notion of integration is well defined.

*Proof.* We need to show that the formula doesn't change form when we change charts. From the calculations done in this lecture, we have

$$\begin{aligned} \int_U f &= \int_{y(U)} d^d \beta \left| \det \left( \frac{\partial x}{\partial y} \right) \right| f_{(y)}(\beta) \cdot \left[ \det \left( \frac{\partial x}{\partial y} \right) \right]^{-1} \omega_{(y)}(y^{-1}(\beta)) \\ &= \int_{y(U)} d^d \beta f_{(y)}(\beta) \cdot \omega_{(y)}(y^{-1}(\beta)), \end{aligned}$$

which is our well definition condition. Note we have used the fact that our atlas is positive orientated to 'remove' the absolute value sign.  $\square$

For the special case of an oriented metric manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}^\uparrow, g)$  our result becomes

$$\int_U f := \int_{x(U)} d^d \alpha \sqrt{\det (g_{(x)ij})} (x^{-1}(\alpha)) f_{(x)}(\alpha).$$

*Notation.* To lighten notation, it is common to denote

$$g := \det (g_{(x)ij}) (x^{-1}(\alpha)),$$

turning the above expression into

$$\int_U f := \int_{x(U)} d^d \alpha \sqrt{g} f_{(x)}(\alpha).$$

*Remark 12.4.2.* In the above we have assumed we are using a Riemannian manifold, in which case  $g > 0$ . The above formula is adapted to pseudo-Riemannian manifolds with  $g < 0$  by simply replacing  $\sqrt{g} \rightarrow \sqrt{-g}$ . We will see this later when, for example, considering Maxwell's action.

<sup>5</sup>We drop the indices in the determinant to lighten the notation, but they are just the indices used for  $x$  and  $y$ .

## 12.5 Integration Over The Entire Manifold

We might be tempted at this point to simply say that the integral over the whole manifold is just given as the sum over the chart domain integrals. That is let  $\{(U_i, x_i)\} \subseteq \mathcal{A}^\dagger$  be a subatlas (i.e.  $\cup_i U_i = \mathcal{M}$ ) then we might be tempted to say

$$\int_{\mathcal{M}} f = \sum_i \left( \int_{x_i(U_i)} d^d \alpha \sqrt{g} f_{(x_i)}(\alpha) \right).$$

However this suffers from the common problem of over counting. That is, as it is defined, the contributions from the overlap regions  $U_i \cap U_j$  are counted in *both* the  $x_i(U_i)$  integral and the  $(U_j, x_j)$  integral. We need a way, therefore, to remove this over counting. This problem is resolved by requiring that the manifold admit a so-called *partition of unity*.

Roughly speaking, for any finite<sup>6</sup> subatlas  $\mathcal{A}' = \{(U_1, x_1), \dots, (U_N, x_N)\} \subseteq \mathcal{A}^\dagger$  there exists continuous functions  $\rho_i : U_i \rightarrow \mathbb{R}$ , such that for all  $p \in \mathcal{M}$

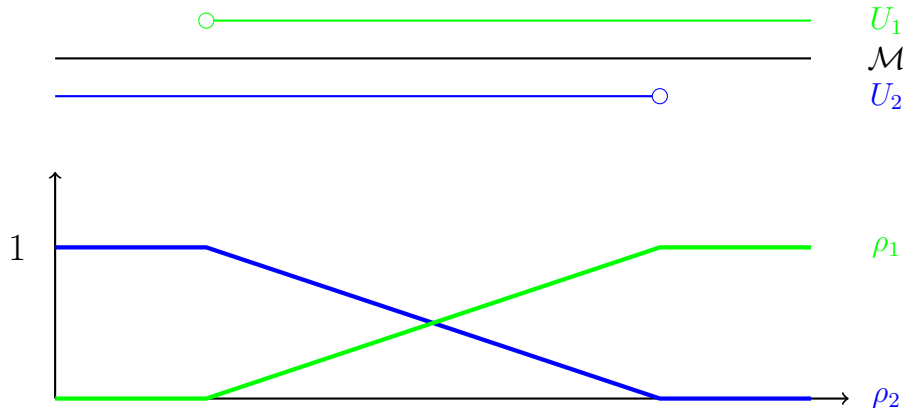
$$\sum_i \rho_i(p) = 1,$$

where the sum is performed such that  $p \in U_i$ .<sup>7</sup> This accounts exactly for this over counting and allows us to define the integral over the whole manifold as

$$\int_{\mathcal{M}} f := \sum_i \left( \int_{U_i} \rho_i \cdot f \right).$$

That is, the  $\rho_i$ s are defined such that their sum at any point in  $\mathcal{M}$  equals unity. Clearly this removes the over counting and just gives one contribution for the overlap region.

*Example 12.5.1.* Let  $\dim \mathcal{M} = 1$  and let it be covered by two charts  $(U_1, x_1)$  and  $(U_2, x_2)$ . Define the  $\rho_i$ s to be such that they change linearly across the overlap region. Then we have something like the below diagram.



<sup>6</sup>Note we require it to finite, as otherwise we would need to check for convergence in order to be able to take the sum out of the integral below.

<sup>7</sup>Alternatively you could just say  $\rho_i(p) = 0$  if  $p \notin U_i$  and let the sum run from  $i = 1$  to  $i = N$ .

# 13 | Relativistic Spacetime

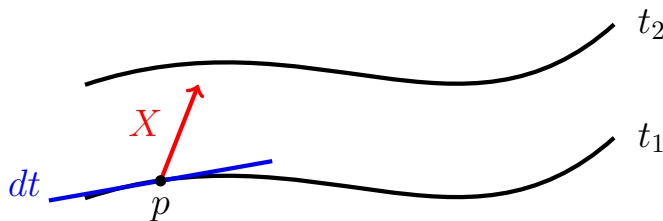
We now start talking about physics. Of course we will use all of the mathematical tools developed so far and so it is important that the reader understands the content up to this point fully.

Recall the definition of Newtonian spacetime from Lecture 9 as the quintuple  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla, t)$  where  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$  is a 4-dimensional smooth manifold,  $\nabla$  is a torsion free connection and an absolute time  $t \in C^\infty(\mathcal{M})$  satisfying  $dt|_p \neq 0$  for all  $p \in \mathcal{M}$  and  $\nabla dt = 0$ .

Recall also the definition given at the very start of the course. We have a 4-dimensional topological manifold with a smooth atlas,  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$ , carrying a torsion free connection,  $\nabla$ , but now we also require the connection be compatible with a Lorentzian metric,  $g$ , and a so-called *time orientation*,  $T$ . So we need the sextuple  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla, g, T)$ .

## 13.1 Time Orientation

The absolute time function in Newtonian spacetime associates to each  $p \in \mathcal{M}$  a time. That is, given any point you can just quote the time of that point unarguably. We used the absolute time function to define a future directed vector field  $X$  as  $dt : X > 0$ . Pictorially this is given by an arrow pointing to the ‘upper side’ of a tangent plane to a constant  $t$  surface.



We don't have an absolute time function for our relativistic spacetime, and so we need some other way to define what a future directed vector field is. We know from the tutorials that a Lorentzian metric structure gives a double cone structure in the tangent space to each point. The question is, "can we use this double cone structure in a similar way to how we use the  $dt$  surfaces in Newtonian physics to define future/past/spatial directed vector fields?" The answer is "yes, but not by itself."

**Definition** (Time Orientation). Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}^\uparrow, g)$  be an oriented Lorentzian manifold. Then a **time orientation** is given by a smooth vector field  $T$  that

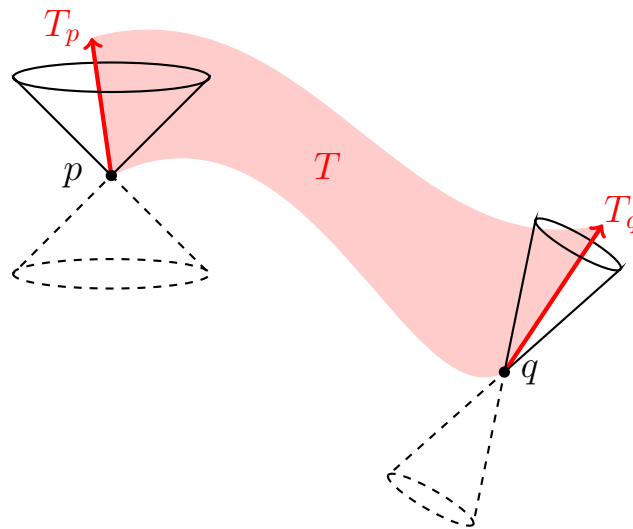
- (i) does not vanish anywhere, and



(ii)  $g(T, T) > 0$ .<sup>1</sup>

**Proposition 13.1.1.** *It is the combination of the metric and the time orientation that allows us to define future/past/spatial directed vector fields in relativistic spacetime.*

The ‘proof’ of the above proposition comes from simply breaking down the definition. The metric structure gives us a double cone structure in the tangent plane to each  $p \in \mathcal{M}$ . We want to identify one of these cones as the future and the other as the past. We know that a vector  $X$  that satisfies  $g(X, X)|_p > 0$  then it lies within *either* one of the two cones tangent to  $p$ . It doesn’t, however, tell us which cone it lies in, and so we don’t know if it’s future directed or past directed. We therefore need some method to select which cone is which. This is exactly what the time orientation does. Condition (i) tells us that it is defined everywhere, and so we can define the future cone at each point, and condition (ii) tells us that  $T$  lies within the cone (an obvious necessity). We then simply say ‘whichever cone  $T$  lies in, that is the future cone’. The final, but very important, property is that  $T$  is a *smooth* vector field. This means that the future cones at separate points are smoothly connected. That is, the selected cone doesn’t suddenly ‘flip’ as you move from point to point.



**Figure 13.1:** *Pictorial representation of the relativistic spacetime. The metric  $g$  produces a double cone structure in the tangent plane to each point of the manifold. In order to differentiate the two cones, a smooth vector field  $T \in \Gamma\mathcal{M}$  is introduced in such a way that, at each point  $p \in \mathcal{M}$ , the vector  $T_p \in T$  points within one of the two cones associated to that point. This cone is then identified as the ‘future’ relative to that point. The smoothness of  $T$  (indicated by the shaded region) ensures a smooth transition from the ‘future’ of one cone to another. Solid lined cones indicate the chosen ‘future’ cones and dashed the ‘past’ cones.*

**Remark 13.1.2.** For the Newtonian spacetime picture, it is always possible to find a so-called *stratified atlas*, in which all of the  $dt$  planes lie horizontally in the charts. For the relativistic

<sup>1</sup>In our signature which is  $(+, -, -, -)$ . For the signature  $(-, +, +, +)$  the condition would be  $g(T, T) < 0$ .

picture, this is not true; that is, we can not in general define an atlas such that all the cones line up. Physically this is not a problem because of course who cares what they look like in a chart, its the physical things that are important. However it can make calculations harder and so it is worth noting.

Notation. We shall now simply refer to relativistic spacetime as just spacetime.

Note for Newtonian spacetime a future directed vector only had to point ‘above’ the  $dt$  tangent surface and no restriction was placed on its steepness (i.e. the angle between it and the  $dt$  plane). Recall that particles are defined to travel along future-directed worldlines. Intuitively, this corresponds to the idea that there is no bound to the speed<sup>2</sup> of a particle, provided it is still future directed. This is obviously in contrast to the idea from special relativity that no massive object can travel at the speed of light (or faster).

In the spacetime picture, though, we require that the future directed lie *within* the cone. They are then bounded by the surface of the cone (which, as we will define, correspond to so-called *null* vectors). If we then identified the surface of this cone with the worldlines of light, this would correspond to exactly the condition that massive particles are bound by the speed of light.<sup>3</sup>

Let’s make this more precise.

**Postulate 1.** *The worldline  $\gamma$  of a massive particle satisfies*

$$(i) \quad g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) > 0, \text{ and}$$

$$(ii) \quad g_{\gamma(\lambda)}(T, v_{\gamma,\gamma(\lambda)}) > 0.$$

**Postulate 2.** *The worldline  $\gamma$  of a massless particle satisfies*

$$(i) \quad g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) = 0, \text{ and}$$

$$(ii) \quad g_{\gamma(\lambda)}(T, v_{\gamma,\gamma(\lambda)}) > 0.$$

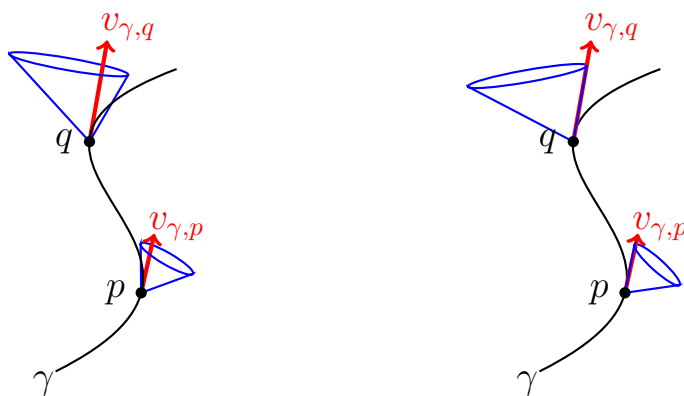
Postulate 1 tells us that (i) a massive particle’s worldline lies inside the cone structure, and (ii) it is future-directed. The only difference with postulate 2 is that the worldline of a massless particle lines on the surface of the future cone. It is at this point that we can identify the surface of the cone as the trajectory of light, as light is massless particle.

Remark 13.1.3. The wording above is a bit sloppy. The trajectory of the light is the worldline, which is defined on the manifold. The surfaces of the light cones live in the tangent spaces. It is therefore none sense to identify the two. What we mean by identify is that the velocity vectors to the worldline of light lie on the cone, which is exactly what condition (i) says.

Remark 13.1.4. Note in Newtonian mechanics, we can’t not talk about massless particles and therefore we can’t define something akin to postulate 2.

<sup>2</sup>Note we don’t have a metric and so can’t actually define a speed here.

<sup>3</sup>Again we should be careful saying speed here because speed is relative in relativity. We simply mean that there is no frame of reference where the speed of a massive particle exceeds the speed of light.



**Figure 13.2:** World lines in spacetime of a massive particle (left) and a massless particle (right). It is important to remember that the cones and velocity vectors live in the tangent space to the point, not on the manifold itself, which the above picture might lead you to believe.

*Remark 13.1.5.* Note we required that the time orientation be a non-vanishing smooth vector field. We have already seen examples of topological manifolds that do not support such things, namely the sphere. We are saved here by the fact that we can also not define a Lorentzian metric on the sphere, and so we can't even begin to try and define a time orientation.

*Example 13.1.6.* Consider the example spacetime given by  $\mathcal{M} = \mathbb{R}^4$ ,  $\mathcal{O} = \mathcal{O}_{st}$  and where the atlas contains the chart  $(\mathbb{R}^4, \mathbb{1}_{\mathbb{R}^4})$ . Let the metric in this chart be given by  $g_{(x)ij} = \eta_{ij}$  and the time orientation be  $T_{(x)} = (1, 0, 0, 0)$ . From the metric components we get vanishing Christoffel symbols  $\Gamma^k_{ij} = 0$  everywhere, and from which, by using the Levi-Civita connection, it follows that the Riemann curvature vanishes. This spacetime is therefore *flat*. This is the spacetime of special relativity and is known as **Minkowski spacetime** (or just Minkowski space). In the chart given, the representations of the light cones all stand up-right, i.e. they make a 45 degree angle to the horizontal plane.

## 13.2 Observers

**Definition** (Observer). An **observer** on a 4-dimensional spacetime  $(\mathcal{M}, \mathcal{O}, \mathcal{A}^\uparrow, g, T)$  is a worldline  $\gamma$  of a massive particle together with a choice of basis  $\{e_0(\lambda), \dots, e_3(\lambda)\}$  in each  $T_{\gamma(\lambda)}\mathcal{M}$ , with

(i)  $g(e_a, e_b) = \eta_{ab}$ , and

(ii)  $e_0(\lambda) = v_{\gamma, \gamma(\lambda)}$ ,

where<sup>4</sup>

$$\eta_{00} = 1, \quad \eta_{11} = \eta_{22} = \eta_{33} = -1, \quad \text{and} \quad \eta_{ab} = 0 \quad \forall a \neq b.$$

*Notation.* We will denote observers by  $(\gamma, e)$  where  $e$  stands for the whole basis selection.

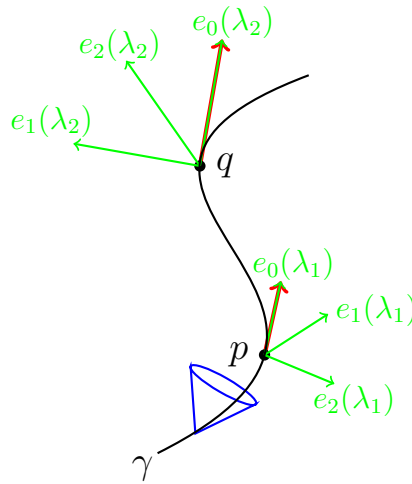
<sup>4</sup>As normal this is just in our signature. If we used  $(-, +, +, +)$  the definition of  $\eta_{ab}$  changes accordingly.

Condition (i) is the condition that the basis in each tangent space be orthonormal (in the Lorentzian sense). The significance of condition (ii) shall be clarified soon, but it is the idea that the observer does not move in space relative to themselves.

There is an alternative, more precise, definition of an observer, which we give below.

**Definition** (Observer (Frame Bundle)). An **observer** is a smooth section in the *frame bundle*  $\mathcal{LM}$  over  $\mathcal{M}$ .

We do not need to go into great detail here about what the frame bundle is, but the basic idea is that the fibres are the space of bases. That is, an element in the fibre is a quadruple of elements corresponding to a basis for that  $p \in \mathcal{M}$ . We take a section so that we have a basis at every point along the worldline and finally require the section to be smooth, so that the bases smoothly transition from one to another as you move along  $\gamma$ ; that is you don't want left to suddenly become down.



**Figure 13.3:** Pictorial representation of an observer,  $(\gamma, e)$ . The curve  $\gamma$  is that of a massive particle, and for each point  $p \in \gamma$ , the observer has a basis for  $T_p\mathcal{M}$ , such that  $e_0$  is the velocity at that point. The bases at different points are related by the smooth curve in the frame bundle — where smoothness ensures a continuous transition from the one at  $p$  to the one at  $q$ .

**Postulate 3.** A **clock** carried by a specific observer  $(\gamma, e)$  will measure a **time**  $\tau$ , known as the proper/eigen-time, between two events  $\gamma(\lambda_1)$  and  $\gamma(\lambda_2)$  as

$$\tau := \int_{\lambda_1}^{\lambda_2} d\lambda \sqrt{g(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)})}.$$

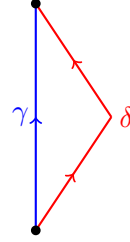
It the combination of this with condition (ii) in the definition of an observer, that tells us that they simply follow time *as they know it*. As the emphasis suggests, this time is defined relative to them. What we are highlighting here is the fact that time is a derived notion on our spacetime. Indeed, a different observer could well disagree with the time and there would be no way to determine who is correct in an absolute way, unlike with Newtonian spacetime,

where the absolute time function would give us our answer. This is the idea that time is relative and the simultaneity is ill-defined.

*Remark 13.2.1.* Note it also only makes sense for an observer to measure the time between events they have passed through. This is a subtle point but actually has far reaching impact, for example when it comes to talking about things like infinite redshift surfaces of black holes.

*Example 13.2.2.* Consider two observers on Minkowski spacetime. In the chart  $(\mathbb{R}^4, \mathbb{1}_{\mathbb{R}^4})$  let these observers be parameterised as

$$\begin{aligned}\gamma_{(x)}(\lambda) &= (\lambda, 0, 0, 0) \\ \delta_{(x)}(\lambda) &= \begin{cases} (\lambda, \alpha\lambda, 0, 0) & \lambda \leq \frac{1}{2} \\ (\lambda, (1-\lambda)\alpha, 0, 0) & \lambda > \frac{1}{2} \end{cases}\end{aligned}$$



for  $\lambda \in (0, 1)$  and  $\alpha$  a constant between 0 and 1. We calculate

$$\tau_\gamma = \int_0^1 d\lambda \sqrt{g_{(x)ij} \dot{\gamma}_{(x)}^i \dot{\gamma}_{(x)}^j} = \int_0^1 d\lambda \sqrt{1} = 1,$$

and

$$\tau_\delta = \int_0^{\frac{1}{2}} d\lambda \sqrt{1 - \alpha^2} + \int_{\frac{1}{2}}^1 d\lambda \sqrt{1 - (-\alpha)^2} = \sqrt{1 - \alpha^2}.$$

So the  $\delta$  observer measures a shorter time. This is the twin paradox and time dilation, where  $\alpha \rightarrow 1$  corresponds to  $v \rightarrow c$ .

**Postulate 4.** Let  $(\gamma, e)$  be an observer and  $\delta$  be a massive particle worldline, that is parameterised such that  $g(v_{\delta, \delta(\lambda)}, v_{\delta, \delta(\lambda)}) = 1$  everywhere along  $\delta$ .<sup>5</sup> Suppose the observer and the particle meet at some  $p \in \mathcal{M}$ , i.e.  $\gamma(\tau_1) = p = \delta(\tau_2)$ . This observer measures the 3-velocity (or spatial velocity) of this particle as

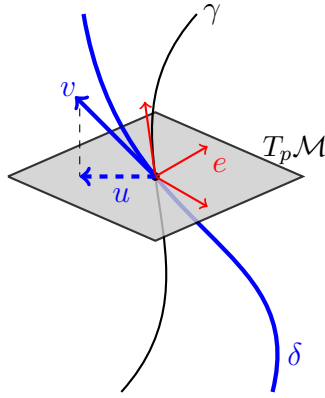
$$u_{\delta(\tau_2)} := (\epsilon^\alpha : v_{\delta, \delta(\tau_2)}) e_\alpha, \quad \alpha = 1, 2, 3,$$

where  $\epsilon^\alpha$  is the  $\alpha^{\text{th}}$  component of the so-called dual basis<sup>6</sup> of  $e$ .

We see the basis dependence clearly in the above postulate and so we know that a different observer, that also meets  $\delta$  at  $p \in \mathcal{M}$ , could get a different measurement for the 3-velocity of the massive particle. This is exactly the idea that 3-velocity is a relative concept. Note that the 4-velocity  $v_{\delta, \delta(\tau_2)}$  is objective; it is only the 3-velocity (which we can think of as a ‘projection’ of the 4-velocity into the spatial plane of the observer) that is ill-defined.

<sup>5</sup>This corresponds to normalising the worldlines to follow the clock that the observer carries. We choose to do this because it makes the following definitions easier.

<sup>6</sup>See Dr. Schuller’s Lectures on the Geometrical Anatomy of Theoretical Physics course for more details.



### 13.3 Role Of Lorentz Transformations

Lorentz transformations emerge as follows: let  $(\gamma, e)$  and  $(\tilde{\gamma}, \tilde{e})$  be observers with  $\gamma(0) = \tilde{\gamma}(0)$ . Now  $\{e_0, \dots, e_3\}$  and  $\{\tilde{e}_0, \dots, \tilde{e}_3\}$  are both bases for the tangent space  $T_{\gamma(0)}\mathcal{M}$ . Thus we can express the latter basis in terms of the former one. That is,

$$\tilde{e}_a = \Lambda^b_a e_b,$$

where  $\Lambda \in GL(4)$ .<sup>7</sup> From the definition of an observer we have

$$\begin{aligned} \eta_{ab} &= g(\tilde{e}_a, \tilde{e}_b) \\ &= g(\Lambda^m_a e_m, \Lambda^n_b e_n) \\ &= \Lambda^m_a \Lambda^n_b g(e_m, e_n) \\ \therefore \eta_{ab} &= \Lambda^m_a \Lambda^n_b \eta_{mn}, \end{aligned}$$

which tells us that the  $\Lambda$ s are elements of the Lorentz transformations,  $\Lambda \in O(1, 3)$ .

So we see that the Lorentz transformations relate the *frames* of two observers at the point that they meet. It is completely meaningless say ‘we use a Lorentz transformation to relate a frame of one observer at  $p \in \mathcal{M}$  to another observer at  $q \neq p \in \mathcal{M}$ ’. As such, Lorentz transformations act on a single tangent space to the manifold, and do *not*, by any stretch of the imagination, act on the spacetime. That is writing things like

$$\tilde{x}^\mu = \Lambda^\mu_\nu x^\nu$$

is utter nonsense. It is true in special relativity, where the spacetime is flat, that we can think of extending the tangent space over the whole manifold, and then you could say ‘ah well now it acts on all the tangent spaces and so now we can think of it as acting on the spacetime.’ This just brings us back to Remark 5.5.1, where we said that by doing so you restrict yourselves firstly to linear transformations between frames and then also to the specific case of Lorentz transformations. Physically this is not a good idea, because the objective world does not care which frames we use and therefore we should be able to transform to any frame and study the physics.

<sup>7</sup>For the unfamiliar reader that is the group of 4x4 invertible matrices, known as the 4-dimensional general linear group. See any group theory course for more details.

# 14 | Matter

There are two types of (theoretical<sup>1</sup>) matter: point matter and field matter. Examples of each are a massive point particle and the electromagnetic field, respectively. As we will see, it is this field matter that generates the curvature of spacetime, and therefore, from general relativity's point of view, field matter is the more fundamental type.

## 14.1 Point Matter

Postulate 1 and Postulate 2 already constrain the possible particle worldlines for massive and massless particles. However, it does not tell us what their precise law of motion, possibly in the presence of forces, is.

### 14.1.1 Without External Forces

We know that the equations of motion for a system can be obtained by varying a suitable action and obtaining the Euler-Lagrange equations. Below we simply provide the actions for massive and massless particles, however we will see later that they actually arise as a consequence of Einstein's field equations.

$$S_{\text{massive}}[\gamma] := m \int d\lambda \sqrt{g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)})},$$
$$S_{\text{massless}}[\gamma, \mu] := \int d\lambda \mu g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}),$$

where  $\mu$  is a Lagrange multiplier, which is introduced so that when you vary w.r.t. it you get  $g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) = 0$ , which is condition (i) in Postulate 2. Of course we also impose the condition  $g_{\gamma(\lambda)}(T, v_{\gamma,\gamma(\lambda)}) > 0$  on our actions.

It is a fair challenge to ask 'why are we starting from actions instead of just starting from the Euler-Lagrange equations?' The answer is simply the fact that we can add different actions together easily and then find the corresponding e.o.m. for that composite system. That is, composite systems have an action which is given by the sum of the constituent actions, possibly including interaction terms, and we then vary this composite action to obtain the complete e.o.m.

---

<sup>1</sup>This is really just an academic distinction, as it is often useful to think about these two separate kinds of matter and treat them accordingly. Of course in the real world we just have matter.

### 14.1.2 Presence of External Forces

Roughly speaking, in special relativity, the reaction of a particle to a force is not instantaneous, but has some time delay. This time delay is explained by the fact that forces are mediated by fields and if the particle is to react to the field it must be *coupled*. So what we really mean by ‘presence of external forces’ is ‘presence of fields to which the particles couple’.

The prime example for action of a particle coupling to an external field is that of a massive charged particle coupling to the electromagnetic field,

$$S[g; A] := \int d\lambda \left[ m \sqrt{g_{\gamma(\lambda)}(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)})} + q(A : v_{\gamma, \gamma(\lambda)}) \right],$$

where  $A \in \Gamma T\mathcal{M}$  is the electromagnetic potential on  $\mathcal{M}$  and  $q \in \mathbb{R}$  is the charge of the particle.

*Notation.* We have used a semi-colon in the arguments of the action to indicate that we treat  $A$  as a fixed quantity, and so we do not vary w.r.t. it.

#### Exercise

Let  $L_{\text{int}} := q(A : v_{\gamma, \gamma(\lambda)})$ . Use a chart  $(U, x)$  to show that the Euler-Lagrange equations of the above action are

$$m(\nabla_{v_\gamma} v_\gamma)^a = -q F^a{}_b \dot{\gamma}^b,$$

where  $F^a{}_b := g^{ac}(A_{c,b} - A_{b,c})$ .

*Hint: If you get stuck, this one is done on the videos.*

The result of the above exercise is the *Lorentz force law* on a charged particle in the electromagnetic field. Note also that the action given above is reparameterisation ( $\lambda \rightarrow \lambda'(\lambda)$ ) invariant, as it must be if it is to be the action for the Lorentz force law.

## 14.2 Field Matter

**Definition** (Classical Field Matter). **Classical<sup>2</sup> field matter** is any tensor field on spacetime whose equations of motion derive from an action.

This definition is of course quite unhelpful, but we use it because its hard to give another definition that does no over or understate what field matter is. We rather see what field matter is by considering Maxwell’s action.<sup>3</sup>

$$S_{\text{Maxwell}}[A; g] := \frac{1}{4} \int_{\mathcal{M}} dx^4 \sqrt{-g} F_{ab} F_{cd} g^{ac} g^{bd},$$

where, for the time being, we have assume the metric to be fixed.

*Remark 14.2.1.* Note that we use  $\sqrt{-g}$  not just  $\sqrt{g}$ . This is because we are looking at a Lorentzian metric which has negative determinant.

<sup>2</sup>As in non-quantum.

<sup>3</sup>In the definition below we have assumed there is a chart that covers the whole spacetime. If this is not the case, the definition holds, but we just need to use the ideas discussed at the end of lecture 12.



Example 14.2.2. If we take our spacetime to be Minkowski spacetime and use the chart  $(\mathbb{R}^4, \mathbb{1}_{\mathbb{R}^4})$ , we have  $g = -1$ ,  $g^{ab} = \eta^{ab}$  and so the Maxwell action just becomes

$$S_{\text{Maxwell}}^{\text{Mink}}[A; g] = \frac{1}{4} \int_{\mathbb{R}^4} dx^4 F_{ab} F^{ab},$$

which may be familiar to the reader. Note, however, that it only takes this form *in this chart*. If we chose to use polar coordinates, we would not have  $g = -1$  nor  $g^{ab} = \eta^{ab}$ .

The Euler-Lagrange equations (in a chart) for a field action are given by

$$0 = \frac{\partial \mathcal{L}}{\partial A_m} - \frac{\partial}{\partial x^s} \left( \frac{\partial \mathcal{L}}{\partial \partial_s A_m} \right) + \frac{\partial}{\partial x^t} \frac{\partial}{\partial x^s} \left( \frac{\partial^2 \mathcal{L}}{\partial \partial_t \partial_s A_m} \right) - \dots,$$

where the trend continues with alternating sign. Calculating the Euler-Lagrange equations for the Maxwell action gives the inhomogeneous Maxwell equations

$$(\nabla_a F)^{ab} = 0.$$

If we had considered the action including a coupling to a current  $j \in \Gamma T\mathcal{M}$ ,

$$S[A; g, j] = \frac{1}{4} \int_{\mathcal{M}} dx^4 \sqrt{-g} (F_{ab} F_{cd} g^{ac} g^{bd} + A : j),$$

the Euler-Lagrange equations become

$$(\nabla_a F)^{ab} = j^b.$$

The remaining two Maxwell equations can be obtained via

$$(\nabla_{[a} F)_{bc]} = 0.$$

Remark 14.2.3. There is a much nicer way (in my opinion) to write Maxwell's equations, but it involves properly introducing the exterior derivative,  $d$ , and the Hodge star,  $\star$ . The formulas are

$$dF = 0 \quad \text{and} \quad d\star F = \star j,$$

where  $F = dA$  is the Faraday tensor and  $J$  is the current density. The interested reader is directed to Example 3.14 and Exercise 3.28 of Renteln's *Manifolds, Tensors, and Forms* textbook (or many other textbooks which will cover it).

There are other well liked (by textbooks) examples, including the *Klein-Gordan* action

$$S_{\text{KG}}[\phi] := \int_{\mathcal{M}} d\phi \sqrt{-g} [g^{ab} \partial_a \phi \partial_b \phi - m^2 \phi^2],$$

where  $\phi \in C^\infty(\mathcal{M})$ , is a scalar field on the spacetime.

### 14.3 Energy-Momentum Tensor Of Matter Fields

So far we have always assumed that we are given the Lorentzian metric for our spacetime. The obvious question is ‘which metric?’ If we are to describe a physical system, e.g. the universe, obviously we want a metric that will give us precisely these physical results. We therefore want to obtain some action for the metric tensor field itself, which we shall denote  $S_{\text{grav}}[g]$ . This action will be added to any matter action  $S_{\text{matter}}[\dots]$ , in order to describe the total system.

*Example 14.3.1.* If we take the Maxwell action we have

$$S_{\text{total}}[g, A] = S_{\text{grav}}[g] + S_{\text{Maxwell}}[A, g],$$

where the metric is no longer taken as fixed in the Maxwell action, i.e. we use a comma not a semi-colon.

Of course, varying the total action w.r.t. the arguments of the matter action ( $A$  in the above example) will just give us the matter e.o.m (Maxwell’s equations for the example). However, now varying w.r.t.  $g$  will give a contribution from both  $S_{\text{grav}}$  and  $S_{\text{matter}}$ ,

$$G_{ab} = 8\pi G_N T_{ab},$$

where  $G_{ab}$  is the contribution from  $S_{\text{grav}}$ ,  $T_{ab}$  is the contribution from  $S_{\text{matter}}$ , and where we have included the factor  $8\pi G_N$ , where  $G_N$  is Newton’s constant, for convention. This is the so-called **Einstein equation**.

Once we have fixed  $S_{\text{grav}}$  we will of course always obtain the same  $G_{ab}$ , but the  $T_{ab}$  depends on which matter action we are using. We can ensure that our e.o.m. are always satisfy the Einstein equation by introducing the following definition.

**Definition** (Energy-Momentum Tensor). Let  $S_{\text{matter}}[\dots, g]$  be any matter action that couples to the metric. Then we define the components of the **energy-momentum tensor** via

$$T^{ab} := \frac{-2}{\sqrt{-g}} \left[ \frac{\partial \mathcal{L}_{\text{matter}}}{\partial g_{ab}} - \frac{\partial}{\partial x^s} \left( \frac{\partial \mathcal{L}_{\text{matter}}}{\partial \partial_s g_{ab}} \right) + \frac{\partial}{\partial x^t} \frac{\partial}{\partial x^s} \left( \frac{\partial^2 \mathcal{L}_{\text{matter}}}{\partial \partial_t \partial_s g_{ab}} \right) - \dots \right],$$

where the terms continue with alternating sign.

*Remark 14.3.2.* In the above definition we said ‘that couples to the metric’. This is true for all of the classical matter fields of the standard model, and so we will always have this in this course.

*Remark 14.3.3.* The minus sign in that above definition is included to ensure  $T(\epsilon^0, \epsilon^0) > 0$ , which tells us that energy is positive.

*Example 14.3.4.* For the Maxwell action, the energy-momentum tensor is

$$T_{ab} = F_{am} F_{bn} g^{mn} - \frac{1}{4} F_{mn} F^{mn} g_{ab}.$$

# 15 | Einstein Gravity

*Remark 15.0.1.* In this lecture (and the proceeding ones) we will use a lot of indices. This obviously implies that we are using charts and so our results could turn out to be physically nonsense in the end. This remark claims that the results, unless otherwise specified, are indeed chart independent, and we simply use indices to make it notationally clearer what we're doing.

Recall that in lecture 9, we were able to reformulate Poisson's equation,  $\nabla^2\phi = 4\pi G_N\rho$ , in terms of the curvature of *Newtonian spacetime*, namely as  $\text{Ric}_{00} = 4\pi G_N\rho$ . This prompted Einstein to postulate that the relativistic field equations for the Lorentzian metric  $g$  of spacetime<sup>1</sup> as

$$\text{Ric}_{ab} = 8\pi G_N T_{ab}.$$

However, this equation suffers from a problem: it can be shown<sup>2</sup> that  $(\nabla_a T)^{ab} = 0$ . This would imply that  $(\nabla_a \text{Ric})^{ab} = 0$ , which is *not* true generically. Einstein tried to argue this problem away, but it turns out that these equations are fundamentally wrong and cannot be upheld, and we to obtain a new set of field equations.

## 15.1 Hilbert

Hilbert was an variation principle specialist and had the brilliant idea to say "The right-hand side of the gravitational field equations come from an action, so why don't we try and obtain the left-hand side from an action too?" He decided to work through the simplest actions<sup>3</sup> he could until he obtained one that worked. His final result was the following:

$$S_H[g] := \int_{\mathcal{M}} \sqrt{-g} R := \int_{\mathcal{M}} \sqrt{-g} \text{Ric}_{ab} g^{ab}.$$

The aim is to vary this action w.r.t.  $g_{ab}$  and obtain some tensor, which we denote  $-G^{ab}$ .<sup>4</sup> The obvious next step is to do this variation and find out what  $-G^{ab}$  is.

<sup>1</sup>Recall that when we say spacetime we mean *relativistic spacetime*.

<sup>2</sup>See my notes for Dr. Shiraz Minwalla's String theory course for an outline of the proof.

<sup>3</sup>That is he asked what combination of objects will give a scalar field.

<sup>4</sup>The minus sign is just a convention choice.

## 15.2 Variation of $S_H$

We have

$$\delta S_H[g] = \int_{\mathcal{M}} [\delta\sqrt{-g} \cdot \text{Ric}_{ab} \cdot g^{ab} + \sqrt{-g} \cdot \delta\text{Ric}_{ab} \cdot g^{ab} + \sqrt{-g} \cdot \text{Ric}_{ab} \cdot \delta g^{ab}].$$

Let's consider this term by term.

First let's consider  $\delta g$ , from  $g := \det g = \exp(\text{Tr}(\ln g))$ , we have

$$\delta g = -\frac{g \cdot g^{ab} \delta g_{ab}}{2\sqrt{-g}} = -\frac{1}{2}\sqrt{-g} \cdot g^{ab} \cdot \delta g_{ab}$$

Next, let's look at  $\delta g^{ab}$ . We know  $g^{ac}g_{cb} = \delta_b^a$ , and so we have

$$\delta g^{ac} \cdot g_{cb} + g^{ac} \cdot \delta g_{cb} = 0 \quad \implies \quad \delta g^{ab} = -g^{am} \cdot g^{bn} \cdot \delta g_{mn},$$

where we have relabelled the indices on the latter equation.

Finally, we have to work out  $\delta\text{Ric}_{ab}$ . This one is a little more tricky, and involves us making some clever steps. We start of by considering normal coordinates,<sup>5</sup> giving us

$$\delta\text{Ric}_{ab} = \delta(\Gamma^m_{am,b}) - \delta(\Gamma^m_{ab,m}) = (\delta\Gamma^m_{am})_{,b} - (\delta\Gamma^m_{ab})_{,m}.$$

This seems like an awful idea because the results depend on the fact that we're in normal coordinates. However we now use a remarkably clever trick. Recall that  $\Gamma$ s are not tensor components because they have a term in their transformation given by second derivatives. We now note that this term does not depend on the  $\Gamma$ s themselves, and therefore if we take the difference of two  $\Gamma$ s this term will vanish in the transformation. That is

$$\Gamma^k_{(x)ij} - \tilde{\Gamma}^k_{(x)ij}$$

transform like the components of a tensor. We then note that the derivative essentially compares two  $\Gamma$ s, and so conclude that the the derivatives of the  $\Gamma$ s are indeed (1,2)-tensor components. This is good, but we then run into the problem that we can't just take the derivative of a tensor. This problem is solved quite easily by the fact that we are considering normal coordinates and so the covariant derivative and the partial derivative coincide (that is the  $\Gamma$ s vanish in normal coordinates). So we have

$$\begin{aligned} \sqrt{-g} \cdot g^{ab} \cdot \delta\text{Ric}_{ab} &= \sqrt{-g} \cdot g^{ab} \cdot [(\delta\Gamma^m_{am})_{,b} - (\delta\Gamma^m_{ab})_{,m}] \\ &= \sqrt{-g} \cdot [(g^{ab}\delta\Gamma^m_{am})_{,b} - (g^{ab}\delta\Gamma^m_{ab})_{,m}] \\ &=: \sqrt{-g} [A^b_{;b} - B^m_{;m}], \end{aligned}$$

where we have defined  $A^b := g^{ab}\delta\Gamma^m_{am}$  and similarly for  $B^b$  we have used the metric compatibility condition (as spacetime is equipped with the Levi-Civita connection) to 'move  $g^{ab}$  inside the covariant derivative'. Next we have

$$\sqrt{-g}_{,b} = -\frac{1}{2}\sqrt{-g} \cdot g^{ac} \cdot g_{ac,b},$$

---

<sup>5</sup>That is a *local* flat space, so the  $\Gamma$ s vanish, but their derivatives need not.

which, using normal coordinates again along with the metric compatibility condition, gives us

$$(\sqrt{-g}A)^b{}_{,b} = \sqrt{-g} \left[ -\frac{1}{2}g^{ac} \cdot g_{ac,b} \cdot A^b + A^b{}_{;b} \right] = \sqrt{-g}A^b{}_{;b}.$$

So we finally arrive at

$$\sqrt{-g} \cdot g^{ab} \cdot \delta \text{Ric}_{ab} = (\sqrt{-g}A)^b{}_{,b} - (\sqrt{-g}B)^b{}_{,b}.$$

Collecting terms, we have

$$\delta S_H[g] = \int_{\mathcal{M}} \left[ \frac{1}{2} \sqrt{-g} g^{cd} (\delta g_{cd}) g^{ab} \text{Ric}_{ab} - \sqrt{-g} g^{ac} g^{bd} \delta g_{cd} \text{Ric}_{ab} + (\sqrt{-g}A)^b{}_{,b} - (\sqrt{-g}B)^b{}_{,b} \right].$$

We now notice that the last two terms are volume integrals over divergences and so, by Stoke's law, are surface terms. These terms will therefore not contribute to the equations of motion, which is what we're interested in, and so we can essentially just drop them. This leaves us finally with

$$0 = \delta S_H = \int_{\mathcal{M}} \sqrt{-g} \delta g_{ab} \left[ \frac{1}{2} g^{ab} R - \text{Ric}^{ab} \right],$$

where we have used  $R := g^{ab} \text{Ric}_{ab}$  and  $\text{Ric}^{cd} := g^{ac} g^{bd} \text{Ric}_{ab}$  and then relabelled the indices. This must hold for an arbitrary variation  $\delta g_{ab}$ , and so we conclude

$$G^{ab} = \text{Ric}^{ab} - \frac{1}{2} g^{ab} R.$$

This expression is known as the (components of the) **Einstein curvature**. They are the field equations for the vacuum spacetime, i.e. one with no matter. If we include matter into our spacetime, our action changes in accordance with the previous lecture, and we obtain<sup>6</sup>

$$G_{ab} = \text{Ric}_{ab} - \frac{1}{2} g_{ab} R = 8\pi G_N T_{ab}.$$

These are known as the **Einstein equations**, as Einstein also arrived at this result using more physical arguments. As such the Hilbert action is often called the Einstein-Hilbert action and is denoted

$$S_{EH}[g] = \int_{\mathcal{M}} \sqrt{-g} R.$$

*Remark 15.2.1.* With the remark made at the start of this lecture in mind, we have a clear way to distinguish between the Riemann curvature, the Ricci curvature and the Ricci scalar, namely the number of indices. We shall therefore write the Einstein equations simply as

$$G^{ab} = R^{ab} - \frac{1}{2} g^{ab} R = 8\pi G_N T^{ab}.$$

We do this both for notational brevity, but also because this is how it appears in basically all textbooks.

---

<sup>6</sup>Note we have moved the indices back down here. It is annoying common to just move the indices in equations up and down like this, however you should be careful when doing this as in order to do it the metric components have been used twice.

### 15.3 Solution Of The $(\nabla_a T)^{ab} = 0$ Issue

Recall the Bianchi identity in components<sup>7</sup>

$$R^a{}_{b[mn;\ell]} = R^a{}_{bmn;\ell} + R^a{}_{b\ell m;n} + R^a{}_{bn\ell;m} = 0.$$

If we then use the metric compatible condition we obtain the so-called *contracted* Bianchi identity

$$R_{ab[mn;\ell]} = R_{abmn;\ell} + R_{ab\ell m;n} + R_{abn\ell;m} = 0.$$

Further contraction (i.e. using the metric components to set indices equal to each other) can be used to give<sup>8</sup>

$$R^\ell{}_{m;\ell} = \frac{1}{2}R_{;m},$$

and so we get

$$G^{ab}{}_{;a} := (\nabla_a G)^{ab} = 0,$$

which resolves our problem.

### 15.4 Variations of The Field Equations

Firstly let's choose units such that  $G_N = 8\pi$ , so the factor in Einstein's equations becomes 1, so we have

$$R_{ab} - \frac{1}{2}g_{ab}R = T_{ab}.$$

We now want to manipulate this a little to express it in different ways.

#### 15.4.1 Ricci Scalar Expression

First consider contracting with  $g^{ab}$ . This gives

$$\begin{aligned} g^{ab}T_{ab} &= g^{ab}R_{ab} - \frac{1}{2}g^{ab}g_{ab}R \\ T &= R - 2R \\ T &= -R, \end{aligned}$$

where we have used  $g^{ab}g_{ab} = \delta_a^a = \dim \mathcal{M} = 4$ , and where we have defined  $T := g^{ab}T_{ab}$ . Substituting this back into the Einstein equations, we get

$$R_{ab} = T_{ab} - \frac{1}{2}g_{ab}T =: \widehat{T}_{ab}.$$

So we have  $R_{ab} = \widehat{T}_{ab}$ , which is of the same form as what Einstein proposed right at the start of this lecture, the only difference being we need to use the modified energy-momentum tensor.

<sup>7</sup>Technically there is a 3! missing here, but the right-hand side vanishes so it's not important.

<sup>8</sup>See tutorial.

### 15.4.2 Cosmological Constant

We could modify the Einstein-Hilbert action to include some constant term  $\Lambda$ , known as the **cosmological constant**. That is

$$S_{EH}[g] = \int_{\mathcal{M}} \sqrt{-g}(R + 2\Lambda).$$

You might ask why we would do such a thing, and the answer is to do with talking about an expanding universe. Einstein initially included it as a negative value in order to ensure the universe was static (i.e. not expanding). Hubble then showed that the universe was indeed expanding and so we could have  $\Lambda = 0$ , which prompted Einstein to call this his ‘greatest blunder’. It now turns out that we know the universe is not only expanding, but it is also accelerating in its expansion and so we require  $\Lambda > 0$ .

The real question is, though, what on Earth is the cosmological constant? Well, if we think of  $\int_{\mathcal{M}} \sqrt{-g}R$  as being gravity, it appears to us that  $\Lambda$  is some matter contribution to the action that is always there. That is it has a contribution to the field equations of the form  $\Lambda g_{ab}$ .

This is rather remarkable though, as  $\Lambda$  is a constant and  $g_{ab} \neq 0$  everywhere<sup>9</sup> and so this matter contribution takes the same value over the *entirety* of the universe! Note it also does not couple to any fields. This is what people refer to as *dark energy*.

The next question is: what causes dark energy? The answer is nobody knows.

Our observations tell us that, although  $\Lambda \neq 0$ , it is very small. This is what troubles us. We need something that exists throughout the whole universe in a constant manner, but that also doesn’t contribute much energy to the universe system. For an idea of how bad this problem is, consider the following proposal.

It was suggested that vacuum fluctuations of quantum field theories could be the root of dark energy. However, the calculation for the contribution to the energy from QCD fluctuations alone gave a value for  $\Lambda$  that was 120 orders of magnitude too big!

---

<sup>9</sup>Well everywhere it’s defined at least. Who knows what values it takes at places like the singularity.

# 16 | Optical Geometry I

*Remark 16.0.1.* This lecture is given by Dr. Werner, and he decides to use the opposite signature to Dr. Schuller, namely he uses  $(-, +, +, +)$ . I shall change to this signature too as Dr. Schuller changes to it anyways in lecture 20 and it is also my preferred signature.

*Notation.* Dr. Werner uses the notation that Greek indices represent spacetime components (i.e.  $\mu = 0, 1, 2, 3$ ), whereas Latin indices represent spatial components (i.e.  $i = 1, 2, 3$ ). We shall use the opposite convention here as that is what we have been using throughout these notes.

We want to look at *gravitational lensing*, which is the bending of light in space. Historically, gravitational lensing played a really important part in the field of general relativity, as it was one of the first proposed predictions of the theory. In order to study gravitational lensing we shall first return to Fermat's principle and try and express it in the context of GR.

## 16.1 Fermat's Principle

Classically, Fermat's principle is the statement that light will follow a path that minimises its time. That is,

$$0 = \delta \int_{\gamma} dt = \delta \int_{\gamma} \frac{1}{v} dl = \delta \int_{\gamma} \frac{n}{c} dl.$$

There is a problem with trying to do this in GR, though; light rays follow null geodesics and so have zero spacetime length. That is  $g(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)}) = 0$  for all  $\gamma$  that represent the path of a light ray.

For us to proceed here, we are going to assume that our spacetime is so-called *stationary*.

**Definition** (Stationary Spacetime). A spacetime  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g, T)$  is called **stationary** if it admits a Killing vector field  $K$  such that  $g(K, K) < 0$ .<sup>1</sup>

*Claim 16.1.1.* A stationary spacetime is one where we can find a chart such that the components of the metric do not depend on time.

*Proof.* Recall a vector field is Killing if  $\mathcal{L}_K g = 0$ . The exercise at the end of lecture 11 shows that in a chart this condition reads

$$T^c g_{ab,c} + g_{cb} T^c{}_{,a} + g_{ca} T^c{}_{,b} = 0.$$

---

<sup>1</sup>Technically all we require is that the spacetime has an asymptotically flat region and that the Killing vector field satisfies  $g(K, K) < 0$  in this region. This distinction does carry forward into some of the next expressions, however we shall ignore it in these notes as the general idea holds.



Now imagine we pick a chart such that  $T = \delta_0^a \partial_a = \partial_0$ , then the second two terms vanish and we are simply left with

$$g_{ab,0} = 0,$$

which is the statement that the metric components are time-independent in this chart.  $\square$

In the chart described above, a general stationary spacetime is one whose metric is of the form

$$g = -dt \otimes dt + \omega_\mu (dt \otimes dx^\mu + dx^\mu \otimes dt) + h_{\mu\nu} dx^\mu \otimes dx^\nu,$$

where  $h_{\mu\nu} = \text{diag}(+, +, +)$ , and where both  $h_{\mu\nu}$  and  $\omega_\mu$  are functions of the  $x$ s only, i.e.  $h_{\mu\nu,0} = 0 = \omega_{\mu,0}$ .

**Definition** (Static Spacetime). A spacetime  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g, T)$  is called **static** if it is stationary and *hypersurface-orthogonal*, which essentially means  $\omega_\mu = 0$  for all  $\mu \in \{1, 2, 3\}$ .

*Remark 16.1.2.* The  $\omega_i$ s have the nice geometrical interpretation of being (the spatial part) of a twisting vector, which corresponds to a rotation of the spacetime. So the difference between a stationary and static spacetime can be thought of as allowing or not rotation.

I will finish typing up this lecture, and the next three later. I have typed up all of Dr. Schuller's lectures though. These lectures are very well taught (I just decided to finish Dr. Schuller's stuff first), so please watch them if you haven't already.

1. Fermat's principle GR — it is the variation of the arrival time that vanishes, not the total time.
2. Finsler-Randers Geometry
3. Optical metrics
4. Schwarzschild
5. Gaussian Curvature

## 17 | Optical Geometry II

1. Geodesic Deviation
2. Show that for Schwarzschild spacetime that geodesics diverge locally everywhere.
3. If that is true, how do we get multiple images of stars? Must be some global property converging them again.
4. Guass-Bonnet

## 18 | Canonical Formulation of GR I

# 19 | The Canonical Formulation of GR II

## 20 | Cosmology: The Early Epoch

Cosmology is the study of the spacetime of the entire universe. As we have seen, Einstein's equations are highly non-linear and so are hard to solve.<sup>1</sup> Nevertheless, they do allow us to ask the scientific question of biggest possible scope: "How did the universe evolve?"

Now it seems like an incredibly bold task to try and solve Einstein's equations for the entire universe, when we have just said they are already very difficult to solve on much smaller scales with restrictive conditions. In fact, solving Einstein's equations for the entire universe can be thought of as the 'most difficult' problem because our energy-momentum tensor must include *all* the matter in the universe!

In order to solve the problem we are going to insert some ideas and then use these recklessly, by which we mean that it is not a priori clear whether this is a valid procedure. We will continue to comment on this as we go along.

*Remark 20.0.1.* As mentioned in Remark 16.0.1, we shall now use the signature to  $(-, +, +, +)$  so that when we restrict ourselves to the spatial part of the universe (which we will do next) we don't have to carry minus signs around.

### 20.1 Assumption of Spatial Homogeneity & Isotropy at Large Scale

The idea is to assume that if we were to 'zoom out' far enough and look at the universe on the large scale, the small scale 'untidiness' would disappear and we would obtain a homogeneous (same at all places) and isotropic (same in all directions) picture.

*Remark 20.1.1.* Note we have said *spatial* homogeneity and isotropy. It would be a bit much to assume that the universe is also homogeneous and isotropic in time. However, relativity is based around the idea that space and time are essentially indistinguishable (in the sense that they are two parts of the same thing) and so we need to clarify what we mean by spatial and temporal. We shall return to this.

These assumptions allow us to make a symmetric ansatz for the metric of the universe spacetime, and in doing so we massively simplify Einstein's equations. It is important to note however that doing this is very reckless. We are not guaranteed that making such an ansatz a priori will give us the same solution we would obtain from first solving the problem and *then* imposing the ansatz. However, this is the method used in mainstream cosmology and so we will adopt it here.

---

<sup>1</sup>Indeed no one has been able to write down a *general* solution.

*Remark 20.1.2.* Note just because an idea is adopted by mainstream research, it need not be true. This is just a remark to highlight the point that when doing research it is not always a bad idea to disagree with mainstream ideas (provided you have evidence to support your claims).

Recalling the discussion at the end of lecture 11, we can formulate the above symmetry assumptions more precisely. We say that our spacetime admits 6 spatial<sup>2</sup> Killing vector fields, which we denote  $J_1, J_2, J_3, P_1, P_2$  and  $P_3$ . The  $J_a$ s correspond to the rotational symmetries (i.e. the isotropy condition) and the  $P_a$ s correspond to the translation symmetries (i.e. the homogeneity condition). They satisfy the following relations

$$[J_a, J_b] = \sum_{c=1}^3 \epsilon_{abc} J_c, \quad [P_a, P_b] = 0, \quad \text{and} \quad [J_a, P_b] = \sum_{c=1}^3 \epsilon_{abc} J_c.$$

*Remark 20.1.3.* Note this is a condition on the whole spacetime, not just on some kind of spatial slice (whatever that would mean). It is only by providing these 6 Killing vector fields that we can work out what we meant by ‘spatial’ homogeneity and isotropy above. That is we look at the ‘planes’ spanned by these 6 Killing vector fields and identify them as spatial planes and the vector fields orthogonal to them as the temporal flow.

It is important to be careful with taking this idea too far. We know that moving observers have ‘tilted time-axis’ relative to each other, and so we could be led to think that their spatial planes, and therefore the Killing vectors, also tilt. Clearly this is unphysical (a symmetry of a metric is independent of an observer noticing it) and so cannot be the case.

Fortunately, for 6 ( $\mathbb{R}$ -linearly independent) Killing vector fields there is a short cut<sup>3</sup> to understand how a metric with such symmetries looks like.

**Lemma 20.1.4.** *On a  $d$ -dimensional manifold the maximal number of  $\mathbb{R}$ -linearly independent Killing vector fields is  $d(d+1)/2$ , which is equal to the number of independent component functions of a metric in  $d$ -dimensions.*

*Proof.* (By example)<sup>4</sup> Recall the Killing vector condition can be expressed as

$$g(\nabla_X K, Y) + g(X, \nabla_Y K) = 0.$$

If we pick a chart where we simply have  $X = \partial_a$  and  $Y = \partial_b$  (that is they point along one of the basis directions each) then this condition becomes<sup>5</sup>

$$K_{b;a} + K_{a;b} = 0.$$

Now recall that the definition of the Riemann tensor (in the absence of Torsion) can be written as

$$\text{Riem}^d_{cab} K_d = \nabla_b \nabla_a K_c - \nabla_a \nabla_b K_c =: K_{c;a;b} - K_{c;b;a}.$$

<sup>2</sup>That is  $g(K, K) > 0$  in our updated signature.

<sup>3</sup>Relative to having to introduce a coordinate chart and work it out from that.

<sup>4</sup>A complete proof is not too much different and can be found in Weinberg’s *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, in Part 4, Chapter 13, Section 1.

<sup>5</sup>Try showing this as an additional exercise.

Putting this into the Bianchi identity

$$\text{Riem}^d_{[abc]} = 0$$

gives us

$$(K_{a;b} - K_{b;a})_{;c} + (K_{b;c} - K_{c;b})_{;a} + (K_{c;a} - K_{a;c})_{;b} = 0,$$

which using the charted Killing condition gives us

$$K_{a;b;c} = K_{c;b;a} - K_{c;a;b} = -R^d_{cab}K_d.$$

Now comes the ‘by example’ part. Consider a  $d$ -dimensional flat space, then the Riemann tensor components all vanish and we can pick a chart such that the  $\Gamma$ s vanish, and so the covariant derivative simply becomes the partial derivative. We therefore have

$$K_{a,b,c} = 0, \quad \iff \quad K_a = \beta_{ab}x^b + \alpha_a,$$

for constants  $\beta_{ab}$  and  $\alpha_a$ .

We now just need to impose the linearly independent condition. Antisymmetry tells us that  $\beta_{ab} = -\beta_{ba}$  and so there are  $d(d-1)/2$  independent  $\beta_{ab}$  components and clearly there are  $d$  independent  $\alpha_a$  components. Adding these together gives

$$\frac{d(d-1)}{2} + d = \frac{d(d+1)}{2}.$$

□

A space with the maximal number of Killing vector fields is called a maximal space and the metric is said to be maximally symmetric.

From the above lemma (and the fact that  $3(3+1)/2 = 6$ ) we see that the spacial metric induced from the spacetime metric on the spatial slices spanned by the Killing vector fields is the metric of a maximally symmetric 3-space.

**Definition** (Sectional Curvature). Given a Riemannian manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, g)$  and two linearly independent tangent vectors to the same point  $X, Y \in T_p\mathcal{M}$ , we can define the **sectional curvature** as

$$\kappa(X, Y) := \frac{g(\text{Riem}(\cdot, Y, X, Y), X)}{g(X, X)g(Y, Y) - [g(X, Y)]^2},$$

where  $\text{Riem}(\cdot, Y, X, Y) = \nabla_X \nabla_Y Y - \nabla_Y \nabla_X Y \in T_p\mathcal{M}$ .

The sectional curvature can be seen geometrically as the product of the curvatures at a point. For example, both the curvature directions a sphere ‘go inwards’ and so they have the same sign and therefore  $\kappa > 0$ . Alternatively, the throat of a wormhole has  $\kappa < 0$ .

*Remark 20.1.5.* Note that the sectional curvature actually only depends on the 2-plane  $\sigma_p \subset T_p\mathcal{M}$  spanned by  $X$  and  $Y$ . For a  $d > 2$  dimensional spaces, the different 2-planes tell us about the product of the different curvatures.

**Definition** (Constant (Sectional) Curvature). A space is said to have **constant (sectional) curvature** if  $\kappa$  takes the same value at every point on  $\mathcal{M}$  and every 2-plane.

**Proposition 20.1.6.** *Riemannian manifolds with constant curvature can be of one of three geometries:*

- (i) flat  $\kappa = 0$ ,
- (ii) spherical  $\kappa > 0$ , or
- (iii) hyperbolic  $\kappa < 0$ .

Such spaces are called **space forms**.

For a spacetime with constant curvature we have

$$\text{Riem}_{\alpha\beta\rho\delta} = \kappa(\gamma_{\alpha\rho}\gamma_{\beta\delta} - \gamma_{\alpha\delta}\gamma_{\beta\rho}),$$

where  $\gamma_{\alpha\beta}$  is the spatial metric which can be written in a certain chart as

$$\gamma_{\alpha\beta}(r, \theta, \varphi) = \begin{pmatrix} \frac{1}{1-kr^2} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}_{\alpha\beta}.$$

The spacetime metric then has the form

$$g_{ab}(t, r, \theta, \varphi) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{a^2(t)}{1-kr^2} & 0 & 0 \\ 0 & 0 & a^2(t)r^2 & 0 \\ 0 & 0 & 0 & a^2(t)r^2 \sin^2 \theta \end{pmatrix}_{ab},$$

where  $a : \mathbb{R} \rightarrow \mathbb{R}$  is called the **scale factor**, which is all the freedom left after the symmetry reduction. Geometrically,  $a(t)$  tells us how the different spatial slices are related. That is if we had a spherical spacial space and  $a(t) = t$  then the spatial spaces would be a set of spheres of increasing radius.

**Lemma 20.1.7.** *We can redefine  $a(t)$  such that our condition for the geometries of constant curvature becomes  $\kappa = 0, \pm 1$ .*

*Remark 20.1.8.* Note that, provided  $a(t)$  is not constant, the time vector field (i.e. the one orthogonal to all the Killing vector fields) is not Killing. That is, in our chart

$$\mathcal{L}_{\frac{\partial}{\partial t}} g \neq 0.$$

This is the statement that the universe need not be stationary.

## 20.2 Einstein Equations

### 20.2.1 Ricci Tensor

Let's find the  $\Gamma$ s for our spacetime metric above. We have

$$\begin{aligned} \Gamma^t_{\alpha\beta} &= \frac{1}{2}g^{t\sigma}(g_{\alpha\sigma,\beta} + g_{\beta\sigma,\alpha} - g_{\alpha\beta,\sigma}) \\ &= -\frac{1}{2}\partial_t \langle a^2(t)\gamma_{\alpha\beta}(r, \theta, \varphi) \rangle \\ &= a\dot{a}\gamma_{\alpha\beta}, \end{aligned}$$



where we have used the fact that  $g_{ab}$  is diagonal so only we must take  $\sigma = t$ . Similarly we have

$$\Gamma^\alpha_{t\beta} = \frac{\dot{a}}{a} \delta^\alpha_\beta.$$

We also have that the all spatial  $\Gamma$ s (i.e. the ones of the form  $\Gamma^\alpha_{\beta\rho}$ ) only depend on the 3-metric  $\gamma$ .

#### Exercise

Show that all the unrelated (i.e. cannot be obtained via symmetry of indices)  $\Gamma$ s to the above all vanish.

*This is a rather tedious one, but it's worth doing for practice.*

We can use these above results to show that the components of the Ricci tensor are

$$\text{Ric}_{tt} = -3\frac{\ddot{a}}{a}, \quad \text{and} \quad \text{Ric}_{\alpha\beta} = (a\ddot{a} + 2\dot{a}^2 + 2\kappa)\gamma_{\alpha\beta}.$$

#### Exercise

Show the above Ricci tensor results.

### 20.2.2 Matter

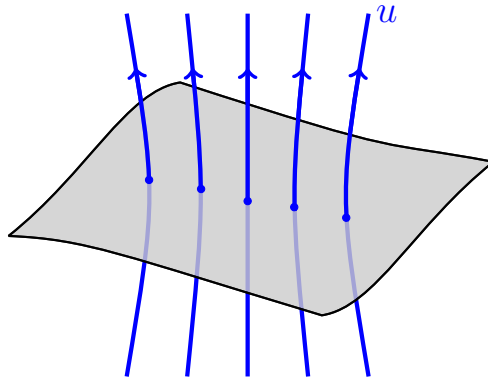
So far we have only used the symmetry conditions to talk about the geometry, and not the actual matter distribution itself. We can now use our symmetry conditions for exactly this, and in doing so obtain the right-hand side of the Einstein equations. That is, we want to find out what kind of matter distributions are allowed such that the symmetry conditions are obeyed.

The trick is to again ‘zoom out’ and only look at the matter at a very large scale. We model the matter in the universe via the following energy-momentum tensor

$$T^{ab} = (\rho + p)u^a u^b + pg^{ab},$$

where  $u^a = (1, 0, 0, 0)^a$  in our coordinates. Such a model is known as a **perfect fluid of density  $\rho$  and pressure  $p$** .

Pictorially this is seen as the idea of the worldlines of large scale structures (e.g. galactic clusters) flowing along some temporal direction, given by  $u$ .



*Remark 20.2.1.* Note that the pressure and density can be functions of  $t$ , but they cannot be functions of  $(r, \theta, \varphi)$ . This is the statement that they can vary through time, but if we want homogeneity and isotropy, they cannot vary through space.

*Terminology.* The vector field  $u$  is often called the **cosmic time**, as it represents how the cosmos flows through time.

### 20.2.3 Reduction of Einstein Equations

Recall that we can write the Einstein equations as

$$\text{Ric}_{ab} = 8\pi G_N \left( T_{ab} - \frac{1}{2} T g_{ab} \right).$$

Inserting our ansatz for  $g_{ab}$  and  $T_{ab}$  we can show

$$\ddot{a} = -\frac{4\pi G_N}{3} (\rho + 3p) a \quad (\text{Acceleration Equation})$$

$$\left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G_N}{3} \rho - \frac{\kappa}{a^2} \quad (\text{Friedmann Equation})$$

**Definition** (Hubble Function). We define the **Hubble function**<sup>6</sup> to be

$$H := \frac{\dot{a}}{a}.$$

#### Exercise

Derive the Acceleration and Friedmann equations, and show that if we include a cosmological constant that the right-hand side of both equations gets a  $+\frac{\Lambda}{3}$ .

*Hint:* For the second part, recall that including the cosmological constant into the Einstein-Hilbert action we get a contribution of  $\Lambda$  to Einstein's equations. That is we have

$$\text{Ric}_{ab} - \frac{1}{2} R g_{ab} + \Lambda g_{ab} = 8\pi G_N T_{ab}$$

Start from here and do the contraction to obtain an expression for  $\text{Ric}_{ab} = \dots$  and then use the above results.

### 20.3 Models of Perfect Fluid Matter

The upshot so far is that for our universe (with the symmetric assumptions) we have two equations for three unknowns, namely  $\rho, p$  and  $a$ . This is obviously a problem.

What do we do? Well if we could somehow obtain another equation relating at least two of these unknowns we would stand a better chance. The two that seem most physical to relate are the density and pressure, and so we want to ask the question "can we obtain a relation between  $\rho$  and  $p$  from more detailed knowledge of what the nature of our perfect fluid is?"

<sup>6</sup>It is often called the Hubble constant, but it need not be a constant and so we call it the Hubble function.

**Definition** (Equation of State). A relation between the momentum and density  $p = \mathcal{P}(\rho)$  is called an **equation of state** for the perfect fluid.

One often looks for a linear relationship, i.e.  $p = \omega \cdot \rho$  for some constant  $\omega \in \mathbb{R}$ .

So what could the fluid be? For now we shall just consider a Universe with only one type of matter in it (next lecture we shall consider multiple kinds). The four main types are:

- (i) A fluid made up of photons.<sup>7</sup> This must obviously satisfy Maxwell theory, which tells us that the energy-momentum tensor must be trace free

$$T^{ab}g_{ab} = 0.$$

We therefore have the condition that<sup>8</sup>

$$p = \frac{1}{3}\rho,$$

which tells us that  $\omega = 1/3$  for the *radiation* fluid. This also turns out to be a good approximation for ultra-relativistic massive particles.

- (ii) Another type of fluid is so-called *dust*. It simply represents a collection of particles which do not interact, and therefore cannot exert a pressure. We conclude, then that  $\omega = 0$  for dust.
- (iii) The case of  $\omega = -1$  corresponds to the equation of motion for the *cosmological constant*. It corresponds to fluid that has everywhere negative pressure.
- (iv) The case for  $\omega = -1/3$  captures the spatial curvature in an equation of state.

*Remark 20.3.1.* For (iii) and (iv) above, what we mean is that we can mimic the behaviour of these quantities by introducing matter into the universe with the respective values of  $\omega$ .

## 20.4 Solutions

Given the acceleration equation, the Friedmann equation and an equation of state for a given matter type, we can solve the system. In the tutorials we will show that the following hold. For  $\kappa = 0 = \Lambda$ :

- (i)  $H^2 \sim \rho \sim a^{-n(\omega)}$ , where  $n(\omega) = 3(1 + \omega)$
- (ii) concretely,

$$a(t) = a_0 \cdot \begin{cases} t^{2/n(\omega)} & \text{if } \omega \neq -1, \\ e^{Ht} & \text{if } \omega = -1. \end{cases}$$

*Remark 20.4.1.* Note the  $\omega = -1$  case for  $a(t)$  tells us that  $H$  must be constant here, as  $H = \dot{a}/a = \dot{H}t + H$ , and so  $\dot{H} = 0$ .

<sup>7</sup>We use the word ‘photon’ here in a rather loose sense. We are discussing classical physics and so photons should not be spoken of.

<sup>8</sup>If you have done the previous exercise, this result should be easy to see.

The  $\omega \neq -1$  condition gives us a very important result:

$$\rho(t) \sim t^{-2}$$

for *all* matter types with  $\omega \neq -1$ . This is an important result because we see as  $t \rightarrow 0$ ,  $\rho$  diverges. That is the density tends to infinity at the start of cosmic time.

Now this might just seem like an artefact of our coordinate choice, but we know that the density appears in Einstein's equations and so if  $\rho$  diverges, the Ricci curvature must also diverge. But the Ricci curvature is a tensor and so if it diverges in one chart it must diverge in all charts, and so we get an infinite curvature of our spacetime at this point. Clearly we can't have this physically and so we must remove this point from our spacetime. Another way to see this last point is that at  $t = 0$ ,  $a = 0$  and so the spacetime metric becomes non-invertible. We clearly cannot have this and so we must exclude this point. So in other words there is no meaning to the question "what happened before  $t = 0$ ?"

Putting all this together we see that what we have just described is the big bang! This is another reason why  $t$  is called the cosmic time, it tells us the age of the cosmos. This result clearly depends explicitly on all of the assumptions we have made so far, namely the perfect symmetry of our universe and the fact that our equations of state are linear. One would be very justified in asking "does this behaviour disappear if we do not make such conditions?" Indeed this is what Hawking, Penrose and others sought to study.

# 21 | Cosmology: The Late Epoch

We now want to consider various matter types simultaneously. We will continue to assume that our equations of state are linear, though. We do this simply because these lectures are meant as an introduction to the field of cosmology and so we need to specialise somewhat. Of course linear equations of state by no means cover every possible situation, and that is part of what the research in cosmology is about; looking for what happens if we change our conditions.

The information we have obtained so far can be summarised in the following table.

$\omega$	$n(\omega)$	$a(t)$	matter type
1/3	4	$t^{1/2}$	radiation
0	3	$t^{2/3}$	dust
-1	0	$e^{Ht}$	cosmological constant
-1/3	2	$t$	spatial curvature

where  $H = \dot{a}/a$  is the Hubble function. We also had that  $\rho \sim t^{-2}$  for all matter types with  $\omega \neq -1$  and  $H^2 \sim \rho$ , which tells us that  $H^{-1}$  is the age of the universe.

## 21.1 Density Parameters

**Definition** (Density Parameter). Let  $\rho_i$  be the density of the  $i^{\text{th}}$  matter type, where  $i = 1, \dots, N$  with  $N$  being the number of matter types we're considering. Then we define, for any non cosmological constant or spatial curvature matter type, the **density parameter**

$$\Omega_i := \frac{8\pi G_N}{3} \frac{\rho_i}{H^2}.$$

For the cosmological constant type matter we define

$$\Omega_\Lambda := \frac{\Lambda}{3H^2},$$

and for the spatial curvature we define

$$\Omega_\kappa := -\frac{\kappa}{H^2 a}.$$

*Remark 21.1.1.* Dr. Schuller likes to refer the the  $\Omega_i$ s are density parameters whereas call the  $\Omega_\kappa$  a 'fake' density parameter. The reason for this is that this type of density has not risen

from some matter contribution to the Einstein equation, but more comes about by saying ‘what matter type would we need to simulate the effects of  $\kappa$ ?’ Similarly you could call  $\Lambda$  a pseudo-fake density parameter, as it does enter the action and we choose to view it as a matter type rather than a curvature. With this idea in mind we shall define  $N$  to include the cosmological constant type matter but not the  $\kappa$  type.

Using the Hubble function and the density parameters, the Friedmann and acceleration equations give us

$$\Omega_\kappa + \sum_{i=1}^N \Omega_i = 1, \quad \text{and} \quad H^{-2} \frac{\ddot{a}}{a} = -\frac{1}{2} \sum_{i=1}^N (1 + 3\omega_i) \Omega_i,$$

respectively.

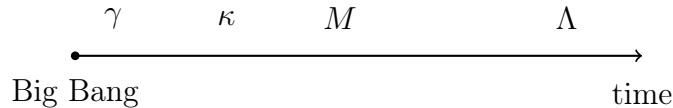
## 21.2 Dominant Matter At Various Epochs

*Terminology.* We shall use  $\gamma$  to denote radiation matter and  $M$  to denote dust matter, e.g.  $n_\gamma = 4$  and  $n_M = 3$ .

Using the above terminology along with the table at the start of this lecture and the result  $\rho_i \sim a^{-n(\omega_i)}$ , which also holds for the  $\kappa$  matter, we conclude that

$$\Omega_\Lambda \sim a^2 \Omega_\kappa \sim a^3 \Omega_M \sim a^4 \Omega_\gamma.$$

This is an important observation and allows us to read off which matter types dominated at which epochs of the universe. An expanding universe is one with  $H > 0$ , corresponding to  $a(t_2) > a(t_1)$  for  $t_2 > t_1$ . We see, therefore, that at later times the matter types to the left become more and more dominating and conversely at early times the ones of the right are more dominant.



Note this result comes from the theory, it is not something we have proposed as a model. That is, given our assumptions, the theory tells us what matter types dominate at what epochs.

## 21.3 A More Realistic Late Universe

We now want to start accounting for having multiple matter types in the universe at the same time. Let's start by considering the example where we have  $\Omega_M$ ,  $\Omega_\kappa$  and  $\Omega_\Lambda$ . We can use the Friedmann equation to express  $\Omega_\kappa = 1 - \Omega_M - \Omega_\Lambda$ , and so the parameter space of our problem is two-dimensional, i.e.  $(\Omega_M, \Omega_\Lambda)$ .

We want to plot this parameter space, but it is worth deriving a few results in order to classify the different regions of the plot.

- (i) We have seen that  $\kappa$  can be positive, negative or zero, so let's try classify these regions. Recall that  $\Omega_\kappa \sim \kappa$  and so if  $\kappa = 0$ ,  $\Omega_\kappa = 0$ . The Friedmann equation then tells us that this corresponds to

$$\Omega_\Lambda = 1 - \Omega_M.$$

By the same method we get  $\Omega_\Lambda > 1 - \Omega_M$  for  $\kappa > 0$  and similarly for  $\kappa < 0$ .

- (ii) Now let's consider the acceleration equation:

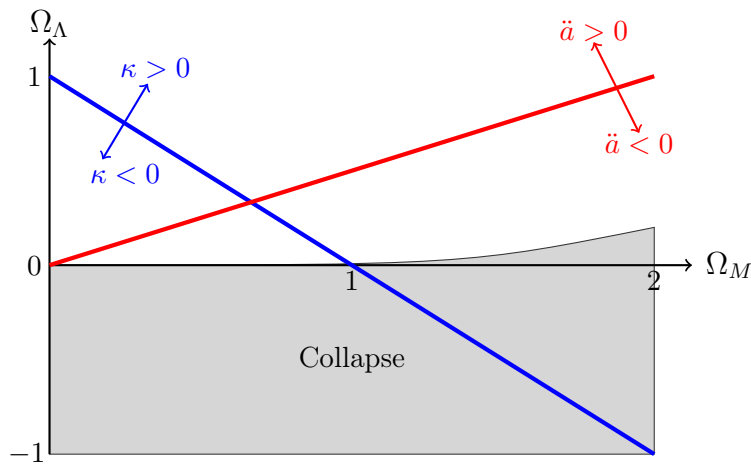
$$H^{-2} \frac{\ddot{a}}{a} = -\frac{1}{2} \sum_{i=1}^N (1 + 3\omega_i) \Omega_i.$$

We know that  $H$  and  $a$  are both positive, and so the sign of the left-hand just depends on the sign of  $\ddot{a}$ . We think of this physically as the acceleration of the expansion of the universe, e.g.  $\ddot{a} > 0$  corresponds to an accelerated expansion. Using  $\omega_M = 0$  and  $\omega_\Lambda = -1$  we get

$$\Omega_\Lambda = \frac{1}{2} \Omega_M$$

for  $\ddot{a} = 0$ . We get analogous results for  $\ddot{a} > 0$  and  $\ddot{a} < 0$ .

- (iii) Now let's consider collapse vs. eternal expansion. That is, we want to ask the question as to whether there is a maximum turning in  $a(t)$ . We formulate this mathematically as looking for a  $t^* \in \mathbb{R}_0^+$  such that  $\dot{a} = 0$  and  $\ddot{a} < 0$ . You can analytically calculate the expression for the turning point (in the sense of the line that separates collapse from eternal expansion), however it's rather complicated. We shall just plot its form on the graph below.



Experimental observation tells us that  $\Omega_\Lambda = 0.7$  and  $\Omega_M = 0.3$ , from which we conclude (up to the uncertainty of the experiment) that  $\kappa = 0$ , i.e. the universe is a flat geometry. It also turns out that the 30% of curvature generated by matter is further split into standard model matter, which is only 5%, and so-called *dark matter*, which is the remaining 25%. So we see, again assuming everything we have done is true and valid, that the standard model of physics only makes up 5% of all the matter *needed* in the universe to explain our observations. This is one of the main driving forces behind research in cosmology: what is this other stuff?!

## 22 | Black Holes

We want to study the Schwarzschild solution to Einstein's equations. It is a vacuum solution with the metric in the Schwarzschild chart, whose coordinates are  $(t, r, \theta, \varphi)$ , given by

$$g = \left(1 - \frac{2m}{r}\right) dt \otimes dt - \frac{1}{1 - \frac{2m}{r}} dr \otimes dr - r^2 (d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi),$$

where  $m = G_N M$  with  $M$  being the mass of the object (in this case the black hole).

*Remark 22.0.1.* Note that Dr. Schuller has gone back to using the  $(+, -, -, -)$  signature here. I will follow this convention in these notes.

*Notation.* The final two terms in the above expression are often grouped into one and we define

$$d\Omega \otimes d\Omega = r^2 (d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi).$$

This notation is very popular in textbooks etc. as it lightens notation, and as we will see, it is mainly the  $dt \otimes dt$  and  $dr \otimes dr$  terms we are concerned with.

The above expression is obviously only for the Schwarzschild coordinates, but the metric itself can of course be expressed in any chart. We may think that the ranges of the Schwarzschild coordinated coordinates are  $t \in (-\infty, \infty)$ ,  $r \in (0, \infty)$ ,  $\theta \in (0, \pi)$  and  $\varphi \in (0, 2\pi)$ . However, after paying closer attention to the metric above we note an immediate problem: what happens at  $r = 2m$ ? The  $dt \otimes dt$  term goes to zero, which is bad enough, but on top of that the  $dr \otimes dr$  time diverges! We must, therefore, remove this point from our domain, i.e.  $r \in (0, 2m) \dot{\cup} (2m, \infty)$ , where the dot denotes the fact that the union is disjoint. We then have to ask the question about what actually happens at  $r = 2m$ ?

The next question we should ask is "is there anything in the real world beyond the points  $t \rightarrow \pm\infty$ ?" This question sounds silly, as what is beyond  $\pm\infty$ , but we need to remind ourselves that  $t : \mathcal{M} \rightarrow \mathbb{R}^4$  is a chart map and we need not cover all of  $\mathcal{M}$  with it. That is, we can parameterise  $t$  such that a finite volume of  $\mathcal{M}$  is mapped to an infinite volume in  $\mathbb{R}^4$ . We can ask a similar question about  $r \rightarrow \infty$ .

The insight to these questions comes from taking a step back and not looking at the above expression itself but looking at objective objects instead, namely geodesics.



## 22.1 Radial Null Geodesics

Consider null<sup>1</sup> geodesics in Schwarzschild spacetime. The action is

$$S[\gamma] = \int d\lambda \left[ \left(1 - \frac{2m}{r}\right) \dot{t}^2 - \left(1 - \frac{2m}{r}\right)^{-1} \dot{r}^2 - r^2(\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2) \right].$$

Let's first find the  $t$  equation of motion, i.e. vary w.r.t.  $\delta t$ . We have

$$\frac{d}{d\lambda} \left[ \left(1 - \frac{2m}{r}\right) \dot{t} \right] = 0 \quad \iff \quad \left(1 - \frac{2m}{r}\right) \dot{t} = k,$$

for some constant  $k$ .

**Definition** (Radial Geodesics). We define a **radial geodesic** to be one that 'follows  $r$ '. In other words, we set  $\theta = \theta_0$  and  $\varphi = \varphi_0$ , for some constants  $\theta_0$  and  $\varphi_0$ .

### Exercise

Use the temporal equation of motion, the null condition and the radial condition to show that we can use  $r$  as an affine parameter.

*Hint: Show that  $r = \pm k\lambda$  and then argue why we can consider  $r$  to be a affine parameter. (If you need help Dr. Schuller explains this argument in the video).*

We express the result of the above exercise as  $\tilde{t}(r) = t(\pm k\lambda)$ . Let's consider each case:

(i) First consider  $\tilde{t}_+(r) = t(k\lambda)$ . The chain rule gives us

$$\frac{d\tilde{t}_+}{dr} = \frac{d\tilde{t}}{d\lambda} \frac{d\lambda}{dr} = \frac{\dot{t}}{\dot{r}} = \frac{k}{\left(1 - \frac{2m}{r}\right)k} = \frac{r}{r - 2m}.$$

Integrating this<sup>2</sup>

$$\tilde{t}_+(r) = r + 2m \ln |r - 2m| + \text{constant}.$$

These are the **outgoing** null geodesics.

(ii) Now consider  $\tilde{t}_-(r) = t(-k\lambda)$ . A similar method to the above gives us

$$\tilde{t}_-(r) = -r - 2m \ln |r - 2m| + \text{constant}.$$

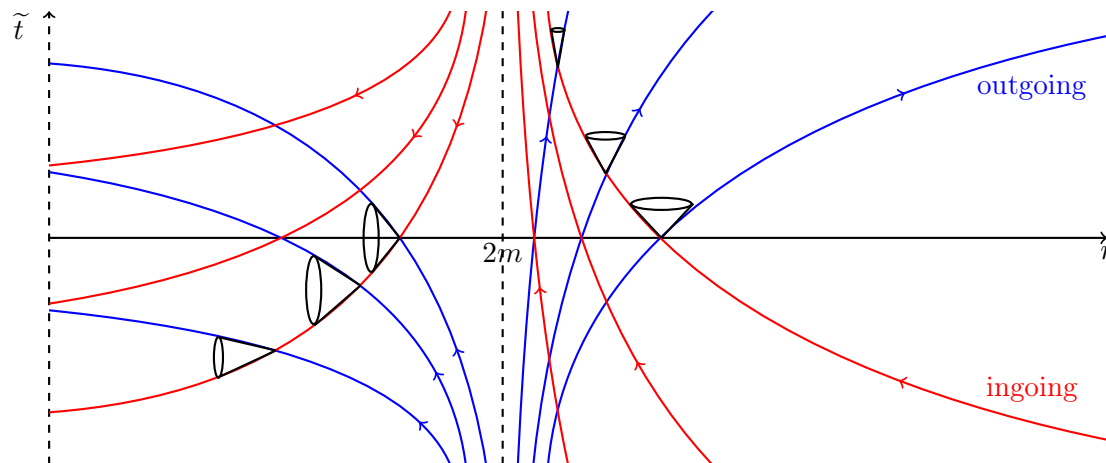
These are the **ingoing** null geodesics.

To see what's happening let's plot the outgoing and ingoing geodesics.<sup>3</sup>

<sup>1</sup>Recall this just means  $g(v_\gamma, v_\gamma) = 0$ , which correspond to the worldlines of massless particles.

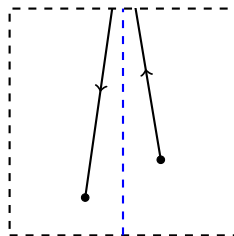
<sup>2</sup>This integral is not actually the easiest to do, but differentiating the result shows you that its true.

<sup>3</sup>This diagram was a pain to draw, so to any readers: I hope you like it!



This diagram can actually be very misleading. Firstly the light cones appear to be closing, which is very strange and then all of a sudden emerge tilted on their side in the region  $r < 2m$ . This problem stems from the fact that we are drawing deciding what the light cones look like based on the charted geodesics, and is actually not a problem at all.<sup>4</sup>

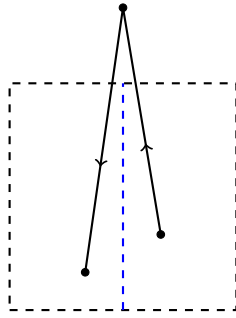
The next problem is it appears that a geodesic that starts in the region  $r > 2m$  cannot get to the region  $r < 2m$ . This seems completely crazy, because haven't we heard before that a Schwarzschild black hole is a dense object who has an event horizon at  $r = 2m$ . So shouldn't a geodesic go through this line? The answer is obviously "yes" and the problem is again an artefact of the coordinate choice, specifically the range of  $t$ . To see why, imagine mapping London<sup>5</sup> into a chart of infinite volume. Consider someone who lives in London but works outside London: in the morning they set off from home to work, and then after work they decide to come back in to London for a meal with a friend. If we plotted their path in our chart it would look like the following:



It is clearly not true that this person didn't cross the blue line, it is simply just not included in the chart. In other words, we expect that the path is really like the following:

<sup>4</sup>I have written my own ideas for what I believe is really happening here and uploaded it to my blog site. Those notes contain many errors I need to go back and fix, but the interested reader will still hopefully find it an enjoyable read.

<sup>5</sup>Dr. Schuller uses Linz, but I am British so I'll use London.



Could the same thing be true for our Schwarzschild picture? The answer is "yes", and we shall explain how in the next section.

## 22.2 Eddington-Finkelstein

The idea is to change the coordinates such that, in our new coordinates the *ingoing* null geodesics appear as straight lines of slope  $-1$ . This is achieved by the coordinate

$$\bar{t}_{\pm}(t, r, \theta, \varphi) := \tilde{t}_{\pm} + 2m \ln |r - 2m|.$$

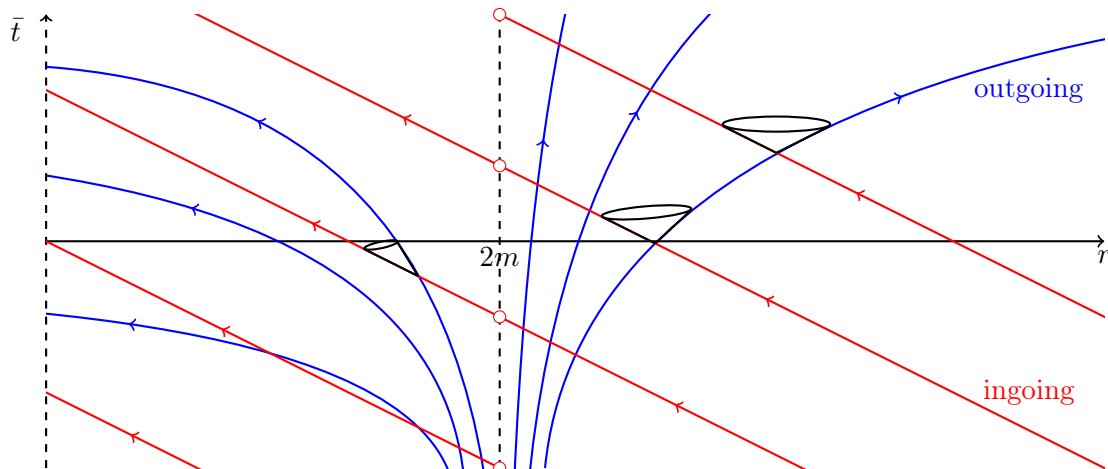
Rearranging this for  $\tilde{t}_-$  and plugging it into the expression for  $\tilde{t}_-(r)$  you get

$$\bar{t}_- = -r + \text{constant},$$

which is exactly what we wanted. We also get

$$\bar{t}_+ = r + 4m \ln |r - 2m| + \text{constant},$$

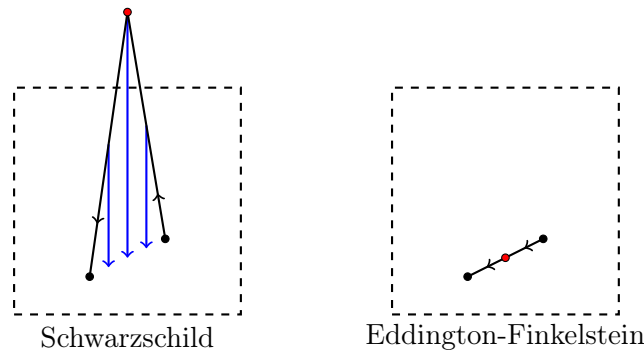
which will have the same kind of shape as before (but slightly scaled). The graph therefore becomes:



Note the points along  $r = 2m$  are still not a part of our chart  $(U, x)$  and so must be excluded. However at this point it becomes clear that there is nothing wrong with this points and we can simply define a new chart domain  $V := U \cup \{r = 2m\}$ . This then gives the same diagram as above but without the white circles.

We won't draw it again just for the sake of removing the circles, however it is worth noting that on this new chart we could plot the line cones at  $r = 2m$ . All the the cones along this line would have their right side vertical. This is the condition that the  $r = 2m$  is the horizon and it corresponds to the point of no return. That is, recalling that the world line of a massive observer must have its tangent vectors within the light cone, at  $r = 2m$  an observer can no longer move away from the black hole and is destined to meet the singularity.

*Remark 22.2.1.* We can think of the Eddington-Finkelstein coordinate transformation as one that 'pulled' points downwards. Using our London commuter as an example, we see this as the following diagram: (the blue arrows represent what that transformation does)



*Remark 22.2.2.* Note to self: Maybe include a comment about maximal extension here. Just roughly what it means etc.

Now let's calculate the Schwarzschild metric  $g$  w.r.t. Eddington-Finkelstein coordinates. Our coordinate transformation is given by

$$\begin{aligned}\bar{t}(r, t, \theta, \varphi) &= t + 2m \ln |r - 2m| \\ \bar{r}(r, t, \theta, \varphi) &= r \\ \bar{\theta}(r, t, \theta, \varphi) &= \theta \\ \bar{\varphi}(r, t, \theta, \varphi) &= \varphi.\end{aligned}$$

If we denote the Schwarzschild coordinates by  $(x^0, x^1, x^2, x^3)$  and the Eddington-Finkelstein coordinates by  $(y^0, y^1, y^2, y^3)$  then our problem is to find

$$g_{(y)ab} = \frac{\partial x^m}{\partial y^a} \frac{\partial x^n}{\partial y^b} g_{(x)mn}.$$

It looks like we need to invert the transformations above to obtain  $x^i(y)$ , however we have shown in the tutorial 5 that

$$\delta_n^m = \left( \frac{\partial x^m}{\partial y^a} \right) \left( \frac{\partial y^a}{\partial x^n} \right) \quad \implies \quad \left( \frac{\partial y^a}{\partial x^m} \right)^{-1} = \left( \frac{\partial x^m}{\partial y^a} \right),$$

and so we can use our transformation equations above and then invert the matrix of results.<sup>6</sup>

<sup>6</sup>Note this is not the same as just doing the reciprocal of the fraction, as if the matrix is not diagonal the inverse elements are not just the reciprocals.

We have

$$\begin{pmatrix} \frac{\partial y^a}{\partial x^m} \end{pmatrix} = \begin{pmatrix} 1 & \frac{2m}{r-2m} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \implies \begin{pmatrix} \frac{\partial x^m}{\partial y^a} \end{pmatrix} = \begin{pmatrix} 1 & \frac{-2m}{r-2m} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Using the result, and dropping the bars on  $r, \theta$  and  $\varphi$ , we get

$$g = \left(1 - \frac{2m}{r}\right) d\bar{t} \otimes d\bar{t} - \frac{2m}{r} (d\bar{t} \otimes dr + dr \otimes d\bar{t}) - \left(1 + \frac{2m}{r}\right) dr \otimes dr - d\Omega \otimes d\Omega.$$

### Exercise

Prove the above result.

*Hint: You can either do it using the transformation above (which is done in the video) or you can use the definition of the exterior derivative  $d$ . The two methods are, of course, equivalent.*

## 22.3 Kruskal-Szekeres

As we have seen in both the Schwarzschild and Eddington-Finkelstein coordinates, the light cones either squash up or rotate and, although these can give some nice insights, we can actually make further coordinate transformations such that the outgoing geodesics also become straight lines of slope  $+1$ . In doing so, our light cones will all sit vertically and will always make a 90 degree angle. Of course this comes at the expense of the coordinates themselves look a bit funny, but that is the trade off. Such coordinates are known as Kruskal-Szekeres coordinates. The coordinate transformatio is given by: for  $r > 2m$

$$\begin{aligned} \bar{\bar{t}}(t, r, \theta, \varphi) &:= \left(\frac{r}{2m} - 1\right)^{1/2} e^{r/4m} \sinh\left(\frac{t}{4m}\right) \\ \bar{\bar{r}}(t, r, \theta, \varphi) &:= \left(\frac{r}{2m} - 1\right)^{1/2} e^{r/4m} \cosh\left(\frac{t}{4m}\right) \end{aligned}$$

and for  $r < 2m$

$$\begin{aligned} \bar{\bar{t}}(t, r, \theta, \varphi) &:= \left(1 - \frac{r}{2m}\right)^{1/2} e^{r/4m} \cosh\left(\frac{t}{4m}\right) \\ \bar{\bar{r}}(t, r, \theta, \varphi) &:= \left(1 - \frac{r}{2m}\right)^{1/2} e^{r/4m} \sinh\left(\frac{t}{4m}\right) \end{aligned}$$

and  $\theta$  and  $\varphi$  are unchanged.

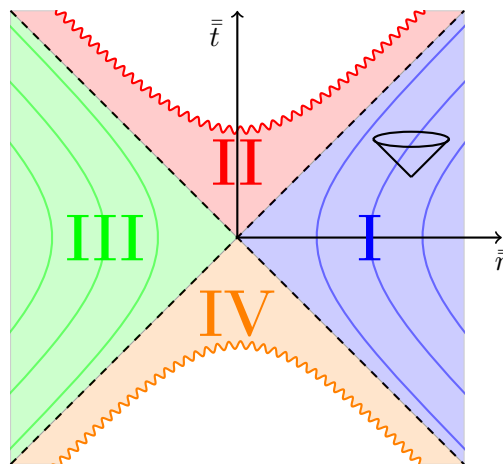
## Exercise

Show that the Kruskal-Szekeres coordinates tell us

$$\bar{t}^2 - \bar{r}^2 = \begin{cases} -k^2 & r > 2m \\ \ell^2 & r < 2m \end{cases}$$

for  $k, \ell \in \mathbb{R}$ .

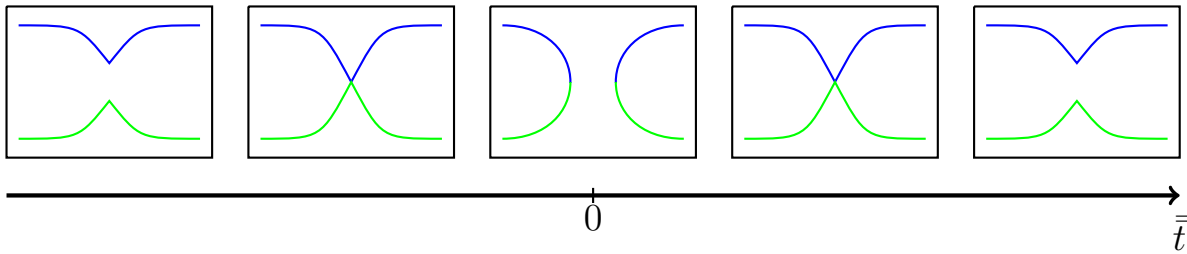
From the exercise above we see that the plot consists of sets of hyperbolas. What is truly surprising about these solutions is that admit new regions to our spacetime, as the following diagram shows.



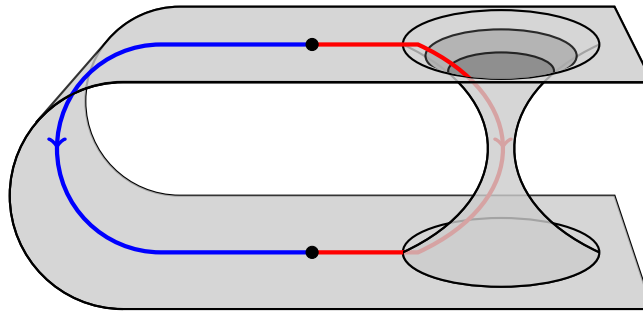
We get four regions: region I is our universe, and lines of constant  $r$  are the hyperbolas drawn; region II is the black hole with the snake-line being the singularity, and it is only the shaded region that is part of the spacetime; region III is completely new and represents another, causally disconnected, universe where again lines of constant  $r$  are the hyperbolas drawn; region IV is what we call a *white hole*, as all casual geodesics (i.e. massive and massless particles) must leave it and enter either region I or region III. Inside the black/white hole the relevant hyperbolas represent lines of constant  $t$ .

Light cones stand 'upright' everywhere on the diagram, which allows us to note that the dashed lines represent the event horizons of the black hole and white hole; any geodesic that passes the dashed line between regions I and II is doomed to meet the singularity.

After getting over the immediate shock of another universe and a white hole, a vital question raises itself: "what on earth happens at the origin (i.e. where the dashed lines cross)?" The answer to this question is quite complicated but is basically that there is a change in topology. If we consider spatial slices moving up the diagram, we get something like the following picture: suppressing the  $\theta$  and  $\varphi$  directions, we get



The structure formed at  $\bar{t} = 0$  is a so-called *wormhole* and it corresponds to a ‘portal’ between regions I and III. The points where the blue lines become green is known as the *throat*<sup>7</sup> of the wormhole. The wormhole corresponds to a spatial slice and so it is not actually something an observer could travel through, but it is an incredibly interesting idea, and it led Einstein and Rosen to try and propose such a ‘bridge’ between spacetime points on a causally connected manifold. The result of this is a so-called *Einstein-Rosen bridge*.



## 22.4 Other Types of Black Hole

**Theorem 22.4.1 (No Hair).** *All black hole solutions to Einstein’s equations and Maxwell’s equations can be completely characterised by their mass, angular momentum and electric charge.*

From the above theorem it is clear that we can have four different types of black hole, summarised in the table below

Name	Mass	Angular Momentum	Electric Charge
Schwarzschild	✓	✗	✗
Kerr	✓	✓	✗
Reissner–Nordström	✓	✗	✓
Kerr–Newman	✓	✓	✓

In this lecture we have only discussed the Schwarzschild black hole, but have not mentioned any of the other three at all. We shall not discuss the other black holes in great detail in these notes, but in order to highlight some quite surprising results we shall make some brief comments on the Kerr black hole.

<sup>7</sup>The name comes from the idea that if we reinsert  $\theta$  that we would get a tube like structure here and it would look like a throat connecting two spaces. See the diagram of the Einstein-Rosen bridge.

We have seen above that the Schwarzschild black hole give rise to a point on the space-time that must be removed, i.e. the singularity. Physically speaking, we imagine a massive spherically symmetric star collapsing down into a *single point* at the centre. This singularity point must, therefore, contain the information about the black hole.<sup>8</sup>

A Kerr black hole, however, is an electrically neutral, rotating black hole. Putting this together with the fact that general relativity is a classical theory, it is clear that we can not have a single point for our singularity. That is, the singularity must contain information about the angular momentum of the black hole, but classically a single point cannot have angular momentum. The next best option is to consider an infinitely thin ring of non-vanishing radius. This is indeed what you get for a Kerr black hole, the result being known as either a **ring singularity** or the composite word: a **ringularity**.

*Claim 22.4.2.* An observer can avoid a ring singularity and pass through the disc bound by it an emerge in what some people call an *antiverse*.

We do not prove the above claim.

### 22.4.1 Event Horizons & Infinite Redshift Surfaces

As well as exhibiting a ring singularity, Kerr black holes also possess another new idea. In order to understand it a bit better, let's take another look at what happens to the Schwarzschild metric (in Schwarzschild coordinates) at  $r = 2m$ . Specifically, let's consider the  $g_{tt}$  and  $g_{rr}$  components:

$$g_{tt} = g(\partial_t, \partial_t) = \left(1 - \frac{2m}{r}\right), \quad \text{and} \quad g_{rr} = g(\partial_r, \partial_r) = \left(1 - \frac{2m}{r}\right)^{-1}.$$

We see straight away that for  $r > 2m$ ,  $g_{tt} > 0$  and  $g_{rr} < 0$ , whereas for  $r < 2m$ ,  $g_{tt} < 0$  and  $g_{rr} > 0$ . So the point  $r = 2m$  seems to correspond to a sign change in these components. We can word this as the statement: "as we move from  $r > 2m$  to  $r < 2m$  the timelike vector field  $\partial_t$  becomes spacelike, whereas the spacelike vector field  $\partial_r$  becomes timelike."

Now these two conditions in themselves need not be related, that is the fact that  $\partial_t$  becomes spacelike is not fundamentally related to  $\partial_r$  becoming timelike. It is true, however, the *something* must become timelike, otherwise the signature of our metric would change — i.e. we would end up with  $(+, +, +, +)$ , in our convention — but  $\partial_r$  is not the only choice, we also have  $\partial_\theta$  and  $\partial_\varphi$ . Before moving on to discuss when one of these latter two might become timelike, let's first try and work out what our two conditions mean physically.

First let's consider  $\partial_r$  becoming timelike. Recall that the interior of the future light cone gave the possible future of a massive observer. We use the word 'future' as clock carried by this observer will increase in time as you follow timelike geodesics. This was given as a definition and so will always hold, regardless of which coordinate vector field is timelike and which are spacelike. However, this latter distinction does have an effect our interactions with external observers. Roughly speaking, the projection of our velocity is non-vanishing only for timelike vector fields, and so our future cannot be orthogonal to these directions. The more 'central' to the cone the timelike vector field, the more our future is determined by it.

<sup>8</sup>Disclaimer: I'm not sure how correct of a statement this is to make. I am not overly fond of attributing information to an absence of a point on the spacetime, but this argument makes explaining the next bit much easier, so we'll assume its ok.



This might sound a bit funny, but it is easily understood by considering time dilation in special relativity. The spacetime of special relativity is flat Minkowski space and all of the cones stand upright and make 90 degrees. Let's now consider the chart with chart maps  $(t, x, y, z)$ . A stationary observer  $(\gamma, e)$  in this frame will follow a geodesic whose tangent vectors are integral curves of the vector field  $\partial_t$ . We can choose to parameterise this curve such that  $g(e^0, \partial_t) = 1$ , that is the clock carried by this observer agrees with the coordinate time  $t$ .

Now consider another observer  $(\delta, f)$  moving relative to the first. They will follow a geodesic that is *not* an integral curve of the  $\partial_t$  vector field. This observer's frame will obey  $0 < g(f^0, \partial_t) < 1$ , where the value in the range depends on the velocity of  $\delta$ . This is what we mean above about having a non-vanishing projection onto our timelike vector field.

Now to the stationary observer, both of them age, as they both have a non-vanishing projection onto  $e^0$  direction,<sup>9</sup> however the second observer seems to age slower, as  $g(f^0, \partial_t) < g(e^0, \partial_t)$ .

With the aside on special relativity in mind, we can see that the condition that  $\partial_r$  become a timelike vector tells us that, from the perspective of an observer stationary w.r.t. the Schwarzschild coordinates, the observer *must* move along with some projection along the  $\partial_r$  axis. But what is stationary w.r.t. the Schwarzschild coordinates? Well the black hole of course! So we see that an observer at  $r < 2m$  must move along the radial direction. The question is "which way?" Again this might sound silly, but it makes a lot more sense when we remember that in the special relativity case, we always move up the  $t$  axis and not backwards into the past.

We shall not discuss this too deeply here<sup>10</sup> (as we are already being rather hand-wavy), but the general idea is that when  $\partial_r$  becomes timelike it actually points into the past light cone, and so our future is determined by moving *towards* the black hole. So we have discovered that  $\partial_r$  becoming timelike corresponds to an event horizon!

What about  $\partial_t$  becoming spacelike, what is that all about? In order to save space, we shall not discuss exactly where this comes from, but it turns out this corresponds to a so-called **infinite redshift surface**. The name comes from the idea that, *to an observer at  $r \rightarrow \infty$* , light emitted at the points where  $\partial_t$  becomes timelike is redshifted so much that it actually 'disappears'. This is a very strange statement to make as firstly it's only true for an observer infinitely far (indeed an observer at finite distance will be able to detect the light, albeit it highly redshifted) away and secondly the idea of completely redshifting something away is strange. This is the reason we are not presenting the mathematical origin of this phenomenon here.

There is, however, a much nicer (in my opinion) physical interpretation to what happens at an infinite redshift surface. As we have explained, for  $\partial_t$  to become spacelike, one of the spacelike vector fields must become timelike. We argued that this in turn causes relative motion between the observer and the black hole along (or against) the direction of this vector field. This is a much nicer idea, and so it is what we shall use.

*Remark 22.4.3.* Note it follows from our arguments above that in order to reach an event

<sup>9</sup>We should be careful here because  $e^0$  is only defined along  $\gamma$ , but it should be relatively clear what we mean.

<sup>10</sup>I discuss this in more detail in the notes I made about light cones and event horizons, which are available on my blog site.

horizon an observer must pass through an infinite redshift surface. That is, if  $\partial_r$  is to become timelike,  $\partial_t$  has to become spacelike before it or at the same time.

*Remark 22.4.4.* It is important not to confuse the coordinate  $t$  with the time measured on an observers clock. Inside an infinite redshift surface  $\partial_t$  is spacelike, and so we can move ‘down’ it. If we took this to be time then this would be the statement that we can travel backwards in time. If this was true time reversal, travelling along this direction would take us backwards and out of the infinite redshift surface. This is not what happens at all, and is seen easily by the fact that we measure time via the clock we carry. What the above does say, though, is that, to an observer stationary w.r.t. the black hole, we can travel backwards in time as we can move along a geodesic that has positive projection along  $-\partial_t$ .

### 22.4.2 Ergoregion

The metric for the Kerr black hole is not particularly insightful to see itself, and so we do not present it here but instead just summarise the results.

It turns out that, unlike for the Schwarzschild black hole, the event horizon and infinite redshift surface for the Kerr black hole do not coincide. That is  $r_{IRS} > r_{EH}$ , where  $r_{IRS}$  ( $r_{EH}$ , respectively) is the  $r$  value of the infinite redshift surface (event horizon, respectively).

**Definition** (Ergoregion). The region  $r_{EH} < r < r_{IRS}$  is called the **ergoregion** (or **ergosphere**).

The obvious question to ask is "what is the timelike vector field in the ergoregion?" The answer to that question is the direction of rotation,  $\partial_\varphi$ . It turns out the  $\partial_\varphi$  lies in the future cone and so an observer in the ergoregion *must* rotate *with* the black hole.

*Remark 22.4.5.* There is an interesting idea to extract energy from the ergoregion of a Kerr black hole known as the **Penrose process**. We shall not discuss it here, but readers are encouraged to search it as it is quite interesting.

### 22.4.3 Multiple Event Horizons

It also turns out to be true that the non-Schwarzschild black holes all have two event horizons and infinite redshift surfaces. The presence of the inner event horizon means that one can traverse a black hole without meeting the singularity and can emerge into another universe! We do not discuss this in more detail here but highlight it in the Penrose diagrams next lecture.

## 23 | Penrose Diagrams

Would it not be nice to be able to draw an informative picture of an *entire* spacetime on a finite portion of paper? For some spacetimes, this is possible and the resulting diagrams are known as **Penrose** (or Penrose-Carter) diagrams. In order for the diagrams to be useful, we will compromise on a number of issues, but we will *not* compromise on the nice property of null geodesics having slope  $\pm 1$ , i.e. the light cones stand upright and make a 90 degree angle everywhere.

### 23.1 Recipe To Construct A Diagram

The ‘recipe’ is as follows:

- (i) Start with a spacetime metric in some chart, and *painfully note the coordinate ranges*.
- (ii) Find coordinates such that two previously non-compact coordinates are replaced by two (possibly still non-compact) *null coordinates*, which we label  $v$  and  $w$ .
- (iii) Compactify the two null coordinates separately, i.e. introduce new coordinates<sup>1</sup>

$$p := \arctan(v), \quad \text{and} \quad q := \arctan(w).$$

Thus  $(p, q)$  will take values in (some subset of)  $(-\pi/2, \pi/2) \times (-\pi/2, \pi/2)$ .

- (iv) Define again a temporal and spatial coordinate

$$T := p + q, \quad \text{and} \quad X := p - q,$$

again keeping track of the ranges.

- (v) Express the metric  $g$  in the coordinates  $(T, X, \dots)$ , where  $\dots$  are the original coordinates which we haven’t changed.
- (vi) *If*<sup>2</sup> the metric in these coordinates takes the form

$$g = \Omega^{-2}(T, X, \dots) \left[ dT \otimes dT - dX \otimes dX - R(T, X)(d\dots \otimes d\dots) \right].$$

<sup>1</sup>You don’t need to use arctan, but just any compactifying function.

<sup>2</sup>It may not be possible!

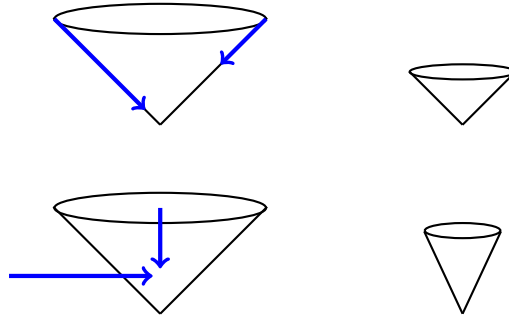
where  $d... \otimes d...$  is meant to indicate the remaining coordinates in diagonal form (e.g.  $d\theta \otimes d\theta + d\varphi \otimes d\varphi$ ), then obtain the non-physical diagram

$$g_{\text{diagram}} = dT \otimes dT - dX \otimes dX,$$

again noting the ranges.

The above result seems rather strange; essentially what we've done is turn everything into what appears to be flat Minkowski space! Well yes and no. Yes because we want the cones to stand upright and at 90 degrees, but no because  $(T, X) \subset \mathbb{R} \times \mathbb{R}$  is not an equality: our diagram is finite in size. We can therefore think about the information of the diagram, and therefore the spacetime, as being contained in where the boundaries are.

*Remark 23.1.1.* Step (ii) in the above is important, as it's what allows us to preserve the cones. This is easily seen pictorially, as compressing along null coordinates doesn't change the angles of the cones, whereas if we had used a temporal coordinate and a spatial coordinate, we would only preserve the angle if we compactified in a 1 : 1 manner everywhere.<sup>3</sup>



*Remark 23.1.2.* In going to step (vi) we seem to have ‘forgotten’ about the  $\Omega^{-2}$  factor and the  $R(T, X)(...)$  terms.

The first is simply a conformal factor,<sup>4</sup> and conformal factors do not change the shape of null geodesics. They will, however, change the shape of others (i.e. timelike and spacelike geodesics). We state this more precisely in the following proposition. Seeing as we are only interested in persevering the null geodesics (i.e. the light cones), we can do this and just accept that the shape of the others will change.

The second point we fix by simply imagining that at each point on our diagram we attach a space whose geometry is given by the  $R(T, X)(...)$  terms. This is why we allowed  $R$  to be a function of  $T$  and  $X$  and also why we denote it  $R$  — we can loosely think of it as being the radius of these geometries we attach.

**Proposition 23.1.3.** *A curve  $\gamma$  is a null geodesic of  $g$  if and only if it is a null geodesic of  $\Omega^2 g$ , where  $\Omega^2 \in C^\infty(\mathcal{M})$  is no-where vanishing.*

*Proof.* Let  ${}^g\nabla$  and  ${}^{\Omega^2 g}\nabla$  denote the connections associated to  $g$  and  $\Omega^2 g$ , respectively. Let  $\gamma : (0, 1) \rightarrow \mathcal{M}$  be a curve and denote the tangent vector field to it by  $X$ . Then:

<sup>3</sup>The below diagram is a bit misleading: the cones are infinitely big so when we draw them smaller to the right we just mean that the part part of the cone has been made smaller; both are still infinite in size.

<sup>4</sup>A lot more information on conformal transformations can be found in my notes on Dr. Shiraz Minwalla's String Theory course.

( $\Rightarrow$ ) Assume  $\gamma$  is a affinely parameterised null geodesic of  $g$ . That is  ${}^g\nabla_X X = 0$ . Now consider the covariant derivative of  $X$  using  $\Omega\nabla$ :

$$\begin{aligned} (\Omega\nabla_X X)^a &= X^b \frac{\partial}{\partial x^b} X^a + \Omega \Gamma^a{}_{cb} X^b X^c \\ &= X^b \frac{\partial}{\partial x^b} X^a + \frac{1}{2\Omega^2} g^{ad} [(\Omega^2 g)_{cd,b} + (\Omega^2 g)_{bd,c} - (\Omega^2 g)_{bc,d}] X^b X^c \end{aligned}$$

Let's just consider the second term,

$$\begin{aligned} \frac{1}{2\Omega^2} g^{ad} [(\Omega^2 g)_{cd,b} + (\Omega^2 g)_{bd,c} - (\Omega^2 g)_{cb,d}] X^b X^c &= \frac{1}{2\Omega^2} g^{ad} [2\Omega^2{}_{,b}g_{cd} - \Omega^2{}_{,d}g_{bc}] X^b X^c \\ &\quad + \frac{1}{2\Omega^2} \Omega^2 g^{ad} [g_{cd,b} + g_{bd,c} - g_{bc,d}] X^b X^c, \end{aligned}$$

where we have used the summation convention to obtain the 2 inside the first square brackets. The second term on the right-hand side goes with the first term of the right-hand side of the first equation to give us  ${}^g\nabla_X X$ , which we assumed vanished, so we are just left with

$$(\Omega\nabla_X X)^a = \frac{1}{2\Omega^2} g^{ad} [2\Omega^2{}_{,b}g_{cd} - \Omega^2{}_{,d}g_{bc}] X^b X^c.$$

Now the second term on the right-hand side contains  $g_{bc} X^b X^c = g(X, X) = 0$ , as our geodesic is null, so we are just left with the first term. We have

$$\begin{aligned} \frac{1}{\Omega^2} g^{ad} g_{cd} \Omega^2{}_{,b} X^b X^c &= \frac{2}{\Omega} \Omega_{,b} X^b X^a \\ &= 2X \langle \ln \Omega \rangle X^a \\ &= A \cdot X^a \end{aligned}$$

where we have used  $g^{ad} g_{cd} = \delta_c^a$  and the fact that  $X \langle \ln \Omega \rangle \in C^\infty(\mathcal{M})$  and denoted it by  $A$ . So we finally have

$$(\Omega\nabla_X X)^a = A \cdot X^a.$$

This is the equation for a geodesic that has not been affinely parameterised, which is why it doesn't vanish.

So we have shown it is a geodesic. We now just need to show it is null. Trivially we have

$$(\Omega^2 g)(X, X) = \Omega^2 \cdot (g(X, X)) = 0.$$

( $\Leftarrow$ ) This is the same calculation as above but made in reverse.

□

*Remark 23.1.4.* Note we had to use the null condition to show that we had a geodesic of  $\Omega^2$  (i.e. to remove the  $g_{bc} X^b X^c$  term). It is for this reason that it is only the null geodesics that are left untouched by our conformal transformation, whereas the shape of our other geodesics change.

## 23.2 Minkowski

The simplest vacuum solution of Einstein's equations is Minkowski space, which in coordinates  $(t, r, \theta, \varphi)$  with  $t \in (-\infty, \infty)$ ,  $r \in (0, \infty)$ ,  $\theta \in (0, \pi)$  and  $\varphi \in (0, 2\pi)$ , has the metric

$$g = dt \otimes dt - dr \otimes dr - r^2(d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi).$$

Our two non-compact coordinates are  $t$  and  $r$  and so it is these we replace by null coordinates. We define<sup>5</sup>

$$v := t + r, \quad \text{and} \quad w = t - r.$$

The range here is  $v, w \in \mathbb{R}$ , but with the condition  $r = \frac{1}{2}(v - w) > 0$ , and so we require  $v > w$ . Now we compactify:

$$p := \arctan(v), \quad \text{and} \quad q := \arctan(w).$$

Our range is now  $p, q \in (-\pi/2, \pi/2)$  with the condition  $p > q$ . Now we construct the new temporal and spatial coordinates

$$T := p + q, \quad \text{and} \quad X := p - q.$$

Using  $p = \frac{1}{2}(T + X)$  and  $q = \frac{1}{2}(T - X)$ , the ranges/condition then become

$$-\pi < T + X < \pi, \quad -\pi < T - X < \pi, \quad \text{and} \quad X > 0.$$

We now need to express our metric in terms of  $(T, X, \theta, \varphi)$ . We need to obtain expressions for  $T$  and  $X$  in terms of these coordinates. Putting the above results together, we have

$$T := \arctan(t + r) + \arctan(t - r), \quad \text{and} \quad X := \arctan(t + r) - \arctan(t - r).$$

The metric in these coordinates is

$$g = \sec^2(T + X) \sec^2(T - X) \left[ dT \otimes dT - dX \otimes dX - R(T, X)(d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi) \right],$$

where

$$R(T, X) := \frac{r^2(T, X)}{\sec^2(T + X) \sec^2(T - X)}.$$

### Exercise

Prove the expression for the metric above is true.

*Hint: use*

$$t + r = \tan(T + X), \quad \text{and} \quad t - r = \tan(T - X).$$

*along with  $d(f(x)) = \partial_i f dx^i$  to find  $dt$  and  $dr$  in terms of  $dT$  and  $dX$ , then multiply out all the terms and cancel.*

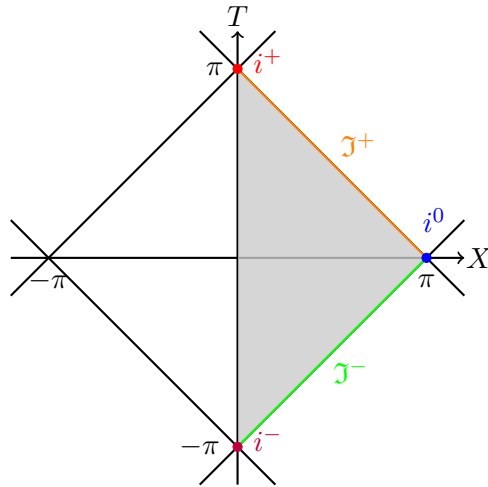
*Hint 2: Before doing the big expansion, look at the expressions for  $dt$  and  $dr$  and argue that its terms containing  $\sec^2(T + X) \sec^2(T - X)$  that will remain.*

We can therefore draw the diagram of

$$g_{\text{diagram}} = dT \otimes dT - dX \otimes dX,$$

with the ranges  $|T - X| < \pi$ ,  $|T + X| < \pi$  and  $X > 0$ .

<sup>5</sup>As an additional exercise, check that these are indeed null.



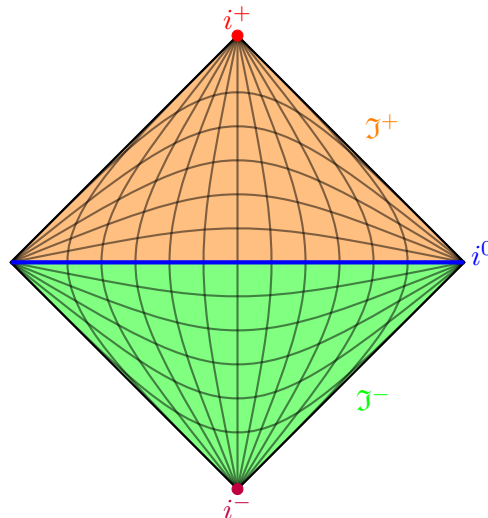
Where we have labelled:

- Spacelike infinity,  $i^0$ ,
- Future timelike infinity,  $i^+$ ,
- Past time like infinity,  $i^-$ ,
- Future null (or lightlike) infinity  $\mathcal{J}^+$ , and
- Past null (or lightlike) infinity  $\mathcal{J}^-$ .

The above points get there name from the following proposition.

- Proposition 23.2.1.**
1. All spacelike geodesics start and end at  $i^0$ ,
  2. All null geodesics start on  $\mathcal{J}^-$  and end at  $\mathcal{J}^+$ , and
  3. All timelike geodesics start at  $i^-$  and end at  $i^+$ .

We then remember that we have suppressed  $\theta$  and  $\varphi$ . So if we reinstate the  $\varphi$ , its like rotating this diagram around the  $T$  axis, and we obtain a diamond shape.

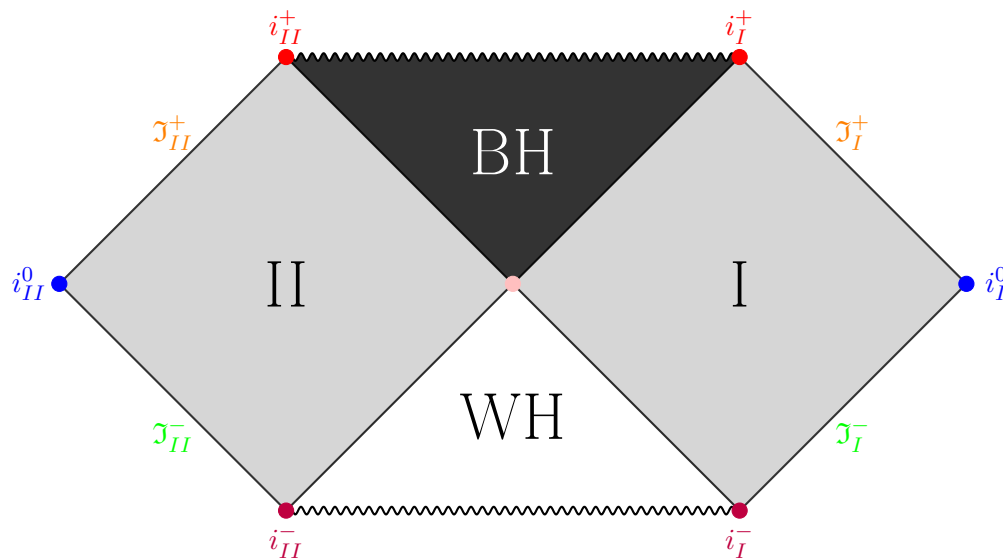


### 23.3 Other Spacetimes

In the following diagrams we shall use light grey to shade the universe(s), yellow to shade antiverse(s), black to shade black hole(s), white to shade white hole(s), pink to shade the wormhole(s), and snake-like lines to indicate singularities, using a broken line for ring singularities to remind us that they can be avoided. We shall also stick to the colours above to label the  $i$ s and  $\mathcal{I}$ s. We shall also use capital Latin numbers (i.e. I, II, etc) to number the universes.

#### 23.3.1 Schwarzschild Black Hole

If you take the maximally extended Kruskal-Szekeres coordinates for the Schwarzschild black hole shown at the end of the last lecture and compactify along the diagonals, you obtain the following Penrose diagram.



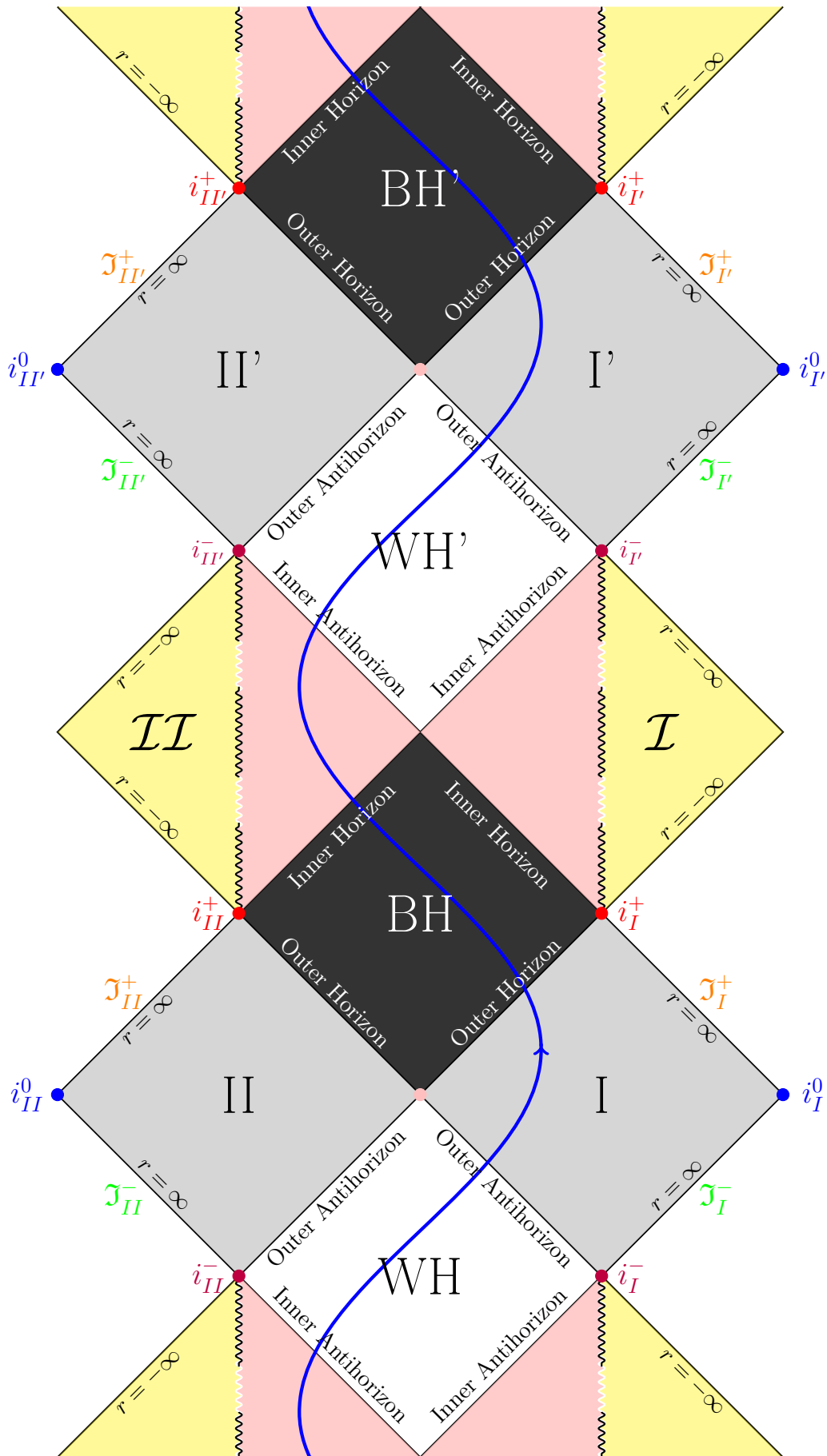
We have already basically discussed this entire diagram when considering Kruskal-Szekeres coordinates, and so we do not make further comments here.

#### 23.3.2 Kerr Black Hole

Recall that a Kerr black hole is a electrically neutral, rotating black hole that has a ring singularity and two event horizons. The presence of the inner horizon (the one at smallest  $r$ ) turns out to result in a passage to a worm hole, which in turn leads to a white hole and then another universe. It also turns out to be true that you can pass through the disc bound by the ring singularity and emerge in what some people refer to as an *antiverse*. The Penrose diagram for a Kerr black hole is the following beast.

We see that we have one universe after another, and unlike the Schwarzschild solution, these universes are causally connected. So an observer could travel into the black hole and out into a worm hole through the inner event horizon and then into another universe. Note that the boundaries of the antiverse correspond to  $r = -\infty$ .

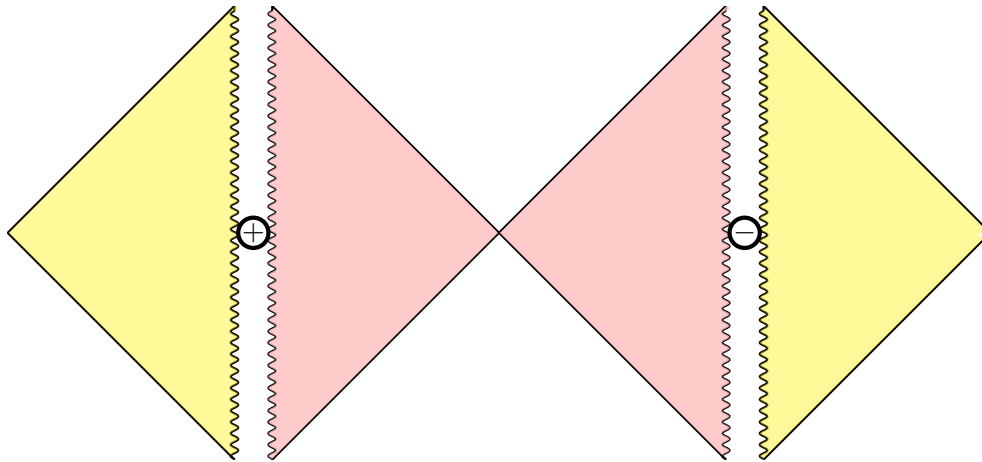




There is an interesting and important point to notice about the antiverse regions  $\mathcal{I}$  and  $\mathcal{II}$ ; there is no event horizon between ‘shielding’ the ring singularity. Such a singularity is known as **naked**. This violates the so-called *weak cosmic censorship hypothesis* which loosely says that singularities shouldn’t be observable to null infinity, hence why we have not labelled the  $r = -\infty$  boundaries with  $\mathcal{I}^+$ s.

### 23.3.3 Reissner-Nordström Black Hole

A Reissner-Nordström black hole is a non-rotating, electrically charged black hole. As the black hole is no longer rotating, it need not have a ring singularity and as such we can no longer avoid it and pass into the antiverses. The Penrose diagram looks basically identical to the Kerr Penrose diagram, but with the broken snake-like lines made continuous. It is also sometimes drawn with the antiverse separated from the wormhole region and an indication of the charge of that side of the black hole given. That is, the relevant parts of the diagram become the following



What we have described above is actually case for a so-called *sub-extremal* Reissner-Nordström black hole, which means that the two horizons do not coincide. If you actually write down the metric and find where these horizons occur, you see that the two can actually coincide if the mass and charge of the black hole are equal. In this case we the topology of the diagram changes. We will not draw the diagram here<sup>6</sup> but we simply make this comment for completeness.

*Remark 23.3.1.* Note because there is a topology change between the extremal and sub-extremal case, it is not believed that we could turn a sub-extremal Reissner-Nordström black hole into an extremal one by simply adding charge to it. Indeed, adding charge would take energy and therefore would also increase the mass of the black hole, keeping the black hole sub-extremal.

### 23.3.4 Kerr-Newmann Black Hole

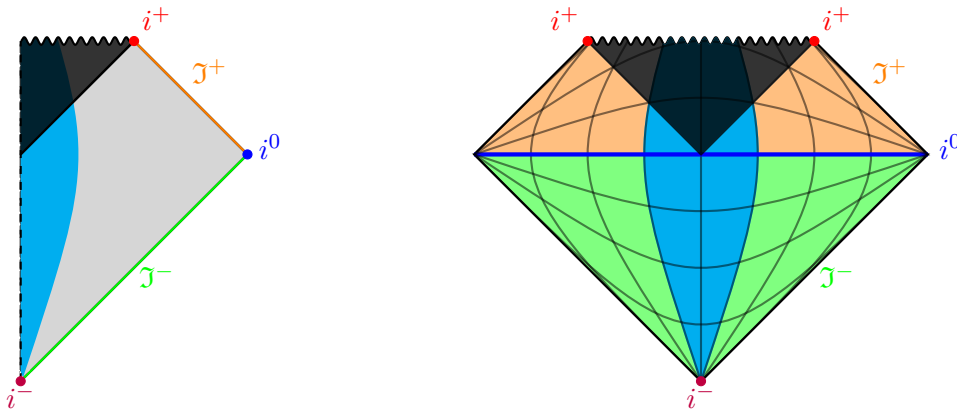
A Kerr-Newmann black hole is both rotating and has electrical charge. We therefore expect the Penrose diagram to be a combination of the previous two. This is indeed the case. Again

<sup>6</sup>It can be found [here](#).

because we have the ring singularity, it is possible to avoid it and enter into the antiverses. The Penrose diagram for the Kerr-Newmann black hole is quite often drawn identically to the one we have presented for the Kerr black hole, but it is important to remember that we should indicate the charge of the black hole somewhere, as we did with the Reissner-Nordström black hole. We shall not draw the diagram here to save space.

### 23.3.5 Gravitational Collapse

All of the Penrose diagrams for black holes assume that the black hole has always existed. That is they were not formed via some physical process such as the collapse of some massive star. We now want to draw such a Penrose diagram. In the following diagram cyan area indicates the collapsing matter, and the dashed line indicates that the line  $r = 0$  is not part of the universe.



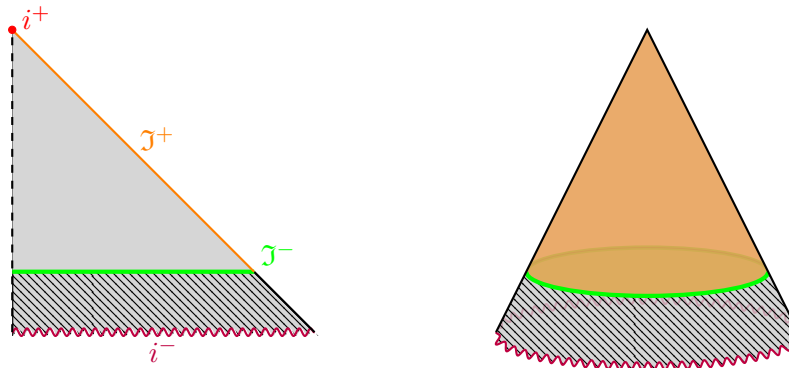
The above diagram might suggest that the matter first spreads out and then comes back together, but we need to remember that the paths of timelike geodesics are affected by our conformal factors. The matter is collapsing as you move up the diagram, and, once the mass is within the Schwarzschild radius  $r = 2m$ , it forms a black hole. Again this is diagram has two dimensions suppressed, and on the right we have tried to draw what it looks after rotated around the vertical line.

Note that by requiring that the Schwarzschild black hole forms in this manner (as opposed to having always existed) has removed the very undesirable white hole region on the diagram.

### 23.3.6 Isotropic & Homogeneous Universe

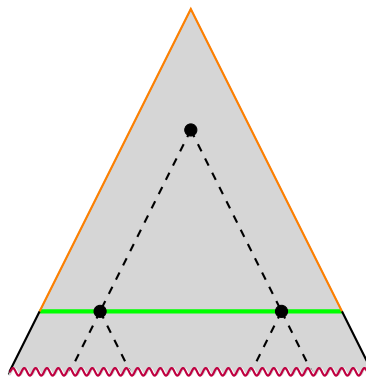
Let's consider the case when the universe is filled only with radiation (i.e.  $\kappa = 0 = \Lambda$ ). For a short period after the Big Bang, certain processes 'held' the light back and so the universe was opaque, and then at some point the light was allowed to propagate, making the universe transparent.

On the suppressed diagram, the points where the universe becomes transparent will be a horizontal line (as its a spatial line). If we then reinsert the suppressed dimensions, this line becomes a ball surrounding us. This is known as *cosmic microwave background*, or CMB for short. This is an important thing to note, as it is something that we should be able to observe experimentally, and further supports the Big Bang theory. Indeed some people even refer to the CMB as being 'the after glow of the Big Bang'.



We have used a lined fill to indicate the region where the universe is opaque. The green line is used to indicate that light rays are free to propagate from that point on-wards.<sup>7</sup> Note that the singularity at the bottom corresponds to the past timelike infinity. This is the statement that all matter (which travels along timelike curves) is created by the Big Bang. On the right we have again included reintroduced one of the suppressed dimensions, and we see that the CMB becomes a disc. There is actually a very important question to ask about the CMB, which we highlight now with use of the next diagram.

Experiments tell us that wherever we look at the CMB we get similar data, e.g. that the temperature is about 3K. This might not seem like a problem, after all we want homogeneity and isotropy, but it does pose a problem. Projecting the rotated diagram onto a 2D-plane to simplify the drawing, consider some point in the universe and look at its past light cone to which points are casually connected, i.e. can influence our point. Take the two points on the opaque-transparent dividing line and consider their past cones, if they do not overlap they are completely independent. That is there is no point that can influence both points.

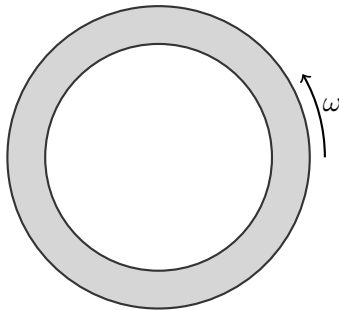


If the points are completely independent, why do we get the same measurements for both of them? This is not just for two special points on the CMB, but *all* the points on the CMB. This problem is known as the *cosmological horizon problem* (or the *homogeneity problem*). The most commonly accepted fix to the problem is to include a so-called *inflaton field*, which gives rise to *cosmic inflation*.

<sup>7</sup>We might actually not be able to call this  $\mathcal{J}^-$  because maybe null geodesics are defined below this line. I don't actually know the answer to this question, if any readers do please feel free to send me an explanation and I'll update it with credit.

## 24 | Perturbation Theory I

We have seen that in order for us to solve Einstein's equations exactly, we require strong symmetry conditions, or, equivalently, simply energy-momentum tensors. This is a shame as we would obviously like to also study the cases of weaker symmetry conditions. An important example would be to find the gravitational field,  $g$ ,<sup>1</sup> within for a rotating shell of evenly distributed mass.



Note that Newtonian gravity, which is given by Poisson's equation

$$\rho = \nabla^2 \phi,$$

would tell us that the gravitational field inside the shell vanishes as  $\rho$  vanishes inside it. Besides this fact, introducing rotation would not effect the gravitational field as all Poisson's equation cares about is the distribution of mass  $\rho$ , and our mass is evenly distributed so there is no change by rotation.

Einstein's equations<sup>2</sup>

$$G_{ab}[g] = T_{ab}[g, \Phi],$$

will encode both a non-vanishing field inside the shell and will also encode the change due to introducing rotation, as the  $T_{0\mu}$  components encode the angular momentum.

### 24.1 Perturbation Of Exact Solutions

Assume that an exact solution  $g$  to some Einstein's equations

$$G_{ab}[g] = T_{ab}[g]$$

<sup>1</sup>Technically  $g$  is the gravitational potential, but this subtlety is not important here.

<sup>2</sup>In units where  $\frac{8\pi G_N}{c^4} = 1$ .

is known in terms of component functions  $g_{ab}$  w.r.t. to some coordinate chart(s). It would be nice to be able to extract the gravitational potential,  $g_{ab} + \delta g_{ab}$ , that solves the equations

$$G_{ab}[g + \delta g] = T_{ab}[g] + \delta T_{ab}[g],$$

where  $\delta T_{ab}[g]$  is a small perturbation of the right-hand side of the Einstein equations. If the perturbation on the right-hand side is small, we can assume that  $\delta g$  itself is small<sup>3</sup> and so we can expand

$$G_{ab}[g] + \delta G[g, \delta g] + \mathcal{O}(\delta g^2) = T_{ab}[g] + \delta T_{ab}[g],$$

where  $\delta G[g, \delta g]$  has linear dependence on  $\delta g$ . Dropping the higher order terms and using the fact that  $g$  is an exact solution, we have

$$\delta G_{ab}[g, \delta g] = \delta T_{ab}[g].$$

The remaining task is to then find  $\delta g_{ab}$ . This method is known as *linear perturbation theory*.

Example 24.1.1. For the shell of mass, we could consider the exact solution to be that where there is no shell at all, i.e.

$$G_{ab}[g] = 0.$$

We know an exact solution to this is given by the flat metric  $g = \eta$ . We can now ask the question about introducing a small mass that rotates slowly and treat it as a perturbation to the energy-momentum tensor,

$$\delta G_{ab}[\eta, \delta g] = \delta T_{ab}[g].$$

Remark 24.1.2. It will turn out to be interesting to also consider cases where the right-hand side of the Einstein equations are not perturbed (i.e.  $\delta T_{ab}[g] = 0$ ). For example this will lead us to so-called *gravitational waves*. This might seem like a strange thing to say, as how can we perturb the metric but not the right-hand side of the Einstein equations? The answer is the idea that the right-hand side encodes the matter in the spacetime, whereas the metric encodes the gravity and curvature. We can think of this as ‘prodding’ the spacetime manifold and getting it to ripple, without introducing any new matter. It is by the same argument that we see that  $\delta T_{ab}$  is only a function of  $g$  and not  $\delta g$ .

## 24.2 The Perturbed Metric

The idea is the following: we are calculating/solving in some chart  $(U, x)$  anyway, so we can equally well arrange for our known exact solution to take a particular form. For example, for a static metric  $g$ , we can always find a chart  $(U, x)$  such that

$$g_{ab} = \left( \begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & & \\ 0 & & -\gamma_{\alpha\beta} \\ 0 & & \end{array} \right),$$

<sup>3</sup>If it turned out not to be small, we should have that a small change to the right-hand side of our equations of motion gives rise to a large change on the left-hand side, indicating that the solution is not stable. We clearly ignore any cases like this on physical grounds.

where  $\gamma_{\alpha\beta}$  is some time-independent Riemannian 3-metric. We can write this more compactly as

$$g = dx^0 \otimes dx^0 - \gamma_{\alpha\beta} dx^\alpha \otimes dx^\beta.$$

This particular case is useful for taking perturbations about Schwarzschild spacetime, for example.

Now it is clever to describe the 10 small fields encoded in  $\delta g_{ab}$  as

$$\delta g = 2a dx^0 \otimes dx^0 - b_\alpha [dx^0 \otimes dx^\alpha + dx^\alpha \otimes dx^0] - [2c\gamma_{\alpha\beta} + e_{\alpha\beta}] dx^\alpha \otimes dx^\beta,$$

for *small*, spatial

- (i) scalar fields  $a$  and  $c$ ,
- (ii) vector field  $b^\alpha = \gamma^{\alpha\beta} b_\beta$ ,
- (iii) symmetric tensor field  $e_{\alpha\beta}$ , which is trace free,  $\gamma^{\alpha\beta} e_{\alpha\beta} = 0$ ,

all of which are allowed to depend on all  $x^a$ s (i.e. can depend on  $x^0$ , even though they are spatial).

Remark 24.2.1. Note this is just the perturbation. That is the complete new metric is  $g + \delta g = (1 + 2a)dx^0 \otimes dx^0 - \dots$

Counting the number of degrees of freedom, we have

- (i)  $1 + 1 = 2$ ,
- (ii)  $3$ ,
- (iii)  $\frac{3(3+1)}{2} - 1 = 5$

to give a total of 10, as required.

Remark 24.2.2. It is simply convenient to think of a general perturbation of the metric in terms of these 10 degrees of freedom.

Remark 24.2.3. We shall use notation such that Greek indices which do not appear in their natural position have been raised/lowered using  $\gamma^{\alpha\beta}/\gamma_{\alpha\beta}$ . We do this just to lighten notation a bit.

### 24.3 Helmholtz-Hodge

It is immensely useful to further decompose the

- (a) vector field  $b_\alpha$  as

$$b_\alpha = D_\alpha B + B_\alpha,$$

where  $D$  is the Levi-Civita covariant derivative of  $\gamma$ ,  $B$  is a scalar field and  $B_\alpha$  is a *divergence-free* vector field,  $D_\alpha B^\alpha = 0$ . This is known as **Helmholtz Theorem**, and actually states that  $B$  and  $B_\alpha$  are unique.

(b) tensor field  $e_{\alpha\beta}$  as

$$e_{\alpha\beta} = \left( 2D_{(\alpha}D_{\beta)} - \frac{1}{3}\gamma_{\alpha\beta}\Delta \right) E + 2D_{(\alpha}E_{\beta)} + E_{\alpha\beta},$$

where  $\Delta := \gamma^{\alpha\beta}D_{\alpha}D_{\beta}$  is the spatial *Laplacian*,  $E$  is a scalar field,  $E_{\alpha}$  is a divergence-free vector field, and  $E_{\alpha\beta}$  is a symmetric, divergence-free,  $D_{\alpha}E^{\alpha\beta} = 0$ , and trace-free tensor field. This decomposition is also unique, and is known as **Hodge Theorem**.

Thus a general perturbation from a static metric uniquely decomposes into three, independent types of perturbation:

$$\delta g = \delta g_{\text{scalar}} + \delta g_{\text{vector}} + \delta g_{\text{tensor}}.$$

If we make the trivial decompositions  $a = A$  and  $c = C$ , the above formula is summarised in the table below, where we have included some common terminology for the categories

Type of perturbation	Contains	Terminology
$\delta g_{\text{scalar}}$	$A, B, C$ and $E$	Scalars
$\delta g_{\text{vector}}$	$B_{\alpha}$ and $E_{\alpha}$	Solenoidal vector fields
$\delta g_{\text{tensor}}$	$E_{\alpha\beta}$	Symmetric, TT <sup>4</sup> tensor fields

People then say "scalar perturbations come from scalars, vector perturbations come from solenoidal vector fields, and tensor perturbations come from symmetric, TT tensor fields."

The rational behind this distinction is that a perturbation  $\delta T_{ab}[g]$  on the right-hand side which is effected only by a scalar fields will *at most* 'switch on' the scalar fields in the metric  $g + \delta g$  on the left-hand side. Similarly for solenoidal and TT perturbations of the energy-momentum tensor.

This means that if we decompose the right-hand side of our perturbed system into the three types of contributions, we can solve each part separately and see its contribution to the system.

*Remark 24.3.1.* Note that the above decomposition only works in *linear* perturbation theory, as higher order terms would contain cross terms when the decompositions are expanded out. So we can only do the above solving independently for in linear perturbation theory.

*Example 24.3.2.* For our rotating shell, the scalar perturbation is given by introducing the mass distribution  $\rho$ , and the solenoidal perturbation is given by introducing rotation, as the vector field associated to it is divergence free (it is essentially a curl field). We could introduce a TT perturbation by applying pressure to the shell.

*Remark 24.3.3.* It is important to consider what is fundamentally contributing to the perturbation. In our shell example, we might think of some vector field which causes the shell to pulse and oscillate in size is a solenoidal vector perturbation. This is not the case as it is not divergence-free and so is not a *solenoidal* vector perturbation. It is, in fact, a scalar contribution as the required vector field can be obtained as the gradient of some scalar field, and it is the scalar field that generates the perturbation.

*Terminology.* Despite the last remark, people do not often say the "solenoidal" and simply say "vector perturbations". Similarly they just say "tensor perturbations".



### 24.4 The Price Paid For The Luxury Of Working In A Chart

We have made the argument again and again that real world objects, like the metric, are independent of which chart you choose to express them in, and that the components of these objects can vary vastly from one chart to another.

So far we have calculated everything in the chart  $(U, x)$  and obtained some  $\delta g_{(x)ab}$ . The obvious question to ask is whether there exists another chart  $(U, y)$  such that

$$g_{(x)ab} + \delta g_{(x)ab} = g_{(y)ab}?$$

If this is the case, we have no choice but to conclude that the metric with components  $g_{(x)ab} + \delta g_{(x)ab}$  is precisely the metric we started with. That is, we have not actually found a *real world* perturbation to the system but instead we have generated a ‘fake’ one at the chart level.

*Example 24.4.1.* Consider an infinitely extended plane. We introduce an evenly distributed matter density across the whole plane. Newtonian theory tells us that the gravitational field is homogeneous and just points orthogonal to the plane. However, we saw when discussing tidal forces that in general relativity such a gravitational field can be removed by transforming to a freely falling frame. We can therefore find a coordinate system in which the contribution made by the matter vanishes, and so it is a ‘fake’ perturbation.

The insight into this problem is that for

- (i) scalar perturbations, only two combinations of  $A, B, C$  and  $E$

$$\Psi := A + \dot{B} - \ddot{E}, \quad \text{and} \quad \Phi := C - \frac{1}{3}\Delta E$$

- (ii) vector perturbations, only one combination of the  $B_\alpha$  and  $E_\alpha$

$$\Theta_\alpha = B_\alpha - \dot{E}_\alpha$$

- (iii) tensor perturbations all the  $E_{\alpha\beta}$

are unaffected (and therefore *not* removable) by general coordinate transformations. They are known as **gauge invariants**. They are the only ones that can be taken seriously. We will derive these results next lecture.

This concept should be familiar from electrodynamics, where we know that the fields  $A_\mu$  themselves are not physically meaningful, but certain combinations are, for example the field strength tensor components  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ .

## 25 | Perturbation Theory II

It is a fact that any small, but otherwise arbitrary, transformation of coordinates is given by choice of vector field  $\xi$ , whose component functions w.r.t. the original chart,  $\xi_{(x)}^m$ , are small. Then the change incurred by the metric components under such a small, but otherwise arbitrary, transformation is given by<sup>1</sup>

$$\Delta_\xi g_{ab} = (\mathcal{L}_\xi g)_{ab},$$

where  $\mathcal{L}$  is the Lie derivative. In other words, if the  $\delta g_{ab}$  takes this form it is a ‘fake’ perturbation, as described at the end of last lecture. But of course we write

$$g_{ab} + \delta g_{ab},$$

and so we must study how  $\delta g_{ab}$  itself changes if we choose a small, arbitrary  $\xi$ . If it is a fake perturbation then we know the change in the metric components is of the above form, but  $\delta g_{ab}$  is the change in the metric, and so taking a further change  $\xi$  will simply give us

$$(\Delta_\xi \delta g)_{ab} = (\mathcal{L}_\xi g)_{ab},$$

where we note that there is no  $\delta$  on the right-hand side. The aim is to find which components are fake and then work out which, if any, combinations of components cancel on the right-hand side giving us a real perturbation.

### 25.1 Calculate $\Delta_\xi a$ , $\Delta_\xi c$ , $\Delta_\xi b_\alpha$ & $\Delta_\xi e_{\alpha\beta}$

Even though we did the decomposition in the last lecture to the capital letters ( $A, B_\alpha$ , etc.), it is actually cleverer to work out the change of the lower case letters first.

Our chart is  $(U, x)$  where  $g = dx^0 \otimes dx^0 - \gamma_{\alpha\beta} dx^\alpha \otimes dx^\beta$ . We can decompose the vector field  $\xi$  in this chart as

$$\xi = T\partial_0 + L^\beta\partial_\beta,$$

where  $T$  and  $L^\mu$  are scalar fields that can depend on all the chart coordinates, including  $x^0$ . We then calculate the change of  $\delta g_{ab}$  incurred by  $\xi$  by considering  $T$  and  $L^\mu$ . We have

$$\begin{aligned} (\Delta_\xi \delta g)_{ab} &= (\mathcal{L}_\xi g)_{ab} \\ &= \xi^m g_{ab,m} + \xi^m{}_{,a} g_{mb} + \xi^m{}_{,b} g_{am} \\ &= T g_{ab,0} + L^\mu g_{ab,\mu} + T_{,a} g_{0b} + L^\mu{}_{,a} g_{\mu b} + T_{,b} g_{a0} + L^\mu{}_{,b} g_{a\mu}. \end{aligned}$$

<sup>1</sup>The  $\Delta_\xi$  here is not the Laplacian, but simply denotes the change under  $\xi$ .

We now consider each of the components separately. Using a dot to indicate a derivative w.r.t.  $x^0$ , we have:

$$(\Delta_\xi \delta g)_{00} = 0 + 0 + \dot{T} + 0 + \dot{T} + 0 = 2\dot{T},$$

where we have used the fact that  $g_{00} = 1$  and  $g_{0\alpha} = 0$ .

Next we have

$$(\Delta_\xi \delta)_{0\beta} = 0 + 0 + 0 - \dot{L}^\mu \gamma_{\mu\beta} + T_{,\beta} + 0 = -\dot{L}^\mu \gamma_{\mu\beta} + T_{,\beta},$$

where we have used  $g_{0\beta} = 0$ ,  $g_{00} = 1$  and  $g_{\alpha\beta} = -\gamma_{\alpha\beta}$ . We can then use the fact that  $\gamma_{\alpha\beta}$  is independent of  $x^0$  to give us

$$(\Delta_\xi \delta)_{0\beta} = T_{,\beta} - (L^\mu \gamma_{\mu\beta})_{,0} = D_\beta T - \dot{L}_\beta,$$

where  $D_\beta$  is the Levi-Civita covariant derivative using  $\gamma$ .

Finally we have

$$(\Delta_\xi \delta g)_{\alpha\beta} = 0 - L^\mu \gamma_{\alpha\beta,\mu} + 0 - L^\mu{}_{,\alpha} \gamma_{\mu\beta} + 0 - L^\mu{}_{,\beta} \gamma_{\alpha\mu}.$$

We then use

$$L^\mu{}_{,\alpha} \gamma_{\mu\beta} = (L^\mu \gamma_{\mu\beta})_{,\alpha} - L^\mu \gamma_{\mu\beta,\alpha}$$

to give

$$(\Delta_\xi \delta g)_{\alpha\beta} = -2 \left[ L_{(\alpha,\beta)} - \frac{1}{2} L^\mu (\gamma_{\mu\beta,\alpha} + \gamma_{\mu\alpha,\beta} - \gamma_{\alpha\beta,\mu}) \right] = -2D_{(\alpha} L_{\beta)},$$

where again  $D_\alpha$  is the Levi-Civita covariant derivative using  $\gamma$ .

Using the expression for  $\delta g$  in terms of  $a$ ,  $b_\alpha$ ,  $c$  and  $e_{\alpha\beta}$  from last lecture, we conclude that

$$\begin{aligned} \Delta_\xi(2a) &= 2\dot{T} \\ \Delta_\xi(-b_\alpha) &= D_\alpha T - \dot{L}_\alpha \\ \Delta_\xi(-2c\gamma_{\alpha\beta} - e_{\alpha\beta}) &= -2D_{(\alpha} L_{\beta)}. \end{aligned}$$

The first two expressions clearly tell us that

$$\Delta_\xi a = \dot{T}, \quad \text{and} \quad \Delta_\xi b_\alpha = \dot{L}_\alpha - D_\alpha T,$$

but the third expression needs a little work. We contract both sides with  $\gamma^{\alpha\beta}$  and use  $\gamma^{\alpha\beta}\gamma_{\alpha\beta} = 3$  along with the fact that  $D$  is  $\gamma$ -metric compatible (so we can ‘take  $\gamma$  inside it’), giving

$$6\Delta_\xi c + \Delta_\xi \gamma^{\alpha\beta} e_{\alpha\beta} = 2D_\alpha (\gamma^{\alpha\beta} L_\beta) \quad \implies \quad \Delta_\xi c = \frac{1}{3} D_\alpha L^\alpha,$$

where we used the trace-free condition of  $e_{\alpha\beta}$ . From this it follows that

$$\Delta_\xi e_{\alpha\beta} = 2D_{(\alpha} L_{\beta)} - \frac{2}{3} D_\mu L^\mu \gamma_{\alpha\beta}.$$

## 25.2 Scalar Perturbations

We now want to consider the decomposed fields, and we start with the scalar perturbations.

*Claim 25.2.1.* If we consider the seemingly restricted case  $L^\alpha = D^\alpha L$  for some scalar field  $L$ , we actually get the same result for the scalar perturbations as if we had done it generally.

We do not prove this claim here but just use it to simplify the following.  
Using the above claim and the results from last lecture we see that

$$\begin{aligned}\Delta_\xi A &= \dot{T}, \\ \Delta_\xi B &= \dot{L} - T \\ \Delta_\xi C &= \frac{1}{3}\Delta L \\ \Delta_\xi E &= L,\end{aligned}$$

where on the right-hand side of the  $C$  equation  $\Delta$  is the Laplacian.

### Exercise

Show the above results hold.

*Hint: If you're stuck Dr. Schuller does this in the videos.<sup>a</sup>*

<sup>a</sup>There is a couple factors of 2 for the  $E$  in my notes that I think Dr. Schuller missed, but they end up cancelling above so I get the same result for  $\Delta_\xi E$ .

We can now use these results to show the results quoted at the end of last lecture, namely

$$\Psi := A + \dot{B} - \ddot{E}, \quad \text{and} \quad \Phi := C - \frac{1}{3}\Delta E$$

are gauge invariants.

### Exercise

Show that above hold.

*Hint: Again this is done in the lectures.*

*Remark 25.2.2.* Note these results tell us that the scalars  $A, B, C$  and  $E$  can not be 'switched on' independently, as the only real perturbations are given as combinations of them.

For convenience only, and precisely because the  $\Phi$  and  $\Psi$  are gauge invariants, we are free to pick any  $(T, L^\alpha)$  we want. Such a choice is known as a **gauge**, and the process **gauge fixing**. Again should be familiar from electromagnetism. We decide to use the gauge

$$T := B - \dot{E}, \quad \text{and} \quad L := -E.$$

We choose this gauge as then

$$\begin{aligned}\Delta_\xi E = L &\implies E = 0 \\ \Delta_\xi B = \dot{L} - T &\implies B = 0,\end{aligned}$$

and so

$$\Psi = A, \quad \text{and} \quad \Phi = C.$$

This is known as the **longitudinal gauge**. Working in this gauge, we finally get the left-hand side for the scalar perturbations for Einsteins equations,  $\delta G_{ab} = T_{ab}$ , as

$$\begin{aligned} \delta G_{00} &= 2\Delta\Phi \\ \delta G_{0\alpha} &= 2D_\alpha\dot{\Phi} \\ \delta G_{\alpha\beta} &= \left[2\ddot{\Phi} - \frac{2}{3}\Delta(\Psi + \Phi)\right]\gamma_{\alpha\beta} + \left[D_{(\alpha}D_{\beta)} - \frac{1}{3}\gamma_{\alpha\beta}\Delta\right](\Psi + \Phi) \end{aligned}$$

### 25.3 Vector Perturbations

*Claim 25.3.1.* As with the scalar perturbations, it turns out we can consider the seemingly restricted cases of

$$T = 0, \quad \text{and} \quad D_\alpha L^\alpha = 0,$$

and get the full result.

Again, we do not prove this claim but just use it.

The results here are

$$\Delta_\xi B_\alpha = \dot{L}_\alpha, \quad \text{and} \quad \Delta_\xi E_\alpha = L_\alpha,$$

and so the only gauge invariant quantities are

$$\Theta := B_\alpha - \dot{E}_\alpha.$$

We then use the **vector gauge**, which sets  $E = 0$ , and obtain  $\Theta_\alpha = B_\alpha$ . We then obtain the left-hand side of the vector perturbations of Einsteins equations,

$$\delta G_{0\alpha} = \frac{1}{2}\Delta\Theta_\alpha, \quad \text{and} \quad \delta G_{\alpha\beta} = D_{(\alpha}\dot{\Theta}_{\beta)}.$$

### 25.4 Tensor Perturbations

*Claim 25.4.1.* Once again we can consider a seemingly restricted case, this time  $T = 0 = L^\alpha$ .

With this claim it follows trivially that

$$\Delta_{xi}E_{\alpha\beta} = 0,$$

and so all  $E_{\alpha\beta}$  are gauge invariants. The left-hand side equations are

$$\delta G_{\alpha\beta} = -\ddot{E}_{\alpha\beta} + \Delta E_{\alpha\beta}.$$

*Remark 25.4.2.* If we start from the exact solution  $G_{ab}[g] = 0$  and do not introduce any mass (so  $\delta T_{ab}[g] = 0$ ) we get for the tensor perturbations

$$-\ddot{E}_{\alpha\beta} + \Delta E_{\alpha\beta} = 0,$$

which is a *wave equation*. These are so-called **gravitational waves**. It is important to note that this wave equation is made from gauge invariant quantities and is a real world wave, not some ‘coordinate wave’ that can be removed by a transformation.

## 26 | Tutorials

The following chapters are the tutorials for the course. I haven't included everything from the tutorials (e.g. the definitions etc are omitted), but I have tried to include the most helpful questions. Some of the answers to the tutorial questions have also been placed within the main content of the lecturers as remarks etc.

### 26.1 Topology

#### 26.1.1 Topologies On A Simple Set

Let  $\mathcal{M} = \{1, 2, 3, 4\}$  be a set.

**Question:** does  $\mathcal{O}_1 := \{\emptyset, \{1\}, \{1, 2, 3, 4\}\}$  constitute a topology on  $\mathcal{M}$ ?

**Solution:** Check the conditions:

- (i)  $\emptyset \in \mathcal{O}_1$  and  $\mathcal{M} \in \mathcal{O}_1$ .
- (ii)  $\emptyset \cap \{1\} = \emptyset \in \mathcal{O}_1$ ,  $\emptyset \cap \{1, 2, 3, 4\} = \emptyset \in \mathcal{O}_1$  and  $\{1\} \cap \{1, 2, 3, 4\} = \{1\} \in \mathcal{O}_1$ .
- (iii)  $\emptyset \cup \{1\} = \{1\} \in \mathcal{O}_1$ ,  $\emptyset \cup \{1, 2, 3, 4\} = \{1, 2, 3, 4\} \in \mathcal{O}_1$ ,  $\{1\} \cup \{1, 2, 3, 4\} = \{1, 2, 3, 4\} \in \mathcal{O}_1$  and  $\emptyset \cup \{1\} \cup \{1, 2, 3, 4\} = \{1, 2, 3, 4\} \in \mathcal{O}_1$ .

So the answer is yes.

**Question:** What about  $\mathcal{O}_2 := \{\emptyset, \{1\}, \{2\}, \{1, 2, 3, 4\}\}$ .

**Solution:** The only new addition is  $\{2\}$  so just need to check its involvement. We see straight away that  $\{1\} \cup \{2\} = \{1, 2\} \notin \mathcal{O}_2$  and so we conclude that it is not a topology.

#### 26.1.2 Continuous Functions

**Question:** Let  $\mathcal{M} = \{1, 2, 3, 4\}$  and consider the identity map  $\text{id}_{\mathcal{M}} : \mathcal{M} \rightarrow \mathcal{M}$  defined by

$$\text{id}_{\mathcal{M}}(1) = 1, \quad \text{id}_{\mathcal{M}}(2) = 2, \quad \text{id}_{\mathcal{M}}(3) = 3, \quad \text{id}_{\mathcal{M}}(4) = 4.$$

Is the map  $\text{id}_{\mathcal{M}}$  continuous if the domain is equipped with the chaotic topology and

the target with the topology  $\mathcal{O}_{\text{target}} := \{\emptyset, \{1\}, \{1, 2, 3, 4\}\}$ ?

**Solution:** The chaotic topology here is

$$\mathcal{O}_{\text{chaotic}} = \{\emptyset, \{1, 2, 3, 4\}\}.$$

We see straight away that the preimage of  $\{1\} \in \mathcal{O}_{\text{target}}$  is  $\{1\}$  which is not in  $\mathcal{O}_{\text{chaotic}}$ , and so the map is not continuous w.r.t. these topologies.

**Question:** Consider the inverse  $\text{id}_{\mathcal{M}}^{-1} : \mathcal{M} \rightarrow \mathcal{M}$  of the identity map  $\text{id}_{\mathcal{M}}$ , such that now the target is equipped with the chaotic topology and the domain with the topology  $\{\emptyset, \{1\}, \{1, 2, 3, 4\}\}$ . Provide the values of the map  $\text{id}_{\mathcal{M}}^{-1}$  and decide whether  $\text{id}_{\mathcal{M}}^{-1}$  is continuous!

**Solution:**

$$\text{id}_{\mathcal{M}}^{-1}(1) = 1, \quad \text{id}_{\mathcal{M}}^{-1}(2) = 2, \quad \text{id}_{\mathcal{M}}^{-1}(3) = 3, \quad \text{id}_{\mathcal{M}}^{-1}(4) = 4$$

Now consider the preimages.

$$\text{preim}_{\text{id}_{\mathcal{M}}^{-1}}(\emptyset) = \emptyset, \quad \text{and} \quad \text{preim}_{\text{id}_{\mathcal{M}}^{-1}}(\{1, 2, 3, 4\}) = \{1, 2, 3, 4\} = \mathcal{M},$$

both of which are in our domain's topology. Therefore the map is continuous w.r.t. to these topologies.

### 26.1.3 The Standard Topology On $\mathbb{R}^d$

I have not included question 4 here because it would be too much drawing on Tikz for me... however the questions are worth looking at, so if you haven't already go [watch the video](#).

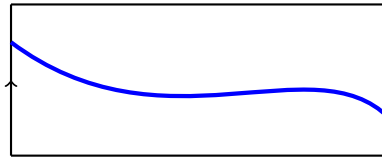
## 26.2 Topological Manifolds

### 26.2.1 An Atlas From A Real World — the Möbius river

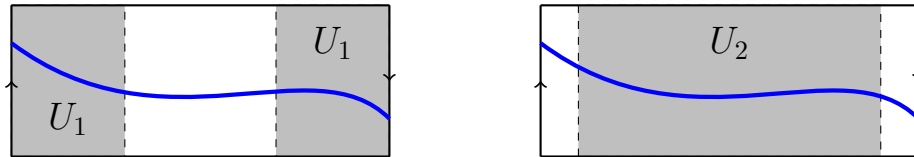
**Question:** Consider a Möbius strip<sup>a</sup> with a river drawn along it. How many charts do you need to cover the Möbius strip? Draw an image of the river and Möbius strip under the chart map(s)!

<sup>a</sup>Please Google it if you don't know what it is.

**Solution:** The Möbius strip can be represented by the following diagram



where the arrows indicate how we identify the two edges together (i.e. the bottom right corner goes to the top left corner). The blue line is the river. It is clear that we will need at least two charts in order to map the whole strip. We could choose them as follows



We then define maps from these shaded regions into  $\mathbb{R}^2$ . For example we can just map them so that they are the shaded regions on the page, taking care to region the two parts of  $U_1$  properly. To save my writing more Tikz code, I'll leave this to your imagination.

### 26.2.2 A Real World From An Atlas

This is worth watching on [the video](#), but would be way too hard for me to do on here. So please go watch the video for this one.

### 26.2.3 Before The Invention Of The Wheel

Consider the set  $F^1 := \{(m, n) \in \mathbb{R}^2 \mid m^4 + n^4 = 1\}$  of pairs of real numbers  $(m, n)$ . Let it be equipped with the topology subset topology  $\mathcal{O}_s|_{F^1}$  inherited from the standard topology on  $\mathbb{R}^2$ .

**Question:** We look at a map  $x : F^1 \rightarrow \mathbb{R}$  that maps a pair in  $F^1$  to the first entry in the pair. Write this in formal mathematical terms! Is the map injective?

**Solution:** The map is  $x : F^1 \rightarrow [-1, 1]$  given by  $x : (m, n) \mapsto m$ . Clearly this map is not injective as  $x((m, n)) = x((m, n'))$  where  $n' = -n$ .

**Question:** This map may be made injective by restricting its domain to either of two maximal open subsets of  $F^1$ . Which ones? Call them  $x_\uparrow$  and  $x_\downarrow$ .

**Solution:** The problem was that  $n' = -n$  gave non-injectivity, so clearly we define

$$U_\uparrow := \{(m, n) \in \mathbb{R}^2 \mid m^4 + n^4 = 1, n > 0\}$$

$$U_\downarrow := \{(m, n) \in \mathbb{R}^2 \mid m^4 + n^4 = 1, n < 0\}$$



and so have the maps

$$x_{\uparrow} : U_{\uparrow} \rightarrow (-1, 1), \quad \text{and} \quad x_{\downarrow} : U_{\downarrow} \rightarrow (-1, 1).$$

It is important that we take the inequalities for  $n$  (i.e.  $n \neq 0$ ) so that our sets are open, as required by the question.

**Question:** Now, construct an injective map  $y_{\uparrow} : F^1 \rightarrow \mathbb{R}$  that maps every pair in a maximal open subset of  $F^1$  to the second entry of the pair.

**Solution:** Same as above we simply define

$$V_{\uparrow} := \{(m, n) \in \mathbb{R}^2 \mid m^4 + n^4 = 1, m > 0\},$$

and the map

$$y_{\uparrow} : V_{\uparrow} \rightarrow (-1, 1),$$

given by  $y_{\uparrow} : (m, n) \mapsto n$ .

**Question:** Is  $y_{\uparrow}$  invertible? If so, construct  $y_{\uparrow}^{-1}$ !

**Solution:** Yes it is invertible as it is bijective. We construct the inverse as

$$y_{\uparrow}^{-1} : (-1, 1) \rightarrow V_{\uparrow},$$

where the action is given by

$$y_{\uparrow}^{-1} : n \mapsto (\sqrt[4]{1 - n^4}, n).$$

**Question:** Do the domains of the maps  $x_{\uparrow}$  and  $y_{\uparrow}$  overlap? If so, construct the transition map  $x_{\uparrow} \circ y_{\uparrow}^{-1}$  and specify its domain and target.

**Solution:** Yes their domains overlap as

$$U_{\uparrow} \cap V_{\uparrow} = \{(m, n) \in \mathbb{R}^2 \mid m^4 + n^4 = 1, n, m > 0\}$$

The transition map is

$$(x_{\uparrow} \circ y_{\uparrow}^{-1}) : (0, 1) \rightarrow (0, 1)$$

with

$$(x_{\uparrow} \circ y_{\uparrow}^{-1})(a) = x_{\uparrow}(\sqrt[4]{1 - a^4}, a) = \sqrt[4]{1 - a^4} \in (0, 1).$$

**Question:** How many maps (constructed this way) do you need for their domains to cover the whole of  $F^1$ ? Does the collection of these domains and maps form an atlas on  $F^1$ ?

**Solution:** The answer is obviously 4 maps,  $x_{\uparrow}, x_{\downarrow}, y_{\uparrow}$  and  $y_{\downarrow}$ , and yes they form an atlas. From this we see that  $F^1$  is a topological manifold of dimension  $d = 1$ .

## 26.3 Multilinear Algebra

### 26.3.1 Vector Spaces

Let  $V := \mathbb{R}^3$  be the set of all real triples.

**Question:** We equip  $V$  with addition  $\oplus : V \times V \rightarrow V$  and s-multiplication  $\odot : \mathbb{R} \times V \rightarrow V$  defined by

$$(a, b, c) \oplus (d, e, f) := (a + d, b + e, c + f)$$

and

$$\lambda \odot (a, b, c) := (\lambda \cdot a, \lambda \cdot b, \lambda \cdot c)$$

where  $+$  and  $\cdot$  are the addition and multiplication on  $\mathbb{R}$ . Check that  $(V, \oplus, \odot)$  is a vector space.

**Solution:** We need to check it meets the 8 axioms.

(i) Commutative w.r.t.  $\oplus$ :

$$(a, b, c) \oplus (d, e, f) = (a + d, b + e, c + f) = (d + a, b + e, f + c) = (d, e, f) \oplus (a, b, c).$$

(ii) Associative w.r.t.  $\oplus$ :

$$\begin{aligned} [(a, b, c) \oplus (d, e, f)] \oplus (h, i, j) &= (a + d, b + e, c + f) \oplus (h, i, j) \\ &= (a + d + h, b + e + i, c + f + j) \\ &= (a, b, c) \oplus (d + h, e + i, f + j) \\ &= (a, b, c) \oplus [(d, e, f) \oplus (h, i, j)]. \end{aligned}$$

(iii) Neutral element w.r.t.  $\oplus$ : this is clearly just  $(0, 0, 0)$ .

(iv) Inverse w.r.t.  $\oplus$ : this is clearly just  $(-a, -b, -c)$ . Note for this to be true it is important that we take the whole real line, not just the positive reals (otherwise  $(-a, -b, -c) \notin V$ ).

(v) Associative w.r.t.  $\odot$ :

$$\begin{aligned} \lambda \odot [\mu \odot (a, b, c)] &= \lambda \odot (\mu \cdot a, \mu \cdot b, \mu \cdot c) \\ &= (\lambda \odot \mu \cdot a, \lambda \odot \mu \cdot b, \lambda \odot \mu \cdot c) \\ &= [\lambda \odot \mu] \odot (a, b, c). \end{aligned}$$

(vi) Distributive 1:

$$\begin{aligned} (\lambda + \mu) \odot (a, b, c) &= ((\lambda + \mu) \cdot a, (\lambda + \mu) \cdot b, (\lambda + \mu) \cdot c) \\ &= (\lambda \cdot a, \lambda \cdot b, \lambda \cdot c) \oplus (\mu \cdot a, \mu \cdot b, \mu \cdot c) \\ &= \lambda \odot (a, b, c) \oplus \mu \odot (a, b, c). \end{aligned}$$

(vii) Distributive 2:

$$\begin{aligned}\lambda \odot [(a, b, c) \oplus (d, e, f)] &= \lambda \odot (a + d, b + e, c + f) \\ &= (\lambda \cdot (a + d), \lambda \cdot (b + e), \lambda \cdot (c + f)) \\ &= (\lambda \cdot a, \lambda \cdot b, \lambda \cdot c) + (\lambda \cdot d, \lambda \cdot e, \lambda \cdot f) \\ &= \lambda \odot (a, b, c) \oplus \lambda \odot (d, e, f).\end{aligned}$$

(viii) Unitary w.r.t.  $\odot$ : Clearly  $1 \odot (a, b, c) = (a, b, c)$ .

**Question:** Consider the map  $d : V \rightarrow V$ ;  $(a, b, c) \mapsto d((a, b, c)) := (b, 2c, 0)$ .  
Is  $d$  linear?

**Solution:** Consider

$$\begin{aligned}d(\lambda \odot (a, b, c) \oplus (d, e, f)) &= d([\lambda \cdot a] + d, [\lambda \cdot b] + e, [\lambda \cdot c] + f) \\ &= ([\lambda \cdot b] + e, 2[\lambda \cdot c] + 2f, 0) \\ &= \lambda \odot (b, 2c, 0) \oplus (e, 2f) && \text{soyesit's linear.} \\ &= \lambda \odot d((a, b, c)) \oplus d((d, e, f)),\end{aligned}$$

**Question:** Show that the map  $d \circ d$  is linear.

**Solution:** This follows immediately from Theorem 3.2.2, however if you want you can show it explicitly (we won't here to save me writing.)

**Question:** Consider the map

$$i : V \rightarrow \mathbb{R}; (a, b, c) \mapsto i((a, b, c)) := a + \frac{1}{2}b + \frac{1}{3}c.$$

Check linearity. Of what set is  $i$  an element?

**Solution:** The linearity calculation is exactly the same method as above, so to save me typing I'm just going to say the answer: yes it is. Being a linear map from  $V$  to  $\mathbb{R}$  it is, by definition, an element of  $V^*$ .

**Question:** Another (multi)linear question. Again I don't want to keep typing out that check so see [the video](#) for this one.

**Question:** Compare the above map  $d : V \rightarrow V$  with the map  $\delta : P \rightarrow P$  from the lecture and construct a bijective linear map  $j : P_2 \rightarrow \mathbb{R}^3$  such that

$$d = j \circ \delta \circ j^{-1}.$$

**Solution:** Recall that  $\delta : P \rightarrow P$  is defined as the derivative, i.e.  $\delta(p) = p'$ . Clearly for this question we want to focus on the map  $\delta : P_2 \rightarrow P_2$ . Recall the definition

$$P_2 := \{p : \mathbb{R} \rightarrow \mathbb{R} \mid p(x) = a + bx + cx^2\},$$

and so we have

$$p'(x) = b + 2cx + 0x^2.$$

We instantly see this resembles the map  $d : V \rightarrow V$  which acts as  $d((a, b, c)) := (b, 2c, 0)$ . We just need to construct the map  $j : P_2 \rightarrow \mathbb{R}^3$ . With a bit of thought it is clear that the answer is simply

$$j(p) := (a, b, c),$$

where  $p = a + bx + cx^2$ . The inverse map  $j^{-1} : \mathbb{R}^3 \rightarrow P_2$  is then simply given by

$$(j^{-1}(a, b, c))(x) := a + bx + cx^2.$$

Direct substitution then gives

$$((j \circ \delta \circ j^{-1})(a, b, c)) := (b, 2c, 0) =: d((a, b, c)).$$

Similar exercises can be done to compare  $i : V \rightarrow \mathbb{R}$  to  $I : P \rightarrow \mathbb{R}$  given in the lecture (and also for the above exercise which I haven't typed here.) See [the video](#) for more details.

### 26.3.2 Indices

Let  $V$  be a  $d$ -dimensional vector space. Consider two maps  $A$  and  $B$ , where

$$A : V^* \times V^* \rightarrow \mathbb{R}$$

$$B : V \times V \rightarrow \mathbb{R}.$$

$V$  has a basis  $e_1, \dots, e_d$  and  $V^*$  has the basis  $\epsilon^1, \dots, \epsilon^d$ .

**Question:** Define the components  $A^{ab}$  of  $A$  and  $B_{ab}$  of  $B$  with respect to the given bases.

**Solution:** We simply uses the duality of the basis elements, namely  $e_i(\epsilon^j) = \delta_i^j = \epsilon^j(e_i)$ . So we have

$$A^{ab} = A(\epsilon^a, \epsilon^b), \quad \text{and} \quad B_{ab} = B(e_a, e_b).$$

**Question:** We define  $A^{[ab]} := \frac{1}{2}(A^{ab} - A^{ba})$ . Show that

$$A^{[ab]} = -A^{[ba]}$$

and also

$$A^{[ab]}B_{ab} = A^{ab}B_{[ab]}.$$

**Solution:** From direct calculation we have

$$A^{[ab]} := \frac{1}{2}(A^{ab} - A^{ba}) = -\frac{1}{2}(A^{ba} - A^{ab}) =: -A^{[ba]}.$$

We also have

$$\begin{aligned} A^{[ab]}B_{ab} &:= \frac{1}{2}(A^{ab} - A^{ba})B_{ab} \\ &= \frac{1}{2}(A^{ab}B_{ab} - A^{ba}B_{ab}) \\ &= \frac{1}{2}(A^{ab}B_{ab} - A^{ab}B_{ba}) \\ &= A^{ab}\frac{1}{2}(B_{ab} - B_{ba}) \\ &=: A^{ab}B_{[ab]}, \end{aligned}$$

where we have used the Einstein summation convention to relabel the dummy indices of the second term to get to the third line.

**Question:** We additionally define  $B_{(ab)} := \frac{1}{2}(B_{ab} + B_{ba})$ . Now, show that

$$B_{(ab)} = B_{(ba)}$$

and again

$$A^{ab}B_{(ab)} = A^{(ab)}B_{ab}.$$

**Solution:** Follows exactly like the previous solution.

**Question:** Using the results from the previous questions, we can easily show

$$A^{[ab]}B_{(ab)} = 0,$$

i.e., the summation (contraction) of symmetric and antisymmetric indices yields zero.

**Solution:** We have

$$\begin{aligned} A^{[ab]}B_{(ab)} &:= \frac{1}{2}(A^{ab}B_{(ab)} - A^{ba}B_{(ab)}) \\ &= \frac{1}{2}(A^{(ab)}B_{ab} - A^{(ba)}B_{ab}) \\ &= \frac{1}{2}(A^{(ab)}B_{ab} - A^{(ab)}B_{ab}) \\ &= 0. \end{aligned}$$

### 26.3.3 Linear Maps As Tensors

**Question:** Given a vector space  $V$  and a linear map  $\phi : V^* \xrightarrow{\sim} V^*$  construct a  $(1, 1)$ -tensor  $T_\phi$ .

**Solution:** We know a  $(1, 1)$ -tensor is a linear map  $T_\phi : V^* \times V \xrightarrow{\sim} \mathbb{R}$ . We also know that  $\psi \in V^*$  is a map  $\psi : V \xrightarrow{\sim} \mathbb{R}$ . So it's clear that we just define

$$T_\phi(\sigma, v) := (\phi(\sigma))(v),$$

where  $\sigma \in V^*$  and  $v \in V$ .

**Question:** Given a  $(1, 1)$ -tensor  $T : V^* \times V \xrightarrow{\sim} \mathbb{R}$ , construct a linear map  $\phi_T : V^* \xrightarrow{\sim} V^*$ .

**Solution:** This is just the previous exercise in reverse. That is we define

$$(\phi_T(\sigma))(v) := T(\sigma, v),$$

which we write as

$$\phi_T(\sigma) := T(\sigma, \cdot).$$

**Question:** Show that

- (a)  $T_{\phi_T} = T$ , and
- (b)  $\phi_{T_\phi} = \phi$ .

**Solution:**

- (a) We have

$$T_{\phi_T}(\sigma, v) := (\phi_T(\sigma))(v) =: T(\sigma, v).$$

- (b) We have

$$(\phi_{T_\phi}(\sigma))(v) := T_\phi(\sigma, v) =: (\phi(\sigma))(v),$$

and so  $\phi_{T_\phi} = \phi$ .

**Question:** Conclude that we can consider a linear map  $\phi : V^* \xrightarrow{\sim} V^*$  as a  $(1, 1)$ -tensor.

**Solution:** The above exercises have just shown us that there is a unique link between  $T$  and  $\phi$  and so we can think of them as being isomorphic as maps and as such can be viewed as the same object.

## 26.4 Differential Manifolds

### 26.4.1 Restricting The Atlas

Let  $(\mathbb{R}, \mathcal{O}_{\text{standard}})$  be a topological space. Let it be further equipped with an atlas  $\mathcal{A} = \{(\mathbb{R}, x), (\mathbb{R}, y)\}$  where  $x : \mathbb{R} \rightarrow \mathbb{R}; a \mapsto x(a) = a$  and  $y : \mathbb{R} \rightarrow \mathbb{R}; a \mapsto y(a) = a^3$ .

**Question:** Construct the chart transition map  $y \circ x^{-1} : \mathbb{R} \rightarrow \mathbb{R}$  and give its differentiability class.

**Solution:** We have  $x^{-1} : \mathbb{R} \rightarrow \mathbb{R}; a \mapsto x^{-1}(a) = a$ , and so  $y \circ x^{-1} : \mathbb{R} \rightarrow \mathbb{R}; a \mapsto a^3$ . This is  $C^\infty(\mathbb{R})$ .

**Question:** Also construct the chart transition map  $x \circ y^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ . Is  $(\mathbb{R}, \mathcal{O}_{\text{standard}}, \mathcal{A})$  a differentiable manifold?

**Solution:** We have  $x \circ y^{-1} : \mathbb{R} \rightarrow \mathbb{R}; a \mapsto \sqrt[3]{a}$ , but this function is not even  $C^1(\mathbb{R} \rightarrow \mathbb{R})$  as the first derivative is  $\frac{1}{3}a^{-2/3}$ , which blows up as  $a \rightarrow 0$ . So no it is not a differentiable manifold.

**Question:** Restrict the atlas  $\mathcal{A}$  to an atlas  $\tilde{\mathcal{A}}$  in order to make  $(\mathbb{R}, \mathcal{O}_{\text{standard}}, \tilde{\mathcal{A}})$  a smooth manifold.

**Solution:** The problem above came from the chart transition maps. Both chart maps themselves are  $C^\infty(\mathbb{R} \rightarrow \mathbb{R})$  and their domain is the whole manifold (namely  $\mathbb{R}$ ) and so if we just remove one of the two charts we get a smooth manifold.

### 26.4.2 Soft Squares on $\mathbb{R} \times \mathbb{R}$

Let  $\mathcal{M} = \mathbb{R} \times \mathbb{R}$  equipped with the soft square topology  $\mathcal{O}_{\text{ssq}}$  and an atlas  $\mathcal{A} = \{(U_n, x_n)\}$ , where  $U_n = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid |x| < n, |y| < n, n \in \mathbb{N}^+\}$  and

$$x_n : U_n \rightarrow x_n(U_n) \subset \mathbb{R}^2; (x, y) \mapsto x_n((x, y)) := \left( \frac{x+y}{2n}, \frac{x-y}{2n} \right).$$

**Question:** Recall the definition of a chart and show that the  $(U_n, x_n)$  are indeed charts.

**Solution:** For  $(U_n, x_n)$  to be charts, we need to show that  $U_n \in \mathcal{O}_{\text{ssq}}$  and that the  $x_n$ s are homeomorphisms (i.e. inevitable and continuous).

The soft square topology is rather self explanatory, it is the set of squares around the origin without the boundary. This is exactly the definition of the  $U_n$ s where the sides of consecutive  $U_n$  increase by  $2n$ . So we have  $U_n \in \mathcal{O}_{\text{ssq}}$ .

Next we need to show that the chart maps are homeomorphisms. It is clear that  $x_n$  is continuous w.r.t.  $\mathcal{O}_{\text{ssq}}$  and  $\mathcal{O}_{\text{standard}}$ . So we just need to check that  $x_n^{-1}$  exists and is continuous. We have

$$x_n^{-1}(a, b) = (n(a+b), n(a-b)),$$

which is again clearly continuous.

Therefore we know that the  $(U_n, x_n)$  are indeed charts.

**Question:** Show that  $\mathcal{A}$  is a  $C^k$ -atlas by explicitly constructing the chart transition maps. What is  $k$ ?

**Solution:** The chart transition maps are given by

$$\begin{aligned} x_m \circ x_n^{-1} : x_n(U_n \cap U_m) &\rightarrow x_m(U_n \cap U_m) \\ (a, b) &\mapsto \left( \frac{na}{m}, \frac{nb}{m} \right) = \frac{n}{m}(a, b). \end{aligned}$$

This tells us that  $x_m \circ x_n^{-1} = \frac{n}{m} \mathbb{1}_{\mathbb{R}^2}$ , which together with  $m \neq 0$  tells us  $k = \infty$ .

**Question:** Construct at least one other chart that would lie in the maximal extension of  $\mathcal{A}$  and prove that it does.

**Solution:** There are many, but we could take the map  $\tilde{x} : U_5 \rightarrow \widetilde{U}_5 \subset \mathbb{R}^2$ ;  $(a, b) \mapsto (a, b)$ , i.e. it is the identity map restricted to  $U_5$ . It is clear that this will be  $C^\infty$  compatible with any overlapping charts and so it lies in the atlas.

### 26.4.3 Undergraduate Multi-Dimensional Analysis

**Question:** There is a question about calculating partial derivatives. I am not including it here as its fairly straight forward.

### 26.4.4 Differentiability On A Manifold

**Question:** There is a question about drawing a diagram to show a bunch of different spaces and maps. It will take a while to draw in Tikz and there are plenty examples in the notes themselves, so I haven't included it here. This is followed by a calculation of a derivative of a map, it is worth seeing this exercise, so please [the video](#).

## 26.5 Tangent Spaces

### 26.5.1 Virtuoso Use Of The Symbol $\left(\frac{\partial}{\partial x^i}\right)_p$

**Question:** Show that, for overlapping charts  $(U, x)$  and  $(V, y)$ , one has

$$\left(\frac{\partial x^a}{\partial y^m}\right)_p \left(\frac{\partial y^m}{\partial x^b}\right)_p = \delta_b^a$$

for any  $p \in U \cap V$ .



**Solution:** We have

$$\begin{aligned}
 \delta_b^a &= \left( \frac{\partial}{\partial x^b} \right)_p (x^a) \\
 &:= \partial_b (x^a \circ x^{-1})(x(p)) \\
 &= \partial_b (x^a \circ (y^{-1} \circ y) \circ x^{-1})(x(p)) \\
 &= \partial_b ((x^a \circ y^{-1}) \circ (y \circ x^{-1}))(x(p)) \\
 &= \partial_b (y^m \circ x^{-1})(x(p)) \cdot \partial_m (x^a \circ y^{-1})(y(p)) \\
 &=: \left( \frac{\partial y^m}{\partial x^b} \right)_p \left( \frac{\partial x^a}{\partial y^m} \right)_p,
 \end{aligned}$$

where we have inserted the identity, used the associativity of the composition of maps, and the multidimensional chain rule along with  $(y \circ x^{-1})(x(p)) = y(p)$ .

**Question:** After inserting  $y^{-1} \circ y$ , where  $y$  is another chart map on the same chart domain  $U$ , at the appropriate position in the definition of the left hand side of

$$\left( \frac{\partial f}{\partial x^i} \right)_p = \left( \frac{\partial y^m}{\partial x^i} \right)_p \left( \frac{\partial f}{\partial y^m} \right)_p,$$

use the multidimensional chain rule to show that it equals the right-hand side.

**Solution:** This follows in a similar manner to the previous question and so I won't type it here.

**Question:** Do the  $\dim \mathcal{M}$  many quantities defined by the left-hand side of the above expression constitute the components of a tensor? If so, what is the valence and rank of the tensor?

**Solution:** The above expression is clearly of the form of

$$T_{(x)i}(p) = \left( \frac{\partial y^m}{\partial x^i} \right)_p T_{(y)m}(p),$$

which is the transformation law for the components of a  $(0, 1)$ -tensor. So the answer is "yes" and the valence is  $(0, 1)$  and the rank is 1.

### 26.5.2 Transformation Of Vector Components

Let the topological space  $(\mathbb{R}^2, \mathcal{O}_{st.})$  be equipped with the atlas  $\mathcal{A} = \{(\mathbb{R}^2, x), (\mathbb{R}^2, y)\}$ , where

$$x : (a, b) \mapsto (a, b) \quad \text{and} \quad y : (a, b) \mapsto (a, b + a^3).$$

**Question:** Calculate the objects  $\left( \frac{\partial x^i}{\partial y^j} \right)_p!$

**Solution:** We have  $(x \circ y^{-1})((u, v)) = (u, v - u^3)$ , and so direct calculation gives

$$\left(\frac{\partial x^1}{\partial y^1}\right)_p := \partial_1(x^1 \circ y^{-1})(y(p)) = 1,$$

where we have used  $\partial_1(x^1 \circ y^{-1})((u, v)) = 1$ , irrespective of the value  $u$  takes.

Next we have,

$$\left(\frac{\partial x^2}{\partial y^1}\right)_p := \partial_1(x^2 \circ y^{-1})(y(p)) = \partial_1(x^2 \circ y^{-1})((a, b + a^3)) = -3a^2,$$

where we have used  $\partial_2(x^1 \circ y^{-1})((u, v)) = -2u^2$  along with  $p = (a, b)$ .

Similar calculations give

$$\left(\frac{\partial x^1}{\partial y^2}\right)_p = 0, \quad \text{and} \quad \left(\frac{\partial x^2}{\partial y^2}\right)_p = 1$$

(Reworded slightly to save typing) Recall that the components of the velocity to a curve  $\gamma$  in a chart  $(U, x)$  at point  $p = \gamma(\lambda_0)$  are given by

$$\dot{\gamma}_x^i(\lambda_0) := (x \circ \gamma)^i(\lambda_0).$$

Now consider the curve

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^2; \quad \lambda \mapsto (\lambda, -\lambda).$$

**Question:** Calculate the components  $\dot{\gamma}_x^i(\lambda_0)$  and  $\dot{\gamma}_y^i(\lambda_0)$ !

**Solution:** We have  $(x \circ \gamma)(\lambda) = (\lambda, -\lambda)$  and  $(y \circ \gamma)(\lambda) = (\lambda, -\lambda + \lambda^3)$ , so we have

$$\begin{aligned} \dot{\gamma}_x^1(\lambda_0) &:= (x \circ \gamma)^1(\lambda_0) = 1, \\ \dot{\gamma}_x^2(\lambda_0) &:= (x \circ \gamma)^2(\lambda_0) = -1, \\ \dot{\gamma}_y^1(\lambda_0) &:= (y \circ \gamma)^1(\lambda_0) = 1, \\ \dot{\gamma}_y^2(\lambda_0) &:= (y \circ \gamma)^2(\lambda_0) = -1 + 3\lambda_0^2. \end{aligned}$$

**Question:** With the above results in mind, how could you have obtained the components of  $\dot{\gamma}_x^i(\lambda_0)$  from the  $\dot{\gamma}_y^i(\lambda_0)$ ?

**Solution:** The answer is clearly to just use the transformation property

$$\dot{\gamma}_x^i(\lambda_0) = \left(\frac{\partial x^i}{\partial y^m}\right)_p \dot{\gamma}_y^m(\lambda_0).$$

That is, we have

$$\begin{aligned} \dot{\gamma}_x^1(\lambda_0) &= \left(\frac{\partial x^1}{\partial y^1}\right)_p \dot{\gamma}_y^1(\lambda_0) + \left(\frac{\partial x^1}{\partial y^2}\right)_p \dot{\gamma}_y^2(\lambda_0) \\ &= 1 \cdot 1 + 0 \cdot (-1 + 3\lambda_0^2) \\ &= 1, \end{aligned}$$

and

$$\begin{aligned}\dot{\gamma}_x^2(\lambda_0) &= \left(\frac{\partial x^2}{\partial y^1}\right)_p \dot{\gamma}_y^1(\lambda_0) + \left(\frac{\partial x^2}{\partial y^2}\right)_p \dot{\gamma}_y^2(\lambda_0) \\ &= -3\lambda_0^2 \cdot 1 + 1 \cdot (-1 + 3\lambda_0^2) \\ &= -1.\end{aligned}$$

### 26.5.3 The Gradient

Given a function  $f$  on a manifold  $\mathcal{M}$ , the level sets of  $f$  for a constant  $c \in \mathbb{R}$  are defined as

$$N_c(f) := \{p \in \mathcal{M} \mid f(p) = c\}.$$

**Question:** Formulate the condition for a curve  $\gamma : \mathbb{R} \rightarrow \mathcal{M}$  to take values solely in one of the level sets of a function  $f$ !

**Solution:** Clearly we just want  $(f \circ \gamma)(\lambda) = c$  for all  $\lambda \in \mathbb{R}$ . We can formulate this as  $f \circ \gamma$  being equivalent to the constant function, which simply obeys the above property.

**Question:** Now show that the gradient of the function annihilates the velocity vector  $v_{\gamma,p}$  for any such  $\gamma$  through  $p$  in  $N_c(f)$ . In other words, show that

$$(df)_p(v_{\gamma,p}) = 0.$$

**Solution:** We have

$$(df)_p(v_{\gamma,p}) := v_{\gamma,p}(f) := (f \circ \gamma)'(\lambda_0) = 0,$$

as the derivative of a constant vanishes.

### 26.5.4 Is There A Well-Defined Sum Of Curves?

Let the topological manifold  $(\mathbb{R}^2, \mathcal{O}_{st.})$  be equipped with the atlas  $\mathcal{A} = \{(\mathbb{R}^2, x), (\mathbb{R}^2, y)\}$  where

$$x : (a, b) \mapsto (a, b) \quad \text{and} \quad y : (a, b) \mapsto (a, b \cdot e^a).$$

**Question:** Is  $\mathcal{A}$  a  $C^\infty$ -atlas?

**Solution:** We need to construct the chart transition maps

$$\begin{aligned}(x \circ y^{-1})(u, v) &= x((u, v \cdot e^{-u})) = (u, v \cdot e^{-u}) \\ (y \circ x^{-1})(u, v) &= y((u, v)) = (u, v \cdot e^u).\end{aligned}$$

These are both infinitely times continuously differentiable and so the answer is "yes".

**Question:** On our  $\mathbb{R}^2$  above, consider two curves  $\gamma, \delta : \mathbb{R} \rightarrow \mathbb{R}^2$  given by

$$\gamma : \lambda \mapsto (\lambda, 1), \quad \text{and} \quad \delta : \lambda \mapsto (1, \lambda).$$

Without referring to any chart, can you give the sum  $\gamma + \delta$  of these curves?

**Solution:** The answer is "no" because our manifold is just the set  $\mathbb{R}^2$  (with a topology) and so carries no vector space structure so we cannot talk about the addition on  $\mathbb{R}^2$ .

**Question:** Calculate the representatives of both curves with respect to both charts. Illustrate the results. Where do the curves in the charts intersect?

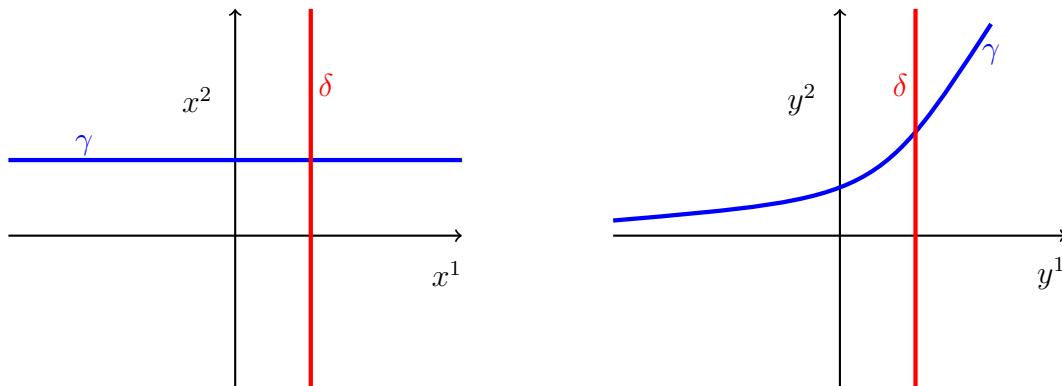
**Solution:** First consider the  $x$  chart. We have

$$(x \circ \gamma)(\lambda) = (\lambda, 1), \quad \text{and} \quad (x \circ \delta)(\lambda) = (1, \lambda).$$

In the  $y$  chart we have

$$(y \circ \gamma)(\lambda) = (\lambda, e^\lambda), \quad \text{and} \quad (y \circ \delta)(\lambda) = (1, \lambda \cdot e).$$

The illustrations are as follows



The intersection points are  $(1, 1)$  in the  $x$  chart and  $(1, e)$  in the  $y$  chart. These both correspond to  $\lambda = 1$ , which it must from the definition of the curves.

(Reworded to save typing) **Question:** Use the formula from the lectures

$$\sigma_x : \mathbb{R} \rightarrow U; \quad \lambda \mapsto x^{-1}((x \circ \gamma)(\lambda + \lambda_0) + (x \circ \delta)(\lambda + \lambda_1) - (x \circ \gamma)(\lambda_0)),$$

where  $\gamma(\lambda_0) = \delta(\lambda_1)$ , to find the sum  $\gamma + \delta$ . Also do the calculation for  $\sigma_y(\lambda)$ .

**Solution:** Using the previous question, direct calculation gives: for the  $x$  chart

$$\begin{aligned} \sigma_x(\lambda) &= x^{-1}((\lambda + 1, 1) + (1, \lambda + 1) - (1, 1)) \\ &= x^{-1}(\lambda + 1, \lambda + 1) \\ &= (\lambda + 1, \lambda + 1). \end{aligned}$$

The calculation in the  $y$  chart gives

$$\sigma_y(\lambda) = (\lambda + 1, 1 + \lambda \cdot e^{-\lambda}).$$

**Question:** Show that – despite the above results – the velocity of  $\sigma_x$  and the velocity  $\sigma_y$  are equal at the intersection point.

**Solution:** The intersection point is  $(1, 1)$  in  $x$  and so we require  $\lambda = 0$ . Using this, we have

$$\dot{\sigma}_{(x)x}^1 = 1 \quad \text{and} \quad \dot{\sigma}_{(x)x}^2 = 1,$$

as both are just the derivative w.r.t.  $\lambda$  of  $\lambda + 1$ .

In the  $y$  chart we have (note we use the  $x$  chart in to find these components in order to compare it to the above ones)

$$\dot{\sigma}_{(x)y}^1 = 1 \quad \text{and} \quad \dot{\sigma}_{(x)y}^2 = 1 \cdot e^0 - 0 \cdot e^0 = 1.$$

So we have

$$v_{\sigma_x, (1,1)} = \dot{\sigma}_{(x)x}^i \left( \frac{\partial}{\partial x^i} \right)_p = \dot{\sigma}_{(x)y}^i \left( \frac{\partial}{\partial x^i} \right)_p = v_{\sigma_y, (1,1)}.$$

## 26.6 Fields

### 26.6.1 Vector Fields For Practitioners

**Question:** Let  $(U, x)$  be a chart of a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$ . Explain why the map

$$\frac{\partial}{\partial x^i} : U \rightarrow TU; \quad p \mapsto \left( \frac{\partial}{\partial x^i} \right)_p$$

is a vector field on  $U$ .

**Solution:** We need to check that it is a section. That is  $\pi \circ \frac{\partial}{\partial x^i} = \mathbb{1}_{\mathcal{M}}$ . This is clearly true as  $\pi : T_p \mathcal{M} \rightarrow \mathcal{M}$  is defined as

$$\pi : \left( \frac{\partial}{\partial x^i} \right)_p \mapsto p.$$

We now need to check if this section is smooth. If we denote the chart map on  $TU$  as  $\xi_x$ , that is we need to check that

$$\xi_x \circ \frac{\partial}{\partial x^i} \circ x^{-1} : x(U) \rightarrow \xi(TU)$$

is smooth. Recalling the definition of  $\xi_x$ , and using the fact that the only non-vanishing component of  $\frac{\partial}{\partial x^i}$  is the  $i^{\text{th}}$  entry, we have

$$\left( \xi_x \circ \frac{\partial}{\partial x^i} \circ x^{-1} \right) (\alpha^1, \dots, \alpha^d) = (\alpha^1, \dots, \alpha^d, 0, \dots, 1, \dots, 0),$$

where the 1 appears at the  $(d + i)^{\text{th}}$  entry. This is clearly smooth (w.r.t. the standard topologies on  $\mathbb{R}^d$  and  $\mathbb{R}^{2d}$ ), and so  $\frac{\partial}{\partial x^i}$  is a vector field on  $U$ .

### 26.6.2 The Cotangent Bundle $T^*\mathcal{M} \xrightarrow{\pi} \mathcal{M}$

We consider the *cotangent bundle* total space  $T^*\mathcal{M}$  as the disjoint union

$$T^*\mathcal{M} := \dot{\bigcup}_{p \in \mathcal{M}} T_p^*\mathcal{M}$$

of all cotangent spaces and define the bundle projection map

$$\begin{aligned} \pi : T^*\mathcal{M} &\rightarrow \mathcal{M} \\ \omega &\mapsto \text{the unique } p \text{ with } \omega \in T_p^*\mathcal{M}. \end{aligned}$$

**Question:** Show that

$$\mathcal{O}_{T^*\mathcal{M}} := \{\text{preim}_\pi(U) \mid U \in \mathcal{O}_\mathcal{M}\}$$

defines a topology on  $T^*\mathcal{M}$ .

**Solution:** We check the three conditions for a topology in the order given in the definition.

(i) We have

$$\text{preim}_\pi(\emptyset) = \emptyset, \quad \text{and} \quad \text{preim}_\pi(\mathcal{M}) = T^*\mathcal{M},$$

and so  $\emptyset, T^*\mathcal{M} \in \mathcal{O}_{T^*\mathcal{M}}$ .

(ii) This just follows from properties of the preimage, namely

$$\text{preim}_f(U \cap V) = \text{preim}_f(U) \cap \text{preim}_f(V).$$

(iii) This just follows from another property of the preimage, namely

$$\bigcup_i \text{preim}_f(U_i) = \text{preim}_f\left(\bigcup_i U_i\right).$$

The rest of this tutorial is basically the same calculations as the ones in the lecture. Specifically find the components of  $\xi_x^*$ , its inverse and showing that the chart transition maps are smooth.

## 26.7 Connections

### 26.7.1 Practical Rules For How $\nabla$ Acts

**Question:** What is the result of the following applications of a torsion free covariant derivative?

- There are some others in the video, but they are basically covered in the lectures so I won't repeat them here.
- $(\nabla_{[m}A)_{n]}$ ,
- $(\nabla_{[m}\omega)_{nr]}$ .

**Solution:** We have

$$\begin{aligned} (\nabla_{[m}A)_{n]} &:= \frac{1}{2} [(\nabla_m A)_n - (\nabla_n A)_m] \\ &= \frac{1}{2} (A_{n,m} - \Gamma^r{}_{nm} A_r - A_{m,n} + \Gamma^r{}_{mn} A_r) \\ &= A_{[n,m]} + \Gamma^r{}_{[mn]} \\ &= \frac{1}{2} F_{mn}, \end{aligned}$$

where we have used the fact that  $\nabla$  is torsion free and so  $\Gamma^a{}_{[bc]} = 0$ .

Next, we have

$$(\nabla_{[m}\omega)_{nr]} = \frac{1}{3!} [(\nabla_m \omega)_{nr} - (\nabla_m \omega)_{rn} + (\nabla_r \omega)_{mn} - (\nabla_r \omega)_{nm} + (\nabla_n \omega)_{rm} - (\nabla_n \omega)_{mr}].$$

If we expand this out we will see that the  $\Gamma$ s all cancel, e.g.

$$\begin{aligned} (\nabla_m \omega)_{nr} - (\nabla_n \omega)_{mr} &= \omega_{nr,m} - \Gamma^s{}_{nm} \omega_{sr} - \Gamma^s{}_{rm} \omega_{ns} - \omega_{mr,n} + \Gamma^s{}_{mn} \omega_{sr} + \Gamma^s{}_{rn} \omega_{ms} \\ &= \omega_{nr,m} - \omega_{mr,n} + 2\Gamma^s{}_{[mn]} \omega_{sr} + 2\Gamma^s{}_{r[n} \omega_{m]s} \\ &= \omega_{nr,m} - \omega_{mr,n} + 2\Gamma^s{}_{r[n} \omega_{m]s}, \end{aligned}$$

and the other  $\Gamma$  will cancel with another term's expansion. We are therefore left with

$$(\nabla_{[m}\omega)_{nr]} = \omega_{[nr,m]}.$$

### 26.7.2 Connection Coefficients

There is a question about stating the chart transformation laws of the  $\Gamma$ s and stating which class of transformations make the  $\Gamma$ s look like tensors. We have discussed this in the notes already, but I have put this box here to remind readers to re-read Remark 7.3.1 as it's an important point often missed.

**Question:** Let  $(\mathcal{M}, \mathcal{O}, \mathcal{A}, \nabla)$  be the flat plane. Consider two charts that both cover the upper half plane, that is all the points  $(a, b) \in \mathbb{R}^2$  with  $b > 0$ , one representing the familiar Cartesian coordinates and the other the familiar polar coordinates on there. We already know the chart transition map from Cartesian to polar coordinates is given by

$$y \circ x^{-1}(a, b) = \left( \sqrt{a^2 + b^2}, \arccos \left( \frac{a}{\sqrt{a^2 + b^2}} \right) \right),$$

while the inverse transition map from polar to Cartesian is given by

$$x \circ y^{-1}(r, \varphi) = (r \cos \varphi, r \sin \varphi), \quad \text{for } r \in \mathbb{R}^+ \text{ and } \varphi \in (0, \pi).$$

Starting from the assumption of vanishing connection coefficient functions in the Cartesian chart, calculate the connection coefficient functions  $\Gamma_{(y)bc}^a$  w.r.t. the polar chart!

**Solution:** Recall that the transformation law is

$$\Gamma_{(y)bc}^a = \frac{\partial y^a}{\partial x^k} \frac{\partial^2 x^k}{\partial y^b \partial y^c} + \frac{\partial y^a}{\partial x^k} \frac{\partial x^n}{\partial y^b} \frac{\partial x^m}{\partial y^c} \Gamma_{(x)mn}^k.$$

The assumption is  $\Gamma_{(x)sp}^q = 0$ , and so we just need to find the first term. We have

$$\left( \frac{\partial x^k}{\partial y^c} \right)_{y^{-1}(r, \varphi)} := \partial_c (x^k \circ y^{-1})(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}_c^k.$$

Now, using the first question in the tangent spaces tutorial above, we have

$$\begin{aligned} \left( \frac{\partial y^a}{\partial x^k} \right)_p &= \left( \frac{\partial x^k}{\partial y^a} \right)^{-1} \\ &= \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}_a^k \\ &= \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{1}{r} \sin \varphi & \frac{1}{r} \cos \varphi \end{pmatrix}_a^k, \end{aligned}$$



where we have used the fact that the determinant of the first matrix is  $r$ . We can also calculate

$$\begin{aligned} \left( \frac{\partial^2 x^1}{\partial y^1 \partial y^1} \right)_{y^{-1}(r, \varphi)} &= 0 \\ \left( \frac{\partial^2 x^1}{\partial y^2 \partial y^1} \right)_{y^{-1}(r, \varphi)} &= -\sin \varphi \\ \left( \frac{\partial^2 x^1}{\partial y^1 \partial y^2} \right)_{y^{-1}(r, \varphi)} &= -\sin \varphi \\ \left( \frac{\partial^2 x^1}{\partial y^2 \partial y^2} \right)_{y^{-1}(r, \varphi)} &= -r \cos \varphi \\ \left( \frac{\partial^2 x^2}{\partial y^1 \partial y^1} \right)_{y^{-1}(r, \varphi)} &= 0 \\ \left( \frac{\partial^2 x^2}{\partial y^2 \partial y^1} \right)_{y^{-1}(r, \varphi)} &= \cos \varphi \\ \left( \frac{\partial^2 x^2}{\partial y^2 \partial y^2} \right)_{y^{-1}(r, \varphi)} &= -r \sin \varphi. \end{aligned}$$

We then simply have to pick the relevant expressions to find the  $\Gamma_{(y)bc}^a$ s. For example

$$\Gamma_{(y)11}^1 = \frac{\partial y^1}{\partial x^1} \frac{\partial^2 x^1}{\partial y^1 \partial y^1} + \frac{\partial y^1}{\partial x^2} \frac{\partial^2 x^2}{\partial y^1 \partial y^1} = 0,$$

and

$$\begin{aligned} \Gamma_{(y)22}^1 &= \frac{\partial y^1}{\partial x^1} \frac{\partial^2 x^1}{\partial y^2 \partial y^2} + \frac{\partial y^1}{\partial x^2} \frac{\partial^2 x^2}{\partial y^2 \partial y^2} \\ &= \cos \varphi (-r \cos \varphi) + \sin \varphi (-r \sin \varphi) \\ &= -r(\cos^2 \varphi + \sin^2 \varphi) \\ &= -r. \end{aligned}$$

## 26.8 Parallel Transport and Curvature

### 26.8.1 Where Connection Coefficients Appear

**Question:** Determine the coefficients of the Riemann tensor with respect to a chart  $(U, x)$  in terms of the connection coefficient functions.

**Solution:** Recall the definition

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

We have already seen<sup>1</sup> that Riem is  $C^\infty$ -linear in all its entries and so we can just consider basis elements. That is, we can set (using the notation  $\partial_j := \frac{\partial}{\partial x^j}$ )

$$\omega = dx^i, \quad Z = \partial_j, \quad X = \partial_k, \quad \text{and} \quad Y = \partial_m.$$

So we have

$$\begin{aligned} \text{Riem}(dx^i, \partial_j, \partial_k, \partial_m) &:= dx^i : \left( \nabla_k \nabla_m \partial_j - \nabla_m \nabla_k \partial_j - \nabla_{[\partial_k, \partial_m]} \partial_j \right) \\ &= dx^i : \left[ \nabla_k (\Gamma_{(x)jm}^r \partial_r) - \nabla_m (\Gamma_{(x)jk}^r \partial_r) \right] \\ &= dx^i : \left[ \Gamma_{(x)jm,k}^r \partial_r + \Gamma_{(x)jm}^r \Gamma_{(x)rk}^s \partial_s - \Gamma_{(x)jk,m}^r \partial_r - \Gamma_{(x)jk}^r \Gamma_{(x)rm}^s \partial_s \right] \\ \text{Riem}^i{}_{jkm} &= \Gamma_{(x)jm,k}^i - \Gamma_{(x)jk,m}^i + \Gamma_{(x)jm}^r \Gamma_{(x)rk}^i - \Gamma_{(x)jk}^r \Gamma_{(x)rm}^i. \end{aligned}$$

We can see the antisymmetry in the last two entries immediately from the above. That is  $\text{Riem}^i{}_{jkm} = -\text{Riem}^i{}_{jmk}$ .

**Question:** Does a one-dimensional manifold with connection have curvature? Why?

**Solution:** No, as if it is one-dimensional we only have one  $\Gamma$ , namely  $\Gamma^1{}_{11}$ , and if we put this into the above definition, all the only component  $\text{Riem}^1{}_{111}$  vanishes and so there is no curvature.

Geometrically this makes sense if we think about embedding the one-dimensional manifold into a higher dimensional space. We can always ‘pull’ the one-dimensional manifold straight and thus show that it has no intrinsic curvature.

### 26.8.2 The Round Sphere

There is a question about finding the components of Riem for given  $\Gamma$ s. I am not going to write it here to save myself time, but I recommend at least watching the video (available [here](#), for some reason this video is not on YouTube) for a worked calculation.

### 26.8.3 How Not To Define Parallel Transport

**Question:** Hänchen defines two vectors  $X \in T_p \mathcal{M}$  and  $Y \in T_q \mathcal{M}$  as parallel if

$$X_{(x)}^i = Y_{(x)}^i$$

with respect to some chart  $(U, x)$  whose domain  $U$  contains both  $p$  and  $q$ . Prove that this notion of parallelity is ill-defined!

<sup>1</sup>Or more correctly, you have shown as an exercise in Lecture 8.

**Solution:** Consider the transformation to another chart with the same domain  $(U, y)$ ,

$$X_{(y)}^i = \left( \frac{\partial y^i}{\partial x^j} \right)_p X_{(x)}^j = \left( \frac{\partial y^i}{\partial x^j} \right)_p Y_{(x)}^j = \left( \frac{\partial y^i}{\partial x^j} \right)_p \left( \frac{\partial x^j}{\partial y^k} \right)_q Y_{(y)}^k,$$

but because  $p \neq q$  we cannot, in general, use

$$\left( \frac{\partial y^i}{\partial x^j} \right)_p \left( \frac{\partial x^j}{\partial y^k} \right)_q = \delta_k^i,$$

and so

$$X_{(y)}^i \neq Y_{(y)}^i.$$

There is then a diagram to show how the above definition of parallelity fails for parallel (as defined in the lecture) vectors around a circle when considered in the Cartesian chart and the polar chart.

To save myself time I have not drawn the diagrams here, but if you can't picture the images in your head, again go see the video!

## 26.10 Metric Manifolds

### 26.10.1 Recognising & Dealing With Different Signatures

**Question:** (reworded) State the possible signatures of a metric  $g$  for the following set of  $g$ -null vectors:

- (i) A cone through the origin,
- (ii) A point at the origin,
- (iii) A straight line through the origin, and
- (iv) A plane through the origin,

where the origin is the origin of the vector space, namely the point  $p \in \mathcal{M}$  that we are tangent to.

**Solution:** W.l.o.g. let's assume the vector space is 3-dimensional and introduce a basis  $\{e_1, e_2, e_3\}$ . The relevant equations giving the correct surfaces are

- (i)  $g(X, X) = -(X^1)^2 + (X^2)^2 + (X^3)^2 = 0$ , so we have  $(-, +, +)$ , or equivalently  $(+, -, -)$ .
- (ii)  $g(X, X) = 0$  only for the zero vector, and so we need  $(+, +, +)$  or  $(-, -, -)$ .
- (iii) From the above case, we just need to have one of the entries to be 0, i.e.  $(0, +, +)/(0, -, -)$ , as then all vectors with only an  $e^1$  component are null.
- (iv) Extending the previous case, we have  $(0, 0, +)/(0, 0, -)$ .

### 26.10.2 Levi-Civita Connection

**Question:** Expand in terms of connection coefficient functions

- (i)  $(\nabla_a g)_{bc}$ ,
- (ii)  $(\nabla_b g)_{ca}$ ,
- (iii)  $(\nabla_c g)_{ab}$ .

**Solution:** We just do the first one, as the other two are obtained by simply relabeling the indices. We have

$$(\nabla_a g)_{bc} = g_{bc,a} - \Gamma^m_{ba} g_{mc} - \Gamma^m_{ca} g_{bm}.$$

**Question:** By adding and/or subtracting (i), (ii) and (iii) above in a clever way, obtain

$$\Gamma^a_{bc} = \frac{1}{2}(g^{-1})^{am}(g_{mc,b} + g_{mb,c} - g_{bc,m})$$

and conclude that  $\nabla g = 0$  and  $T = 0$  (torsion) uniquely determine the connection coefficient functions in terms of the metric.

**Solution:** Consider (i)+(ii)-(iii),

$$g_{bc,a} - \Gamma^m_{ba} g_{mc} - \Gamma^m_{ca} g_{bm} + g_{ca,b} - \Gamma^m_{cb} g_{ma} - \Gamma^m_{ab} g_{cm} - g_{ab,c} + \Gamma^m_{ac} g_{mb} + \Gamma^m_{bc} g_{am}.$$

Now if we consider a metric compatible connection and vanishing torsion, we have that the above vanishes (as each of (i), (ii) and (iii) vanish themselves) and that the  $\Gamma$ s are symmetric in the lower two indices. Also using the fact that the metric is symmetric, we have

$$0 = g_{bc,a} + g_{ac,b} - g_{ab,c} - 2\Gamma^m_{ba} g_{mc},$$

which after rearranging and using  $g_{mc} g^{-1}{}^{cn} = \delta_m^n$  we have

$$\Gamma^n_{ba} = \frac{1}{2}(g^{-1})^{cn}(g_{bc,a} + g_{ac,b} - g_{ab,c}),$$

which, relabelling  $n \rightarrow a \rightarrow c \rightarrow m$  gives

$$\Gamma^a_{bc} = \frac{1}{2}(g^{-1})^{ma}(g_{bm,c} + g_{cm,b} - g_{cb,m})$$

which is the result (when you use the symmetries). So we see the connection coefficient functions are uniquely determined by the metric components given the above conditions.

### 26.10.3 Massaging The Length Functional

**Question:** Let  $\gamma : (0, 1) \rightarrow \mathcal{M}$  be a smooth curve on a smooth manifold  $(\mathcal{M}, \mathcal{O}, \mathcal{A})$ . Now consider a second curve  $\tilde{\gamma} : (0, 1) \rightarrow \mathcal{M}$  defined by

$$\tilde{\gamma}(\lambda) = \gamma(\sigma(\lambda)),$$

where  $\sigma : (0, 1) \rightarrow (0, 1)$  is an increasing bijective smooth function. Show that the length of both curves is the same:

$$L[\tilde{\gamma}] = L[\gamma].$$

**Solution:** Using

$$\mathcal{L}[\gamma] = \sqrt{g(v_\gamma, v_\gamma)}, \quad \text{and} \quad \dot{\tilde{\gamma}}^a(\lambda) = (x^a \circ \gamma)'(\lambda),$$

and introducing the notation

$$\tilde{\lambda} := \sigma(\lambda), \quad \implies \quad \tilde{\gamma}(\lambda) = \gamma(\tilde{\lambda}),$$

we have

$$\begin{aligned} L[\tilde{\gamma}] &:= \int_0^1 d\lambda \sqrt{g(v_{\tilde{\gamma}}, v_{\tilde{\gamma}})} \\ &= \int_0^1 d\lambda \sqrt{g_{ab}(\tilde{\gamma}(\lambda)) \cdot \dot{\tilde{\gamma}}^a(\lambda) \cdot \dot{\tilde{\gamma}}^b(\lambda)} \\ &= \int_0^1 d\lambda \sqrt{g_{ab}(\gamma(\tilde{\lambda})) \cdot (x^a \circ \gamma \circ \sigma)'(\lambda) \cdot (x^b \circ \gamma \circ \sigma)'(\lambda)} \\ &= \int_0^1 d\lambda \sqrt{g_{ab}(\gamma(\tilde{\lambda})) \cdot (x^a \circ \gamma)'(\tilde{\lambda}) \cdot \dot{\sigma}(\lambda) \cdot (x^b \circ \gamma)'(\tilde{\lambda}) \cdot \dot{\sigma}(\lambda)} \\ &= \int_0^1 d\lambda \dot{\sigma}(\lambda) \sqrt{g_{ab}(\gamma(\tilde{\lambda})) \cdot \dot{\gamma}^a(\tilde{\lambda}) \cdot \dot{\gamma}^b(\tilde{\lambda})} \\ &= \int_0^1 d\tilde{\lambda} \sqrt{g_{ab}(\gamma(\tilde{\lambda})) \cdot \dot{\gamma}^a(\tilde{\lambda}) \cdot \dot{\gamma}^b(\tilde{\lambda})} \\ &= L[\gamma], \end{aligned}$$

where we have used the chain rule, and the result  $d\tilde{\lambda} = \dot{\sigma}d\lambda$ .

**Question:** Show that the Euler-Lagrange equations for a Lagrangian  $\mathcal{T}$  have precisely the same solutions as the Euler-Lagrange equations for the Lagrangian  $\mathcal{L} := \sqrt{\mathcal{T}}$ , if of the latter one only selects those solutions that satisfy the condition  $\mathcal{T} = 1$  on their parameterisation.

**Solution:** Let's denote the canonical variables at  $t$  and  $q$ , then the Euler-Lagrange equations for  $\mathcal{L}$  read

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \right) - \frac{\partial \mathcal{L}}{\partial q^a} = 0.$$

Substituting in  $\mathcal{L} := \sqrt{\mathcal{T}}$ , and using the assumption that we only consider solutions where  $\mathcal{T} = 1$ , and so it is a constant w.r.t.  $t$ , we have

$$\begin{aligned} 0 &= \frac{d}{dt} \left( \frac{\partial \sqrt{\mathcal{T}}}{\partial \dot{q}^a} \right) - \frac{\partial \sqrt{\mathcal{T}}}{\partial q^a} \\ &= \frac{d}{dt} \left( \frac{1}{2\sqrt{\mathcal{T}}} \frac{\partial \mathcal{T}}{\partial \dot{q}^a} \right) - \frac{1}{2\sqrt{\mathcal{T}}} \frac{\partial \mathcal{T}}{\partial q^a} \\ &= \frac{1}{2\sqrt{\mathcal{T}}} \left[ \frac{d}{dt} \left( \frac{\partial \mathcal{T}}{\partial \dot{q}^a} \right) - \frac{\partial \mathcal{T}}{\partial q^a} \right], \end{aligned}$$

which multiplying out the  $1/2\sqrt{\mathcal{T}}$  gives the answer.

#### 26.10.4 A Practical Way To Quickly Determine Christoffel Symbols

**Question:** Derive the geodesic equation for the two-dimensional round sphere of radius  $R$ , whose metric in some chart  $(U, x)$  is given by

$$g_{ab}(x^{-1}(\vartheta, \phi)) = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \vartheta \end{pmatrix}$$

via a convenient Euler-Lagrange equation. In order to lighten notation, you may define

$$\vartheta(\lambda) := (x^1 \circ \gamma)(\lambda), \quad \text{and} \quad \phi(\lambda) := (x^2 \circ \gamma)(\lambda).$$

**Solution:** We have  $\mathcal{L}[\gamma] := \sqrt{g(v_\gamma, v_\gamma)}$ , but the previous question showed us we can instead consider  $\mathcal{T} := g(v_\gamma, v_\gamma) = g_{ab} \dot{\gamma}^a \dot{\gamma}^b$  if we restrict ourselves to  $\mathcal{T} = 1$  on the parameterisation. For our metric components, we have

$$\mathcal{T} = R^2 \cdot \dot{\vartheta}(\lambda) \cdot \dot{\vartheta}(\lambda) + R^2 \sin^2 \vartheta \cdot \dot{\phi}(\lambda) \cdot \dot{\phi}(\lambda).$$

Plugging this into our Euler-Lagrange equations, we have

$$\begin{aligned} \frac{d}{d\lambda} \left( \frac{\partial \mathcal{T}}{\partial \dot{\vartheta}} \right) - \frac{\partial \mathcal{T}}{\partial \vartheta} &= 2R^2 (\ddot{\vartheta}(\lambda) - \sin \vartheta \cos \vartheta \cdot \dot{\phi}^2(\lambda)) \\ \frac{d}{d\lambda} \left( \frac{\partial \mathcal{T}}{\partial \dot{\phi}} \right) - \frac{\partial \mathcal{T}}{\partial \phi} &= 2R^2 \sin \vartheta (\sin \vartheta \cdot \ddot{\phi}(\lambda) + 2 \cos \vartheta \cdot \dot{\vartheta} \cdot \dot{\phi}) \end{aligned}$$

which simplify to

$$\begin{aligned} \ddot{\vartheta}(\lambda) - \sin \vartheta \cos \vartheta \cdot \dot{\phi}^2(\lambda) &= 0 \\ \ddot{\phi} + 2 \cot \vartheta \cdot \dot{\vartheta} \cdot \dot{\phi} &= 0. \end{aligned}$$

These are the geodesic equations for our round sphere of radius  $R$ .

**Question:** Read off the metric-induced connection coefficient functions for the round sphere.

**Solution:** Comparing the above result to the geodesic equation for the metric induced connection coefficients,

$$\ddot{\gamma}^a + \Gamma^a_{bc} \dot{\gamma}^b \dot{\gamma}^c,$$

we see straight away that

$$\Gamma^1_{22} = -\sin \vartheta \cos \vartheta, \quad \text{and} \quad \Gamma^2_{12} = \Gamma^2_{21} = \cot \vartheta,$$

with all other  $\Gamma$ s vanishing. Note in the second expression there is no 2 as we distribute it across the  $\Gamma^2_{12}$  and  $\Gamma^2_{21}$ .

### 26.10.5 Properties Of The Riemann-Christoffel Tensor

**Question:** Show that the chart-induced basis fields act on the coefficient functions as

$$\frac{\partial}{\partial x^c} (g^{-1})^{ab} = -(g^{-1})^{ar} (g^{-1})^{bs} \frac{\partial}{\partial x^c} g_{rs}.$$

**Solution:** Using  $(g^{-1})^{ab} g_{bs} = \delta_s^a$ , we have

$$\begin{aligned} 0 &= \frac{\partial}{\partial x^c} \delta_s^a \\ &= \frac{\partial}{\partial x^c} \left( (g^{-1})^{ab} g_{bs} \right) \\ &= \left( \frac{\partial}{\partial x^c} (g^{-1})^{ab} \right) g_{bs} + (g^{-1})^{ab} \left( \frac{\partial}{\partial x^c} g_{bs} \right) \\ &= \frac{\partial}{\partial x^c} (g^{-1})^{ar} + (g^{-1})^{sr} (g^{-1})^{ab} \left( \frac{\partial}{\partial x^c} g_{bs} \right) \end{aligned}$$

which, after relabelling and rearranging gives the result.

**Question:** Use normal coordinates to find an expression for the Riemann-Christoffel tensor

$$R_{abcd} = g_{ak} R^k_{bcd}$$

at a given point  $p$  in terms of  $g_{ab}$  and its first and second order derivatives at that very point.

**Solution:** This is just a long calculation involving the product rule and using the fact that in normal coordinates all the  $\Gamma$ s vanish. The full calculation is given in the video. The result is

$$R_{abcd} = \frac{1}{2} (g_{ad,bc} - g_{bd,ac} + g_{ac,bd} - g_{bc,ad}),$$

where

$$g_{ab,cd} := \frac{\partial^2 g_{ab}}{\partial x^c \partial x^d}.$$

**Question:** Show — in normal coordinates — that  $R_{abcd} = -R_{bacd}$ .

**Solution:** Switching the indices  $a \leftrightarrow b$  in the result of the previous exercise we have

$$R_{bacd} = \frac{1}{2}(g_{bd,ac} - g_{ad,bc} + g_{bc,ad} - g_{ac,bd}) = -\frac{1}{2}(g_{ad,bc} - g_{bd,ac} + g_{ac,bd} - g_{bc,ad}) = -R_{abcd}.$$

**Question:** Similarly, show that  $R_{abcd} = R_{cdab}$ .

**Solution:** Again just switch the indices and get the result.

**Question:** Show that  $R_{a[bcd]} = 0$  for the Riemann-Christoffel tensor.

**Solution:** We have

$$R_{a[bcd]} := \frac{1}{3!}(R_{abcd} - R_{abdc} + R_{acdb} - R_{acbd} + R_{adbc} - R_{adcb}).$$

Then using the previous two questions, we also have

$$R_{abcd} = R_{cdab} = -R_{dcab} = -R_{abdc}.$$

Following this with the other indice arrangements, we get

$$R_{a[bcd]} = \frac{1}{3}(R_{abcd} + R_{acdb} + R_{adbc}).$$

If you then plug in the expansion in terms of  $g_{ab}$  and its derivatives, you can show everything cancels and you get the result.

## 26.11 Symmetry

### 26.11.1 Pull-Back & Push-Forward

**Question:** Consider a smooth map  $\phi : \mathcal{M} \rightarrow \mathcal{N}$  between two differential manifolds. Show that a function  $f \in C^\infty(\mathcal{N})$ , the pull-back of the gradient of  $f$  is the same as the gradient of the pull-back of  $f$ , i.e.

$$\phi^*(df) = d(\phi^*f).$$

**Solution:** By definition, we have

$$(\phi_*X)\langle f \rangle = X\langle f \circ \phi \rangle,$$



for  $X \in T\mathcal{M}$ , giving

$$\begin{aligned}\phi^*(df) : X &:= df : \phi_*(X) \\ &= \phi_*X \langle f \rangle \\ &= X \langle f \circ \phi \rangle \\ &= d(f \circ \phi) : X \\ &=: d(\phi^*f) : X,\end{aligned}$$

which holds for arbitrary  $X$  and therefore proves the result.

**Question:** The push-forward  $\phi_* : T\mathcal{M} \rightarrow T\mathcal{N}$  is a linear map between tangent bundles. Calculate its component functions

$$\phi_*^a_b := dy^a : \phi_* \left( \frac{\partial}{\partial x^b} \right)$$

with respect to charts  $(U \subset \mathcal{M}, x)$  and  $(V \subset \mathcal{N}, y)$ !

**Solution:** If we considered a general vector and gradient, we would have

$$df : \phi_*X = \phi_*X \langle f \rangle = X \langle f \circ \phi \rangle = X^i \left( \frac{\partial(f \circ \phi)}{\partial x^i} \right).$$

Now we have

$$\begin{aligned}\left( \frac{\partial(f \circ \phi)}{\partial x^i} \right)_p &:= \partial_i(f \circ \phi \circ x^{-1})|_{x(p)} \\ &= \partial_i(f \circ y^{-1} \circ y \circ \phi \circ x^{-1})|_{x(p)} \\ &= \partial_j(f \circ y^{-1})|_{(y \circ \phi \circ x^{-1} \circ x)(p)} \cdot \partial_i(y^j \circ \phi \circ x^{-1})|_{x(p)} \\ &=: \left( \frac{\partial f}{\partial y^j} \right)_q \cdot \left( \frac{\partial(y^j \circ \phi)}{\partial x^i} \right)_p,\end{aligned}$$

where  $q := \phi(p) \in \mathcal{N}$ . Now put in  $f = y^a$  and  $X = \frac{\partial}{\partial x^b}$ , giving

$$\begin{aligned}\phi_*^a_b(p) &= \left( \frac{\partial y^a}{\partial y^j} \right)_q \cdot \left( \frac{\partial(y^j \circ \phi)}{\partial x^b} \right)_p \\ &= \delta_j^a \cdot \left( \frac{\partial(y^j \circ \phi)}{\partial x^b} \right)_p \\ &= \left( \frac{\partial(y \circ \phi)^a}{\partial x^b} \right)_p.\end{aligned}$$

**Question:** Show that the component functions of the pull back  $\phi^*g$  of the metric

tensor field are obtained from the component functions of  $g$  by

$$(\phi^*g)_{ab}(p) = \left( \frac{\partial(y \circ \phi)^m}{\partial x^a} \right)_p \left( \frac{\partial(y \circ \phi)^n}{\partial x^b} \right)_p g_{mn}(\phi(p)).$$

**Solution:** With the definition of the induced metric in mind, we have

$$\begin{aligned} (\phi^*g)(X, Y) &= g(\phi_*X, \phi_*Y) \\ &= g_{ab}(\phi_*X)^a (\phi_*Y)^b \\ &= g_{ab}(dy^a : \phi_*X)(dy^b : \phi_*Y). \end{aligned}$$

Then if we use  $g_{ab} = g\left(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right)$  and the results of the previous exercise we get

$$\begin{aligned} (\phi^*g)_{ab}(p) &= g_{mn}(q) \cdot \left[ dy^m : \phi_*\left(\frac{\partial}{\partial x^a}\right) \right](p) \cdot \left[ dy^n : \phi_*\left(\frac{\partial}{\partial x^b}\right) \right](p) \\ &= \left( \frac{\partial(y \circ \phi)^m}{\partial x^a} \right)_p \left( \frac{\partial(y \circ \phi)^n}{\partial x^b} \right)_p g_{mn}(\phi(p)). \end{aligned}$$

### 26.11.2 Lie Derivative — The Pedestrian Way

**Question:** Consider the smooth embedding  $\iota : S^2 \rightarrow \mathbb{R}^3$  of  $(S^2, \mathcal{O}, \mathcal{A})$  into  $(\mathbb{R}^3, \mathcal{O}_{st}, \mathcal{B})$ , which for the familiar chart  $(U, x) \in \mathcal{A}$  and  $(\mathbb{R}^3, y = \mathbb{1}_{\mathbb{R}^3}) \in \mathcal{B}$  is given by

$$y \circ \iota \circ x^{-1} : (\vartheta, \varphi) \mapsto (a \cos \varphi \sin \vartheta, b \sin \varphi \sin \vartheta, c \cos \vartheta),$$

where  $a, b$  and  $c$  are positive real numbers. What can you say about the shape of  $\iota(S^2)$ ?

**Solution:** Nothing as in order to talk about shape you need either a covariant derivative or a metric, neither of which we have. I guess you could say it is some closed and compact 2-dimensional shape, but you could not specify which one, i.e. if its a round sphere or an ellipsoid or a potato.

**Question:** Now assume  $(\mathbb{R}^3, \mathcal{O}_{st}, \mathcal{B})$  is additionally equipped with the Euclidean metric  $g$ , whose components with respect to the chart  $(\mathbb{R}^3, y)$  are given by

$$g_{ab}(p) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for any } p \in U.$$

Write down the component functions of  $g^{\text{ellipsoid}} := \iota^*g$  with respect to the chart  $(U, x)$ !

**Solution:** Using the result of the previous questions, we need to find

$$\left(\frac{\partial(y \circ \iota)^m}{\partial x^a}\right)_p := \partial_a(y^m \circ \iota \circ x^{-1})|_{x(p)}.$$

Using the definition given, we have

$$\begin{aligned} \left(\frac{\partial(y \circ \iota)^1}{\partial x^1}\right)_p &= a \cos \varphi \cos \vartheta, & \left(\frac{\partial(y \circ \iota)^1}{\partial x^2}\right)_p &= -a \sin \varphi \sin \vartheta \\ \left(\frac{\partial(y \circ \iota)^2}{\partial x^1}\right)_p &= b \sin \varphi \cos \vartheta, & \left(\frac{\partial(y \circ \iota)^2}{\partial x^2}\right)_p &= b \cos \varphi \sin \vartheta \\ \left(\frac{\partial(y \circ \iota)^3}{\partial x^1}\right)_p &= -c \sin \vartheta, & \left(\frac{\partial(y \circ \iota)^3}{\partial x^2}\right)_p &= 0. \end{aligned}$$

You then just plug in the relevant terms, giving

$$\begin{aligned} g_{11}^{\text{ellipsoid}} &= a^2 \cos^2 \varphi \cos^2 \vartheta + b^2 \sin^2 \varphi \cos^2 \vartheta + c^2 \sin^2 \vartheta \\ g_{22}^{\text{ellipsoid}} &= a^2 \sin^2 \varphi \sin^2 \vartheta + b^2 \cos^2 \varphi \sin^2 \vartheta, \end{aligned}$$

and  $g_{12}^{\text{ellipsoid}} = 0 = g_{21}^{\text{ellipsoid}}$ .

**Question:** For convenience, denote by  $(\vartheta, \varphi)$  the coordinate functions  $(x^1, x^2)$ . Check the vector fields

$$\begin{aligned} X_1(p) &= -\sin \varphi(p) \left(\frac{\partial}{\partial \vartheta}\right)_p - \cot \vartheta(p) \cos \varphi(p) \left(\frac{\partial}{\partial \varphi}\right)_p \\ X_2(p) &= \cos \varphi(p) \left(\frac{\partial}{\partial \vartheta}\right)_p - \cot \vartheta(p) \sin \varphi(p) \left(\frac{\partial}{\partial \varphi}\right)_p \\ X_3(p) &= \left(\frac{\partial}{\partial \varphi}\right)_p \end{aligned}$$

constitute a Lie subalgebra of  $(\Gamma TS^2, [\cdot, \cdot])$  and determine the structure constants!

**Solution:** Let  $f \in C^\infty(\mathcal{M})$  be an arbitrary smooth function. We need to consider the action of the Lie bracket expressions on  $f$ , e.g.  $[X_1, X_2]\langle f \rangle$ . We use the clever trick that in this expansion only the terms where a derivative acts on a term in the  $X$ s will remain. That is, any terms that are second order derivative of  $f$  will vanish because it will appear in both  $X_1\langle X_2\langle f \rangle \rangle$  and  $X_2\langle X_1\langle f \rangle \rangle$  which the order switched, but partial derivative commute and so these terms cancel.

We then have (dropping the  $ps$  for notational reasons)

$$\begin{aligned}
 [X_1, X_2]\langle f \rangle &:= X_1\langle X_2\langle f \rangle \rangle - X_2\langle X_1\langle f \rangle \rangle \\
 &= \left[ (-\operatorname{cosec}^2 \vartheta \sin^2 \varphi + \cot^2 \vartheta \cos^2 \varphi) \left( \frac{\partial f}{\partial \varphi} \right) + \cot \vartheta \cos \varphi \sin \varphi \left( \frac{\partial f}{\partial \vartheta} \right) \right] \\
 &\quad - \left[ (\operatorname{cosec}^2 \vartheta \cos^2 \varphi - \cot^2 \vartheta \sin^2 \varphi) \left( \frac{\partial f}{\partial \varphi} \right) + \cot \vartheta \sin \varphi \cos \varphi \left( \frac{\partial f}{\partial \vartheta} \right) \right] \\
 &= (\cot^2 \vartheta - \operatorname{cosec}^2 \vartheta) \left( \frac{\partial f}{\partial \varphi} \right) \\
 &= - \left( \frac{\partial}{\partial \varphi} \right) \langle f \rangle \\
 &= -X_3 \langle f \rangle \\
 \implies [X_2, X_1] &= X_3,
 \end{aligned}$$

where on the last line we have used the antisymmetry of the Lie bracket.

Next we have

$$\begin{aligned}
 [X_1, X_3]\langle f \rangle &= 0 - \left[ -\cos \varphi \left( \frac{\partial f}{\partial \vartheta} \right) + \cot \vartheta \sin \varphi \left( \frac{\partial f}{\partial \varphi} \right) \right] \\
 &= X_2 \langle f \rangle \\
 \implies [X_1, X_3] &= X_2,
 \end{aligned}$$

and

$$\begin{aligned}
 [X_3, X_2] &= \left[ -\sin \varphi \left( \frac{\partial f}{\partial \vartheta} \right) - \cot \vartheta \cos \varphi \left( \frac{\partial f}{\partial \varphi} \right) \right] - 0 \\
 &= X_1 \langle f \rangle \\
 \implies [X_3, X_2] &= X_1.
 \end{aligned}$$

So we see that  $\{X_1, X_2, X_3\}$  is closed under the Lie bracket, and so forms a Lie subalgebra. The structure constants are

$$C^3_{21} = C^2_{13} = C^1_{32} = 1,$$

and all other non-related (i.e. not  $C^3_{12} = -C^3_{21}$ , etc.) structure constants vanish.

Note this result tells us that  $\{X_1, X_2, X_3\}$  is a 3-dimensional rotation algebra, as defined in the lecture. We therefore expect it to be a symmetry of  $S^2$ , which we show explicitly below for  $X_3$ .

**Question:** Calculate the integral curve of  $X_3$  through the point  $p = x^{-1}(\vartheta_0, \varphi_0)$ , i.e. the curve  $\gamma_p$  satisfying

$$\gamma_p(0) = p, \quad \text{and} \quad v_{\gamma_p, \gamma_p(\lambda)} = (X_3)_{\gamma_p(\lambda)}$$

in the chart  $(U, x)$ !

**Solution:** Using

$$v_{\gamma_p, \gamma_p(\lambda)} = (X_3)_{\gamma_p(\lambda)} \iff \dot{\gamma}_{p(x)}^i \left( \frac{\partial}{\partial x^i} \right)_{\gamma_p(\lambda)} = \left( \frac{\partial}{\partial \varphi} \right)_{\gamma_p(\lambda)},$$

we have

$$\dot{\gamma}_{p(x)}^1(\lambda) := (x^1 \circ \gamma_p)'(\lambda) = 0, \quad \text{and} \quad \dot{\gamma}_{p(x)}^2(\lambda) := (x^2 \circ \gamma_p)'(\lambda) = 1,$$

from which it follows that

$$\vartheta(\gamma_p) = a, \quad \text{and} \quad \varphi(\gamma_p) = \lambda + b,$$

for constants  $a$  and  $b$ . We see from the question that  $a = \vartheta_0$  and  $b = \varphi_0$ . So we have

$$\gamma_{p(x)}(\lambda) = (\vartheta_0, \lambda + \varphi_0),$$

which satisfies

$$\gamma_p(0) = x^{-1}(\gamma_{p(x)}(0)) = x^{-1}(\vartheta_0, \varphi_0) = p.$$

**Question:** The integral curves  $\gamma_p$  give rise to a one-parameter family of smooth maps  $h_\lambda^{X_3} : S^2 \rightarrow S^2$ . Calculate the pull-back

$$\left( h_\lambda^{X_3} \right)^* g^{\text{ellipsoid}}$$

of the metric on  $S^2$ . What can you conclude about the Lie derivative  $\mathcal{L}_{X_3} g^{\text{ellipsoid}}$ ?

**Solution:** The flow is

$$h_\lambda^{X_3} : p \mapsto \gamma_p(\lambda),$$

so we have

$$\left( x^m \circ h_\lambda^{X_3} \circ x^{-1} \right) : (\vartheta, \varphi) \mapsto \gamma_p^m(\lambda),$$

and so

$$\begin{aligned} \left( \frac{\partial(x \circ h_\lambda^{X_3})^1}{\partial x^1} \right)_p &:= \partial_1 \left( x^1 \circ h_\lambda^{X_3} \circ x^{-1} \right) \Big|_{x(p)} = \vartheta_0 \\ \left( \frac{\partial(x \circ h_\lambda^{X_3})^1}{\partial x^2} \right)_p &:= \partial_2 \left( x^1 \circ h_\lambda^{X_3} \circ x^{-1} \right) \Big|_{x(p)} = 0 \\ \left( \frac{\partial(x \circ h_\lambda^{X_3})^2}{\partial x^1} \right)_p &:= \partial_1 \left( x^2 \circ h_\lambda^{X_3} \circ x^{-1} \right) \Big|_{x(p)} = 0 \\ \left( \frac{\partial(x \circ h_\lambda^{X_3})^2}{\partial x^2} \right)_p &:= \partial_2 \left( x^2 \circ h_\lambda^{X_3} \circ x^{-1} \right) \Big|_{x(p)} = \lambda + \varphi_0. \end{aligned}$$

This gives us

$$\begin{aligned} \left[ \left( h_\lambda^{X_3} \right)^* g^{\text{ellipsoid}} \right]_{11}(p) &= \vartheta_0^2 g_{11}^{\text{ellipsoid}}(\gamma_p(\lambda)) \\ \left[ \left( h_\lambda^{X_3} \right)^* g^{\text{ellipsoid}} \right]_{22}(p) &= (\lambda + \varphi_0)^2 g_{22}^{\text{ellipsoid}}(\gamma_p(\lambda)) \end{aligned}$$

and again the other two components vanish. If we then take the coordinate transformation

$$\vartheta \rightarrow \frac{1}{\vartheta_0} \vartheta, \quad \text{and} \quad \varphi \rightarrow \frac{1}{\lambda + \varphi_0} \varphi,$$

which we can do as  $\vartheta_0, \varphi_0 > 0$  (as the ranges of  $\vartheta$  and  $\varphi$  are positive), we then get

$$\left[ \left( h_\lambda^{X_3} \right)^* g^{\text{ellipsoid}} \right]_{ab} (p) = g_{ab}^{\text{ellipsoid}} (\gamma_p(\lambda)),$$

or more nicely

$$\left( h_\lambda^{X_3} \right)^* g^{\text{ellipsoid}} = g^{\text{ellipsoid}}.$$

This tells us that  $X_3$  is a symmetry of the metric, and so  $\mathcal{L}_{X_3} g^{\text{ellipsoid}} = 0$ .

## 26.12 Integration

### 26.12.1 Integrals & Volumes

**Question:** Calculate the volume of the round sphere  $S^2$  of radius  $R$ , i.e.,

$$\text{vol}(S^2) = \int_{S^2} 1.$$

**Solution:** The first thing we have to not is that ‘volume’ here does not mean what we intuitively think, i.e. the Euclidean 3-volume, but it means what we would normally call the ‘surface area’. This distinction comes from which metric we are using, the metric on  $S^2$  itself or the Euclidean metric with a sphere of radius  $R$  embedded into it. Once this distinction is made the calculation is trivial, consider the chart with  $(x^1, x^2) = (\vartheta, \varphi)$ , then

$$g_{(x)ab}(x^{-1}(\vartheta, \varphi)) = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \vartheta \end{pmatrix}, \quad \implies \quad g := \det(g_{(x)ab})(x^{-1}(\vartheta, \varphi)) = R^4 \sin^2 \vartheta$$

and so

$$\begin{aligned} \int_{S^2} 1 &:= \int_{x(S^2)} d^2x \sqrt{g} 1 \\ &= \int_0^\pi d\varphi \int_0^{2\pi} d\vartheta |R^2 \sin^2 \vartheta| \\ &= 4\pi R^2, \end{aligned}$$

which is what we expect.

Technically we need to include another chart, as  $x(S^2)$  will miss two antipodal points and a geodesic connecting them, however this will contribute nothing to the volume as, we would only consider this line (the partition of unity removing the overlap region), which has no ‘thickness’ and so no volume. For example if the line missing was the line of longitude connecting the North and South poles, we would have

$$\int_{\varphi_0}^{\varphi_0} d\varphi \int_0^\pi d\vartheta R^2 \sin \vartheta = 0,$$

where  $\varphi_0$  is the value of  $\varphi$  along the line of longitude.

## 26.13 Schwarzschild Spacetime

### 26.13.1 Geodesics In A Schwarzschild Spacetime

The Schwarzschild metric is given and we are told to use the light hand notation

$$t(\lambda) := (x^0 \circ \gamma)(\lambda),$$

and similarly for  $r(\lambda)$ ,  $\theta(\lambda)$  and  $\varphi(\lambda)$ , where  $\gamma : \mathbb{R} \rightarrow U$  is some curve.

**Question:** Write down the Lagrangian  $\mathcal{L} := g_{ab}\dot{\gamma}^a\dot{\gamma}^b$ !

**Solution:** Using the metric given in the question (see the video if you don't know it) we have

$$\mathcal{L} = \left(1 - \frac{2GM}{r}\right)\dot{t}^2 - \left(1 - \frac{2GM}{r}\right)^{-1}\dot{r}^2 - r^2\dot{\theta}^2 - r^2\sin^2\theta\dot{\varphi}^2.$$

**Question:** Find the Euler-Lagrange equation with respect to  $t(\lambda)$ !

**Solution:** We see straight away that

$$\frac{\partial \mathcal{L}}{\partial t} = 0.$$

We also have

$$\frac{d}{d\lambda} \left( \frac{\partial \mathcal{L}}{\partial \dot{t}} \right) = 2 \left( 1 - \frac{2GM}{r} \right) \dot{t} + \frac{4GM}{r^2} \dot{r} \dot{t},$$

which gives the Euler-Lagrange equation

$$\dot{t} + \frac{2GM}{r^2 \left( 1 - \frac{2GM}{r} \right)} \dot{r} \dot{t} = 0$$

**Question:** Show that the Lie derivative of  $g$  with respect to the vector field  $K_t := \frac{\partial}{\partial t}$  vanishes. What does this mean?

**Solution:** We have

$$(\mathcal{L}_{K_t} g)_{ab} = K_t \langle g_{ab} \rangle + g_{mb} \frac{\partial}{\partial x^a} (K_t)^m + g_{ab} \frac{\partial}{\partial x^b} (K_t)^m = 0,$$

as all three terms vanish. This tells us that  $K_t$  is a symmetry of the metric. Indeed the Schwarzschild spacetime is stationary (and even static), the definitions for which are given in lecture 16. We will see this symmetry is the conservation of energy.

**Question:** The exact form of the conserved quantity is given by  $(K_t)_a(x^a)'(\lambda) = \text{const.}$  (without proof). Derive an expression for the quantity  $t'(\lambda)$  appearing in the Lagrangian!

**Solution:** Using  $(K_t)_a := g_{ab}(K_t)^b$ , we have

$$g_{ab}(K_t)^b(x^a)'(\lambda) = g_{00}t'(\lambda) = \left(1 - \frac{2GM}{r}\right)t'(\lambda) = \text{const.}$$

Letting  $\sqrt{E}$  be the constant, we have

$$t'(\lambda) = \frac{r\sqrt{E}}{r - 2GM}$$

**Question:** Moreover, we can find so-called "spherical symmetry", that is, the Lie derivative of  $g$  with respect to the already known vector fields

$$\begin{aligned} X_1 &= \sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \\ X_2 &= \cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \\ X_3 &= \frac{\partial}{\partial \varphi} \end{aligned}$$

vanishes. What physical quantity is conserved by this symmetry?

**Solution:** Angular momentum.

**Question:** Due to  $X_1$  and  $X_2$  (without proof), one can fix the motion to a plane of constant  $\theta = \frac{\pi}{2}$ . How can you derive an expression for the remaining term  $\varphi'(\lambda)$ ?

**Solution:** From two questions above, we have

$$\text{const.} = g_{ab}X_3^b(x^a)'(\lambda) = g_{33}\varphi'(\lambda) = -r^2 \sin^2 \theta \varphi'(\lambda),$$

so using  $\theta = \frac{\pi}{2}$  and labelling the constant  $J$ , we have

$$\varphi'(\lambda) = \frac{J}{r^2}.$$

**Question:** Use all the fact that  $\mathcal{L} = 1$  on the parameterisation. Insert the previously obtained results and take all terms not containing  $E$  to one side!



**Solution:** First note that we have replaced the dots with primes, and also note that  $\theta' = 0$  as  $\theta$  is a constant. We therefore have

$$1 = \left(1 - \frac{2GM}{r}\right) \frac{E}{\left(1 - \frac{2GM}{r}\right)^2} - \frac{1}{1 - \frac{2GM}{r}} (r')^2 - r^2 \frac{J^2}{r^4},$$

which can be rearranged to

$$E = (r')^2 + 1 - \frac{2GM}{r} + \frac{J^2}{r^2} - \frac{2GMJ^2}{r^3}.$$

**Question:** Can you interpret the terms appearing in this expression?

**Solution:** If we then consider a particle of mass  $m = 1$ , we see the above formula as representing

- $E$  is total energy,
- $(r')^2$  is the kinetic energy,
- 1 is the mass,
- $-\frac{2GM}{r}$  is the Newtonian gravitational potential,
- $\frac{J^2}{r^2}$  is some angular momentum contribution, and
- $-\frac{2GMJ^2}{r^3}$  is some GR correction term.

### 26.13.2 Gravitational Redshift

Consider a spacetime equipped with the Schwarzschild metric as well as two observers 1 and 2 at rest in their respective system of reference ( $\dot{r} = 0$ ,  $\dot{\theta} = 0$ ,  $\dot{\varphi} = 0$ ). The observers sit at the same  $\theta$  and  $\varphi$  while  $r_1 < r_2$ .

**Question:** Derive an expression for  $t'(\lambda)$  using the Lagrangian from the previous exercise!

**Solution:** Using  $1 = \mathcal{L} = g_{ab} \dot{\gamma}^a \dot{\gamma}^b$ , we get

$$t'(\lambda) = \left(1 - \frac{2GM}{r}\right)^{-1/2}.$$

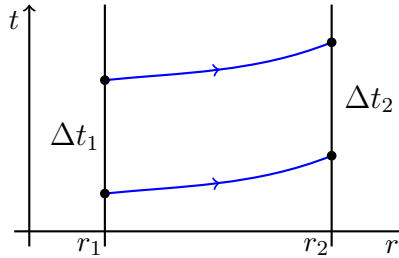
**Question:** Observer 1 emits photons that observer 2 detects. The gap between the two photon emissions is  $\Delta\lambda_1$ . Find the gap  $\Delta\lambda_2$  seen by observer 2!

**Solution:** We have

$$\Delta t_1 = \left(1 - \frac{2GM}{r_1}\right)^{-1/2} \Delta \lambda_1$$

$$\Delta t_2 = \left(1 - \frac{2GM}{r_2}\right)^{-1/2} \Delta \lambda_2.$$

Then, we use the fact that  $K_t$  was a Killing vector field, and so the path taken by the two emitted photons is the same, apart from a constant time shift.



As the above diagram shows, the Killing condition basically tells us that  $\Delta t_1 = \Delta t_2$ , and so we get

$$\Delta \lambda_2 = \sqrt{\frac{1 - \frac{2GM}{r_2}}{1 - \frac{2GM}{r_1}}} \Delta \lambda_1$$

**Question:** Consider the ratio of frequencies  $\frac{\omega_1}{\omega_2}$ . What happens for observer 2 being approximately at infinity? What happens when sending  $r_1$  to the Schwarzschild radius  $r_s = 2GM$ ?

**Solution:** We can think of  $\Delta \lambda$  being the time period (time between photons emitted) and so the frequency (which is the reciprocal of the time period) is

$$\frac{\omega_1}{\omega_2} = \frac{\Delta \lambda_2}{\Delta \lambda_1} = \sqrt{\frac{1 - \frac{2GM}{r_2}}{1 - \frac{2GM}{r_1}}}.$$

As  $r_1 < r_2$ , the above tells us that  $\omega_2 < \omega_1$ . In terms of wavelength, this is  $\mu_1 < \mu_2$  (we have used  $\mu$  for wavelength as  $\lambda$  is already used above), which tells us that the light has been *red-shifted*. This is seen also in the definition of redshift,

$$1 + z = \frac{\omega_1}{\omega_2},$$

where  $z > 0$  is redshift and  $z < 0$  is blueshift.

For  $r_2 \rightarrow \infty$  we have

$$\frac{\omega_1}{\omega_2} \rightarrow \left(1 - \frac{2GM}{r_1}\right)^{-1/2}.$$

For  $r_1 \rightarrow r_s$ , we have

$$\frac{\omega_1}{\omega_2} \rightarrow \infty,$$

which tells us  $\omega_1 \rightarrow \infty$ . In terms of wavelengths, this says  $\mu_2 \rightarrow \infty$ , and so the light is infinitely redshifted. This is a so-called *infinite redshift surface*.

This result is actually misleading as it really only holds because of the choice of reference frame. We took the coordinate time to be that of the black hole's rest frame. If we use an in-falling observer's clock to define the coordinate time, this infinite redshift behaviour disappears. I have discussed this in more detail in the notes I have put on my blog site.

## 26.14 Relativistic Spacetime, Matter & Gravitation

### 26.14.1 Lorentz Force Law

**Question:** Recall from the lecture that for a particle coupling to the electromagnetic potential, we have

$$m(\nabla_{v_\gamma} v_\gamma)^a = qF^a{}_b v_\gamma^b,$$

where  $v_\gamma$  is the velocity of a particle of mass  $m$  and charge  $q$ .

Now "1 + 3"-decompose this equation in components with respect to the frame of an observer.

**Solution:** The observer  $(\delta, e)$  has a frame such that

$$g(e_a, e_b) = \eta_{ab}, \quad \text{and} \quad e_0(\lambda) = v_{\delta, \delta(\lambda)}.$$

We thus have

$$m(\nabla_{v_\gamma} v_\gamma)^a = qF^{ac} \eta_{cb} v_\gamma^b,$$

and so

$$m(\nabla_{v_\gamma} v_\gamma)^0 = qF^{00} v_\gamma^0 + \sum_{\beta=1}^3 qF^{0\beta} v_\gamma^\beta$$

$$m(\nabla_{v_\gamma} v_\gamma)^\alpha = -qF^{\alpha 0} v_\gamma^0 - \sum_{\beta=1}^3 qF^{\alpha\beta} v_\gamma^\beta.$$

**Question:** Using the definitions  $E_\alpha := F_{\alpha 0}$  for the electric field and  $B^\alpha := \frac{1}{2} \varepsilon^{\alpha\rho\sigma} F_{\rho\sigma}$  for the magnetic 2 field seen by an observer, bring the right hand side of the above equation to the familiar form of the Lorentz force law for a particle of charge  $q$  and spatial velocity

$$\mathbf{v} := \left( \frac{e^\alpha : v_\delta}{e^0 : v_\delta} \right) e_\alpha \quad (\alpha = 1, 2, 3 \text{ and careful: the denominator was forgotten in lectures})$$

that the observer detects for the particle.

*Hint:*  $(a \times b)^\alpha = g^{\alpha\mu} \varepsilon_{\mu\rho\sigma} a^\rho b^\sigma$ ,  $\varepsilon_{123} = 1$  and  $\varepsilon^{123} = 1$ .

**Solution:** We set

$$\mathcal{F}^\alpha = m(\nabla_{v_\delta} v_\delta)^\alpha = qF^\alpha_b v_\delta^b,$$

where  $\mathcal{F}$  is the Lorentz force, up to some factors. We can split the right-hand side into two terms, one for  $b = 0$  and the other for  $b = \beta$ . For the former we note that

$$\eta^{\alpha\beta} E_\beta = \eta^{\alpha\beta} F_{\beta 0} = F^\alpha_0,$$

so, using  $\epsilon^0 : (v_\delta) = v_\delta^0$ , the first term is simply

$$q\eta^{\alpha\beta} E_\beta(\epsilon^0 : v_\delta).$$

The second term need a little more work, but we start using the hint: using  $g^{ab} = \eta^{ab}$  in the observers frame, we have

$$(v_\delta \times B)^\alpha = \eta^{\alpha\mu} \varepsilon_{\mu\rho\sigma} v_\delta^\rho B^\sigma := \frac{1}{2} \eta^{\alpha\mu} \varepsilon_{\mu\rho\sigma} \varepsilon^{\sigma\nu\tau} v_\delta^\rho F_{\nu\tau}.$$

We see that the  $\eta^{\alpha\mu}$  tells us that  $\mu = \alpha$  in the Levi-Civita symbol. By definition, the only non-vanishing terms on the right, then, have  $\rho \neq \sigma \neq \alpha$ . Let's consider the case for  $\alpha = 1$  (the other two follow analogously)

$$\begin{aligned} \frac{1}{2} \eta^{1\mu} \varepsilon_{\mu\rho\sigma} \varepsilon^{\sigma\nu\tau} F_{\nu\tau} v_\delta^\rho &= \frac{1}{2} \varepsilon_{123} (\varepsilon^{312} F_{12} + \varepsilon^{321} F_{21}) v_\delta^2 + \frac{1}{2} \varepsilon_{132} (\varepsilon^{213} F_{13} + \varepsilon^{231} F_{31}) v_\delta^3 \\ &= \frac{1}{2} \varepsilon_{123} (\varepsilon^{123} F_{12} + (-\varepsilon^{123})(-F_{12})) v_\delta^2 + \frac{1}{2} (-\varepsilon_{123}) ((-\varepsilon^{123}) F_{13} + \varepsilon^{123} (-F_{13})) v_\delta^3 \\ &= F_{12} v_\delta^2 + F_{13} v_\delta^3, \end{aligned}$$

where we have indicated where the antisymmetries of  $\varepsilon$  and  $F$  have been used, and we have also used  $\varepsilon_{123} = 1 = \varepsilon^{123}$ . This generalises to

$$\frac{1}{2} \eta^{\alpha\mu} \varepsilon_{\mu\rho\sigma} \varepsilon^{\sigma\nu\tau} F_{\nu\tau} v_\delta^\rho = \eta^{\alpha\sigma} F_{\sigma\beta} v_\delta^\beta = F^\alpha_\beta v_\delta^\beta,$$

where we note the fact that  $F_{\alpha\alpha} = 0$  due to antisymmetry.

So the second term in our Lorentz force equation is simply

$$qF^\alpha_\beta v_\delta^\beta = q(v_\delta \times B)^\alpha,$$

giving us

$$\mathcal{F}^\alpha = q\eta^{\alpha\beta} E_\beta(\epsilon^0 : v_\delta) + q(v_\delta \times B)^\alpha.$$

Finally, we note that

$$\frac{1}{(\epsilon^0 : v_\delta)} (v_\delta \times B)^\alpha = (\mathbf{v} \times B)^\alpha,$$

as everywhere in the calculation above we'll get terms like

$$\frac{v_\delta^\rho}{(\epsilon^0 : v_\delta)} = \frac{(\epsilon^\rho : v_\delta)}{(\epsilon^0 : v_\delta)} =: \mathbf{v}^\rho.$$

This gives us the form we want on the right-hand side, i.e.

$$\frac{1}{(\epsilon^0 : v_\delta)} \mathcal{F}^\alpha = q\eta^{\alpha\beta} E_\beta + q(\mathbf{v} \times B)^\alpha.$$

### 26.14.2 Which Curvature Can Feature In Einstein's Equations?

This questions asks to show that the differential Bianchi identity holds for the Riemann tensor. This was given as an exercise in the lecture.

If the reader couldn't work it out there, [this link](#) should prove helpful. Note the notation is slightly different in the link, but of course the answer is the same.

I didn't include this link in the notes to hopefully avoid temptation of just looking up the answer.

**Question:** The above component-free version can equivalently be written as

$$R^w{}_{zab;c} + R^w{}_{zbc;a} + R^w{}_{zca;b} = 0.$$

Using this result, show that by appropriate contractions one obtains

$$(\nabla_a G)^{ab} = 0.$$

**Solution:** First use the antisymmetry to give

$$R^w{}_{zab;c} + R^w{}_{zbc;a} + R^w{}_{zca;b} = R^w{}_{zab;c} + R^w{}_{zbc;a} - R^w{}_{zac;b}.$$

Now, we can use the fact that  $\nabla$  is metric-compatible and the result

$$g^{av} g_{vw} = \delta_w^a$$

to give

$$\begin{aligned} g^{av} g_{vw} (R^w{}_{zab;c} + R^w{}_{zbc;a} + R^w{}_{zca;b}) &= R^a{}_{zab;c} + R^a{}_{zbc;a} - R^a{}_{zac;b} \\ &= R_{zb;c} - R_{zc;b} + R^a{}_{zbc;a}. \end{aligned}$$

Then contract with  $g^{bz}$  along with the results  $R_z{}^a{}_{bc} = -R^a{}_{zbc}$ <sup>2</sup> and  $R := g^{ab} R_{ab}$  to give

$$\begin{aligned} g^{bz} (R_{zb;c} - R_{zc;b} + R^a{}_{zbc;a}) &= R_{;c} - R^b{}_{c;b} - R^{ba}{}_{bc;a} \\ &= R_{;c} - R^b{}_{c;b} - R^a{}_{c;a} \\ &= R_{;c} - 2R^b{}_{c;b}. \end{aligned}$$

Finally using the fact that the Ricci tensor is symmetric,<sup>3</sup> and contracting with  $g^{ac}$ , we have (after relabelling)

$$(\nabla_a G)^{ab} = R^{ab}{}_{;a} - \left( \frac{1}{2} g^{ab} R \right)_{;a} = 0.$$

## 26.20 Cosmology

### 26.20.1 Killing's Equation

<sup>2</sup>We essentially showed this result in tutorial 9 (we showed  $R_{abcd} = -R_{bacd}$ ).

<sup>3</sup>This result is obtained from  $R_{abcd} = R_{cdab}$ , then setting  $a = c$  and raising the first index back up.

**Question:** Show that a vector field  $K$  is Killing if, and only if,

$$(\nabla_a K)_b + (\nabla_b K)_a = 0.$$

**Solution:** If you have done the exercise in the notes to show

$$g(\nabla_X K, Y) + g(X, \nabla_Y K) = 0,$$

this question follows trivially by setting  $X = \partial_a$  and  $Y = \partial_b$ :

$$0 = g_{cb}(\nabla_a K)^c + g_{ac}(\nabla_b K)^c =: (\nabla_a K)_b + (\nabla_b K)_a.$$

If you didn't do that exercise, go back and do it, but also you can see [the video](#) for a method considering the components.

This result is known as Killing's equation.

### 26.20.2 Age Of The Universe...

The energy-momentum tensor for a perfect fluid is

$$T^{ab} := [\rho(t) + p(t)]u^a u^b + g^{ab}p(t),$$

where  $u^a = (1, 0, 0, 0)^a$  are the components functions of a smooth vector field and  $g_{ab}$  are those of a FRW metric w.r.t. the coordinate chart  $(t, r, \vartheta, \phi)$  employed in the lectures.

**Question:** Derive the conservation equation

$$\dot{\rho}(t) = -3\frac{\dot{a}}{a}(\rho(t) + p(t))$$

by evaluating the condition

$$(\nabla_a T)^{ab} u_b = 0,$$

which follows from the Einstein equations by virtue of the differential Bianchi identity.

**Solution:** We have

$$(\nabla_a T)^{ab} u_b = (\nabla_a T)^{a0},$$

as  $u_b = (1, 0, 0, 0)_b$ . Then using  $T^{a0} = 0$  and  $T^{00} = \rho(t)$  we have

$$\begin{aligned} (\nabla_a T)^{ab} u_b &= T^{a0}_{,a} + \Gamma^a_{na} T^{n0} + \Gamma^0_{na} T^{an} \\ &= \dot{\rho}(t) + \Gamma^a_{0a} \rho(t) + \Gamma^0_{na} T^{an}. \end{aligned}$$

We now use the results from the lecture,

$$\Gamma^\alpha_{0\alpha} = \frac{\dot{a}}{a} \delta^\alpha_\alpha = 3\frac{\dot{a}}{a}, \quad \Gamma^0_{\alpha\beta} = a\dot{a}\gamma_{\alpha\beta},$$

and the other  $\Gamma$ s vanishing. This gives

$$(\nabla_a T)^{ab} u_b = \dot{\rho}(t) + 3\frac{\dot{a}}{a}\rho(t) + a\dot{a}\gamma_{\alpha\beta}T^{\alpha\beta}.$$

Then using  $T^{\alpha\beta} = p(t)$  and  $g^{\alpha\beta} = \frac{1}{a^2}\gamma^{\alpha\beta}$  we get

$$0 = \dot{\rho}(t) + 3\frac{\dot{a}}{a}\rho(t) + \frac{\dot{a}}{a}\gamma_{\alpha\beta}\gamma^{\alpha\beta}p(t),$$

which gives the result.

**Question:** For  $p(t) = \omega\rho(t)$ , solve the conservation equation above for  $\rho$  and use the Friedmann equation with vanishing spatial curvature

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho,$$

to derive an autonomous differential equation for  $a$ .

**Solution:** We have

$$\begin{aligned}\dot{\rho}(t) &= -3\frac{\dot{a}(t)}{a(t)}(1+\omega)\rho(t) \\ \frac{\dot{\rho}(t)}{\rho(t)} &= -3\frac{\dot{a}(t)}{a(t)}(1+\omega) \\ \frac{d}{dt}\ln(\rho(t)) &= -3(1+\omega)\frac{d}{dt}\ln(a(t)) \\ \rho(t) &= Ba^{-3(1+\omega)},\end{aligned}$$

for some constant  $B = e^A$ , where  $A$  is the constant of integration. If we plug this into the expression given in the question we have

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi GB}{3}a^{-3(1+\omega)},$$

which is an autonomous differential equation for  $a$ .

**Question:** Show that

$$a(t) = C \cdot t^\alpha, \quad C = \text{const}$$

solves the autonomous differential equation equation for the scale factor  $a$  if  $\omega \neq 1$  and a suitably chosen  $\alpha$ .

**Solution:** By direct calculation, we have

$$\alpha^2 t^{-2} = \frac{8\pi GBC}{3}t^{-3\alpha(1+\omega)},$$

from which we see

$$\alpha = \frac{2}{3(1+\omega)}.$$

**Question:** Use the result of the previous question to write down an equation for  $H(t) := \frac{\dot{a}}{a}$  and estimate the age of the universe only filled with dust for today's value of the Hubble constant being given by  $\frac{1}{H_0} \approx 13 \times 10^9 \text{ yrs}$ . Repeat the calculation for a universe containing only radiation.

**Solution:** We have

$$a(t) = C \cdot t^{\frac{2}{3(1+\omega)}}, \quad \implies \quad H(t) = \frac{2}{3(1+\omega)} t^{-1}.$$

So the age is given by

$$t_0 = \frac{2}{3(1+\omega)} \frac{1}{H_0}.$$

For dust  $\omega = 0$  and so we have

$$t_0^{\text{dust}} \approx 8.6 \times 10^9 \text{ yrs}.$$

For radiation,  $\omega = \frac{1}{3}$ , giving

$$t_0^{\text{radiation}} \approx 6.5 \times 10^9 \text{ yrs}.$$

**Question:** Consider a universe filled with only one type of matter characterised by a linear equation of state with constant  $\omega$ . For which values of the latter is the expansion of the universe accelerating?

**Solution:** We have

$$\ddot{a}(t) = \left( \frac{2}{3(1+\omega)} \right) \left( \frac{2}{3(1+\omega)} - 1 \right) \cdot a \cdot t^{-2}.$$

An expanding universe means  $a(t) > 0$  (or, equivalently,  $C > 0$ ), and so the turning point for accelerated expansion is the condition

$$\frac{2}{3(1+\omega)} - 1 = 0 \quad \implies \quad \omega = -\frac{1}{3}.$$

We therefore get accelerated expansion for  $\omega < -\frac{1}{3}$  and decelerated expansion for  $\omega > -\frac{1}{3}$ .

## 26.23 Diagrams

### 26.23.1 Penrose Diagram Of A Radiation-Filled Universe

**Question:** Find a differential equation for radial null geodesics in a spatially flat FRW universe filled with radiation, using the chart  $(t, r, \vartheta, \phi)$  introduced in the lectures. Explicitly write down the precise range of the chart variables.



**Solution:** A spatially flat FRW universe has  $\kappa = 0$  and a metric with components

$$g_{ab}(t, r, \vartheta, \varphi) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a^2(t) & 0 & 0 \\ 0 & 0 & a^2(t)r^2 & 0 \\ 0 & 0 & 0 & a^2(t)r^2 \sin^2 \vartheta \end{pmatrix}_{ab},$$

in the chart given. If we have a radiation-filled universe, we have  $\omega = -\frac{1}{3}$ , and from the last tutorial we have

$$a(t) = C \cdot t^{1/2}.$$

Radially null geodesics have  $\dot{\vartheta} = 0 = \dot{\varphi}$ , so our Lagrangian reads

$$0 = -\dot{t}^2 + a^2 \cdot \dot{r}^2,$$

which, subbing in our  $a(t)$  expression and using the chain rule backwards, gives the differential equation

$$\frac{dt}{dr} = \pm C\sqrt{t}.$$

We have seen that in the FRW universe, there is a beginning time (the Big Bang), and so our  $t$  coordinate must be lower bounded. We can choose to parameterise it such that  $t = 0$  is the Big Bang value, giving us the coordinate ranges

$$t \in (0, \infty), \quad r \in (0, \infty), \quad \vartheta \in (0, \pi), \quad \text{and} \quad \phi \in (0, 2\pi),$$

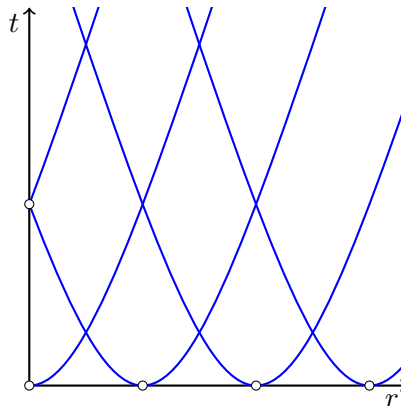
where the 0 point is removed from  $t$ 's range as it is not actually a point in our spacetime (it's a singularity).

**Question:** Determine the  $t$ -coordinate of a geodesic in terms of the  $r$  coordinate. Draw some of the null geodesics in the underlying chart.

**Solution:** Solving the differential equation from the previous question gives us

$$t_{\pm}(r) = \frac{1}{4}(A \pm Cr)^2,$$

for some integration constant  $A$ . So we have a series of squared curves, shifted for different values of  $A$ .



The white circles indicated that these points are not part of our diagram as  $t$  and  $r$  cannot take the value 0. As these parts are not included, the lines to either side actually represent two separate geodesics, just as we saw with the Schwarzschild drawings in the lectures.

**Question:** Find a chart in which the geodesics are lines of constant slope  $\pm 1$ . Determine the range of the coordinates.

**Solution:** We can rearrange the expression for  $t_{\pm}$  in terms of  $r$  to give

$$r = \pm \frac{2}{C} \sqrt{t_{\pm}} - \frac{A}{C}, \quad r = \mp \frac{2}{C} \sqrt{t_{\pm}} + \frac{A}{C}.$$

If we then define

$$\bar{t}_{\pm} := \frac{2}{C} \sqrt{t_{\pm}},$$

we get

$$r = \pm \bar{t}_{+} - B, \quad r = \mp \bar{t}_{-} + B,$$

where  $B = \frac{A}{C}$ . All of these plots are just lines of constant slope  $\pm 1$ . With a bit of thought, it's clear that we can just consider either  $\bar{t}_{+}$  or  $\bar{t}_{-}$  and obtain the other results using different values of  $B$ . So we shall just write

$$r_{\pm} = \pm \bar{t} - B.$$

The ranges are

$$\bar{t} \in (0, \infty), \quad r \in (0, \infty), \quad \vartheta \in (0, \pi), \quad \text{and} \quad \phi \in (0, 2\pi),$$

**Question:** Choose the so-called null coordinates  $u$  and  $v$  in which the null geodesics of positive slope are parallel to the  $u$ -axis and the ones of negative slope are parallel to the  $v$ -axis. Determine the range of the coordinates.

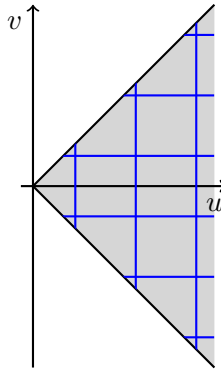
**Solution:** We define

$$u := \bar{t} + r, \quad \text{and} \quad v := \bar{t} - r,$$

and their ranges are

$$u \in (0, \infty), \quad \text{and} \quad v \in (-\infty, \infty).$$

We then have the conditions  $u + v = 2\bar{t} > 0$  and  $u - v = 2r > 0$ , so we need to exclude the regions  $u < 0$ ,  $v + u < 0$  and  $u - v < 0$  from our diagram. We therefore get



where the shaded area is the only area we consider.

The blue lines are the geodesics in this chart. Note that by going to this chart we can compactify without ruining the 90-degree nature of the cone structure. This is exactly why this step is included, and is what Remark 23.1.1 is talking about.

**Question:** Compactify, i.e., rescale to finite ranges, each of the two null coordinates by an appropriate transformation. Determine the range of the coordinates.

**Solution:** We define

$$p := \arctan(u), \quad \text{and} \quad q := \arctan(v),$$

which have ranges

$$p \in (0, \pi/2), \quad \text{and} \quad q \in (-\pi/2, \pi/2).$$

Our other range conditions still hold, namely  $p + q > 0$  and  $p - q > 0$ .

Our plot looks the same as in the previous question, apart from now the right-hand edge is bounded at value  $p = \pi/2$ .

**Question:** By final transformation, recover the notion of temporal and radial coordinates. Determine the ranges of those coordinates. Draw the Penrose-Carter diagram.

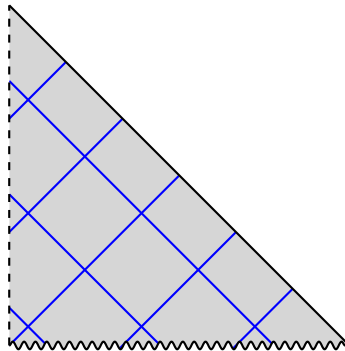
**Solution:** We define

$$T := p + q, \quad \text{and} \quad R := p - q,$$

whose ranges are

$$T \in (0, \pi), \quad \text{and} \quad R \in (0, \pi).$$

We have the further constraint  $T + R = 2U < \pi$ , so our Penrose-Carter diagram looks like



We have used the snake-like line to indicate the Big Bang singularity, and a dashed line on the left-hand side to remind us that there is nothing wrong here (i.e. lines that go off to the left just come back from the dashed line — think about rotating the diagram by reinstating  $\phi$ ).

## 26.24 Perturbation Theory

To come later.

# Useful Texts & Further Readings

## GENERAL RELATIVITY

- R. M. Wald, *General Relativity*, The University of Chicago Press, 1984.
- L. Ryder, *Introduction to General Relativity*, Cambridge University Press, 2009.
- M. P. Hobson, *General Relativity: An Introduction for Physicists*, Cambridge University Press, 2006.
- W. Rindler, *Relativity: Special, General and Cosmological*, Oxford University Press, 2006.
- S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, John Wiley and Sons, Inc., 1972.

## DIFFERENTIAL GEOMETRY

- P. Renteln, *Manifolds, Tensors, and Forms: An Introduction for Mathematicians and Physicists*, Cambridge University Press, 2014.
- L. W. Tu, *An Introduction to Manifolds* (Second addition), Springer, 2011.