Preprint typeset in JHEP style - HYPER VERSION

Michaelmas Term, 2019

# General Relativity

## University of Cambridge Part III Mathematical Tripos

**David Tong**

*Department of Applied Mathematics and Theoretical Physics,*
*Centre for Mathematical Sciences,*
*Wilberforce Road,*
*Cambridge, CB3 OBA, UK*

http://www.damtp.cam.ac.uk/user/tong/gr.html
d.tong@damtp.cam.ac.uk

**Recommended Books and Resources**

There are many decent text books on general relativity. Here are a handful that I like:

- Sean Carroll, "*Spacetime and Geometry*"

A straightforward and clear introduction to the subject.

- Bob Wald, "*General Relativity*"

The go-to relativity book for relativists.

- Steven Weinberg, "*Gravitation and Cosmology*"

The go-to relativity book for particle physicists.

- Misner, Thorne and Wheeler, "*Gravitation*"

Extraordinary and ridiculous in equal measure, this book covers an insane amount of material but with genuinely excellent explanations. Now, is that track 1 or track 2?

- Tony Zee, "*Einstein Gravity in a Nutshell*"

Professor Zee likes a bit of a chat. So settle down, prepare yourself for more tangents than $T_p(M)$, and enjoy this entertaining, but not particularly concise, meander through the subject.

- Nakahara, "*Geometry, Topology and Physics*"

A really excellent book that will satisfy your geometrical and topological needs for this course and much beyond. It is particularly useful for Sections 2 and 3 of these lectures where we cover differential geometry.

A number of excellent lecture notes are available on the web, including an early version of Sean Carroll's book. Links can be found on the course webpage: http://www.damtp.cam.ac.uk/user/tong/gr.html.

## Contents


**0. Introduction** **1**

**1. Geodesics in Spacetime** **6**
- 1.1 Non-Relativistic Particles 7
  - 1.1.1 The Geodesic Equation 8
- 1.2 Relativistic Particles 11
  - 1.2.1 A Particle in Minkowski Spacetime 12
  - 1.2.2 Why You Get Old 14
  - 1.2.3 Rediscovering the Forces of Nature 16
  - 1.2.4 The Equivalence Principle 19
  - 1.2.5 Gravitational Time Dilation 24
  - 1.2.6 Geodesics in Spacetime 26
- 1.3 A First Look at the Schwarzchild Metric 29
  - 1.3.1 The Geodesic Equations 30
  - 1.3.2 Planetary Orbits in Newtonian Mechanics 32
  - 1.3.3 Planetary Orbits in General Relativity 34
  - 1.3.4 The Pull of Other Planets 39
  - 1.3.5 Light Bending 43

**2. Introducing Differential Geometry** **49**
- 2.1 Manifolds 49
  - 2.1.1 Topological Spaces 50
  - 2.1.2 Differentiable Manifolds 51
  - 2.1.3 Maps Between Manifolds 55
- 2.2 Tangent Spaces 56
  - 2.2.1 Tangent Vectors 56
  - 2.2.2 Vector Fields 61
  - 2.2.3 Integral Curves 63
  - 2.2.4 The Lie Derivative 65
- 2.3 Tensors 69
  - 2.3.1 Covectors and One-Forms 69
  - 2.3.2 The Lie Derivative Revisited 71
  - 2.3.3 Tensors and Tensor Fields 72
- 2.4 Differential Forms 76

2.4.1 The Exterior Derivative 77
2.4.2 Forms You Know and Love 80
2.4.3 A Sniff of de Rham Cohomology 83
2.4.4 Integration 85
2.4.5 Stokes' Theorem 87

**3. Introducing Riemannian Geometry 89**
3.1 The Metric 89
3.1.1 Riemannian Manifolds 90
3.1.2 Lorentzian Manifolds 91
3.1.3 The Joys of a Metric 92
3.1.4 A Sniff of Hodge Theory 97
3.2 Connections and Curvature 98
3.2.1 The Covariant Derivative 99
3.2.2 Torsion and Curvature 103
3.2.3 The Levi-Civita Connection 106
3.2.4 The Divergence Theorem 109
3.2.5 The Maxwell Action 111
3.3 Parallel Transport 116
3.3.1 Geodesics Revisited 117
3.3.2 Normal Coordinates 117
3.3.3 Path Dependence: Curvature and Torsion 121
3.3.4 Geodesic Deviation 125
3.4 More on the Riemann Tensor and its Friends 127
3.4.1 The Ricci and Einstein Tensors 129
3.4.2 Connection 1-forms and Curvature 2-forms 130
3.4.3 An Example: the Schwarzchild Metric 134
3.4.4 The Relation to Yang-Mills Theory 136

**4. The Einstein Equations 138**
4.1 The Einstein-Hilbert Action 138
4.1.1 An Aside on Dimensional Analysis 142
4.1.2 The Cosmological Constant 143
4.1.3 Diffeomorphisms Revisited 144
4.2 Some Simple Solutions 147
4.2.1 de Sitter Space 148
4.2.2 Anti-de Sitter Space 153
4.3 Symmetries 157

4.3.1 Isometries 158
4.3.2 A First Look at Conserved Quantities 161
4.4 Asymptotics of Spacetime 163
4.4.1 Conformal Transformations 163
4.4.2 Penrose Diagrams 165
4.5 Coupling Matter 175
4.5.1 Field Theories in Curved Spacetime 175
4.5.2 The Einstein Equations with Matter 177
4.5.3 The Energy-Momentum Tensor 178
4.5.4 Perfect Fluids 180
4.5.5 The Slippery Business of Energy Conservation 181
4.5.6 Spinors 186
4.5.7 Energy Conditions 188
4.6 A Taste of Cosmology 193
4.6.1 The FRW Metric 193
4.6.2 The Friedmann Equations 196

**5. When Gravity is Weak 199**
5.1 Linearised Theory 199
5.1.1 Gauge Symmetry 201
5.1.2 The Newtonian Limit 203
5.2 Gravitational Waves 204
5.2.1 Solving the Wave Equation 205
5.2.2 Bobbing on the Waves 207
5.2.3 Exact Solutions 211
5.3 Making Waves 212
5.3.1 The Green's Function for the Wave Equation 212
5.3.2 An Example: Binary Systems 216
5.3.3 Comparison to Electromagnetism 217
5.3.4 Power Radiated: The Quadrupole Formula 219
5.3.5 Gravitational Wave Sources on the ✉ 225

**Acknowledgements**

These lectures were given to masters (Part 3) students. No prior knowledge of general relativity is assumed, but it's fair to say that you'll find the going easier if you've been exposed to the subject previously. The lectures owe a debt to previous incarnations of this course and, in particular, the excellent lectures of Harvey Reall. My thanks to Wanli Xing for superhuman typo spotting. I'm supported by the Royal Society, the Simons Foundation, and Alex Considine Tong.

**Conventions**

We use the metric with signature $(-+++)$. This is the opposite convention to my lecture notes on Special Relativity and Quantum Field Theory, but it does agree with the lecture notes on Cosmology and on String Theory. There is some mild logic behind this choice. When thinking about geometry, the choice $(-+++)$ is preferable as it ensures that length distances are positive; when thinking about quantum physics, the choice $(+---)$ is preferable as it ensures that frequencies and energies are positive. Ultimately you just need to get used to both conventions.

When dealing with physics, spacetime indices are greek $\mu, \nu = 0, 1, 2, 3$, spatial indices are roman $i, j = 1, 2, 3$.

# 0. Introduction

General relativity is the theory of space and time and gravity. The essence of the theory is simple: gravity is geometry. The effects that we attribute to the force of gravity are due to the bending and warping of spacetime, from falling cats, to orbiting spinning planets, to the motion of the cosmos on the grandest scale. The purpose of these lectures is to explain this.

Before we jump into a description of curved spacetime, we should first explain why Newton's theory of gravity, a theory which served us well for 250 years, needs replacing. The problems arise when we think about disturbances in the gravitational field. Suppose, for example, that the Sun was to explode. What would we see? Well, for 8 glorious minutes – the time that it takes light to reach us from the Sun – we would continue to bathe in the Sun's light, completely oblivious to the fate that awaits us. But what about the motion of the Earth? If the Sun's mass distribution changed dramatically, one might think that the Earth would start to deviate from its elliptic orbit. But when does this happen? Does it occur immediately, or does the Earth continue in its orbit for 8 minutes before it notices the change?

Of course, the theory of special relativity tells us the answer. Since no signal can propagate faster than the speed of light, the Earth must continue on its orbit for 8 minutes. But how is the information that the Sun has exploded then transmitted? Does the information also travel at the speed of light? What is the medium that carries this information? As we will see throughout these lectures, the answers to these questions forces us to revisit some of our most basic notions about the meaning of space and time and opens to door to some of the greatest ideas in modern physics such as cosmology and black holes.

### A Field Theory of Gravity

There is a well trodden path in physics when trying to understand how objects can influence other objects far away. We introduce the concept of a *field*. This is a physical quantity which exists everywhere in space and time; the most familiar examples are the electric and magnetic fields. When a charge moves, it creates a disturbance in the electromagnetic field, ripples which propagate through space until they reach other charges. The theory of general relativity is a relativistic field theory of gravity.

It's a simple matter to cast Newtonian gravity in terms of a field theory. A particle of mass $m$ experiences a force that can be written as

$$\mathbf{F} = -m\nabla\Phi$$

where the gravitational field $\Phi(\mathbf{r}, t)$ is governed by the surrounding matter distribution which is described by the mass density $\rho(\mathbf{r}, t)$. If the matter density is static, so that $\rho(\mathbf{r})$ is independent of time, then the gravitational field obeys

$$\nabla^2 \Phi = 4\pi G \rho \tag{0.1}$$

with Newton's constant $G$ given by

$$G \approx 6.67 \times 10^{-11} \text{ m}^3 \, \text{kg}^{-1} \, \text{s}^{-2}$$

This equation is simply a rewriting of the usual inverse square law of Newton. For example, if a mass $M$ is concentrated at a single point we have

$$\rho(\mathbf{r}) = M\delta^3(\mathbf{r}) \quad \Rightarrow \quad \Phi = -\frac{GM}{r}$$

which is the familiar gravitational field for a point mass.

The question that we would like to answer is: how should we modify (0.1) when the mass distribution $\rho(\mathbf{r}, t)$ changes with time? Of course, we could simply postulate that (0.1) continues to hold even in this case. A change in $\rho$ would then immediately result in a change of $\Phi$ throughout all of space. Such a theory clearly isn't consistent with the requirement that no signal can travel faster than light. Our goal is to figure out how to generalise (0.1) in a manner that is compatible with the postulates of special relativity.

### The Analogy with Electromagnetism

The goal that we've set ourselves above looks very similar to the problem of finding a relativistic generalization of electrostatics. After all, we learn very early in our physics lives that when objects are stationary, the force due to gravity takes exactly the same inverse-square form as the force due to electric charge. It's worth pausing to see why this analogy does not continue when objects move and the resulting Einstein equations of general relativity are considerably more complicated than the Maxwell equations of electromagnetism.

Let's start by considering the situation of electrostatics. A particle of charge $q$ experiences a force

$$\mathbf{F} = -q\nabla\phi$$

where the electric potential $\phi$ is governed by the surrounding charge distribution. Let's call the charge density $\rho_e(\mathbf{r})$ (with the subscript $e$ to distinguish it from the matter distribution). Then the electric potential is given by

$$\nabla^2 \phi_e = -\frac{\rho_e}{\epsilon_0}$$

Apart from a minus sign and a relabelling of the coupling constant ($G \to 1/4\pi\epsilon_0$), this formulation looks identical to the Newtonian gravitational potential (0.1). Yet there is a crucial difference that is all important when it comes to making these equations consistent with special relativity. This difference lies in the objects which source the potential.

For electromagnetism, the source is the charge density $\rho_e$. By definition, this is the electric charge per spatial volume, $\rho_e \sim Q/\text{Vol}$. The electric charge $Q$ is something all observers can agree on. But observers moving at different speeds will measure different spatial volumes due to Lorentz contraction. This means that $\rho_e$ is not itself a Lorentz invariant object. Indeed, in the full Maxwell equations $\rho_e$ appears as the component in a 4-vector, accompanied by the charge density current $\mathbf{j}_e$,

$$J^\mu = \begin{pmatrix} \rho_e c \\ \mathbf{j}_e \end{pmatrix}$$

If you want a heuristic argument for why the charge density $\rho_e$ is the temporal component of the 4-vector, you could think of spatial volume as a four-dimensional volume divided by time: $\text{Vol}_3 \sim \text{Vol}_4/\text{Time}$. The four-dimensional volume is a Lorentz invariant which means that under a Lorentz transformation, $\rho_e$ should change in the same way as (inverse) time.

The fact that the source $J^\mu$ is a 4-vector is directly related to the fact that the fundamental field in electromagnetism is also a 4-vector

$$A_\mu = \begin{pmatrix} \phi/c \\ \mathbf{A} \end{pmatrix}$$

where $\mathbf{A}$ a 3-vector potential. From this we can go on to construct the familiar electric and magnetic fields. More details can be found in the lectures on Electromagnetism.

Now let's see what's different in the case of gravity. The gravitational field is sourced by the mass density $\rho$. But we know that in special relativity mass is just a form of energy. This suggests, correctly, that the gravitational field should be sourced by energy density. However, in contrast to electric charge, energy is not something that all observers can agree on. Instead, energy is itself the temporal component of a 4-vector which also includes momentum. This means that if energy sources the gravitational field, then momentum must too.

Yet now we have to also take into account that it is the energy *density* and momentum *density* which are important. So each of these four components must itself be the temporal component of a four-vector! The energy density $\rho$ is accompanied by an energy density current that we'll call $\mathbf{j}$. Meanwhile, the momentum density in the $i^{\text{th}}$ direction – let's call it $p^i$ – has an associated current $\mathbf{T}^i$. These $i = 1, 2, 3$ vectors $\mathbf{T}^i$ can also be written as a $3 \times 3$ matrix $T^{ij}$. The end result is that if we want a theory of gravity consistent with special relativity, then the object that sources the gravitational field must be a $4 \times 4$ matrix, known as a tensor,

$$T^{\mu\nu} \sim \begin{pmatrix} \rho c & \mathbf{p}c \\ \mathbf{j} & T \end{pmatrix}$$

Happily, a matrix of this form is something that arises naturally in classical physics. It has different names depending on how lazy people are feeling. It is sometimes known as the *energy-momentum* tensor, sometimes as the *energy-momentum-stress* tensor or sometimes just the *stress* tensor. We will describe some properties of this tensor in Section 4.5.

In some sense, all the beautiful complications that arise in general relativity can be traced back to the fact that the source for gravity is a matrix $T^{\mu\nu}$. In analogy with electromagnetism, we may expect that the associated gravitational field is also a matrix, $h_{\mu\nu}$, and this is indeed the case. The Newtonian gravitational field $\Phi$ is merely the upper-left component of this matrix, $h_{00} \sim \Phi$.

However, not all of general relativity follows from such simple considerations. The wonderful surprise awaiting us is that the matrix $h_{\mu\nu}$ is, at heart, a geometrical object: it describes the curvature of spacetime.

#### When is a Relativistic Theory of Gravity Important

Finally, we can simply estimate the size of relativistic effects in gravity. What follows is really nothing more than dimensional analysis, with a small story attached to make it sound more compelling. Consider a planet in orbit around a star of mass $M$. If we assume a circular orbit, the speed of the planet is easily computed by equating the gravitational force with the centripetal force,

$$\frac{v^2}{r} = \frac{GM}{r^2}$$

Relativistic effects become important when $v^2/c^2$ gets close to one. This tells us that the relevant, dimensionless parameter that governs relativistic corrections to Newton's law of gravity is $\sim GM/rc^2$.

A slightly better way of saying this is as follows: the fundamental constants $G$ and $c^2$ allow us to take any mass $M$ and convert it into a distance scale. As we will see later, it is convenient to define this to be

$$R_s = \frac{2GM}{c^2}$$

This is known as the *Schwarzchild radius*. Relativistic corrections to gravity are then governed by $R_s/r$.

In most situations, relativistic corrections to the gravitational force are very small. For our planet Earth, $R_s \approx 10^{-2}$ m. The radius of the Earth is around 6000 km, which means that relativistic effects give corrections to Newtonian gravity on the surface of Earth of order $10^{-8}$. Satellites orbit at $R_s/r \approx 10^{-9}$. These are small numbers. For the Sun, $R_s \approx 3$ km. At the surface of the run, $r \approx 7 \times 10^5$ km, and $R_s/r \approx 10^{-6}$. Meanwhile, the typical distance of the inner planets is $\sim 10^8$ km, giving $R_s/r \approx 10^{-8}$. Again, these are small numbers. Nonetheless, in both cases there are beautiful experiments that confirm the relativistic theory of gravity. We shall meet some of these as we proceed.

There are, however, places in Nature where large relativistic effects are important. One of the most striking is the phenomenon of black holes. As observational techniques improve, we are gaining increasingly more information about these most extreme of environments.

# 1. Geodesics in Spacetime

Classical theories of physics involve two different objects: particles and fields. The fields tell the particles how to move, and the particles tell the fields how to sway. For each of these, we need a set of equations.

In the theory of electromagnetism, the swaying of the fields is governed by the Maxwell equations, while the motion of test particles is dictated by the Lorentz force law. Similarly, for gravity we have two different sets of equations. The swaying of the fields is governed by the *Einstein equations*, which describe the bending and curving of spacetime. We will need to develop some mathematical machinery before we can describe these equations; we will finally see them in Section 4.

Our goal in this section is to develop the analog of the Lorentz force law for gravity. As we will see, this is the question of how test particles move in a fixed, curved spacetime. Along the way, we will start to develop some language to describe curved spacetime. This will sow some intuition which we will then make mathematically precise in later sections.

### The Principle of Least Action

Our tool of choice throughout these lectures is the action. The advantage of the action is that it makes various symmetries manifest. And, as we shall see, there are some deep symmetries in the theory of general relativity that must be maintained. This greatly limits the kinds of equations which we can consider and, ultimately, will lead us inexorably to the Einstein equations.

We start here with a lightening review of the principle of least action. (A more detailed discussion can be found in the lectures on Classical Dynamics.) We describe the position of a particle by coordinates $x^i$ where, for now, we take $i = 1, 2, 3$ for a particle moving in three-dimensional space. Importantly, there is no need to identify the coordinates $x^i$ with the $(x, y, z)$ axes of Euclidean space; they could be any coordinate system of your choice.

We want a way to describe how the particle moves between fixed initial and final positions,

$$x^i(t_1) = x^i_{\text{initial}} \qquad \text{and} \qquad x^i(t_2) = x^i_{\text{final}} \tag{1.1}$$

To do this, we consider all possible paths $x^i(t)$, subject to the boundary conditions above. To each of these paths, we assign a number called the *action* $S$. This is defined as

$$S[x^i(t)] = \int_{t_1}^{t_2} dt\ L(x^i(t), \dot{x}^i(t))$$

where the function $L(x^i, \dot{x}^i)$ is the *Lagrangian* which specifies the dynamics of the system. The action is a functional; this means that you hand it an entire function worth of information, $x^i(t)$, and it spits back only a single number.

The *principle of least action* is the statement that the true path taken by the particle is an extremum of $S$. Although this is a statement about the path as a whole, it is entirely equivalent to a set of differential equations which govern the dynamics. These are known as the *Euler-Lagrange equations*.

To derive the Euler-Lagrange equations, we think about how the action changes if we take a given path and vary it slightly,

$$x^i(t) \to x^i(t) + \delta x^i(t)$$

We need to keep the end points of the path fixed, so we demand that $\delta x^i(t_1) = \delta x^i(t_2) = 0$. The change in the action is then

$$\begin{aligned}\delta S = \int_{t_1}^{t_2} dt\ \delta L &= \int_{t_1}^{t_2} dt\ \left(\frac{\partial L}{\partial x^i}\delta x^i + \frac{\partial L}{\partial \dot{x}^i}\delta \dot{x}^i\right) \\ &= \int_{t_1}^{t_2} dt\ \left(\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\right)\delta x^i + \left[\frac{\partial L}{\partial \dot{x}^i}\delta x^i\right]_{t_1}^{t_2}\end{aligned}$$

where we have integrated by parts to go to the second line. The final term vanishes because we have fixed the end points of the path. A path $x^i(t)$ is an extremum of the action if and only if $\delta S = 0$ for all variations $\delta x^i(t)$. We see that this is equivalent to the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0 \qquad (1.2)$$

Our goal in this section is to write down the Lagrangian and action which govern particles moving in curved space and, ultimately, curved spacetime.

## 1.1 Non-Relativistic Particles

Let's start by forgetting about special relativity and spacetime and focus instead on the non-relativistic motion of a particle in curved space. Mathematically, these spaces are known as manifolds, and the study of curved manifolds is known as Riemannian geometry. However, for much of this section we will dispense with any formal mathematical definitions and instead focus attention on the physics.

### 1.1.1 The Geodesic Equation

We begin with something very familiar: the non-relativistic motion of a particle of mass $m$ in flat Euclidean space $\mathbf{R}^3$. For once, the coordinates $x^i = (x, y, z)$ actually are the usual Cartesian coordinates. The Lagrangian that describes the motion is simply the kinetic energy,

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \tag{1.3}$$

The Euler-Lagrange equations (1.2) applied to this Lagrangian simply tell us that $\ddot{x}^i = 0$, which is the statement that free particles move at constant velocity in straight lines.

Now we want to generalise this discussion to particles moving on a curved space. First, we need a way to describe curved space. We will develop the relevant mathematics in Sections 2 and 3 but here we offer a simple perspective. We describe curved spaces by specifying the infinitesimal distance between any two points, $x^i$ and $x^i + dx^i$, known as the *line element*. The most general form is

$$ds^2 = g_{ij}(x)\, dx^i dx^j \tag{1.4}$$

Where the $3 \times 3$ matrix $g_{ij}$ is called the *metric*. The metric is symmetric: $g_{ij} = g_{ji}$ since the anti-symmetric part drops out of the distance when contracted with $dx^i dx^j$. We further assume that the metric is positive definite and non-degenerate, so that its inverse exists. The fact that $g_{ij}$ is a function of the coordinates $x$ simply tells us that the distance between the two points $x^i$ and $x^i + dx^i$ depends on where you are.

Before we proceed, a quick comment: it matters in this subject whether the indices $i, j$ are up or down. We'll understand this better in Section 2 but, for now, remember that coordinates have superscripts while the metric has two subscripts.

We'll see plenty of examples of metrics in this course. Before we introduce some of the simpler metrics, let's first push on and understand how a particle moves in the presence of a metric. The Lagrangian governing the motion of the particle is the obvious generalization of (1.3)

$$L = \frac{m}{2}\, g_{ij}(x)\dot{x}^i\dot{x}^j \tag{1.5}$$

It is a simple matter to compute the Euler-Lagrange equations (1.2) that arise from this action. It is really just an exercise in index notation and, in particular, making sure that we don't inadvertently use the same index twice. Since it's important, we proceed slowly. We have

$$\frac{\partial L}{\partial x^i} = \frac{m}{2}\frac{\partial g_{jk}}{\partial x^i}\,\dot{x}^j\dot{x}^k$$

where we've been careful to relabel the indices on the metric so that the $i$ index matches on both sides. Similarly, we have

$$\frac{\partial L}{\partial \dot{x}^i} = m g_{ik}\dot{x}^k \quad \Rightarrow \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = m\frac{\partial g_{ik}}{\partial x^j}\dot{x}^j\dot{x}^k + m g_{ik}\ddot{x}^k$$

Putting these together, the Euler-Lagrange equation (1.2) becomes

$$g_{ik}\ddot{x}^k + \left(\frac{\partial g_{ik}}{\partial x^j} - \frac{1}{2}\frac{\partial g_{jk}}{\partial x^i}\right)\dot{x}^j\dot{x}^k = 0$$

Because the term in brackets is contracted with $\dot{x}^j\dot{x}^k$, only the symmetric part contributes. We can make this obvious by rewriting this equation as

$$g_{ik}\ddot{x}^k + \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{ij}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^i}\right)\dot{x}^j\dot{x}^k = 0 \tag{1.6}$$

Finally, there's one last manoeuvre. we multiply the whole equation by the inverse metric, $g^{-1}$, so that we get an equation of the form $\ddot{x}^k = \ldots$. We denote the inverse metric $g^{-1}$ simply by raising the indices on the metric, from subscripts to superscripts. This means that the inverse metric is denoted $g^{ij}$. By definition, it satisfies

$$g^{ij}g_{jk} = \delta^i_k$$

Finally, taking the opportunity to relabel some of the indices, the equation of motion for the particle is written as

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0 \tag{1.7}$$

where

$$\Gamma^i_{jk}(x) = \frac{1}{2}g^{il}\left(\frac{\partial g_{lj}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l}\right) \tag{1.8}$$

These coefficients are called the *Christoffel symbols*. By construction, they are symmetric in their lower indicies: $\Gamma^i_{jk} = \Gamma^i_{kj}$. They will play a very important role in everything that follows. The equation of motion (1.7) is the *geodesic equation* and solutions to this equation are known as *geodesics*.

### A Trivial Example: Flat Space Again

Let's start by considering flat space $\mathbf{R}^3$. Pythagoras taught us how to measure distances using his friend, Descartes' coordinates,

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{1.9}$$

Suppose that we work in polar coordinates rather than Cartesian coordinates. The relationship between the two is given by

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}$$

In polar coordinates, the infinitesimal distance between two points can be simply derived by substituting the above relations into (1.9). A little algebra yields,

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2$$

In this case, the metric (and therefore also its inverse) are diagonal. They are

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} \quad \text{and} \quad g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & (r^2\sin^2\theta)^{-1} \end{pmatrix}$$

where the matrix components run over $i, j = r, \theta, \phi$. From this we can easily compute the Christoffel symbols. The non-vanishing components are

$$\begin{aligned} \Gamma^r_{\theta\theta} = -r \quad , \quad \Gamma^r_{\phi\phi} = -r\sin^2\theta \quad , \quad \Gamma^\theta_{\theta r} = \Gamma^\theta_{r\theta} = \frac{1}{r} \\ \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta \quad , \quad \Gamma^\phi_{\phi r} = \Gamma^\phi_{r\phi} = \frac{1}{r} \quad , \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \frac{\cos\theta}{\sin\theta} \end{aligned} \tag{1.10}$$

There are some important lessons here. First, $\Gamma \neq 0$ does not necessarily mean that the space is curved. Non-vanishing Christoffel symbolds can arise, as here, simply from a change of coordinates. As the course progresses, we will develop a diagnostic to determine whether space is really curved or whether it's an artefact of the coordinates we're using.

The second lesson is that it's often a royal pain to compute the Christoffel symbols using (1.8). If we wished, we could substitute the Christoffel symbols into the geodesic equation (1.7) to determine the equations of motion. However, it's typically easier to revert back to the original action and determine the equations of motion directly. In the present case, we have

$$S = \frac{m}{2}\int dt\ \left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right) \tag{1.11}$$

and the resulting Euler-Lagrange equations are

$$\ddot{r} = r\dot{\theta}^2 + r\sin^2\theta\dot{\phi}^2 \quad , \quad \frac{d}{dt}(r^2\dot{\theta}) = r^2\sin\theta\cos\theta\dot{\phi}^2 \quad , \quad \frac{d}{dt}(r^2\sin^2\theta\dot{\phi}) = = 0 \tag{1.12}$$

These are nothing more than the equations for a straight line described in polar coordinates. The quickest way to extract the Christoffel symbols is usually to compute the equations of motion from the action, and then compare them to the geodesic equation (1.7), taking care of the symmetry properties along the way.

#### A Slightly Less Trivial Example: $\mathbf{S}^2$

The above description of $\mathbf{R}^3$ in polar coordinates allows us to immediately describe a situation in which the space is truly curved: motion on the two-dimensional sphere $\mathbf{S}^2$. This is achieved simply by setting the radial coordinate $r$ to some constant value, say $r = R$. We can substitute this constraint into the action (1.11) to get the action for a particle moving on the sphere,

$$S = \frac{mR^2}{2}\int dt\ \left(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2\right)$$

Similarly, the equations of motion are given by (1.12), with the restriction $r = R$ and $\dot{r} = 0$. The solutions are great circles, which are geodesics on the sphere. To see this in general is a little complicated, but we can use the rotational invariance to aid us. We rotate the sphere to ensure that the starting point is $\theta_0 = \pi/2$ and the initial velocity is $\dot{\theta} = 0$. In this case, it is simple to check that solutions take the form $\theta = \pi/2$ and $\phi = \Omega t$ for some $\Omega$, which are great circles running around the equator.

### 1.2 Relativistic Particles

Having developed the tools to describe motion in curved space, our next step is to consider the relativistic generalization to curved spacetime. But before we get to this, we first need to see how to extend the Lagrangian method to be compatible with special relativity. An introduction to special relativity can be found in the lectures on Dynamics and Relativity.

### 1.2.1 A Particle in Minkowski Spacetime

Let's start by by considering a particle moving in Minkowski spacetime $\mathbf{R}^{1,3}$. We'll work with Cartesian coordinates $x^\mu = (ct, x, y, z)$ and the Minkowski metric

$$\eta_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$$

This distance between two neighbouring points labelled by $x^\mu$ and $x^\mu + dx^\mu$ is then given by

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$$

Pairs of points with $ds^2 < 0$ are said to be *timelike separated*; those for which $ds^2 > 0$ are *spacelike separated*; and those for which $ds^2 = 0$ are said to be *lightlike* separated or, more commonly, *null*.

Consider the path of a particle through spacetime. In the previous section, we labelled positions along the path using the time coordinate $t$ for some inertial observer. But to build a relativistic description of the particle motion, we want time to sit on much the same footing as the spatial coordinates. For this reason, we will introduce a new parameter – let's call it $\sigma$ – which labels where we are along the worldline of the trajectory. For now it doesn't matter what parameterisation we choose; we will only ask that $\sigma$ increases monotonically along the trajectory. We'll label the start and end points of the trajectory by $\sigma_1$ and $\sigma_2$ respectively, with $x^\mu(\sigma_1) = x^\mu_{\text{initial}}$ and $x^\mu(\sigma_2) = x^\mu_{\text{final}}$.

The action for a relativistic particle has a nice geometric interpretation: it extremises the distance between the starting and end points in Minkowski space. A particle with rest mass $m$ follows a timelike trajectory, for which any two points on the curve have $ds^2 < 0$. We therefore take the action to be

$$\begin{aligned} S &= -mc \int_{x_{\text{initial}}}^{x_{\text{final}}} \sqrt{-ds^2} \\ &= -mc \int_{\sigma_1}^{\sigma_2} d\sigma \, \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma}} \end{aligned} \tag{1.13}$$

The coefficients in front ensure that the action has dimensions $[S] = \text{Energy} \times \text{Time}$ as it should. (The action always has the same dimensions as $\hbar$. If you work in units with $\hbar = 1$ then the action should be dimensionless.)

The action (1.13) has two different symmetries, with rather different interpretations.

- Lorentz Invariance: Recall that a Lorentz transformation is a rotation in spacetime. This acts as

$$x^\mu \to \Lambda^\mu{}_\rho x^\rho \tag{1.14}$$

where the matrix $\Lambda^\mu{}_\nu$ obeys $\Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}$, which is the definition of a Lorentz transformation, encompassing both rotations in space and boosts. Equivalently, $\Lambda \in O(1,3)$. This is a symmetry in the sense that if we find a solution to the equations of motion, then we can act with a Lorentz transformation to generate a new solution.

- Reparameterisation invariance: We introduced $\sigma$ as an arbitrary parameterisation of the path. But we don't want the equations of motion to depend on this choice. Thankfully all is good, because the action itself does not depend on the choice of parameterisation. To see this, suppose that we picked a different parameterisation of the path, $\tilde{\sigma}$, related to the first parameterization by a monotonic function $\tilde{\sigma}(\sigma)$. Then we could equally as well construct an action $\tilde{S}$ using this new parameter, given by

$$\begin{aligned}
\tilde{S} &= -m \int_{\tilde{\sigma}_1}^{\tilde{\sigma}_2} d\tilde{\sigma} \, \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\tilde{\sigma}} \frac{dx^\nu}{d\tilde{\sigma}}} \\
&= -m \int_{\sigma_1}^{\sigma_2} d\sigma \, \frac{d\tilde{\sigma}}{d\sigma} \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} \left(\frac{d\sigma}{d\tilde{\sigma}}\right)^2} \\
&= S
\end{aligned}$$

As promised, the action takes the same form regardless of whether we choose to parameterise the path in terms of $\sigma$ or $\tilde{\sigma}$. This is *reparameterisation invariance.*

This is not a symmetry, in the sense that it does not generate new solutions from old ones. Instead, it is a redundancy in the way we describe the system. It is similar to the gauge "symmetry" of Maxwell and Yang-Mills theory which, despite the name, is also a redundancy rather than a symmetry.

It is hard to overstate the importance of the concept of reparameterisation invariance. A major theme of these lectures is that our theories of physics should not depend on the way we choose to parameterise them. We'll see this again when we come to describe the field equations of general relativity. For now, we'll look at a couple of implications of reparameterisation on the worldline.

### Proper Time

Because the action is independent of the parameterisation of the worldline, the value of the action evaluated between two points on a given path has an intrinsic meaning. We call this value *proper time*. For a given path $x^\mu(\sigma')$, the proper time between two points, say $\sigma' = 0$ and $\sigma' = \sigma$, is

$$\tau(\sigma) = \frac{1}{c}\int_0^\sigma d\sigma' \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{d\sigma'}\frac{dx^\nu}{d\sigma'}} \tag{1.15}$$

From our first foray into Special Relativity, we recognise this as the time experienced by the particle itself.

Identifying the action with the proper time means that the particle takes a path that extremises the proper time. In Minkowski space, it is simple to check that the proper time between two timelike-separated points is maximised by a straight line, a fact known as the twin paradox.

### 1.2.2 Why You Get Old

There's a crucial difference between moving in Euclidean space and moving in Minkowski spacetime. You're not obliged to move in Euclidean space. You can just stop if you want to. In contrast, you can never stop moving in a timelike direction in Minkowski spacetime. You will, sadly, always be dragged inexorably towards the future.

Any relativistic formulation of particle mechanics must capture this basic fact. To see how it arises from the action (1.13), we can compute the momentum conjugate to $x^\mu$,

$$p_\mu = \frac{dL}{d\dot{x}^\mu} \tag{1.16}$$

with $\dot{x}^\mu = dx^\mu/d\sigma$. For the action $L = mc\sqrt{-\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}$, we have

$$p_\mu = \frac{mc}{L}\eta_{\mu\nu}\dot{x}^\nu \tag{1.17}$$

But not all four components of the momentum are independent. To see this, we need only compute the square of the 4-momentum to find

$$p \cdot p \equiv \eta_{\mu\nu}p^\mu p^\nu = \frac{m^2c^2}{L^2}\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = -1 \tag{1.18}$$

Rearranging gives

$$(p^0)^2 = \mathbf{p}^2 + 1$$

In particular, we see that we must have $p^0 \neq 0$: the particle is obliged to move in the time direction.

Part of this story is familiar. The condition (1.18) is closely related to the usual condition on the 4-momentum that we met in our earlier lectures on Special Relativity. There, we defined the 4-velocity $U^\mu$ and 4-momentum $P^\mu$ as

$$U^\mu = \frac{dx^\mu}{d\tau} \quad \text{and} \quad P^\mu = m\frac{dx^\mu}{d\tau}$$

This is a special case of (1.16), where we choose to parameterise the worldline by the proper time $\tau$ itself. The definition of the proper time (1.15) means that $d\tau/d\sigma = L/mc^2$. Comparing to the canonical momentum (1.16), we learn that it differs from our previous definition of 4-momentum only by an overall scaling: $P^\mu = mcp^\mu$.

However, part of this story is likely unfamiliar. Viewed from the perspective of classical dynamics, it is perhaps surprising to see that the momenta $p^\mu$ are not all independent. After all, this didn't arise in any of the examples of Lagrangians that we met in our previous course on Classical Dynamics. This novel feature can be traced to the existence of reparameterisation invariance, meaning that there was a redundancy in our original descirption. Indeed, whenever theories have such a redundancy there will be some constraint analogous to (1.18). (In the context of electromagnetism, this constraint is called Gauss law.)

There is another way to view this. The relativistic action (1.13) appears to have four dynamical degrees of freedom, $x^\mu(\sigma)$. This should be contrasted with the three degrees of freedom $x^i(t)$ in the non-relativistic action (1.5). Yet the number of degrees of freedom is one of the most basic ways to characterise a system, with physical consequences such as the heat capacity of gases. Why should we suddenly increase the number of degrees of freedom just because we want our description to be compatible with special relativity? The answer is that, because of reparameterisation invariance, not all four degrees of freedom $x^\mu$ are physical. To see this, suppose that you solve the equations of motion to find the path $x^\mu(\sigma)$ (as we will do shortly). In most dynamical systems, each of these four functions would tell you something about the physical trajectory. But, for us, reparameterisation invariance means that there no actual information in the value of $\sigma$. To find the physical path, we should eliminate $\sigma$ to find the relationship between the $x^\mu$. The net result is that the relativistic system only has three physical degrees of freedom after all.

As an example, we are perfectly at liberty to choose the parameterisation of the path to coincide with the time $t$ for some *inertial* observer: $\sigma = t$. The action (1.13) then becomes

$$S = -mc^2 \int_{t_1}^{t_2} dt \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} \tag{1.19}$$

where here $\dot{\mathbf{x}} = d\mathbf{x}/dt$. This is the action for a relativistic particle in some particular inertial frame, which exhibits the famous factor

$$\gamma = \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}}$$

that is omnipresent in formulae in special relativity. We now see clearly that the action has only three degrees of freedom, $\mathbf{x}(t)$. However, the price we've paid is that the Lorentz invariance (1.14) is now rather hidden, since space $\mathbf{x}$ and time $t$ sit on very different footing.

### 1.2.3 Rediscovering the Forces of Nature

So far, we've only succeeded in writing down the action for a free relativistic particle (1.13). We would now like to add some extra terms to the action to describe a force acting on the particle. In the non-relativistic context, we do this by adding a potential

$$S_{\rm non-rel} = \int dt\ \frac{m}{2}\dot{\mathbf{x}}^2 - V(\mathbf{x})$$

However, now we want to write down an action for a relativistic particle that depends on $x^\mu(\sigma)$. But it's crucial that we retain reparameterisation invariance, since we want to keep the features that this brings. This greatly limits the kind of terms that we can add to the action. It turns out that there are two, different ways to introduce forces that preserve our precious reparameterisations.

#### Rediscovering Electromagnetism

Rather than jumping straight into the reparameterisation invariant action (1.13), we instead start by modifying the action (1.19). We'll then try to guess a reparameterisation invariant form which gives the answer we want. To this end, we consider

$$S_1 = \int_{t_1}^{t_2} dt\ \left[ -mc^2\sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} - V(\mathbf{x}) \right]$$

and ask: how can this come from a reparameterisation invariant action?

We can't just add a term $\int d\sigma\ V(x)$ to the relativistic action (1.13); this is not invariant under reparameterisations. To get something that works, we have to find a way to cancel the Jacobian factor that comes from reparameterisations of the $d\sigma$ measure. One option that we could explore is to introduce a term linear in $\dot{x}^\mu$. But then, to preserve Lorentz invariance, we need to contract the $\mu$ index on $\dot{x}^\mu$ with something. This motivates us to introduce four functions of the spacetime coordinates $A_\mu(x)$. We then write the action

$$S_1 = \int_{\sigma_1}^{\sigma_2} d\sigma \;\left[-mc\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}} - qA_\mu(x)\dot{x}^\mu\right] \tag{1.20}$$

where $q$ is some number, associated to the particle, that characterises the strength with which it couples to the new term $A_\mu(x)$. It's simple to check that the action (1.20) does indeed have reparameterisation invariance.

To understand the physics of this new term, we again pick the worldline parameter to coincide with the time of some inertial observer, $\sigma = t$ so that $dx^0/d\sigma = c$. If we write $A_\mu(x) = (\phi(x)/c, \mathbf{A}(x))$, then we find

$$S_1 = \int_{\sigma_1}^{\sigma_2} d\sigma \;\left[-mc^2\sqrt{1-\frac{\dot{\mathbf{x}}^2}{c^2}} - q\phi(x) - q\mathbf{A}(x)\cdot\dot{\mathbf{x}}\right]$$

We see that the $A_0$ term gives us a potential $V(x) = q\phi(x)$ of the kind we wanted. But Lorentz invariance means that this is accompanied by an additional $\mathbf{A}\cdot\dot{\mathbf{x}}$ term. We have, of course, met both of these terms previously: they describe a particle of electric charge $q$ moving in the background of an electromagnetic field described by gauge potentials $\phi(x)$ and $\mathbf{A}(x)$. In other words, we have rediscovered the Lorentz force law of electromagnetism.

There is a slight generalisation of this argument, in which the particle carries some extra internal degrees of freedom, that results in the mathematical structure of Yang-Mills, the theory that underlies the weak and strong nuclear force. You can read more about this in the lecture notes on Gauge Theory.

**Rediscovering Gravity**

To describe the force of gravity, we must make a rather different modification to our action. This time we consider the generalisation of (1.19) given by the action

$$S_2 = \int_{t_1}^{t_2} dt \;\left[-mc^2\sqrt{1+\frac{2\Phi(\mathbf{x})}{c^2} - \frac{\dot{\mathbf{x}}^2}{c^2}}\right] \tag{1.21}$$

If we Taylor expand the square-root, assuming that $|\dot{\mathbf{x}}| \ll c^2$ and that $2\Phi(\mathbf{x}) \ll c^2$, then the leading terms give

$$S_2 = \int_{t_1}^{t_2} dt \;\left[-mc^2 + \frac{m}{2}\dot{\mathbf{x}}^2 - m\Phi(\mathbf{x}) + \ldots\right] \tag{1.22}$$

The first term is an irrelevant constant. (It is the rest mass energy of the particle.) But the next two terms describe the non-relativistic motion of a particle moving in a potential $V(\mathbf{x}) = m\Phi(\mathbf{x})$.

Why should we identify this potential with the force of gravity, rather than some other random force? It's because the strength of the force is necessarily proportional to the mass $m$ of the particle, which shows up as the coefficient in the $m\Phi(\mathbf{x})$ term. This is the defining property of gravity.

In fact, something important but subtle has emerged from our simple discussion: the same mass $m$ appears in both the kinetic term and the potential term. In the framework of Newtonian mechanics there is no reason that these coefficients should be the same. Indeed, careful treatments refer to the coefficient of the kinetic term as the *inertial mass* $m_I$ and the coefficient of the potential term as the *gravitational mass* $m_G$. It is then an experimentally observed fact that

$$m_I = m_G \tag{1.23}$$

to astonishing accuracy (around $10^{-13}$). This is known as the *equivalence principle*. But our simple-minded discussion above has offered a putative explanation for the equivalence principle, since the mass $m$ sits in front of the entire action (1.21), ensuring that both terms have the same origin.

An aside: you might wonder why the function $\Phi(x)$ does not scale as, say, $1/m$, in which case the potential that arises in (1.22) would appear to be independent of $m$. This is not allowed. This is because the mass $m$ is a property of the test particle whose motion we're describing. Meanwhile the potential $\Phi(x)$ is some field set up by the background sources, and should be independent of $m$, just as $A_\mu(x)$ is independent of the charge $q$ of the test particle.

The equality (1.23) is sometimes called the weak equivalence principle. A stronger version, known as the *Einstein equivalence principle* says that in any metric there exist *local inertial frames*. This is the statement that you can always find coordinates so that, in some small patch, the metric looks like Minkowski space, and there is no way to detect the effects of the gravitational field. We will describe this more below and again in Section 3.3.2.

Finally, we ask: how can we write down a reparameterisation invariant form of the action (1.21)? To answer this, note that the 1 in $\sqrt{1+\ldots}$ came from the $\eta_{00}$ term in the action. If we want to turn this into $1 + 2\Phi(\mathbf{x})/c^2$, then we should promote $\eta_{00}$ to a function of $x$. But if we're going to promote $\eta_{00}$ to a function, we should surely do the same to all metric components. This means that we introduce a curved spacetime metric

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu$$

The metric is a symmetric $4\times4$ matrix, which means that it is specified by 10 functions. We can then write down the reparameterisation invariant action

$$S_2 = -mc\int_{\sigma_1}^{\sigma_2} d\sigma\ \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}}$$

This describes a particle moving in curved spacetime.

In general, the components of the metric will be determined by the Einstein field equations. This is entirely analogous to the way in which the gauge potential $A_\mu(x)$ in (1.20) is determined by the Maxwell equation. We will describe the Einstein equations in Section 4. However, even before we get to the Einstein equations, the story above tell us that, for weak gravitational fields where the Newtonian picture is valid, we should identify

$$g_{00}(x) \approx 1 + \frac{2\Phi(x)}{c^2} \tag{1.24}$$

where $\Phi(x)$ is the Newtonian gravitational field.

### 1.2.4 The Equivalence Principle

A consequence of the weak equivalence principle (1.23) is that it's not possible to tell the difference between constant acceleration and a constant gravitational field. Suppose, for example, that you one day wake up to find yourself trapped inside a box that looks like an elevator. The equivalence principle says that there's no way tell whether you are indeed inside an elevator on Earth, or have been captured by aliens and are now in the far flung reaches of the cosmos in a spaceship, disguised as an elevator, and undergoing constant acceleration. (Actually there are two ways to distinguish between these possibilities. One is common sense. The other is known as tidal forces and will be described below.)

Conversely, if you wake in the elevator to find yourself weightless, the equivalence principle says that there is no way to tell whether the engines on your spaceship have turned themselves off, leaving you floating in space, or whether you are still on Earth, plummeting towards certain death. Both of these are examples of inertial frames.

We can see how the equivalence principle plays out in more detail in the framework of spacetime metrics. We will construct a set of coordinates adapted to a uniformly accelerating observer. We'll see that, in these coordinates, the metric takes the form (1.24) but with a linear gravitational potential $\Phi$ of the kind that we would invoke for a constant gravitational force.

First we need to determine the trajectory of a constantly accelerating observer. This was a problem that we addressed already in our first lectures on Special Relativity (see Section 7.4.6 of those notes). Here we give a different, and somewhat quicker, derivation.

We will view things from the perspective of an inertial frame, with coordinates $(ct, x, y, z)$. The elevator will experience a constant acceleration $a$ in the $x$ direction. We want to know what this looks in the inertial frame; clearly the trajectory is not just $x = \frac{1}{2}at^2$ since this would soon exceed the speed of light. Instead we need to be more careful.

Recall that if we do a boost by $v_1$, followed by a boost by $v_2$, the resulting velocity is

$$v = \frac{v_1 + v_2}{1 + v_1 v_2/c^2}$$

This motivates us to defined the *rapidity* $\varphi$, defined in terms of the velocity $v$ by

$$v = c \tanh \varphi$$

The rapidity has the nice property that is adds linearly under successive boosts: a boost $\varphi_1$ followed by a boost $\varphi_2$ is the same as a boost $\varphi = \varphi_1 + \varphi_2$.

A constant acceleration means that the rapidity increases linearly in time, where here "time" is the accelerating observer's time, $\tau$. We have $\varphi = a\tau/c$ and so, from the perspective of the inertial frame, the velocity of the constantly-accelerating elevator

$$v(\tau) = \frac{dx}{dt} = c \tanh\left(\frac{a\tau}{c}\right)$$

To determine the relationship between the observer's time and the time $t$ in the inertial frame, we use

$$\frac{dt}{d\tau} = \gamma(\tau) = \sqrt{\frac{1}{1 - v^2/c^2}} = \cosh\left(\frac{a\tau}{c}\right) \quad \Rightarrow \quad t = \frac{c}{a}\sinh\left(\frac{a\tau}{c}\right)$$

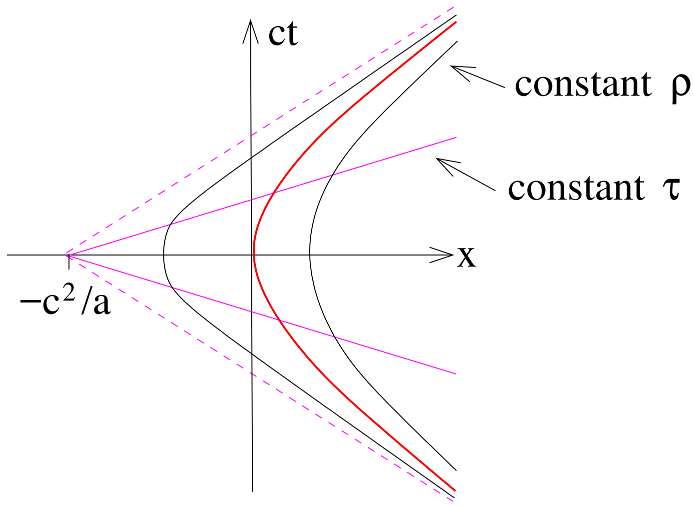



**Figure 1:** A coordinate system for a uniformly accelerating observer.

where we've chosen the integration constant so that $\tau = 0$ corresponds to $t = 0$. Then, to determine the distance travelled in the inertial frame, we use

$$v(\tau) = \frac{dx}{dt} = \frac{dx}{d\tau}\frac{d\tau}{dt} \quad \Rightarrow \quad \frac{dx}{d\tau} = c\sinh\left(\frac{a\tau}{c}\right) \quad \Rightarrow \quad x = \frac{c^2}{a}\cosh\left(\frac{a\tau}{c}\right) - \frac{c^2}{a}$$

where this time we've chosen the integration constant so that the trajectory passes through the origin. The resulting trajectory is a hyperbola in spacetime, given by

$$\left(x + \frac{c^2}{a}\right)^2 - c^2t^2 = \frac{c^4}{a^2}$$

This trajectory is shown in red in Figure 1. As $\tau \to \pm\infty$, the trajectory asymptotes to the straight lines $ct = \pm(x + c^2/a)$. These are the dotted lines shown in the figure.

Now let's consider life from the perspective of guy in the accelerating elevator. What are the natural coordinates that such an observer would use to describe events elsewhere in spacetime? Obviously, for events that happen on his own worldline, we can use the proper time $\tau$. But we would like to extend the definition to assign a time to points in the whole space. Furthermore, we would like to introduce a spatial coordinate, $\rho$, so that the elevator sits at $\rho = 0$. How to do this?

There is, it turns out, a natural choice of coordinates. First, we draw straight lines connecting the point $(ct, x) = (0, -c^2/a)$ to the point on the trajectory labelled by $\tau$ and declare that these are lines of constant $\tau$; these are the pink lines shown in the figure. Next we note that, for any given $\tau$, there is a Lorentz transformation that maps the $x$-axis to the pink line of constant $\tau$. We can use this to define the spatial coordinate $\rho$. The upshot is that we have a map between coordinates $(ct, x)$ in the inertial frame and coordinates $(c\tau, \rho)$ in the accelerating frame given by

$$\begin{aligned} ct &= \left(\rho + \frac{c^2}{a}\right)\sinh\left(\frac{a\tau}{c}\right) \\ x &= \left(\rho + \frac{c^2}{a}\right)\cosh\left(\frac{a\tau}{c}\right) - \frac{c^2}{a} \end{aligned} \tag{1.25}$$

As promised, the line $\rho = 0$ coincides with the trajectory of the accelerating observer. Moreover, lines of constant $\rho \neq 0$ are also hyperbolae.

The coordinates $(c\tau, \rho)$ do not cover all of Minkowski space, but only the right-hand quadrant as shown in Figure 1. This reflects the fact that signals from some regions will never reach the guy in the elevator. This is closely related to the idea of horizons in general relativity, a topic we'll look explore more closely in later sections.

Finally, we can look at the metric experienced by the accelerating observer, using coordinates $\rho$ and $\tau$. We simply substitute the transformation (1.25) into the Minkowski metric to find

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = -\left(1 + \frac{a\rho}{c^2}\right)^2 c^2d\tau^2 + d\rho^2 + dy^2 + dz^2$$

This is the metric of (some part of ) Minkowski space, now in coordinates adapted to an accelerating observer. These are known as *Kottler-Möller coordinates*. (They are closely related to the better known *Rindler* coordinates.) The spatial part of the metric remains flat, but the temporal component is given by

$$g_{00} = \left(1 + \frac{a\rho}{c^2}\right)^2 = 1 + \frac{2a\rho}{c^2} + \ldots$$

where the $\ldots$ is simply $a^2\rho^2/c^4$, but we've hidden it because it is sub-leading in $1/c^2$. If we compare this metric with the expectation (1.24), we see that the accelerated observer feels an effective gravitational potential given by

$$\Phi(\rho) = a\rho$$

This is the promised manifestation of the equivalence principle: from the perspective of an uniformly accelerating observer, the acceleration feels indistinguishable from a linearly increasing gravitational field, corresponding to a constant gravitational force.

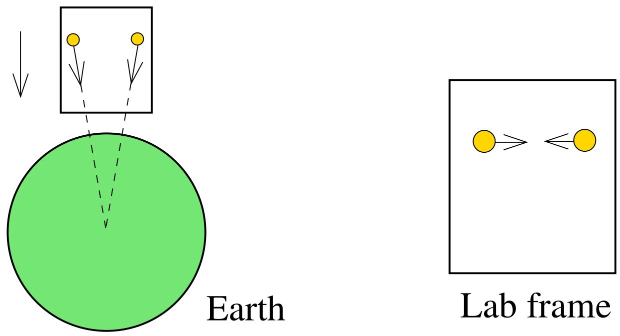



**Figure 2:** The bad news is that you were, in fact, plummeting to your death after all. This is shown on the left (not to scale). The good news is that you get to measure a tidal force on the way down. This is shown on the right.

### The Einstein Equivalence Principle

The weak equivalence principle tells us that uniform acceleration is indistinguishable from a uniform gravitational field. In particular, there is a choice of inertial frame (i.e. free-fall) in which the effect of the gravitational field vanishes. But what if the gravitational field is non-uniform?

The *Einstein equivalence principle* states that there exist *local inertial frames*, in which the effects of any gravitational field vanish. Mathematically, this means that there is always a choice of coordinates — essentially those experienced by a freely falling observer – which ensures that the metric $g_{\mu\nu}$ looks like Minkowski space about a given point. (We will exhibit these coordinates and be more precise about their properties in Section 3.3.2.) The twist to the story is that if the metric looks like Minkowski space about one point, then it probably won't look like Minkowski space about a different point. This means that if you can do experiments over an extended region of space, then you can detect the presence of non-uniform gravitational field.

To illustrate this, let's return to the situation in which you wake, weightless in an elevator, trying to figure out if you're floating in space or plummeting to your death. How can you tell?

Well, you could wait and find out. But suppose you're impatient. The equivalence principle says that there is no local experiment you can do that will distinguish between these two possibilities. But there is a very simple "non-local" experiment: just drop two test masses separated by some distance. If you're floating in space, the test masses will simply float there with you. Similarly, if you're plummeting towards your death then the test masses will plummet with you. However, they will each be attracted to the centre of the Earth which, for two displaced particles, is in a slightly different direction as shown in Figure 2. This means that the trajectories followed by the particles will slightly converge. From your perspective, this will mean that the two test masses will get closer. This is not due to their mutual gravitational attraction. (The fact they're test masses means we're ignoring this). Instead, it is an example of a *tidal force* that signifies you're sitting in a non-uniform gravitational field. We will meet the mathematics behind these tidal forces in Section 3.3.4.

### 1.2.5 Gravitational Time Dilation

Even before we solve the Einstein equations, we can still see build some intuition for the spacetime metric. As we've seen, for weak gravitational fields $\Phi(\mathbf{x})$, we should identify the temporal component of the metric as

$$g_{00}(x) = 1 + \frac{2\Phi(x)}{c^2} \tag{1.26}$$

This is telling us something profound: there is a connection between time and gravity.

To be concrete, we'll take the Newtonian potential that arises from a spherical object of mass $M$,

$$\Phi(r) = -\frac{GM}{r}$$

The resulting shift in the spacetime metric $g_{00}$ means that observer sitting at a fixed distance $r$ will measure a time interval,

$$d\tau^2 = g_{00}\, dt^2 = \left(1 - \frac{2GM}{rc^2}\right) dt^2$$

This means that if an asymptotic observer, at $r \to \infty$, measures time $t$, then an observer at distance $r$ will measure time $T$ given by

$$T(r) = t\sqrt{1 - \frac{2GM}{rc^2}}$$

We learn that time goes slower in the presence of a massive, gravitating object.

We can make this more quantitative. Consider two observers. The first, Alice, is relaxing with a picnic on the ground at radius $r_A$. The second, Bob, is enjoying a romantic trip for one in a hot air balloon, a distance $r_B = r_A + \Delta r$ higher. The time measured by Bob is

$$
\begin{aligned}
T_B = t\sqrt{1 - \frac{2GM}{(r_A + \Delta r)c^2}} &\approx t\sqrt{1 - \frac{2GM}{r_A c^2} + \frac{2GM\Delta r}{r_A^2 c^2}} \\
&\approx t\sqrt{1 - \frac{2GM}{r_A c^2}}\left(1 + \frac{GM\Delta r}{r_A^2 c^2}\right) = T_A\left(1 + \frac{GM\Delta r}{r_A^2 c^2}\right)
\end{aligned}
$$

where we've done a double expansion, assuming both $\Delta r \ll r_A$ and $2GM/r_A c^2 \ll 1$. If the hot air balloon flies a distance $\Delta r = 1000$ m above the ground then, taking the radius of the Earth to be $r_A \approx 6000$ km, the difference in times is of order $10^{-12}$. This means that, over the course of a day, Bob ages by an extra $10^{-8}$ seconds or so.

This effect was first measured by Hafele and Keating in the 1970s by flying atomic clocks around the world on commercial airlines, and has since been repeated a number of times with improved accuracy. In all cases the resultant time delay, which in the experiments includes effects from both special and general relativity, was in agreement with theoretical expectations.

The effect is more pronounced in the vicinity of a black hole. We will see in Section 1.3 that the closest distance that an orbiting planet can come to a black hole is $r = 3GM/c^2$. (Such orbits are necessarily highly elliptical.) In this case, someone on the planet experiences time at the rate $T = \sqrt{1/3}t \approx 0.6t$, compared to an asymptotic observer at $t \to \infty$. This effect, while impressive, is unlikely to make a really compelling science fiction story. For more dramatic results, our bold hero would have to fly her spaceship close to the Schwarzchild radius $R_s = 2GM/c^2$, later returning to $r \to \infty$ to find herself substantially younger than the friends and family she left behind.

### Gravitational Redshift

There is another measurable consequence of the gravitational time dilation. To see this, let's return to Alice on the ground and Bob, above, in his hot air balloon. Bob is kind of annoying and starts throwing peanuts at Alice. He throws peanuts at time intervals $\Delta T_B$. Alice receives these peanuts (now travelling at considerable speed) at time intervals $\Delta T_A$ where, as above,

$$
\Delta T_A = \sqrt{\frac{1 + 2\Phi(r_A)/c^2}{1 + 2\Phi(r_B)/c^2}}\,\Delta T_B \approx \left(1 + \frac{\Phi(r_A)}{c^2} - \frac{\Phi(r_B)}{c^2}\right)\Delta T_B
$$

We have $r_A < r_B$, so $\Phi(r_A) < \Phi(r_B) < 0$ and, hence, $\Delta T_A < \Delta T_B$. In other words, Alice receives the peanuts at a higher frequency than Bob threw them.

The story above doesn't only hold for peanuts. If Bob shines light down at Alice with frequency $\omega_B \sim 1/\Delta T_B$, then Alice will receive it at frequency $\omega_A$ given by

$$\omega_A \approx \left(1 + \frac{\Phi(r_A)}{c^2} - \frac{\Phi(r_B)}{c^2}\right)^{-1} \omega_B$$

This is a higher frequency, $\omega_A > \omega_B$, or shorter wavelength. We say that the light has been *blueshifted*. In contrast, if Alice shines light up at Bob, then the frequency decreases, and the wavelength is stretched. In this case, we say that the light has been *redshifted* as it escapes the gravitational pull. This effect was measured for the first time by Pound and Rebka in 1959, providing the first earthbound precision test of general relativity.

There is a cosmological counterpart of this result, in which light is redshifted in a background expanding space. You can read more about this in the lectures on Cosmology.

### 1.2.6 Geodesics in Spacetime

So far, we have focussed entirely on the actions describing particles, and have have yet to write down an equation of motion, let alone solve one. Now it's time to address this.

We work with the relativistic action for a particle moving in spacetime

$$S = -mc\int_{\sigma_1}^{\sigma_2} d\sigma\ L \quad \text{with} \quad L = \sqrt{-g_{\mu\nu}(x)\dot{x}^\mu\dot{x}^\nu} \tag{1.27}$$

with $\dot{x}^\mu = dx^\mu/d\sigma$. This is similar to the non-relativistic action that we used in Section 1.1.1 when we first introduced geodesics. It differs by the square-root factor. As we now see, this introduces a minor complication.

To write down Euler-Lagrange equations, we first compute

$$\frac{\partial L}{\partial x^\rho} = -\frac{1}{2L}\frac{\partial g_{\mu\nu}}{\partial x^\rho}\dot{x}^\mu\dot{x}^\nu \quad \text{and} \quad \frac{\partial L}{\partial \dot{x}^\rho} = -\frac{1}{L}g_{\rho\nu}\dot{x}^\nu$$

The equations of motion are then

$$\frac{d}{d\sigma}\left(\frac{\partial L}{\partial \dot{x}^\rho}\right) - \frac{\partial L}{\partial x^\rho} = 0 \quad \Rightarrow \quad \frac{d}{d\sigma}\left(\frac{1}{L}g_{\rho\nu}\dot{x}^\nu\right) - \frac{1}{2L}\frac{\partial g_{\mu\nu}}{\partial x^\rho}\dot{x}^\mu\dot{x}^\nu = 0$$

This is almost the same as the equations that led us to the geodesics in Section 1.1.1. There is just one difference: the differentiation $d/d\sigma$ can hit the $1/L$, giving an extra term beyond what we found previously. This can be traced directly to the fact we have a square-root in our original action.

Following the same steps that we saw in Section 1.1.1, and relabelling the indices, the equation of motion can be written as

$$g_{\mu\rho}\ddot{x}^\rho + \frac{1}{2}\left(\frac{\partial g_{\mu\rho}}{\partial x^\nu} + \frac{\partial g_{\mu\nu}}{\partial x^\rho} - \frac{\partial g_{\nu\rho}}{\partial x^\mu}\right)\dot{x}^\nu\dot{x}^\rho = \frac{1}{L}\frac{dL}{d\sigma}g_{\mu\rho}\dot{x}^\rho \tag{1.28}$$

This is the relativistic version of the geodesic equation (1.6). We see that the square-root factor in the action results in the extra term on the right-hand side.

Life would be much nicer if there was some way to ignore this extra term. This would be true if, for some reason, we could set

$$\frac{dL}{d\sigma} \stackrel{?}{=} 0$$

Happily, this is within our power. We simply need to pick a choice of parameterisation of the worldline to make it hold! All we have to do is figure out what parameterisation makes this work.

In fact, we've already met the right choice. Recall that the proper time $\tau(\sigma)$ is defined as (1.15)

$$c\tau(\sigma) = \int_0^\sigma d\sigma'\ L(\sigma') = \int_0^\sigma d\sigma'\ \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{d\sigma'}\frac{dx^\nu}{d\sigma'}} \tag{1.29}$$

This means that, by construction,

$$c\frac{d\tau}{d\sigma} = L(\sigma)$$

If we then choose to parameterise the path by $\tau$ itself, the Lagrangian is

$$L(\tau) = \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}} = \frac{d\sigma}{d\tau}\sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}} = c$$

The upshot of this discussion is that if we parameterise the worldline by proper time then $L = c$ is a constant and, in particular, $dL/d\tau = 0$. In fact this holds for any parameter related to proper time by

$$\tilde{\tau} = a\tau + b$$

with $a$ and $b$ constants. These are said to be *affine parameters* of the worldline.

Whenever we pick such an affine parameter to label the worldline of a particle, the right-hand side of the equation of motion (1.28) vanishes. In this case, we are left with the obvious extension of the geodesic equation (1.7) to curved spacetime

$$\frac{d^2x^\mu}{d\tilde{\tau}^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\tilde{\tau}}\frac{dx^\rho}{d\tilde{\tau}} = 0 \tag{1.30}$$

where the Christoffel symbols are given, as in (1.8), by

$$\Gamma^\mu_{\nu\rho}(x) = \frac{1}{2}g^{\mu\sigma}\left(\frac{\partial g_{\sigma\nu}}{\partial x^\rho} + \frac{\partial g_{\sigma\rho}}{\partial x^\nu} - \frac{\partial g_{\nu\rho}}{\partial x^\sigma}\right) \tag{1.31}$$

**A Useful Trick**

We've gone on something of a roundabout journey. We started in Section 1.1.1 with a non-relativistic action

$$S = \int dt\ \frac{m}{2}\, g_{ij}(x)\dot{x}^i\dot{x}^j$$

and found that it gives rise to the geodesic equation (1.7).

However, to describe relativistic physics in spacetime, we've learned that we need to incorporate reparameterisation invariance into our formalism resulting in the action

$$S = -mc\int d\sigma\ \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}}$$

Nonetheless, when we restrict to a very particular parameterisation – the proper time $\tau$ – we find exactly the same geodesic equation (1.30) that we met in the non-relativistic case.

This suggests something of a shortcut. If all we want to do is derive the geodesic equation for some metric, then we can ignore all the shenanigans and simply work with the action

$$S_{\text{useful}} = \int d\tau\ g_{\mu\nu}(x)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \tag{1.32}$$

This will give the equations of motion that we want, provided that they are supplemented with the constraint

$$g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = -c^2 \tag{1.33}$$

This is the requirement that the geodesic is timelike, with $\tau$ the proper time. This constraint now drags the particle into the future. Note that neither (1.32) nor (1.33) depend on the mass $m$ of the particle. This reflects the equivalence principle, which tells us that each particle, regardless of its mass, follows a geodesic.

Moreover, we can also use (1.32) to calculate the geodesic motion of light, or any other massless particle. These follow null geodesics, which means that we simply need to replace (1.33) with

$$g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0 \tag{1.34}$$

While the action $S_{\text{useful}}$ is, as the name suggests, useful, you should be cautious in how you wield it. It doesn't, as written, have the right dimensions for an action. Moreover, if you try to use it to do quantum mechanics, or statistical mechanics, then it might lead you astray unless you are careful in how you implement the constraint.

## 1.3 A First Look at the Schwarzchild Metric

Physics was born from our attempts to understand the motion of the planets. The problem was largely solved by Newton, who was able to derive Kepler's laws of planetary motion from the gravitational force law. This was described in some detail in our first lecture course on Dynamics and Relativity.

Newton's law are not the end of the story. There are relativistic corrections to the orbits of the planets that can be understood by computing the geodesics in the background of a star.

To do this, we first need to understand the metric created by a star. This will be derived in Section **??**. For now, we simply state the result: a star of mass $M$ gives rise to a curved spacetime given by

$$ds^2 = -\left(1 - \frac{2GM}{rc^2}\right)dt^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

This is the *Schwarzchild metric*. The coordinates $\theta$ and $\phi$ are the usual spherical polar coordinates, with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$.

We will have to be patient to fully understand all the lessons hiding within this metric. But we can already perform a few sanity checks. First, note that far from the star, as $r \to \infty$, it coincides with the Minkowski metric as it should. Secondly, the $g_{00}$ component is given by

$$g_{00} = 1 + \frac{2\Phi}{c^2} \quad \text{with} \quad \Phi(r) = -\frac{GM}{r}$$

which agrees with our expectation (1.24) with $\Phi = -GM/r$ the usual Newtonian potential for an object of mass $M$.

The Schwarzchild metric also has some strange things going on. In particular, the $g_{rr}$ component diverges at $r = R_s$ where

$$R_s = \frac{2GM}{c^2}$$

is called the *Schwarzchild radius*. This is the event horizon of a black hole and will be explored more fully in Section **??**. However, it turns out that space around any spherically symmetric object, such as a star, is described by the Schwarzchild metric, now restricted to $r > R_{\rm star}$, with $R_{\rm star}$ the radius of the star.

In what follows we will mostly view the Schwarzchild metric as describing the spacetime outside a star, and treat the planets as test particles moving along geodesics in this metric. We will also encounter a number of phenomenon that happen close to $r = R_s$; these are relevant only for black holes, since $R_{\rm star} \gg R_s$. However we will, for now, avoid any discussion of what happens if you venture past the event horizon.

### 1.3.1 The Geodesic Equations

Our first task is to derive the equations for a geodesic in the Schwarzchild background. To do this, we use the quick and easy method of looking at the action (1.32) for a particle moving in the Schwarzchild spacetime,

$$\begin{aligned} S_{\rm useful} = \int d\tau\ L &= \int d\tau\ g_{\mu\nu}(x)\dot{x}^\mu \dot{x}^\nu \\ &= \int d\tau\ \Big[ -A(r)c^2\dot{t}^2 + A^{-1}(r)\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2)\Big] \end{aligned} \tag{1.35}$$

with $A(r) = 1 - R_s/r$ and $\dot{x}^\mu = dx^\mu/d\tau$.

When we solved the Kepler problem in Newtonian mechanics, we started by using the conservation of angular momentum to restrict the problem to a plane. We can use the same trick here. We first look at the equation of motion for $\theta$,

$$\frac{d}{d\tau}\left(\frac{dL}{d\dot{\theta}}\right) - \frac{dL}{d\theta} = 0 \quad \Rightarrow \quad \frac{d}{d\tau}(r^2\dot{\theta}) = r^2\sin\theta\cos\theta\,\dot{\phi}^2$$

This tells us that if we kick the particle off in the $\theta = \pi/2$ plane, with $\dot{\theta} = 0$, then it will remain there for all time. This is the choice we make.

We still have to compute the magnitude of the angular momentum. Like many conserved quantities, this follows naturally by identifying the appropriate ignorable coordinate. Recall that if the Lagrangian is independent of some specific coordinate $x$ then the Euler-Lagrange equations immediately give us a conserved quantity,

$$\frac{dL}{dx} = 0 \quad \Rightarrow \quad \frac{d}{d\tau}\left(\frac{dL}{d\dot{x}}\right) = 0$$

This is a baby version of Noether's theorem.

The action (1.35) has two such ignorable coordinates, $t$ and $\phi$. The conserved quantity associated to $\phi$ is the magnitude of the angular momentum, $l$. (Strictly, the angular momentum per unit mass.) Restricting to the $\theta = \pi/2$ plane, we define this to be

$$2l = \frac{dL}{d\dot{\phi}} = 2r^2\dot{\phi} \tag{1.36}$$

where the factor of 2 on the left-hand side arises because the kinetic terms in (1.35) don't come with the usual factor of 1/2. Meanwhile, the conserved quantity associated to $t(\tau)$ is

$$-2E = \frac{dL}{d\dot{t}} = -2A(r)c^2\dot{t} \tag{1.37}$$

The label $E$ is not coincidence: it should be interpreted as the energy of the particle (or, strictly, the energy divided by the rest mass). To see this, we look far away: as $r \to \infty$ we have $A(r) \approx 1$ and we return to Minkowski space. Here, we know from our lectures on Special Relativity that $dt/d\tau = \gamma$. We then have $E \to \gamma c^2$ as $r \to \infty$. But this is precisely the energy per unit rest mass of a particle in special relativity.

We should add to these conservation laws the constraint (1.33) which tells us that the geodesic is parameterised by proper time. Restricting to $\theta = \pi/2$ and $\dot{\theta} = 0$, this becomes

$$-A(r)c^2\dot{t}^2 + A^{-1}(r)\dot{r}^2 + r^2\dot{\phi}^2 = -c^2 \tag{1.38}$$

If we now substitute in the expressions for the conserved quantities $l$ and $E$, this constraint can be rewritten as

$$\frac{1}{2}\dot{r}^2 + V_{\text{eff}}(r) = \frac{1}{2}\frac{E^2}{c^2} \tag{1.39}$$

The effective potential $V_{\text{eff}}(r)$ includes the factor $A(r)$ which we now write out in full,

$$V_{\text{eff}}(r) = \frac{1}{2}\left(c^2 + \frac{l^2}{r^2}\right)\left(1 - \frac{2GM}{rc^2}\right) \tag{1.40}$$

Our goal is to solve for the radial motion (1.39). We subsequently use the expression (1.36) to solve for the angular motion and, in this way, determine the orbit.

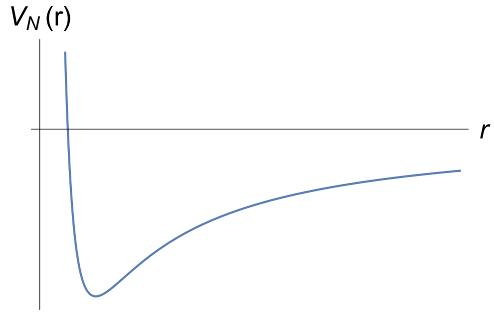



**Figure 3:** The effective potential for Newtonian gravity.

### 1.3.2 Planetary Orbits in Newtonian Mechanics

Before we solve the full geodesic equations, it is useful to first understand how they differ from the equations of Newtonian gravity. To see this, we write

$$V_{\text{eff}}(r) = \frac{c^2}{2} - \frac{GM}{r} + \frac{l^2}{2r^2} - \frac{l^2 GM}{r^3 c^2}$$

The non-relativistic limit is, roughly, $c^2 \to \infty$. This means that we drop the final term in the potential that scales as $1/r^3$. (Since $c$ is dimensionful, it is more accurate to say that we restrict to situations with $l^2 GM/r^3 \ll c^2$.) Meanwhile, we expand the relativistic energy per unit mass in powers of $1/c^2$,

$$E = c^2 + E_N + \ldots$$

where $E_N$ is the non-relativistic energy and $\ldots$ are terms suppressed by $1/c^2$. Substituting these expressions into (1.39), we find

$$\frac{1}{2}\dot{r}^2 + V_N(r) = E_N$$

where $V_N$ is the non-relativistic potential which includes both the Newtonian gravitational potential and the angular momentum barrier,

$$V_N(r) = -\frac{GM}{r} + \frac{l^2}{2r^2}$$

These are precisely the equations that we solved in our first course on classical mechanics. (See Section 4.3 of the lectures on Dynamics and Relativity.) The only difference is that $r(\tau)$ is parameterised by proper time $\tau$ rather than the observers time $t$. However, these coincide in the non-relativistic limit that we care about.

We can build a lot of intuition for the orbits by looking at the potential $V_N(r)$, as shown in Figure 3. At large distances, the attractive $-1/r$ gravitational potential dominates, while the angular momentum prohibits the particles from getting too close to the origin, as seen in the $+1/r^2$ term which dominates at short distances. The potential has a minimum at

$$V'(r_\star) = \frac{GM}{r_\star^2} - \frac{l^2}{r_\star^3} = 0 \quad \Rightarrow \quad r_\star = \frac{l^2}{GM}$$

A particle can happily sit at $r = r_\star$ for all time. This circular orbit always has energy $E_N < 0$, reflecting the fact that $V_N(r_\star) < 0$.

Alternatively, the particle could oscillate back and forth about the minima. This happens provided that $E_N < 0$ so that the particle is unable to escape to $r \to \infty$. This motion describes an orbit in which the distance to the origin varies; we'll see below that the shape of the orbit is an ellipse. Finally, trajectories with $E_N \geq 0$ describe fly-bys, in which the particle approaches the star, but gets only so close before flying away never to be seen again.

The discussion above only tells us about the radial motion. To determine the full orbit, we need to us the angular momentum equation $\dot{\phi} = l/r^2$. Let's remind ourselves how we solve these coupled equations. We start by employing a standard trick of working with the new coordinate

$$u = \frac{1}{r}$$

We then view this inverse radial coordinate as a function of the angular variable: $u = u(\phi)$. This works out nicely, since we have

$$\dot{u} = \frac{du}{d\phi}\dot{\phi} = lu^2\frac{du}{d\phi}$$

where in the last equality, we've used used angular momentum conservation (1.36) to write $\dot{\phi} = lu^2$. Using this, we have

$$\dot{r} = -\frac{1}{u^2}\dot{u} = -l\frac{du}{d\phi} \tag{1.41}$$

The equation giving conservation of energy is then

$$\left(\frac{du}{d\phi}\right)^2 - \left(u - \frac{GM}{l^2}\right)^2 = \frac{2E_N}{l^2} + \frac{G^2M^2}{l^4} \tag{1.42}$$

But this is now straightforward to solve. We choose to write the solution as

$$u(\phi) = \frac{GM}{l^2}(1 + e\cos\phi) \tag{1.43}$$

Back in our original radial variable, we have

$$r(\phi) = \frac{l^2}{GM}\frac{1}{1 + e\cos\phi} \tag{1.44}$$

This is the equation for a conic section, with the eccentricity given by

$$e = 1 + \frac{2E_N l^2}{G^2M^2}$$

The shape of the orbit depends on $e$. A particle with $E_N \geq 0$ is not in a bound orbit, and traces out a hyperbola for $e > 1$ and a parabola for $e = 1$. Planets, in contrast, have energy $E_N < 0$ and, correspondingly, eccentricity $e < 1$. In this case, the orbits are ellipses.

To compare with the relativistic result later, we note an important feature of the Newtonian orbit: it does not precess. To see this, note that for our solution (1.44) the point at which the planet is closest to the origin – known as the *perihelion* – always occurs at the same point $\phi = 0$ in the orbit.

### 1.3.3 Planetary Orbits in General Relativity

We now repeat this analysis for the full relativistic motion of a massive particle moving along a geodesic in the Schwarzchild metric. We have seen that the effective potential takes the form (1.40)

$$V_{\text{eff}}(r) = \frac{c^2}{2} - \frac{GM}{r} + \frac{l^2}{2r^2} - \frac{GMl^2}{r^3c^2}$$

The relativistic correction scales as $-r/r^3$ and changes the Newtonian story at short distances, since it ensures that the potential $V_{\text{eff}}(r) \to -\infty$ as $r \to 0$. Indeed, the potential always vanishes at the Schwarzchild radius $r = R_s = 2GM/c^2$, with $V_{\text{eff}}(R_s) = 0$.

The potential takes different shapes, depending on the size of the angular momentum. To see this, we compute the critical points

$$V'_{\text{eff}}(r) = \frac{GM}{r^2} - \frac{l^2}{r^3} + \frac{3GMl^2}{r^4c^2} = 0 \quad \Rightarrow \quad GMr^2 - l^2r + \frac{3GMl^2}{c^2} = 0 \tag{1.45}$$

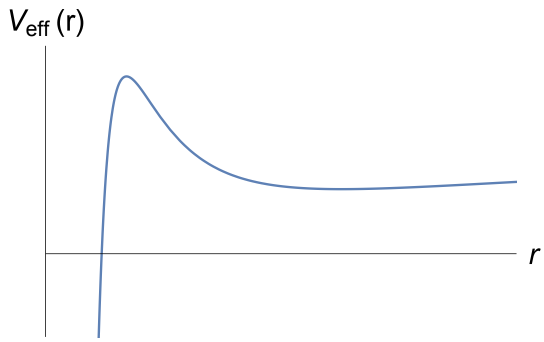



**Figure 4:** The effective potential for a massive particle when $l^2c^2 > 12G^2M^2$.

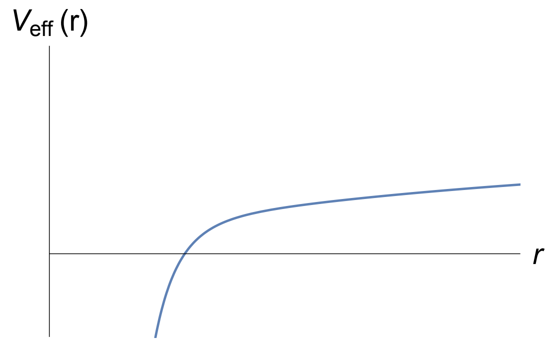



**Figure 5:** ... and when $l^2c^2 < 12G^2M^2$.

If the discriminant is positive, then this quadratic equation has two solutions. This occurs when the angular momentum is suitably large.

$$l^2 > \frac{12G^2M^2}{c^2}$$

In this case, the potential looks like the figure shown on the left. We call the two solutions to the quadratic equation (1.45), $r_+$ and $r_-$ with $r_+ > r_-$. The outermost solution $r_+$ is a minimum of the potential and corresponds to a stable circular orbit; the innermost solution $r_-$ is a maximum of the potential and corresponds to an unstable circular orbit.

As in the Newtonian setting, there are also non-circular orbits in which the particle oscillates around the minimum. However, there is no reason to think these will, in general, remain elliptical. We will study some of their properties below.

Note also that, in contrast to the Newtonian case, the angular momentum barrier is now finite: no matter how large the angular momentum, a particle with enough energy (in the form of ingoing radial velocity) will always be able to cross the barrier, at which point it plummets towards $r = 0$. We will say more about this in Section **??** when we discuss black holes.

If the angular momentum is not large enough,

$$l^2 < \frac{12G^2M^2}{c^2}$$

then the potential $V_{\text{eff}}(r)$ has no turning points and looks like the right-hand figure. In this case, there are no stable orbits; all particles will ultimately fall towards the origin.

The borderline case is $l^2 = 12G^2M^2/c^2$. In this case the turning point is a saddle at

$$r_{\rm ISCO} = \frac{6GM}{c^2} \tag{1.46}$$

This is the *innermost stable circular orbit.* There can be no circular orbits at distances $r < r_{\rm ISCO}$, although it is possible for the non-circular orbits to extend into distances $r < r_{\rm ISCO}$.

The innermost stable orbit plays an important role in black hole astrophysics, where it marks the inner edge of the accretion disc which surrounds the black hole. Roughly speaking, this is seen in the famous photograph captured by the Event Horizon Telescope. Here, the "roughly speaking" is because the light emitted from the accretion disc is warped in a dramatic fashion, so what we see is very different from what is there! (Furthermore, the black hole in the picture almost certainly rotating. This makes $r_{\rm ISCO}$ smaller than $6GM/c^2$ and the picture significantly harder to interpret.)

**Figure 6:**

We could also ask: how close can a non-circular orbit get? This occurs in the limit $l \to \infty$, where a quick calculation shows that the maximum of $V_{\rm eff}$ tends to $r_- \to 3GM/c^2$. This is the closest that any timelike geodesic can get if it wishes to return.

### Perihelion Precession

To understand the orbits in more detail, we can attempt to solve the equations of motion. We follow our Newtonian analysis, introducing the inverse parameter $u = 1/r$ and converting $\dot{r}$ into $du/d\phi$. Our equation (1.39) becomes

$$\left(\frac{du}{d\phi}\right)^2 + u^2 - \frac{2GM}{l^2}u - \frac{2GM}{c^2}u^3 = \frac{E^2}{l^2c^2} - \frac{2c^2}{l^2}$$

This equation is considerably harder than our Newtonian orbit equation (1.42). To proceed, it's simplest to first differentiate again with respect to $\phi$. This gives

$$\frac{d^2u}{d\phi^2} + u - \frac{GM}{l^2} - \frac{3GM}{c^2}u^2 = 0$$

where we have assumed that $du/d\phi \neq 0$, which means that we are neglecting the simple circular solution. The equation above differs from the analogous Newtonian equation by the final term (which indeed vanishes if we take $c^2 \to \infty$). There is no closed-form

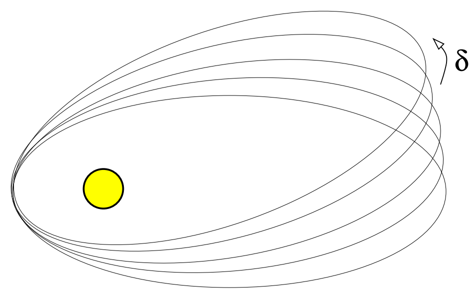



**Figure 7:** The precession of the perihelion (or aphelion) of an almost elliptical orbit.

solution to this equation, but we can make progress by working perturbatively. To this end, we define the dimensionless parameter

$$\beta = \frac{3G^2M^2}{l^2c^2}$$

and write the orbit equation as

$$\frac{d^2u}{d\phi^2} + u - \frac{GM}{l^2} = \beta\,\frac{l^2u^2}{GM} \tag{1.47}$$

We will assume $\beta \ll 1$ and look for series solutions of the form

$$u = u_0 + \beta u_1 + \beta^2 u_2 + \ldots$$

To leading order, we can ignore the terms proportional to $\beta$ on the right-hand-side of (1.47). This gives us an equation for $u_0$ which is identical to the Newtonian orbit

$$\frac{d^2u_0}{d\phi^2} + u_0 - \frac{GM}{l^2} = 0 \quad \Rightarrow \quad u_0(\phi) = \frac{GM}{l^2}(1 + e\cos\phi)$$

We now feed this back into the equation (1.47) to get an equation for $u_1$,

$$\frac{d^2u_1}{d\phi^2} + u_1 = \frac{l^2}{GM}u_0^2 = \frac{GM}{l^2}\left[\left(1 + \frac{e^2}{2}\right) + 2e\cos\phi + \frac{e^2}{2}\cos 2\phi\right]$$

You can check that this is solved by

$$u_1 = \frac{GM}{l^2}\left[\left(1 + \frac{e^2}{2}\right) + e\phi\sin\phi - \frac{e^2}{6}\cos 2\phi\right]$$

We could proceed to next order in $\beta$, but the first correction $u_1$ will be sufficient for our purposes.

The interesting term is the $\phi\sin\phi$ in $u_1$. This is not periodic in $\phi$ and it means that the orbit no longer closes: it sits at a different radial value at $\phi = 0$ and $\phi = 2\pi$. To illustrate this, we ask: when is the particle closest to origin? This is the *perihelion* of the orbit. It occurs when

$$\frac{du}{d\phi} = 0 \quad \Rightarrow \quad -e\sin\phi + \beta\left(e\sin\phi + e\phi\cos\phi - e^2\sin 2\phi\right) = 0$$

Clearly this is solved by $\phi = 0$. The next solution is at $\phi = 2\pi + \delta$ where, due to our perturbative expansion, $\delta$ will be small. Expanding our expression above, and dropping terms of order $\delta^2$ and $\beta\delta$, we find the precession of the perihelion given by

$$\delta = 2\pi\beta = 6\pi\frac{G^2M^2}{l^2c^2} \tag{1.48}$$

For planets orbiting the Sun, the perihelion shift depends only on the angular momentum $l$ of the planet and the mass of the Sun, denoted $M_\odot$. The latter is $M_\odot \approx 2\times 10^{30}$ kg, corresponding to the length scale

$$\frac{GM_\odot}{c^2} \approx 1.5\times 10^3 \text{ m}$$

If a planet on an almost-circular orbit of radius $r$ orbits the sun in a time $T$, then the angular momentum (1.36) is

$$l = \frac{2\pi r^2}{T}$$

Recall that Kepler's third law (which follows from the inverse square law) tells us that $T \propto r^{3/2}$. This means that $l \propto r^{1/2}$ and, correspondingly, the perihelion shift (1.48) is proportional to $\delta \propto 1/r$. We learn that the effect should be more pronounced for planets closest to the Sun.

The closest planet to the Sun is Mercury which, happily, is also the only planet whose orbit differs significantly from a circle; it has eccentricity $e \approx 0.2$, the radius varying from 4.6 to $7\times 10^{10}$ m. Mercury orbits the Sun once every 88 days but, in fact, we don't need to use this to compute the angular momentum and precession. Instead, we can invoke the elliptic formula (1.44) which tells us that the minimum $r_-$ and maximum distance $r_+$ is given by

$$r_\pm = \frac{l^2}{GM}\frac{1}{1\mp e} \quad \Rightarrow \quad l^2 = GMr_+(1-e) \tag{1.49}$$

from which we get the precession

$$\delta = \frac{6\pi GM}{c^2}\frac{1}{r_+(1-e)}$$

Plugging in the numbers gives $\delta \approx 5.0 \times 10^{-7}$. This is rather small. However, the perihelion precession is cumulative. Over a century, Mercury completes 415 orbits, giving the precession of $2.1 \times 10^{-4}$ per century.

The result above is quoted in radians. Astronomers prefer units of arcseconds, with 3600 arcseconds (denoted as $3600''$) in a degree and, of course, 360 degrees in $2\pi$ radians. This means that $1'' \approx 4.8 \times 10^{-6}$ radians. Our calculation from general relativity gives $43''$ per century as the shift in the perihelion. This was one of the first successful predictions of the theory. Subsequently, the perihelion shift of Venus and Earth has been measured and is in agreement with the predictions of general relativity.

### 1.3.4 The Pull of Other Planets

The general relativistic contribution of $43''$ per century is not the full story. In fact the observed perihelion shift of Mercury is much larger, at around $575''$. The vast majority of this is due to the gravitational force of other planets and can be understood entirely within the framework of Newtonian gravity. For completeness, we now give an estimate of these effects.

We start be considering the effect of single, heavy planet with mass $M'$, orbiting at a distance $R$ from the Sun. Of course, the 3-body problem in Newtonian gravity is famously hard. However, there is an approximation which simplifies the problem tremendously: we consider the outer planet to be a circular ring, with mass per unit length given by $M'/2\pi R$.

It's not obvious that this is a good approximation. Each of the outer planets takes significantly longer to orbit the Sun than Mercury. This suggests for any given orbit of Mercury, it would be more appropriate to treat the position of the outer planets to be fixed. (For example, it takes Jupiter 12 years to orbit the Sun, during which time Mercury has completed 50 orbits.) This means that the perihelion shift of Mercury depends on the position of these outer planets and that's a complicated detail that we're happy to ignore. Instead, we want only to compute the total perihelion shift of Mercury averaged over a century. And for this, we may hope that the ring approximation, in which we average over the orbit of the outer planet first, suffices.

In fact, as we will see, the ring approximation is not particularly good: the calculation is non-linear and averaging over the position of the outer planet first does not commute with averaging over the orbits of Mercury. This means that we will get a ballpark figure for the perihelion precession of Mercury but, sadly, not one that is accurate enough to test relativity.

We would like to determine the Newtonian potential felt by a planet which orbits a star of mass $M$ and is surrounded, in the same plane, by a ring of density $M'/2\pi R$. The geometry is shown in the figure. Obviously, the potential (per unit mass) from the star is

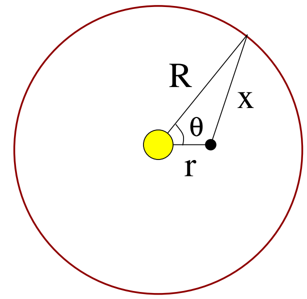



**Figure 8:**

$$V_{\rm star}(r) = -\frac{GM}{r}$$

We need to calculate the potential (per unit mass) from the ring. This is

$$V_{\rm ring}(r) = -\frac{GM'}{2\pi R}\int_0^{2\pi} d\theta\ \frac{1}{x} \quad \text{with} \quad x^2 = R^2 + r^2 - 2Rr\cos\theta \tag{1.50}$$

We use the fact that Mercury is much closer to the Sun than the other planets and Taylor expand the integral in $r^2/R^2$. To leading order we have

$$V_{\rm ring}(r) = -\frac{GM'}{R}\left[1 + \frac{1}{4}\left(\frac{r}{R}\right)^2 + \ldots\right]$$

Dropping constant terms, we learn that the effective potential (per unit mass) experienced by Mercury is, to leading order,

$$V_{\rm eff}(r) = -\frac{GM}{r} + \frac{l^2}{2r^2} - \sum_i \frac{GM_i'}{4}\frac{r^2}{R_i^3} + \ldots$$

where we've included the angular momentum barrier and the sum is over all the outer planets. In what follows, we must assume that the $r^2$ correction term is suitably small so that it doesn't destabilise the existence of orbits. Obviously, this is indeed the case for Mercury.

Now we can follow our calculation for the perihelion precession in general relativity. Conservation of energy tells us

$$\frac{1}{2}\dot{r}^2 + V_{\rm eff}(r) = E$$

Working with everyone's favourite orbit variable, $u = 1/r$, viewed as $u = u(\phi)$, the general relativistic equation (1.47) is replaced by

$$\frac{d^2u}{d\phi^2} + u - \frac{GM}{l^2} = -\alpha\frac{(GM)^4}{l^8u^3} \tag{1.51}$$

where this time our small dimensionless parameter is

$$\alpha = \frac{l^6}{2G^3M^4}\sum_i \frac{M_i'}{R_i^3} = \frac{(1-e)^3}{2}\sum_i \frac{M_i'}{M}\left(\frac{r_+}{R_i}\right)^3$$

where, in the second equality, we've used (1.49); here $M$ is the mass of the Sun, $r_+$ is the outermost radius of Mercury's orbit, and $e \approx 0.2$ is the eccentricity of Mercury's orbit. We safely have $\alpha \ll 1$ and so we look for series solutions of the form

$$u = u_0 + \alpha u_1 + \alpha^2 u_2 + \ldots$$

We've already met the leading order solution, $u_0(\phi) = (GM/l^2)(1 + e\cos\phi)$ with $e$ the eccentricity of the planet's orbit. Feeding this into (1.51), we get an equation for the first correction

$$\frac{d^2u_1}{d\phi^2} + u_1 = -\frac{(GM)^4}{l^8}\frac{1}{u_0^3} = -\frac{GM}{l^2}\frac{1}{(1+e\cos\phi)^3}$$

This equation is somewhat harder to solve than the general relativistic counterpart. To proceed, we will assume that the eccentricity is small, $e \ll 1$, and solve the equation to leading order in $e$. Then this equation becomes

$$\frac{d^2u_1}{d\phi^2} + u_1 = -\frac{GM}{l^2}(1 - 3e\cos\phi)$$

which has the solution

$$u_1 = \frac{GM}{l^2}\left(-1 + \frac{3e}{2}\phi\sin\phi\right)$$

The precession of the perihelion occurs when

$$\frac{du}{d\phi} = 0 \quad \Rightarrow \quad -e\sin\phi + \frac{3e\alpha}{2}\left(\sin\phi + \phi\cos\phi\right) = 0$$

As in the relativistic computation, this is solved by $\phi = 0$ and by $\phi = 2\pi + \delta$ where, to leading order, the shift of the perihelion is given by

$$\delta = 3\pi\alpha = \frac{3\pi}{2}\sum_i \frac{M_i'}{M}\left(\frac{r_0}{R_i}\right)^3$$

with $r_0 = (1-e)r_+$. Once again, we can put the numbers in. The mass of the Sun is $M = M_\odot \approx 2 \times 10^{30}$ kg. The formula is very sensitive to the radius of Mercury's orbit: we use $r_0 \approx 5.64 \times 10^{10}$ m. The relevant data for the other planets is then

| Planet | Mass ($10^{24}$ kg) | Distance ($10^{11}$ m) | $\frac{M}{M_\odot}\left(\frac{r_0}{R}\right)^3$ |
|---|---|---|---|
| Venus | 4.9 | 1.1 | $3.6 \times 10^{-7}$ |
| Earth | 6.0 | 1.5 | $1.7 \times 10^{-7}$ |
| Mars | 0.64 | 2.3 | $5.1 \times 10^{-9}$ |
| Jupiter | 1900 | 7.8 | $3.9 \times 10^{-7}$ |
| Saturn | 570 | 14 | $2.0 \times 10^{-8}$ |

A quick glance at this table shows that the largest contributions come from Jupiter (because of its mass) and Venus (because of it proximity), with the Earth in third place. (The contributions from Uranus, Neptune and Pluto are negligible.)

Adding these contributions, we find $\delta \approx 40\times 10^{-7}$ radians per orbit. This corresponds to $344''$ per century, significantly larger than the $43''$ per century arising from general relativity but not close to the correct Newtonian value of $532''$.

### Higher Order Contributions

Our analysis above gave us a result of $380''$ per century for the perihelion shift of Mercury. A more precise analysis gives $532''$ coming from the Newtonian pull of the other planets.

We made a number of different approximations in the discussion above. But the one that introduced the biggest error turns out to be truncating the ring potential (1.50) at leading order in $r/R$. This, it turns out, is particularly bad for Venus since its orbit compared to Mercury is only $(r_0/R) \approx 0.5$. To do better, we can expand the potential (1.50) to higher orders. We have

$$\begin{aligned} V_{\rm ring}(r) &= -\frac{GM'}{2\pi R}\int_0^{2\pi} d\theta\ \frac{1}{\sqrt{R^2+r^2-2Rr\cos\theta}} \\ &= -\frac{GM'}{R}\left[1+\frac{1}{4}\left(\frac{r}{R}\right)^2+\frac{9}{64}\left(\frac{r}{R}\right)^4+\frac{25}{256}\left(\frac{r}{R}\right)^6+\ldots\right] \end{aligned}$$

An identical calculation to the one above now gives a corresponding perturbative expansion for the perihelion shift,

$$\delta = \pi\sum_i \frac{M_i'}{M}\left[\frac{3}{2}\left(\frac{r_0}{R_i}\right)^3+\frac{45}{16}\left(\frac{r_0}{R_i}\right)^5+\frac{525}{128}\left(\frac{r_0}{R_i}\right)^7+\ldots\right]$$

with $r_0 = (1-e)r_+$ the mean orbit of Mercury. The extra terms give significant contributions for Venus, and smaller for Earth. Using the value of $r_0 \approx 5.64\times 10^{10}$ m and

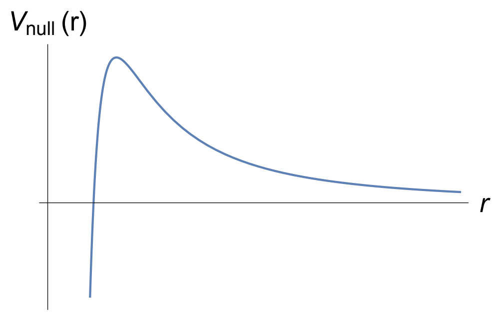



**Figure 9:** The effective potential for null geodesics in the Schwarzchild metric.

the slightly more accurate $R' \approx 10.6 \times 10^{10}$ m for Venus, the sum of the contributions gives $\delta \approx 59 \times 10^{-7}$ radians per orbit, or $507''$ per century, somewhat closer to the recognised value of $532''$ per century but still rather short.

### 1.3.5 Light Bending

It is straightforward to extend the results above to determine the null geodesics in the Schwarzchild metric. We continue to use the equations of motion derived from $S_{\text{useful}}$ in (1.35). But this time we replace the constraint (1.38) with the null version (1.34), which reads

$$-A(r)c^2\dot{t}^2 + A^{-1}(r)\dot{r}^2 + r^2\dot{\phi}^2 = 0$$

The upshot is that we can again reduce the problem to radial motion,

$$\frac{1}{2}\dot{r}^2 + V_{\text{null}}(r) = \frac{1}{2}\frac{E^2}{c^2} \tag{1.52}$$

but now with the effective potential (1.40) replaced by

$$V_{\text{null}}(r) = \frac{l^2}{2r^2}\left(1 - \frac{2GM}{rc^2}\right)$$

A typical potential is shown in Figure 9. Note that, as $r \to \infty$, the potential asymptotes to zero from above, while $V_{\text{null}} \to -\infty$ as $r \to 0$. The potential has a single maximum at

$$V'_{\text{null}}(r_\star) = -\frac{l^2}{r_\star^3} + \frac{3GMl^2}{r_\star^4 c^2} = 0 \quad \Rightarrow \quad r_\star = \frac{3GM}{c^2}$$

We learn that there is a distance, $r_\star$, at which light can orbit a black hole. This is known as the *photon sphere*. The fact that this sits on a maximum of the potential means that this orbit is unstable. In principle, focussing effects mean that much of the light emitted from an accretion disc around a non-rotating black hole emerges from the photon sphere. In practice, it seems likely that photograph of the Event Horizon Telescope does not have the resolution to see this.

The fate of other light rays depends on the relative value of their energy $E$ and angular momentum $l$. To see this, note that the maximum value of the potential is

$$V_{\rm null}(r_\star) = \frac{l^2}{54}\frac{c^4}{G^2M^2}$$

The physics depends on how this compares to the right-hand side of (1.52), $E^2/2c^2$. There are two possibilities

- $E < lc^3/\sqrt{27}GM$: In this case, the energy of light is lower than the angular momentum barrier. This means that light emitted from $r < r_\star$ cannot escape to infinity; it will orbit the star, before falling back towards the origin. The flip side is that light coming from infinity will not fall into the star; instead it will bounce off the angular momentum barrier and return to infinity. In other words, the light will be scattered. We will compute this in more detail below.

- $E < lc^3/\sqrt{27}GM$: Now the energy of the light is greater than the angular momentum barrier. This means that light emitted from $r < r_\star$ can escape to infinity. (We will see in Section **??** that this is only true for light in the region $R_s < r < r_\star$.) Meanwhile, light coming in from infinity is captured by the black hole and asymptotes to $r \to 0$.

To understand the trajectories of light-rays in more detail, we again adopt the inverse parameter $u = 1/r$. The equation of motion (1.52) then becomes

$$\left(\frac{du}{d\phi}\right)^2 + u^2\left(1 - \frac{2GM}{c^2}u\right) = \frac{E^2}{l^2c^2}$$

If we now differentiate again, we get

$$\frac{d^2u}{d\phi^2} + u = \frac{3GM}{c^2}u^2 \tag{1.53}$$

We will again work perturbatively. First, suppose that we ignore the $GM$ term on the right-hand side. We have

$$\frac{d^2u}{d\phi^2} + u = 0 \quad \Rightarrow \quad u = \frac{1}{b}\sin\phi$$

for constant $b$. The meaning of this solution becomes clearer if we write it as $r\sin\phi = b$: this is the equation of a horizontal straight line, a distance $b$ above the origin as shown by the dotted line in Figure 10. The distance $b$ is called the *impact parameter*.

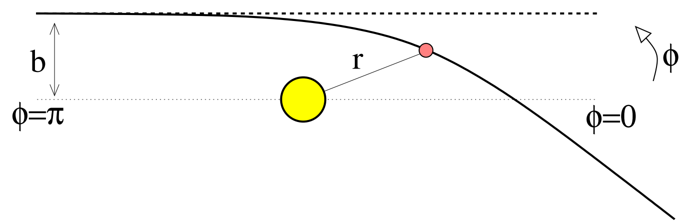



**Figure 10:** Light bending in the Schwarzchild metric.

We will solve the full equation (1.53) perturbatively in the small parameter

$$\tilde{\beta} = \frac{GM}{c^2 b}$$

We then look for solutions of the form

$$u = u_0 + \tilde{\beta} u_1 + \ldots$$

We start with the straight line solution $u_0 = (1/b)\sin\phi$. At leading order, we then have

$$\frac{d^2 u_1}{d\phi^2} + u_1 = \frac{3\sin^2\phi}{b} = \frac{3(1-\cos 2\phi)}{2b}$$

The general solution is

$$u_1 = A\cos\phi + B\sin\phi + \frac{1}{2b}(3+\cos 2\phi)$$

where the first two terms are the complimentary solution, with $A$ and $B$ integration constants. We pick them so that the initial trajectory at $\phi = \pi$ agrees with the straight line $u_0$. This holds if we choose $B = 0$ and $A = 2/b$, so that $u_1 \to 0$ as $\phi \to \infty$. To leading order in $\tilde{\beta}$, the solution is then

$$u = \frac{1}{b}\sin\phi + \frac{GM}{2b^2c^2}\left(3 + 4\cos\phi + \cos 2\phi\right)$$

The question now is: when at what angle does the particle escape to $r = \infty$ or, equivalently, $u = 0$? Before we made the correction this happened at $\phi = 0$. Within our perturbative approach, we can approximate $\sin\phi \approx \phi$ and $\cos\phi \approx 1$ to find that the particle escapes at

$$\phi \approx -\frac{4GM}{bc^2} \tag{1.54}$$

This light bending is known as *gravitational lensing*.

**Figure 11:** Gravitational lensing, as seen by Eddington's 1919 eclipse expedition

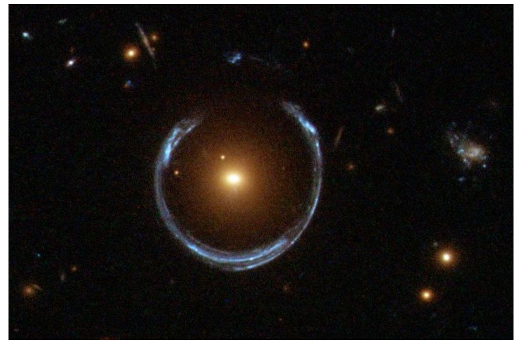


**Figure 12:** Gravitational lensing, as seen by the Hubble space telescope.

For the Sun, $GM_\odot/c^2 \approx 1.5$ km. If light rays just graze the surface, then the impact parameter $b$ coincides with the radius of the Sun, $R_\odot \approx 7 \times 10^5$ km. This gives a scattering angle of $\phi \approx 8.6 \times 10^{-5}$ radians, or $\phi \approx 1.8''$.

There is a difficulty in testing this prediction: things behind the Sun are rarely visible. However, Nature is kind to us because the size of the moon as seen in the sky is more or less the same as the size of the Sun. (This random coincidence would surely make our planet a popular tourist destination for alien hippies if only it wasn't such a long way to travel.) This means that during a solar eclipse the light from the Sun is blocked, allowing us to measure the positions of stars whose light passes nearby the Sun. This can then be compared to the usual positions of these stars.

This measurement was first carried out in May 1919, soon after cessation of war, in two expeditions led by Arthur Eddington, one to the island of Principe and the other to Brazil. The data is shown in the figure above. In the intervening century, we have much more impressive evidence of light bending, in which clusters of galaxies distort the light from a background source, often revealing a distinctive ring-like pattern as shown in the right-hand figure.

### Newtonian Scattering of Light

Before we claim success, we should check to see if the relativistic result (1.54) differs from the Newtonian prediction for light bending. Strictly speaking, there's an ambiguity in the Newtonian prediction for the gravitational force on a massless particle. However, we can invoke the principle of equivalence which tells us that trajectories are independent of the mass. We then extrapolate this result, strictly derived for massive particles, to the massless case.

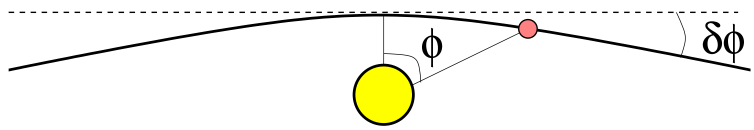



**Figure 13:** The scattering of light using Newtonian gravity.

Scattering under Newtonian gravity follows a hyperbola (1.43)

$$\frac{1}{r} = \frac{GM}{l^2}(1 + e\cos\phi)$$

with $e > 1$. The parameterisation of the trajectory is a little different from the relativistic result, as the light ray asymptotes to infinity at $\cos\phi = -1/e$. For $e \gg 1$, where the trajectory is close to a straight line, the asymptotes occur at $\phi = \pm(\pi + \delta\phi)$ as shown in Figure 13. The scattering angle is then $2\delta\phi$. This is what we wish to compute.

Using (1.41), the speed of light along the trajectory is

$$\dot{r} = -l\frac{du}{d\phi} = \frac{GM}{l}e\sin\phi$$

This is one of the pitfalls of applying Newtonian methods to light bending: we will necessarily find that the speed of light changes as it moves in a gravitational field. The best we can do is ensure that light travels at speed $c$ asymptotically, when $\cos\phi = -1/e$ and $\sin\phi = \sqrt{1 - 1/e^2}$. This gives

$$c^2 = \frac{G^2M^2}{l^2}(e^2 - 1)$$

Meanwhile the angular momentum is $l = bc$, with $b$ the impact parameter. Rearranging, we have

$$e^2 = \frac{b^2c^4}{G^2M^2} + 1 \quad \Rightarrow \quad e \approx \frac{bc^2}{GM}$$

where, in the second equation, we have used the fact we are interested in trajectories close to a straight line with $e \gg 1$. As we mentioned above, the trajectory asymptotes to infinity at $\cos\phi = -1/e$. This occurs at $\phi = \pi/2 + \delta\phi$ and $\phi = -\pi/2 - \delta\phi$ with

$$\delta\phi \approx \frac{1}{e} \approx \frac{GM}{bc^2}$$

The resulting scattering angle is

$$2\delta\phi \approx \frac{2GM}{bc^2}$$

We see that this is a factor of 2 smaller than the relativistic prediction (1.54)

The fact that relativistic light bending is twice as large as the Newtonian answer can be traced to the fact that both $g_{00}$ and $g_{rr}$ components of the Schwarzchild metric are non-vanishing. In some sense, the Newtonian result comes from the $g_{00}$ term, while the contribution from $g_{rr}$ is new. We'll discuss this more in Section 5.1 where we explain how to derive Newtonian gravity from general relativity.

# 2. Introducing Differential Geometry

Gravity is geometry. To fully understand this statement, we will need more sophisticated tools and language to describe curved space and, ultimately, curved spacetime. This is the mathematical subject of *differential geometry* and will be introduced in this section the next. Armed with these new tools, we will then return to the subject of gravity in Section 4.

Our discussion of differential geometry is not particularly rigorous. We will not prove many big theorems. Furthermore, a number of the statements that we make can be checked straightforwardly but we will often omit this. We will, however, be careful about building up the mathematical structure of curved spaces in the right logical order. As we proceed, we will come across a number of mathematical objects that can live on curved spaces. Many of these are familiar – like vectors, or differential operators – but we'll see them appear in somewhat unfamiliar guises. The main purpose of this section is to understand what kind of objects can live on curved spaces, and the relationships between them. This will prove useful for both general relativity and other areas of physics.

Moreover, there is a wonderful rigidity to the language of differential geometry. It sometimes feels that any equation that you're allowed to write down within this rigid structure is more likely than not to be true! This rigidity is going to be enormous help when we return to discuss theories of gravity in Section 4.

## 2.1 Manifolds

The stage on which our story will play out is a mathematical object called a *manifold*. We will give a precise definition below, but for now you should think of a manifold as a curved, $n$-dimensional space. If you zoom in to any patch, the manifold looks like $\mathbf{R}^n$. But, viewed more globally, the manifold may have interesting curvature or topology.

To begin with, our manifold will have very little structure. For example, initially there will be no way to measure distances between points. But as we proceed, we will describe the various kinds of mathematical objects that can be associated to a manifold, and each one will allow us to do more and more things. It will be a surprisingly long time before we can measure distances between points! (Not until Section 3.)

You have met many manifolds in your education to date, even if you didn't call them by name. Some simple examples in mathematics include Euclidean space $\mathbf{R}^n$, the sphere $\mathbf{S}^n$, and the torus $\mathbf{T}^n = \mathbf{S}^1 \times \ldots \mathbf{S}^1$. Some simple examples in physics include the configuration space and phase space that we use in classical mechanics and the state space of thermodynamics. As we progress, we will see how familiar ideas in these subjects can be expressed in a more formal language. Ultimately our goal is to explain how spacetime is a manifold and to understand the structures that live on it.

### 2.1.1 Topological Spaces

Even before we get to a manifold, there is some work to do in order to define the underlying object. What follows is the mathematical equivalent of reading a biography about an interesting person and having to spend the first 20 pages wading through a description of what their grandparents did for a living. This backstory will not be particularly useful for our needs and we include it here only for completeness. We'll keep it down to one page.

Our backstory is called a *topological space*. Roughly speaking, this is a space in which each point can be viewed as living in a neighbourhood of other points, in a manner that allows us to define concepts such as continuity and convergence.

**Definition**: A *topological space* $M$ is a set of points, endowed with a *topology* $\mathcal{T}$. This is a collection of open subsets $\{\mathcal{O}_\alpha \subset M\}$ which obey:

i) Both the set $M$ and the empty set $\emptyset$ are open subsets: $M \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$.

ii) The intersection of a finite number of open sets is also an open set. So if $\mathcal{O}_1 \in \mathcal{T}$ and $\mathcal{O}_2 \in \mathcal{T}$ then $\mathcal{O}_1 \cap \mathcal{O}_2 \in \mathcal{T}$.

iii) The union of any number (possibly infinite) of open sets is also an open set. So if $\mathcal{O}_\gamma \in \mathcal{T}$ then $\cup_\gamma \mathcal{O}_\gamma \in \mathcal{T}$.

Given a point $p \in M$, we say that $\mathcal{O} \in \mathcal{T}$ is a *neighbourhood* of $p$ if $p \in \mathcal{O}$. This concept leads us to our final requirement: we require that, given any two distinct points, there is a neighbourhood which contains one but not the other. In other words, for any $p, q \in M$ with $p \neq q$, there exists $\mathcal{O}_1, \mathcal{O}_2 \in \mathcal{T}$ such that $p \in \mathcal{O}_1$ and $q \in \mathcal{O}_2$ and $\mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$. Topological spaces which obey this criterion are called *Hausdorff*. It is like a magic ward to protect us against bad things happening.

An example of a good Hausdorff space is the real line, $M = \mathbf{R}$, with $\mathcal{T}$ consisting of all open intervals $(a, b)$, with $a < b \in \mathbf{R}$, and their unions. An example of a non-Hausdorff space is any $M$ with $\mathcal{T} = \{M, \emptyset\}$.

**Definition**: One further definition (it won't be our last). A *homeomorphism* between topological spaces $(M, \mathcal{T})$ and $(\tilde{M}, \tilde{\mathcal{T}})$ is a map $f : M \to \tilde{M}$ which is

i) Injective (or one-to-one): for $p \neq q$, $f(p) \neq f(q)$.

ii) Surjective (or onto): $f(M) = \tilde{M}$, which means that for each $\tilde{p} \in \tilde{M}$ there exists a $p \in M$ such that $f(p) = \tilde{p}$.

Functions which are both injective and surjective are said to be bijective. This ensures that they have an inverse

iii) Bicontinuous. This means that both the function and its inverse are continuous. To define a notion of continuity, we need to use the topology. We say that $f$ is continuous if, for all $\tilde{\mathcal{O}} \in \tilde{\mathcal{T}}$, $f^{-1}(\tilde{\mathcal{O}}) \in \mathcal{T}$.

There's an animation of a donut morphing into a coffee mug and back that is often used to illustrate the idea of topology. If you want to be fancy, you can say that a donut is homeomorphic to a coffee mug.

### 2.1.2 Differentiable Manifolds

We now come to our main character: an $n$-dimensional manifold is a space which, locally, looks like $\mathbf{R}^n$. Globally, the manifold may be more interesting that $\mathbf{R}^n$, but the idea is that we can patch together these local descriptions to get an understanding for the entire space.

**Definition**: An $n$-dimensional *differentiable manifold* is a Hausdorff topological space $M$ such that

i) $M$ is locally homeomorphic to $\mathbf{R}^n$. This means that for each $p \in M$, there is an open set $\mathcal{O}$ such that $p \in \mathcal{O}$ and a homeomorphism $\phi : \mathcal{O} \to U$ with $U$ an open subset of $\mathbf{R}^n$.

ii) Take two open subsets $\mathcal{O}_\alpha$ and $\mathcal{O}_\beta$ that overlap, so that $\mathcal{O}_\alpha \cap \mathcal{O}_\beta \neq \emptyset$. We require that the corresponding maps $\phi_\alpha : \mathcal{O}_\alpha \to U_\alpha$ and $\phi_\beta : \mathcal{O}_\beta \to U_\beta$ are *compatible*, meaning that the map $\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(\mathcal{O}_\alpha \cap \mathcal{O}_\beta) \to \phi_\alpha(\mathcal{O}_\alpha \cap \mathcal{O}_\beta)$ is smooth (also known as infinitely differentiable or $C^\infty$). This is depicted in Figure 14.

The maps $\phi_\alpha$ are called *charts* and the collection of charts is called an *atlas*. You should think of each chart as providing a coordinate system to label the region $\mathcal{O}_\alpha$ of $M$. The coordinate associated to $p \in \mathcal{O}_\alpha$ is

$$\phi_\alpha(p) = (x^1(p), \ldots, x^n(p))$$

We write the coordinate in shorthand as simply $x^\mu(p)$, with $\mu = 1, \ldots, n$. Note that we use a superscript $\mu$ rather than a subscript: this simple choice of notation will prove useful as we go along.

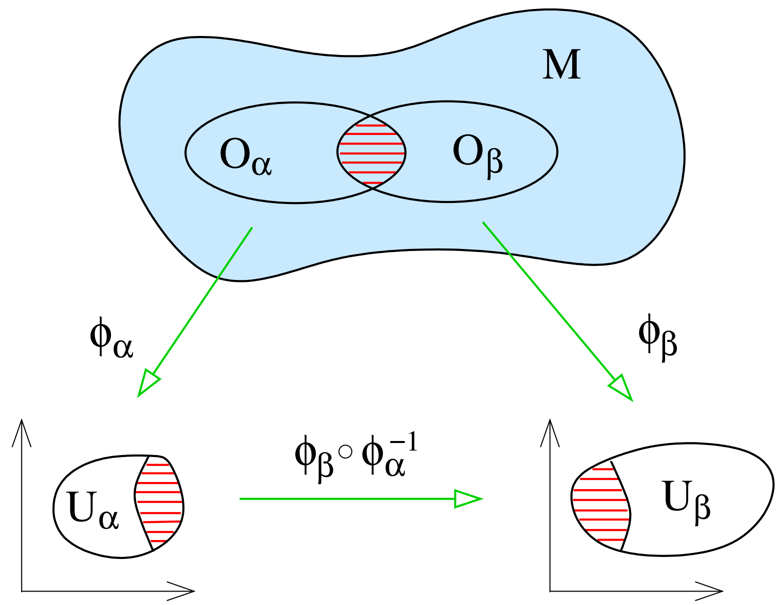



**Figure 14:** Charts on a manifold.

If a point $p$ is a member of more than one subset $\mathcal{O}$ then it may have a number of different coordinates associated to it. There's nothing to be nervous about here: it's entirely analogous to labelling a point using either Euclidean coordinate or polar coordinates.

The maps $\phi_\alpha \circ \phi_\beta^{-1}$ take us between different coordinate systems and are called *transition functions*. The compatibility condition is there to ensure that there is no inconsistency between these different coordinate systems.

Any manifold $M$ admits many different atlases. In particular, nothing stops us from adding another chart to the atlas, provided that it is compatible with all the others. Two atlases are said to be compatible if every chart in one is compatible with every chart in the other. In this case, we say that the two atlases define the same *differentiable structure* on the manifold.

**Examples**

Here are a few simple examples of differentiable manifolds:

- $\mathbf{R}^n$: this looks locally like $\mathbf{R}^n$ because it is $\mathbf{R}^n$. You only need a single chart with the usual Euclidean coordinates. Similarly, any open subset of $\mathbf{R}^n$ is a manifold.

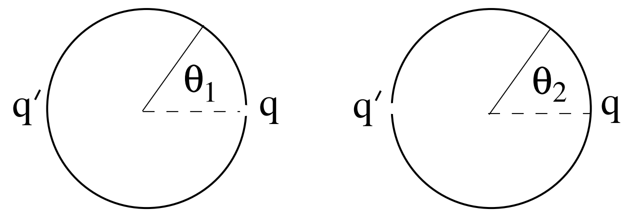



**Figure 15:** Two charts on a circle. The figures are subtly different! On the left, point $q$ is removed and $\theta_1 \in (0, 2\pi)$. On the right, point $q'$ is removed and $\theta_2 \in (-\pi, \pi)$.

- $\mathbf{S}^1$: The circle can be defined as a curve in $\mathbf{R}^2$ with coordinates $(\cos\theta, \sin\theta)$. Until now in our physics careers, we've been perfectly happy taking $\theta \in [0, 2\pi)$ as the coordinate on $\mathbf{S}^1$. But this coordinate does not meet our requirements to be a chart because it is not an open set. This causes problems if we want to differentiate functions at $\theta = 0$; to do so we need to take limits from both sides but there is no coordinate with $\theta$ a little less than zero.

  To circumvent this, we need to use at least two charts to cover $\mathbf{S}^1$. For example, we could identify two antipodal points, say $q = (1, 0)$ and $q' = (-1, 0)$. We take the first chart to cover $\mathcal{O}_1 = \mathbf{S}^1 - \{q\}$ with the map $\phi_1 : \mathcal{O}_1 \to (0, 2\pi)$ defined by $\phi_1(p) = \theta$ as shown in the left-hand of Figure 15. We take the second chart to cover $\mathcal{O}_2 = \mathbf{S}^1 - \{q'\}$ with the map $\phi_2 : \mathcal{O}_2 \to (-\pi, \pi)$ defined by $\phi_2(p) = \theta'$ as shown in the right-hand figure.

  The two charts overlap on the upper and lower semicircles. The transition function is given by

$$\theta' = \phi_2(\phi_1^{-1}(\theta)) = \begin{cases} \theta & \text{if } \theta \in (0, \pi) \\ \theta - 2\pi & \text{if } \theta \in (\pi, 2\pi) \end{cases}$$

  The transition function isn't defined at $\theta = 0$, corresponding to the point $q_1$, nor at $\theta = \pi$, corresponding to the point $q_2$. Nonetheless, it is smooth on each of the two open intervals as required.

- $\mathbf{S}^2$: It will be useful to think of the sphere as the surface $x^2+y^2+z^2 = 1$ embedded in Euclidean $\mathbf{R}^3$. The familiar coordinates on the sphere $\mathbf{S}^2$ are those inherited from spherical polar coordinates of $\mathbf{R}^3$, namely

$$x = \sin\theta\cos\phi \quad , \quad y = \sin\theta\sin\phi \quad , \quad z = \cos\theta \tag{2.1}$$

  with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$. But as with the circle $\mathbf{S}^1$ described above, these are not open sets so will not do for our purpose. In fact, there are two distinct

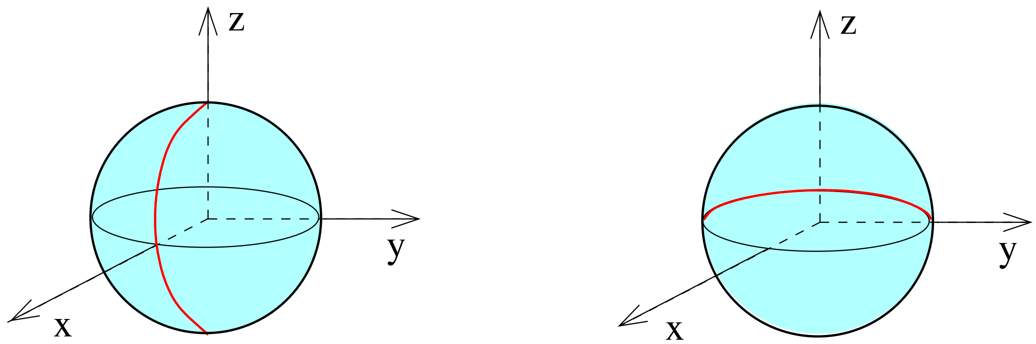



**Figure 16:** Two charts on a sphere. In the left-hand figure, we have removed the half-equator defined as $y = 0$ with $x > 0$, shown in red. In right-figure, we have removed the half-equator $z = 0$ with $x < 0$, again shown in red.

issues. If we focus on the equator at $\theta = \pi/2$, then the coordinate $\phi \in [0, 2\pi)$ parameterises a circle and suffers the same problem that we saw above. On top of this, at the north pole $\theta = 0$ and south pole $\theta = \pi$, the coordinate $\phi$ is not well defined, since the value of $\theta$ has already specified the point uniquely. This manifests itself on Earth by the fact that all time zones coincide at the North pole. It's one of the reasons people don't have business meetings there.

Once again, we can resolve these issues by introducing two charts covering different patches on $\mathbf{S}^2$. The first chart applies to the sphere $\mathbf{S}^2$ with a line of longitude removed, defined by $y = 0$ and $x > 0$, as shown in Figure 16. (Think of this as the dateline.) This means that neither the north nor south pole are included in the open set $\mathcal{O}_1$. On this open set, we define a map $\phi_1 : \mathcal{O}_1 \to \mathbf{R}^2$ using the coordinates (2.1), now with $\theta \in (0, \pi)$ and $\phi \in (0, 2\pi)$, so that we have a map to an open subset of $\mathbf{R}^2$.

We then define a second chart on a different open set $\mathcal{O}_2$, defined by $\mathbf{S}^2$, with the line $z = 0$ and $x < 0$ removed. Here we define the map $\phi_2 : \mathcal{O}_2 \to \mathbf{R}^2$ using the coordinates

$$x = -\sin\theta' \cos\phi' \quad , \quad y = \cos\theta' \quad , \quad z = \sin\theta' \sin\phi'$$

with $\theta' \in (0, \pi)$ and $\phi \in (0, 2\pi)$. Again this is a map to an open subset of $\mathbf{R}^2$. We have $\mathcal{O}_1 \cup \mathcal{O}_2 = \mathbf{S}^2$ while, on the overlap $\mathcal{O}_1 \cap \mathcal{O}_2$, the transition functions $\phi_1 \circ \phi_2^{-1}$ and $\phi_2 \circ \phi_1^{-1}$ are smooth. (We haven't written these functions down explicitly, but it's clear that they are built from cos and sin functions acting on domains where their inverses exist.)

Note that for both $\mathbf{S}^1$ and $\mathbf{S}^2$ examples above, we made use of the fact that they can be viewed as embedded in a higher dimensional $\mathbf{R}^{n+1}$ to construct the charts. However, this isn't necessary. The definition of a manifold makes no mention of a higher dimensional embedding and these manifolds should be viewed as having an existence independent of any embedding.

As you can see, there is a level of pedantry involved in describing these charts. (Mathematicians prefer the word "rigour".) The need to deal with multiple charts arises only when we have manifolds of non-trivial topology; the manifolds $\mathbf{S}^1$ and $\mathbf{S}^2$ that we met above are particularly simple examples. When we come to discuss general relativity, we will care a lot about changing coordinates, and the limitations of certain coordinate systems, but our manifolds will turn out to be simple enough that, for all practical purposes, we can always find a single set of coordinates that tells us what we need to know. However, as we progress in physics, and topology becomes more important, so too does the idea of different charts. This is necessary, for example, when we discuss the magnetic monopole. (See the lectures on Gauge Theory.)

### 2.1.3 Maps Between Manifolds

The advantage of locally mapping a manifold to $\mathbf{R}^n$ is that we can now import our knowledge of how to do maths on $\mathbf{R}^n$. For example, we know how to differentiate functions on $\mathbf{R}^n$, and what it means for functions to be smooth. This now translates directly into properties of functions defined over the manifold.

We say that a function $f : M \to \mathbf{R}$ is smooth, if the map $f \circ \phi^{-1} : U \to \mathbf{R}$ is smooth for all charts $\phi$.

Similarly, we say that a map $f : M \to N$ between two manifolds $M$ and $N$ (which may have different dimensions) is smooth if the map $\psi \circ f \circ \phi^{-1} : U \to V$ is smooth for all charts $\phi : M \to U \subset \mathbf{R}^{\dim(M)}$ and $\psi : N \to V \subset \mathbf{R}^{\dim(N)}$

A *diffeomorphism* is defined to be a smooth homeomorphism $f : M \to N$. In other words it is an invertible, smooth map between manifolds $M$ and $N$ that has a smooth inverse. If such a diffeomorphism exists then the manifolds $M$ and $N$ are said to be *diffeomorphic*. The existence of an inverse means $M$ and $N$ necessarily have the same dimension.

Manifolds which are homeomorphic can be continuously deformed into each other. But diffeomorphism is stronger: it requires that the map and its inverse are smooth. This gives rise to some curiosities. For example, it turns out that the sphere $\mathbf{S}^7$ can be covered by a number of different, incompatible atlases. The resulting manifolds are homeomorphic but not diffeomorphic. These are referred to as exotic spheres. Similarly, Euclidean space $\mathbf{R}^n$ has a unique differentiable structure, except for $\mathbf{R}^4$ where there are an infinite number of inequivalent structures. I don't know of any applications of these facts to physics. Certainly they will not play any role in these lectures.

## 2.2 Tangent Spaces

Our next task is to understand how to do calculus on manifolds. We start here with differentiation; it will take us a while longer to get to integration, which we will finally meet in Section 2.4.4.

Consider a function $f : M \to \mathbf{R}$. To differentiate the function at some point $p$, we introduce a chart $\phi = (x^1, \ldots, x^n)$ in a neighbourhood of $p$. We can then construct the map $f \circ \phi^{-1} : U \to \mathbf{R}$ with $U \subset \mathbf{R}^n$. But we know how to differentiate functions on $\mathbf{R}^n$ and this gives us a way to differentiate functions on $M$, namely

$$\left.\frac{\partial f}{\partial x^\mu}\right|_p := \left.\frac{\partial (f \circ \phi^{-1})}{\partial x^\mu}\right|_{\phi(p)} \tag{2.2}$$

Clearly this depends on the choice of chart $\phi$ and coordinates $x^\mu$. We would like to give a coordinate independent definition of differentiation, and then understand what happens when we choose to describe this object using different coordinates.

### 2.2.1 Tangent Vectors

We will consider smooth functions over a manifold $M$. We denote the set of all smooth functions as $C^\infty(M)$.

**Definition**: A *tangent vector* $X_p$ is an object that differentiates functions at a point $p \in M$. Specifically, $X_p : C^\infty(M) \to \mathbf{R}$ satisfying

i) Linearity: $X_p(f + g) = X_p(f) + X_p(g)$ for all $f, g \in C^\infty(M)$.

ii) $X_p(f) = 0$ when $f$ is the constant function.

iii) Leibnizarity: $X_p(fg) = f(p)\, X_p(g) + X_p(f)\, g(p)$ for all $f, g \in C^\infty(M)$. This, of course, is the product rule.

Note that ii) and iii) combine to tell us that $X_p(af) = aX_p(f)$ for $a \in \mathbf{R}$.

This definition is one of the early surprises in differential geometry. The surprise is really in the name "tangent vector". We know what vectors are from undergraduate physics, and we know what differential operators are. But we're not used to equating the two. Before we move on, it might be useful to think about how this definition fits with other notions of vectors that we've met before.

The first time we meet a vector in physics is usually in the context of Newtonian mechanics, where we describe the position of a particle as a vector $\mathbf{x}$ in $\mathbf{R}^3$.This concept of a vector is special to flat space and does not generalise to other manifolds. For example, a line connecting two points on a sphere is not a vector and, in general, there is no way to think of a point $p \in M$ as a vector. So we should simply forget that points in $\mathbf{R}^3$ can be thought of as vectors.

The next type of vector is the velocity of a particle, $\mathbf{v} = \dot{\mathbf{x}}$. This is more pertinent. It clearly involves differentiation of some object, and is tangent to the curve traced out by the particle. As we will see below, velocities of particles are indeed examples of tangent vectors in differential geometry. More generally, tangent vectors tell us how things change in given direction. They do this by differentiating.

It is simple to check that the object

$$\partial_\mu\Big|_p := \frac{\partial}{\partial x^\mu}\Big|_p$$

which acts on functions as shown in (2.2) obeys all the requirements of a tangent vector.

Note that the index $\mu$ is now a subscript, rather than superscript that we used for the coordinates $x^\mu$. (On the right-hand-side, the superscript in $\partial/\partial x^\mu$ is in the denominator and counts as a subscript.) We will adopt the summation convention, where repeated indices are summed. But, as we will see, the placement of indices up or down will tell us something and all sums will necessarily have one index up and one index down. This is a convention that we met already in Special Relativity where the up/downness of the index changes minus signs. Here it has a more important role that we will see as we go on: the placement of the index tells us what kind of mathematical space the object lives in. For now, you should be aware that any equation with two repeated indices that are both up or both down is necessarily wrong, just as any equation with three or more repeated indices is wrong.

**Theorem**: The set of all tangent vectors at point $p$ forms an $n$-dimensional vector space. We call this the *tangent space* $T_p(M)$. The tangent vectors $\partial_\mu\big|_p$ provide a basis for $T_p(M)$. This means that we can write any tangent vector as

$$X_p = X^\mu\, \partial_\mu\Big|_p$$

with $X^\mu = X_p(x^\mu)$ the components of the tangent vector in this basis.

**Proof**: Much of the proof is just getting straight what objects live in what spaces. Indeed, getting this straight is a large part of the subject of differential geometry. To start, we need a small lemma. We define the function $F = f \circ \phi^{-1} : U \to \mathbf{R}$, with $\phi = (x^1, \ldots, x^n)$ a chart on a neighbourhood of $p$. Then, in some (perhaps smaller) neighbourhood of $p$ we can always write the function $F$ as

$$F(x) = F(x^\mu(p)) + (x^\mu - x^\mu(p))F_\mu(x) \tag{2.3}$$

where we have introduced $n$ new functions $F_\mu(x)$ and used the summation convention in the final term. If the function $F$ has a Taylor expansion then we can trivially write it in the form (2.3) by repackaging all the terms that are quadratic and higher into the $F_\mu(x)$ functions, keeping a linear term out front. But in fact there's no need to assume the existence of a Taylor expansion. One way to see this is to note that for any function $G(t)$ we trivially have $G(1) = G(0) + \int_0^1 dt\; G'(t)$. But now apply this formula to the function $G(t) = F(tx)$ for some fixed $x$. This gives $F(x) = F(0) + x\int_0^1 dt\; F'(xt)$ which is precisely (2.3) for a function of a single variable expanded about the origin. The same method holds more generally.

Given (2.3), we act with $\partial_\mu$ on both sides, and then evaluate at $x^\mu = x^\mu(p)$. This tells us that the functions $F_\mu$ must satisfy

$$\left.\frac{\partial F}{\partial x^\mu}\right|_{x(p)} = F_\mu(x(p)) \tag{2.4}$$

We can translate this into a similar expression for $f$ itself. We define $n$ functions on $M$ by $f_\mu = F_\mu \circ \phi$. Then, for any $q \in M$ in the appropriate neighbourhood of $p$, (2.3) becomes

$$f \circ \phi^{-1}(x^\mu(q)) = f \circ \phi^{-1}(x^\mu(p)) + (x^\mu(q) - x^\mu(p))\; \left[f_\mu \circ \phi^{-1}(x^\mu(q))\right]$$

But $\phi^{-1}(x^\mu(q)) = q$. So we find that, in the neighbourhood of $p$, it is always possible to write a function $f$ as

$$f(q) = f(p) + (x^\mu(q) - x^\mu(p))f_\mu(q)$$

for some $f_\mu(q)$. Note that, evaluated at $q = p$, we have

$$f_\mu(p) = F_\mu \circ \phi(p) = F_\mu(x(p)) = \left.\frac{\partial F}{\partial x^\mu}\right|_{x(p)} = \left.\frac{\partial f}{\partial x^\mu}\right|_p$$

where in the last equality we used (2.2) and in the penultimate equality we used (2.4).

Now we can turn to the tangent vector $X_p$. This acts on the function $f$ to give

$$X_p(f) = X_p\Big(f(p) + (x^\mu - x^\mu(p))f_\mu\Big)$$

where we've dropped the arbitrary argument $q$ in $f(q)$, $x^\mu(q)$ and $f_\mu(q)$; these are the functions on which the tangent vector is acting. Using linearity and Leibnizarity, we have

$$X_p(f) = X_p\Big(f(p)\Big) + X_p\Big((x^\mu - x^\mu(p))\Big)f_\mu(p) + (x^\mu(p) - x^\mu(p))X_p\Big(f_\mu\Big)$$

The first term vanishes because $f(p)$ is just a constant and all tangent vectors are vanishing when acting on a constant. The final term vanishes as well because the Leibniz rule tells us to evaluate the function $(x^\mu - x^\mu(p))$ at $p$. Finally, by linearity, the middle term includes a $X_p(x^\mu(p))$ term which vanishes because $x^\mu(p)$ is just a constant. We're left with

$$X_p(f) = X_p(x^\mu)\left.\frac{\partial f}{\partial x^\mu}\right|_p$$

This means that the tangent vector $X_p$ can be written as

$$X_p = X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p$$

with $X^\mu = X_p(x^\mu)$ as promised. To finish, we just need to show that $\partial_\mu\big|_p$ provide a basis for $T_p(M)$. From above, they span the space. To check linear independence, suppose that we have vector $\alpha = \alpha_\mu \partial_\mu\big|_p = 0$. Then acting on $f = x^\nu$, this gives $\alpha(x^\nu) = \alpha_\mu(\partial_\mu x^\nu)\big|_p = \alpha_\nu = 0$. This concludes our proof. □

**Changing Coordinates**

We have an ambivalent relationship with coordinates. We can't calculate anything without them, but we don't want to rely on them. The compromise we will come to is to consistently check that nothing physical depends on our choice of coordinates.

The idea is a given tangent vector $X_p$ exists independent of the choice of coordinate. However, the chosen basis $\{\partial_\mu\big|_p\}$ clearly depends on our choice of coordinates: to define it we had to first introduce a given chart $\phi$ and coordinates $x^\mu$. A basis defined in this way is called, quite reasonably, a *coordinate basis*. At times we will work with other bases, $\{e_\mu\}$ which are not defined in this way. Unsurprisingly, these are referred to as *non-coordinate bases*. A particularly useful example of a non-coordinate basis, known as *vielbeins*, will be introduced in Section 3.4.2.

Suppose that we picked a different chart $\tilde{\phi}$, with coordinates $\tilde{x}^\mu$ in the neighbourhood of $p$. We then have two different bases, and can express the tangent vector $X_p$ in terms of either,

$$X_p = X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p = \tilde{X}^\mu \left.\frac{\partial}{\partial \tilde{x}^\mu}\right|_p$$

The vector is the same, but the components of the vector change: they are $X^\mu$ in the first set of coordinates, and $\tilde{X}^\mu$ in the second. It is straightforward to determine the relationship between $X^\mu$ and $\tilde{X}^\mu$. To see this, we look at how the tangent vector $X_p$ acts on a function $f$,

$$X_p(f) = X^\mu \left.\frac{\partial f}{\partial x^\mu}\right|_p = X^\mu \left.\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right|_{\phi(p)} \left.\frac{\partial f}{\partial \tilde{x}^\nu}\right|_p$$

where we've used the chain rule. (Actually, we've been a little quick here. You can be more careful by introducing the functions $F = f \circ \phi^{-1}$ and $\tilde{F} = f \circ \tilde{\phi}^{-1}$ and using (2.2) to write $\partial f/\partial x^\mu = \partial F(\tilde{x}(x))/\partial x^\mu$. The end result is the same. We will be similarly sloppy in the same way as we proceed, often conflating $f$ and $F$.) You can read this equation in one of two different ways. First, we can view this as a change in the basis vectors: they are related as

$$\left.\frac{\partial}{\partial x^\mu}\right|_p = \left.\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right|_{\phi(p)} \left.\frac{\partial}{\partial \tilde{x}^\nu}\right|_p \tag{2.5}$$

Alternatively, we can view this as a change in the components of the vector, which transform as

$$\tilde{X}^\nu = X^\mu \left.\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right|_{\phi(p)} \tag{2.6}$$

Components of vectors that transform this way are sometimes said to be *contravariant*. It's annoying language that I can never remember. A more important point is that the form of (2.6) is essentially fixed once you remember that the index on $X^\mu$ sits up rather than down.

#### What Is It Tangent To?

So far, we haven't really explained where the name "tangent vector" comes from. Consider a smooth curve in $M$ that passes through the point $p$. This is a map $\sigma : I \to M$, with $I$ an open interval $I \subset \mathbf{R}$. We will parameterise the curve as $\sigma(t)$ such that $\sigma(0) = p \in M$.

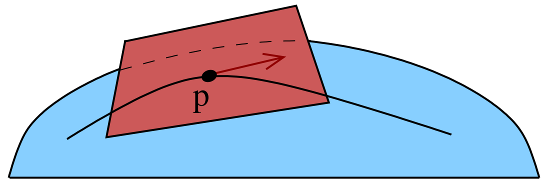



**Figure 17:** The tangent space at a point $p$.

With a given chart, this curve becomes $\phi \circ \sigma : \mathbf{R} \mapsto \mathbf{R}^n$, parameterised by $x^\mu(t)$. We usually say that the tangent vector to the curve at $t = 0$ is

$$X^\mu = \left.\frac{dx^\mu(t)}{dt}\right|_{t=0}$$

But we can take these to be the components of the tangent vector $X_p$, now defined as

$$X_p = \left.\frac{dx^\mu(t)}{dt}\right|_{t=0} \left.\frac{\partial}{\partial x^\mu}\right|_p$$

Our tangent vector now acts on functions $f \in C^\infty(M)$. It is telling us how fast any function $f$ changes as we move along the curve.

Any tangent vector $X_p$ can be written in this form. This gives meaning to the term "tangent space" for $T_p(M)$. It is, literally, the space of all possible tangents to curves passing through the point $p$. For example, a two dimensional manifold, embedded in $\mathbf{R}^3$ is shown in Figure 17. At each point $p$, we can identify a vector space which is the tangent plane: this is $T_p(M)$.

As an aside, note that the mathematical definition of a tangent space makes no reference to embedding the manifold in some higher dimensional space. The tangent space is an object intrinsic to the manifold itself. (In contrast, in the picture it was unfortunately necessary to think about the manifold as embedded in $\mathbf{R}^3$.)

The tangent spaces $T_p(M)$ and $T_q(M)$ at different points $p \neq q$ are different. There's no sense in which we can add vectors from one to vectors from the other. In fact, at this stage there no way to even compare vectors in $T_p(M)$ to vectors in $T_q(M)$. They are simply different spaces. As we proceed, we will make some effort to figure ways to get around this.

### 2.2.2 Vector Fields

So far we have only defined tangent vectors at a point $p$. It is useful to consider an object in which there is a choice of tangent vector for every point $p \in M$. In physics, we call objects that vary over space *fields*.

A *vector field* $X$ is defined to be a smooth assignment of a tangent vector $X_p$ to each point $p \in M$. This means that if you feed a function to a vector field, then it spits back another function, which is the differentiation of the first. In symbols, a vector field is therefore a map $X : C^\infty(M) \to C^\infty(M)$. The function $X(f)$ is defined by

$$\big(X(f)\big)(p) = X_p(f)$$

The space of all vector fields on $M$ is denoted $\mathfrak{X}(M)$.

Given a coordinate basis, we can expand any vector field as

$$X = X^\mu \frac{\partial}{\partial x^\mu} \tag{2.7}$$

where the $X^\mu$ are now smooth functions on $M$.

Strictly speaking, the expression (2.7) only defines a vector field on the open set $\mathcal{O} \subset M$ covered by the chart, rather than the whole manifold. We may have to patch this together with other charts to cover all of $M$.

### The Commutator

Given two vector fields $X, Y \in \mathfrak{X}(M)$, we can't multiply them together to get a new vector field. Roughly speaking, this is because the product $XY$ is a second order differential operator rather than a first order operator. This reveals itself in a failure of Leibnizarity for the object $XY$,

$$XY(fg) = X(fY(g) + Y(f)g) = X(f)Y(g) + fXY(g) + gXY(f) + X(g)Y(f)$$

This is not the same as $fXY(g) + gXY(f)$ that Leibniz requires.

However, we can build a new vector field by taking the *commutator* $[X, Y]$, which acts on functions $f$ as

$$[X, Y](f) = X(Y(f)) - Y(X(f))$$

This is also known as the *Lie bracket*. Evaluated in a coordinate basis, the commutator is given by

$$\begin{aligned}
[X, Y](f) &= X^\mu \frac{\partial}{\partial x^\mu}\left(Y^\nu \frac{\partial f}{\partial x^\nu}\right) - Y^\mu \frac{\partial}{\partial x^\mu}\left(X^\nu \frac{\partial f}{\partial x^\nu}\right) \\
&= \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\right)\frac{\partial f}{\partial x^\nu}
\end{aligned}$$

This holds for all $f \in C^\infty(M)$, so we're at liberty to write

$$[X, Y] = \left( X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu} \right) \frac{\partial}{\partial x^\nu} \tag{2.8}$$

It is not difficult to check that the commutator obeys the Jacobi identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

This ensures that the set of all vector fields on a manifold $M$ has the mathematical structure of a Lie algebra.

### 2.2.3 Integral Curves

There is a slightly different way of thinking about vector fields on a manifold. A *flow* on $M$ is a one-parameter family of diffeomorphisms $\sigma_t : M \to M$ labelled by $t \in \mathbf{R}$. These maps have the properties that $\sigma_{t=0}$ is the identity map, and $\sigma_s \circ \sigma_t = \sigma_{s+t}$. These two requirements ensure that $\sigma_{-t} = \sigma_t^{-1}$. Such a flow gives rise to streamlines on the manifold. We will further require that these streamlines are smooth.

We can then define a vector field by taking the tangent to the streamlines at each point. In a given coordinate system, the components of the vector field are

$$X^\mu(x^\mu(t)) = \frac{dx^\mu(t)}{dt} \tag{2.9}$$

where I've abused notation a little and written $x^\mu(t)$ rather than the more accurate but cumbersome $x^\mu(\sigma_t)$. This will become a habit, with the coordinates $x^\mu$ often used to refer to the point $p \in M$.

A flow gives rise a vector field. Alternatively, given a vector field $X^\mu(x)$, we can integrate the differential equation (2.9), subject to an initial condition $x^\mu(0) = x^\mu_{\text{initial}}$ to generate streamlines which start at $x^\mu_{\text{initial}}$. These streamlines are called *integral curves*, generated by $X$.

In what follows, we will only need the infinitesimal flow generated by $X$. This is simply

$$x^\mu(t) = x^\mu(0) + tX^\mu(x) + \mathcal{O}(t^2) \tag{2.10}$$

Indeed, differentiating this obeys (2.9) to leading order in $t$.

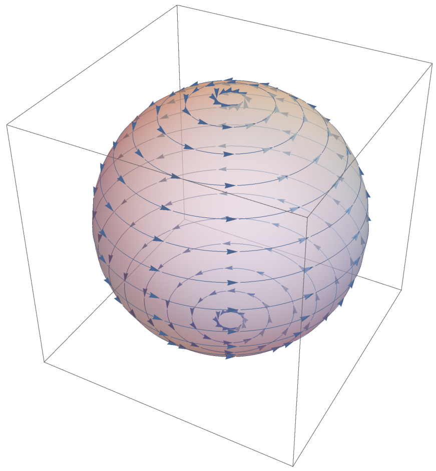


**Figure 18:** Flows on a sphere.

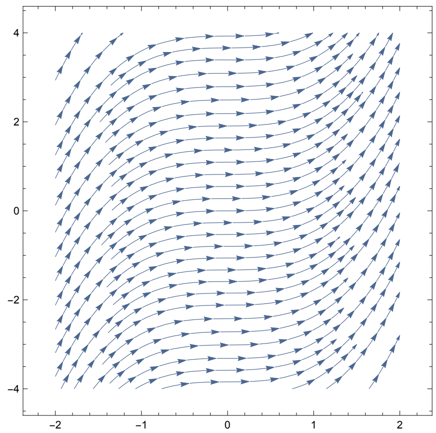



**Figure 19:** Flows in the plane.

(An aside: Given a vector field $X$, it may not be possible to integrate (2.9) to generate a flow defined for all $t \in \mathbf{R}$. For example, consider $M = \mathbf{R}$ with the vector field $X = x^2$. The equation $dx/dt = x^2$, subject to the initial condition $x(0) = a$, has the unique solution $x(t) = a/(1 - at)$ which diverges at $t = 1/a$. Vector fields which generate a flow for all $t \in \mathbf{R}$ are called *complete*. It turns out that all vector fields on a manifold $M$ are complete if $M$ is *compact*. Roughly speaking, "compact" means that $M$ doesn't "stretch to infinity". More precisely, a topological space $M$ is compact if, for any family of open sets covering $M$ there always exists a finite sub-family which also cover $M$. So $\mathbf{R}$ is not compact because the family of sets $\{(-n, n), n \in \mathbf{Z}^+\}$ covers $\mathbf{R}$ but has no finite sub-family. Similarly, $\mathbf{R}^n$ is non-compact. However, $\mathbf{S}^n$ and $\mathbf{T}^n$ are compact manifolds.)

We can look at some examples.

- Consider the sphere $\mathbf{S}^2$ in polar coordinates with the vector field $X = \partial_\phi$. The integral curves solve the equation (2.9), which are

$$\frac{d\phi}{dt} = 1 \quad \text{and} \quad \frac{d\theta}{dt} = 0$$

  This has the solution $\theta = \theta_0$ and $\phi = \phi_0 + t$. The associated one-parameter diffeomorphism is $\sigma_t : (\theta, \phi) \to (\theta, \phi + t)$, and the flow lines are simply lines of constant latitude on the sphere and are shown in the left-hand figure above.

- Alternatively, consider the vector field on $\mathbf{R}^2$ with Cartesian components $X^\mu = (1, x^2)$. The equation for the integral curves is now

$$\frac{dx}{dt} = 1 \quad \text{and} \quad \frac{dy}{dt} = x^2$$

which has the solution $x(t) = x_0 t$ and $y(t) = y_0 + \frac{1}{3}(x_0 + t)^3$. The associated flow lines are shown in the right-hand figure.

### 2.2.4 The Lie Derivative

So far we have learned how to differentiate a function. This requires us to introduce a vector field $X$, and the new function $X(f)$ can be viewed as the derivative of $f$ in the direction of $X$.

Next we ask: is it possible to differentiate a vector field? Specifically, suppose that we have a second vector field $Y$. How can we differentiate this in the direction of $X$ to get a new vector field? As we've seen, we can't just write down $XY$ because this doesn't define a new vector field.

To proceed, we should think more carefully about what differentiation means. For a function $f(x)$ on $\mathbf{R}$, we compare the values of the function at nearby points, and see what happens as those points approach each other

$$\frac{df}{dx} = \lim_{t\to 0} \frac{f(x+t) - f(x)}{t}$$

Similarly, to differentiate a vector field, we need to subtract the tangent vector $Y_p \in T_p(M)$ from the tangent vector at some nearby point $Y_q \in T_q(M)$, and then see what happens in the limit $q \to p$. But that's problematic because, as we stressed above, the vector spaces $T_p(M)$ and $T_q(M)$ are different, and it makes no sense to subtract vectors in one from vectors in the other. To make progress, we're going to have to find a way to do this. Fortunately, there is a way.

#### Push-Foward and Pull-Back

Suppose that we have a map $\varphi : M \to N$ between two manifolds $M$ and $N$. This allows us to import various structures on one manifold to the other.

For example, if we have a function on $f : N \to \mathbf{R}$, then we can construct a new function that we denote $(\varphi^* f) : M \to \mathbf{R}$,

$$(\varphi^* f)(p) = f(\varphi(p))$$

Using the map in this way, to drag objects originally defined on $N$ onto $M$ is called the *pull-back*. If we introduce coordinates $x^\mu$ on $M$ and $y^\alpha$ on $N$, then the map $\varphi(x) = y^\alpha(x)$, and we can write

$$(\varphi^* f)(x) = f(y(x))$$

Some objects more naturally go the other way. For example, given a vector field $Y$ on $M$, we can define a new vector field $(\varphi_* Y)$ on $N$. If we are given a function $f : N \to \mathbf{R}$, then the vector field $(\varphi_* Y)$ on $N$ acts as

$$(\varphi_* Y)(f) = Y(\varphi^* f)$$

where I've been a little sloppy in the notation here since the left-hand side is a function on $N$ and the right-hand side a function on $M$. The equality above holds when evaluated at the appropriate points: $[(\varphi_* Y)(f)](\varphi(p)) = [Y(\varphi^* f)](p)$. Using the map to push objects on $M$ onto $N$ is called the *push-forward*.

If $Y = Y^\mu \partial/\partial x^\mu$ is the vector field on $M$, we can write the induced vector field on $N$ as

$$(\varphi_* Y)(f) = Y^\mu \frac{\partial f(y(x))}{\partial x^\mu} = Y^\mu \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial f(y)}{\partial y^\alpha}$$

Written in components, $(\varphi_* Y) = (\varphi_* Y)^\alpha \partial/\partial y^\alpha$, we then have

$$(\varphi_* Y)^\alpha = Y^\mu \frac{\partial y^\alpha}{\partial x^\mu} \tag{2.11}$$

Given the way that the indices are contracted, this is more or less the only thing we could write down.

We'll see other examples of these induced maps later in the lectures. The push-forward is always denoted as $\varphi_*$ and goes in the same way as the original map. The pull-back is always denoted as $\varphi^*$ and goes in the opposite direction to the original map. Importantly, if our map $\varphi : M \to N$ is a diffeomorphism, then we also have $\varphi^{-1} : N \to M$, so we can transport any object from $M$ to $N$ and back again with impunity.

### Constructing the Lie Derivative

Now we can use these ideas to help build a derivative. Suppose that we are given a vector field $X$ on $M$. This generates a flow $\sigma_t : M \to M$, which is a map between manifolds, now with $N = M$. This means that we can use (2.11) to generate a push-forward map from $T_p(M)$ to $T_{\sigma_t(p)}(M)$. But this is exactly what we need if we want to compare tangent vectors at neighbouring points. The resulting differential operator is called the *Lie derivative* and is denoted $\mathcal{L}_X$.

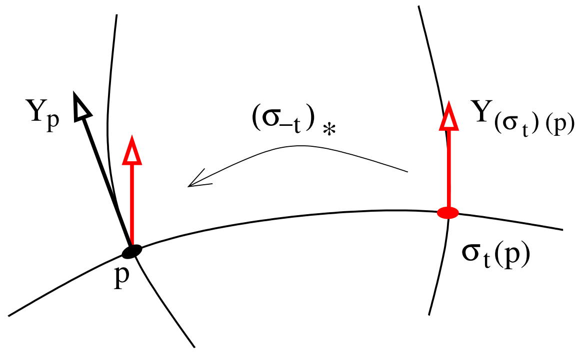



**Figure 20:** To construct the Lie derivative, we use the push-forward $(\sigma_{-t})_\star$ to map the vector $Y_{\sigma_t(p)}$ back to $p$. The resulting vector, shown in red, is $(\sigma_{-t})_\star Y)_p$.

It will turn out that we can use these ideas to differentiate many different kinds of objects. As a warm-up, let's first see how an analogous construction allows us to differentiate functions. Now the function

$$\mathcal{L}_X f = \lim_{t\to 0} \frac{f(\sigma_t(x)) - f(x)}{t} = \left.\frac{df(\sigma_t(x))}{dt}\right|_{t=0} = \left.\frac{\partial f}{\partial x^\mu}\,\frac{dx^\mu}{dt}\right|_{t=0}$$

But, using (2.9), we know that $dx^\mu/dt = X^\mu$. We then have

$$\mathcal{L}_X f = X^\mu(x)\frac{\partial f}{\partial x^\mu} = X(f) \tag{2.12}$$

In other words, acting on functions with the Lie derivative $\mathcal{L}_X$ coincides with action of the vector field $X$.

Now let's look at the action of $\mathcal{L}_X$ on a vector field $Y$. This is defined by

$$\mathcal{L}_X Y = \lim_{t\to 0} \frac{((\sigma_{-t})_* Y)_p - Y_p}{t}$$

Note the minus sign in $\sigma_{-t}$. This reflects that fact that vector fields are pushed, rather than pulled. The map $\sigma_t$ takes us from the point $p$ to the point $\sigma_t(p)$. But to push a tangent vector $Y_{\sigma_t(p)} \in T_{\sigma_t(p)}(M)$ to a tangent vector in $T_p(M)$, where it can be compared to $Y_p$, we need to push with the inverse map $(\sigma_{-t})_*$. This is shown Figure 20.

Let's first calculate the action of $\mathcal{L}_X$ on a coordinate basis $\partial_\mu = \partial/\partial x^\mu$. We have

$$\mathcal{L}_X \partial_\mu = \lim_{t\to 0} \frac{(\sigma_{-t})_*\partial_\mu - \partial_\mu}{t} \tag{2.13}$$

We have an expression for the push-forward of a tangent vector in (2.11), where the coordinates $y^\alpha$ on $N$ should now be replaced by the infinitesimal change of coordinates induced by the flow $\sigma_t$ which, from (2.10) is $x^\mu(t) = x^\mu(0) - tX^\mu + \ldots$. Note the minus sign, which comes from the fact that we have to map back to where we came from as shown in Figure 20. We have, for small $t$,

$$(\sigma_{-t})_*\partial_\mu = \left(\delta^\nu_\mu - t\frac{\partial X^\nu}{\partial x^\mu} + \ldots\right)\partial_\nu$$

Acting on a coordinate basis, we then have

$$\mathcal{L}_X\partial_\mu = -\frac{\partial X^\nu}{\partial x^\mu}\partial_\nu \tag{2.14}$$

To determine the action of $\mathcal{L}_X$ on a general vector field $Y$, we use the fact that the Lie derivative obeys the usual properties that we expect of a derivative, including linearity, $\mathcal{L}_X(Y_1 + Y_2) = \mathcal{L}_X Y_1 + \mathcal{L}_X Y_2$ and Leibnizarity $\mathcal{L}_X(fY) = f\mathcal{L}_X Y + (\mathcal{L}_X f)Y$ for any function $f$, both of which follow from the definition. The action on a general vector field $Y = Y^\mu(x)\partial/\partial x^\mu$ can then be written as

$$\mathcal{L}_X(Y^\mu\partial_\mu) = (\mathcal{L}_X Y^\mu)\partial_\mu + Y^\mu(\mathcal{L}_X\partial_\mu)$$

we're we've simply viewed the components $Y^\mu(x)$ as $n$ functions. We can use (2.12) to determine $\mathcal{L}_X Y^\mu$ and we've computed $\mathcal{L}_X\partial_\mu$ in (2.14). We then have

$$\mathcal{L}_X(Y^\mu\partial_\mu) = X^\nu\frac{\partial Y^\mu}{\partial x^\nu}\partial_\mu - Y^\mu\frac{\partial X^\nu}{\partial x^\mu}\partial_\nu$$

But this is precisely the structure of the commutator. We learn that the Lie derivative acting on vector fields is given by

$$\mathcal{L}_X Y = [X, Y]$$

A corollary of this is

$$\mathcal{L}_X\mathcal{L}_Y Z - \mathcal{L}_Y\mathcal{L}_X Z = \mathcal{L}_{[X,Y]}Z \tag{2.15}$$

which follows from the Jacobi identity for commutators.

The Lie derivative is just one of several derivatives that we will meet in this course. As we introduce new objects, we will learn how to act with $\mathcal{L}_X$ on them. But we will also see that we can endow different meanings to the idea of differentiation. In fact, the Lie derivative will take something of a back seat until Section 4.3 when we will see that it is what we need to understand symmetries.

### 2.3 Tensors

For any vector space $V$, the dual vector space $V^*$ is the space of all linear maps from $V$ to $\mathbf{R}$

This is a standard mathematical construction, but even if you haven't seen it before it should resonate with something you know from quantum mechanics. There we have states in a Hilbert space with with kets $|\psi\rangle \in \mathcal{H}$ and a dual Hilbert space with bras $\langle\phi| \in \mathcal{H}^*$. Any bra can be viewed as a map $\langle\phi| : \mathcal{H} \to \mathbf{R}$ defined by $\langle\phi|(|\psi\rangle) = \langle\phi||\psi\rangle$.

In general, suppose that we are given a basis $\{e_\mu, \mu = 1, \ldots, n\}$ of $V$. Then we can introduce a dual basis $\{f^\mu, \mu = 1, \ldots, n\}$ for $V^*$ defined by

$$f^\nu(e_\mu) = \delta^\nu_\mu$$

A general vector in $V$ can be written as $X = X^\mu e_\mu$ and $f^\nu(X) = X^\mu f^\nu(e_\mu) = X^\nu$. Given a basis, this construction provides an isomorphism between $V$ and $V^*$ given by $e_\mu \to f^\mu$. But the isomorphism is basis dependent. Pick a different basis, and you'll get a different map.

We can repeat the construction and consider $(V^*)^*$, which is the space of all linear maps from $V^*$ to $\mathbf{R}$. But this space is naturally isomorphic to $V$, meaning that the isomorphism is independent of the choice of basis. To see this, suppose that $X \in V$ and $\omega \in V^*$. This means that $\omega(X) \in \mathbf{R}$. But we can equally as well view $X \in (V^*)^*$ and define $X(\omega) = \omega(X) \in \mathbf{R}$. In this sense, $(V^*)^* = V$.

#### 2.3.1 Covectors and One-Forms

At each point $p \in M$, we have a vector space $T_p(M)$. The dual of this space, $T^*_p(M)$ is called the *cotangent space* at $p$, and an element of this space is called a *cotangent vector*, sometimes shortened to *covector*. Given a basis $\{e_\mu\}$ of $T_p(M)$, we can introduce the dual basis $\{f^\mu\}$ for $T^*_p(M)$ and expand any co-vector as $\omega = \omega_\mu f^\mu$.

We can also form fields of cotangent vectors, by picking a member of $T^*_p(M)$ for each point $p$ in a smooth manner. Such a *cotangent field* is better known as a *one-form*; they map vector fields to real numbers. The set of all one-forms on $M$ is denoted $\Lambda^1(M)$.

There is a particularly simple way to construct a one-form. Take a function $f \in C^\infty(M)$ and define $df \in \Lambda^1(M)$ by

$$df(X) = X(f) \tag{2.16}$$

We can use this method to build a basis for $\Lambda^1(M)$. If we introduce coordinates $x^\mu$ on $M$ with the corresponding coordinate basis $e_\mu = \partial/\partial_\mu$ of vector fields, which we often write in shorthand as $\partial/\partial_\mu \equiv \partial_\mu$. We then we simply take the functions $f = x^\mu$ which, from (2.16), gives

$$dx^\mu(\partial_\nu) = \partial_\nu(x^\mu) = \delta^\mu_\nu$$

This means that $f^\mu = dx^\mu$ provides a basis for $\Lambda^1(M)$, dual to the coordinate basis $\partial/\partial_\mu$. In general, an arbitrary one-form $\omega \in \Lambda^1(M)$ can then be expanded as

$$\omega = \omega_\mu \, dx^\mu$$

In such a basis the one-form $df$ takes the form

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu \tag{2.17}$$

To see this, we simply need to evaluate

$$df(X) = \frac{\partial f}{\partial x^\mu} dx^\mu(X^\nu \partial_\nu) = X^\mu \frac{\partial f}{\partial x^\mu} = X(f)$$

which agrees with the expected answer (2.16).

As with vector fields, we can look at what happens if we change coordinates. Given two different charts, $\phi = (x^1, \ldots, x^n)$ and $\tilde{\phi} = (\tilde{x}^1, \ldots, \tilde{x}^n)$, we know that the basis for vector fields changes as (2.5),

$$\frac{\partial}{\partial \tilde{x}^\mu} = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu}$$

We should take the basis of one-forms to transform in the inverse manner,

$$d\tilde{x}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu \tag{2.18}$$

This then ensures that

$$d\tilde{x}^\mu \left( \frac{\partial}{\partial \tilde{x}^\nu} \right) = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \, dx^\rho \left( \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \frac{\partial}{\partial x^\sigma} \right) = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \, dx^\rho \left( \frac{\partial}{\partial x^\sigma} \right) = \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \delta^\rho_\sigma$$

But this is just the multiplication of a matrix and its inverse,

$$\frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial x^\rho}{\partial \tilde{x}^\nu} = \delta^\mu_\nu$$

So we find that

$$d\tilde{x}^\mu \left( \frac{\partial}{\partial \tilde{x}^\nu} \right) = \delta^\mu_\nu$$

as it should. We can then expand a one-form $\omega$ in either of these two bases,

$$\omega = \omega_\mu dx^\mu = \tilde{\omega}_\mu d\tilde{x}^\mu \quad \text{with} \quad \tilde{\omega}_\mu = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \omega_\nu \tag{2.19}$$

In the annoying language that I can never remember, components of vectors that transform this way are said to be *covariant*. Note that, as with vector fields, the placement of the indices means that (2.18) and (2.19) are pretty much the only things that you can write down that make sense.

### 2.3.2 The Lie Derivative Revisited

In Section 2.2.4, we explained how to construct the Lie derivative, which differentiates a vector field in the direction of a second vector field $X$. This same idea can be adapted to one-forms.

Under a map $\varphi : M \to N$, we saw that a vector field $X$ on $M$ can be pushed forwards to a vector field $\varphi_* X$ on $N$. In contrast, one-forms go the other way: given a one-form $\omega$ on $N$, we can pull this back to a one-form $(\varphi^* \omega)$ on $M$, defined by

$$(\varphi^* \omega)(X) = \omega(\varphi_* X)$$

If we introduce coordinates $x^\mu$ on $M$ and $y^\alpha$ on $N$ then the components of the pull-back are given by

$$(\varphi^* \omega)_\mu = \omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu} \tag{2.20}$$

We now define the Lie derivative $\mathcal{L}_X$ acting on one-forms. Again, we use $X$ to generate a flow $\sigma_t : M \to M$ which, using the pull-back, allows us to compare one-forms at different points. We will denote the cotangent vector $\omega(p)$ as $\omega_p$. The Lie derivative of a one-form $\omega$ is then defined as

$$\mathcal{L}_X \omega = \lim_{t \to 0} \frac{(\sigma_t^* \omega)_p - \omega_p}{t} \tag{2.21}$$

Note that we pull-back with the map $\sigma_t$. This is to be contrasted with (2.13) where we pushed forward the tangent vector with the map $\sigma_{-t}$ and, as we now show, this difference in minus sign manifests itself in the expression for the Lie derivative. The infinitesimal map $\sigma_t$ acts on coordinates as $x^\mu(t) = x^\mu(0) + tX^\mu + \ldots$ so, from (2.20), the pull-back of a basis vector $dx^\mu$ is

$$\sigma_t^*\, dx^\mu = \left(\delta^\mu_\nu + t\frac{\partial X^\mu}{\partial x^\nu} + \ldots\right) dx^\nu$$

Acting on the coordinate basis, we then have

$$\mathcal{L}_X(dx^\mu) = \frac{\partial X^\mu}{\partial x^\nu} dx^\nu$$

which indeed differs by a minus sign from the corresponding result (2.14) for tangent vectors. Acting on a general one-form $\omega = \omega_\mu dx^\mu$, the Lie derivative is

$$\begin{aligned} \mathcal{L}_X\omega &= (\mathcal{L}_X\omega_\mu)dx^\mu + \omega_\nu\mathcal{L}_X(dx^\nu) \\ &= (X^\nu\partial_\nu\omega_\mu + \omega_\nu\partial_\mu X^\nu)\, dx^\mu \end{aligned} \tag{2.22}$$

We'll return to discuss one-forms (and other forms) more in Section 2.4.

### 2.3.3 Tensors and Tensor Fields

A *tensor of rank* $(r, s)$ at a point $p \in M$ is defined to be a multi-linear map

$$T : \overbrace{T_p^*(M) \times \ldots \times T_p^*(M)}^{r} \times \overbrace{T_p(M) \times \ldots \times T_p(M)}^{s} \to \mathbf{R}$$

Such a tensor is said to have total rank $r + s$.

We've seen some examples already. A cotangent vector in $T_p^*(M)$ is a tensor of type $(0, 1)$, while a tangent vector in $T_p(M)$ is a tensor of type $(1, 0)$ (using the fact that $T_p(M) = T_p^{**}(M)$).

As before, we define a *tensor field* to be a smooth assignment of an $(r, s)$ tensor to every point $p \in M$.

Given a basis $\{e_\mu\}$ for vector fields and a dual basis $\{f^\mu\}$ for one-forms, the components of the tensor are defined to be

$$T^{\mu_1\ldots\mu_r}{}_{\nu_1\ldots\nu_s} = T(f^{\mu_1}, \ldots, f^{\mu_r}, e_{\nu_1}, \ldots, e_{\mu_s})$$

Note that we deliberately write the string of lower indices after the upper indices. In some sense this is unnecessary, and we don't lose any information by writing $T^{\mu_1\ldots\mu_r}_{\nu_1\ldots\nu_s}$. Nonetheless, we'll see later that it's a useful habit to get into.

On a manifold of dimension $n$, there are $n^{r+s}$ such components. For a tensor field, each of these is a function over $M$.

As an example, consider a rank $(2,1)$ tensor. This takes two one-forms, say $\omega$ and $\eta$, together with a vector field $X$, and spits out a real number. In a given basis, this number is

$$T(\omega,\eta,X) = T(\omega_\mu f^\mu, \eta_\nu f^\nu, X^\rho e_\rho) = \omega_\mu \eta_\nu X^\rho\, T(f^\mu, f^\nu, e_\rho) = T^{\mu\nu}{}_\rho\, \omega_\mu \eta_\nu X^\rho$$

Every manifold comes equipped with a natural $(1,1)$ tensor called $\delta$. This takes a one-form $\omega$ and a vector field $X$ and spits out the real number

$$\delta(\omega, X) = \omega(X) \quad \Rightarrow \quad \delta(f^\mu, e_\nu) = f^\mu(e_\nu) = \delta^\mu_\nu$$

which is simply the Kronecker delta.

As with vector fields and one-forms, we can ask how the components of a tensor transform. We will work more generally than before. Consider two bases for the vector fields, $\{e_\mu\}$ and $\{\tilde{e}_\mu\}$, not necessarily coordinate bases, related by

$$\tilde{e}_\nu = A^\mu{}_\nu e_\mu$$

for some invertible matrix $A$. The respective dual bases are $\{f^\mu\}$ and $\{\tilde{f}^\mu\}$ are then related by

$$\tilde{f}^\rho = B^\rho{}_\sigma f^\sigma$$

such that

$$\tilde{f}^\rho(\tilde{e}_\nu) = A^\mu{}_\nu B^\rho{}_\sigma f^\sigma(e_\mu) = A^\mu{}_\nu B^\rho{}_\mu = \delta^\rho_\nu \quad \Rightarrow \quad B^\rho{}_\mu = (A^{-1})^\rho{}_\mu$$

The lower components of a tensor then transform by multiplying by $A$, and the upper components by multiplying by $B = A^{-1}$. So, for example, a rank $(1,2)$ tensor transforms as

$$\tilde{T}^\mu{}_{\rho\nu} = B^\mu{}_\sigma A^\tau{}_\rho A^\lambda{}_\nu T^\sigma{}_{\tau\lambda} \tag{2.23}$$

When we change between coordinate bases, we have

$$A^\mu{}_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \quad \text{and} \quad B^\mu{}_\nu = (A^{-1})^\mu{}_\nu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu}$$

You can check that this coincides with our previous results (2.6) and (2.19).

**Operations on Tensor Fields**

There are a number of operations that we can do on tensor fields to generate further tensors.

First, we can add and subtract tensors fields, or multiply them by functions. This is the statement that the set of tensors at a point $p \in M$ forms a vector space.

Next, there is a way to multiply tensors together to give a tensor of a different type. Given a tensor $S$ of rank $(p,q)$ and a tensor $T$ of rank $(r,s)$, we can form the *tensor product*, $S \otimes T$ which a new tensor of rank $(p+r, q+s)$, defined by

$$\begin{aligned} S \otimes T(\omega_1, \ldots, \omega_p, \eta_1, \ldots, \eta_r, X_1, \ldots X_q, Y_1, \ldots, Y_s) \\ = S(\omega_1, \ldots, \omega_p, X_1, \ldots X_q) T(\eta_1, \ldots, \eta_r, Y_1, \ldots, Y_s) \end{aligned}$$

In terms of components, this reads

$$(S \otimes T)^{\mu_1 \ldots \mu_p \nu_1 \ldots \nu_r}{}_{\rho_1 \ldots \rho_q \sigma_1 \ldots \sigma_s} = S^{\mu_1 \ldots \mu_p}{}_{\rho_1 \ldots \rho_q} T^{\nu_1 \ldots \nu_r}{}_{\sigma_1 \ldots \sigma_s} \tag{2.24}$$

Given an $(r,s)$ tensor $T$, we can also construct a tensor of lower rank $(r-1, s-1)$ by contraction. To do this, simply replace one of $T_p^*(M)$ entries with a basis vector $f^\mu$, and the corresponding $T_p(M)$ entry with the dual vector $e_\mu$ and then sum over $\mu = 1, \ldots, n$. So, for example, given a rank $(2,1)$ tensor $T$ we can construct a rank $(1,0)$ tensor $S$ by

$$S(\omega) = T(\omega, f^\mu, e_\mu)$$

Alternatively, we could construct a (typically) different $(1,0)$ tensor by contracting the other argument, $S'(\omega) = T(f^\mu, \omega, e_\mu)$. Written in terms of components, contraction simply means that we put an upper index equal to a lower index and sum over them,

$$S^\mu = T^{\mu\nu}{}_\nu \quad \text{and} \quad S'^{\,\mu} = T^{\nu\mu}{}_\nu$$

Our next operation is symmetrisation and anti-symmetrisation. For example, given a $(0,2)$ tensor $T$ we decompose it into two $(0,2)$ tensors, in which the arguments are either symmetrised or anti-symmetrised,

$$\begin{aligned} S(X,Y) &= \frac{1}{2}\Big(T(X,Y) + T(Y,X)\Big) \\ A(X,Y) &= \frac{1}{2}\Big(T(X,Y) - T(Y,X)\Big) \end{aligned}$$

In index notation, this becomes

$$S_{\mu\nu} = \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu}) \quad \text{and} \quad A_{\mu\nu} = \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$$

which is just like taking the symmetric and anti-symmetric part of a matrix. We will work these operations frequently enough to justify introducing some new notation. We define

$$T_{(\mu\nu)} = \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu}) \quad \text{and} \quad T_{[\mu\nu]} = \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$$

These operations generalise to other tensors. For example,

$$T^{(\mu\nu)\rho}{}_{\sigma} = \frac{1}{2}\Big(T^{\mu\nu\rho}{}_{\sigma} + T^{\nu\mu\rho}{}_{\sigma}\Big)$$

We can also symmetrise or anti-symmetrise over multiple indices, provided that these indices are either all up or all down. If we (anti)-symmetrise over $p$ objects, then we divide by $p!$, which is the number of possible permutations. This normalisation ensures that if we start with a tensor which is already, say, symmetric then further symmetrising doesn't affect it. In the case of anti-symmetry, we weight each term with the sign of the permutation. So, for example,

$$T^{\mu}{}_{(\nu\rho\sigma)} = \frac{1}{3!}\Big(T^{\mu}{}_{\nu\rho\sigma} + T^{\mu}{}_{\rho\nu\sigma} + T^{\mu}{}_{\rho\sigma\nu} + T^{\mu}{}_{\sigma\rho\nu} + T^{\mu}{}_{\sigma\nu\rho} + T^{\mu}{}_{\nu\sigma\rho}\Big)$$

and

$$T^{\mu}{}_{[\nu\rho\sigma]} = \frac{1}{3!}\Big(T^{\mu}{}_{\nu\rho\sigma} - T^{\mu}{}_{\rho\nu\sigma} + T^{\mu}{}_{\rho\sigma\nu} - T^{\mu}{}_{\sigma\rho\nu} + T^{\mu}{}_{\sigma\nu\rho} - T^{\mu}{}_{\nu\sigma\rho}\Big)$$

There will times when, annoyingly, we will wish to symmetrise (or anti-symmetrise) over indices which are not adjacent. We introduce vertical bars to exclude certain indices from the symmetry procedure. So, for example,

$$T^{\mu}{}_{[\nu|\rho|\sigma]} = \frac{1}{2}(T^{\mu}{}_{\nu\rho\sigma} - T^{\mu}{}_{\sigma\rho\nu})$$

Finally, given a smooth tensor field $T$ of any rank, we can always take the Lie derivative with respect to a vector field $X$. As we've seen previously, under a map $\varphi : M \to N$, vector fields are pushed forwards and one-forms are pulled-back. In general, this leaves a tensor of mixed type unsure where to go. However, if $\varphi$ is a diffeomorphism then we also have $\varphi^{-1} : N \to M$ and this allows us to define the push-forward of a tensor $T$ from $M$ to $N$. This acts on one-forms $\omega \in \Lambda^1(N)$ and vector fields $X \in \mathfrak{X}(N)$ and is given by

$$(\varphi_* T)(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) = T(\varphi^*\omega_1, \ldots, \varphi^*\omega_r, (\varphi_*^{-1}X_1), \ldots, (\varphi_*^{-1}X_s))$$

Here $\varphi^*\omega$ are the pull-backs of $\omega$ from $N$ to $M$, while $\varphi_*^{-1}X$ are the push-forwards of $X$ from $N$ to $M$.

The Lie derivative of a tensor $T$ along $X$ is then defined as

$$\mathcal{L}_X T = \lim_{t \to 0} \frac{((\sigma_{-t})_* T)_p - T_p}{t}$$

where $\sigma_t$ is the flow generated by $X$. This coincides with our earlier definitions for vector fields in (2.13) and for one-forms in (2.21). (The difference in the $\sigma_{-t}$ vs $\sigma_t$ minus sign in (2.13) and (2.21) is now hiding in the inverse push-forward $\varphi_*^{-1}$ that appears in the definition $\varphi_* T$.)

### 2.4 Differential Forms

Some tensors are more interesting than others. A particularly interesting class are totally anti-symmetric $(0, p)$ tensors fields. These are called *p-forms*. The set of all $p$-forms over a manifold $M$ is denoted $\Lambda^p(M)$.

We've met some forms before. A 0-form is simply a function. Meanwhile, as we saw previously, a 1-form is another name for a covector. The anti-symmetry means that we can't have any form of degree $p > n = \dim(M)$. A $p$-form has $\binom{n}{p}$ different components. Forms in $\Lambda^n(M)$ are called *top forms*.

Given a $p$-form $\omega$ and a $q$-form $\eta$, we can take the tensor product (2.24) to construct a $(p+q)$-tensor. If we anti-symmetrise this, we then get a $(p+q)$-form. This construction is called the *wedge product*, and is defined by

$$(\omega \wedge \eta)_{\mu_1 \ldots \mu_p \nu_1 \ldots \nu_q} = \frac{(p+q)!}{p!q!} \omega_{[\mu_1 \ldots \mu_p} \eta_{\nu_1 \ldots \nu_q]}$$

where the $[\ldots]$ in the subscript tells us to anti-symmetrise over all indices. For example, given $\omega, \eta \in \Lambda^1(M)$, we can construct a 2-form

$$(\omega \wedge \eta)_{\mu\nu} = \omega_\mu \eta_\nu - \omega_\nu \eta_\mu$$

For one forms, the anti-symmetry ensures that $\omega \wedge \omega = 0$. In general, if $\omega \in \Lambda^p(M)$ and $\eta \in \Lambda^q(M)$, then one can show that

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$$

This means that $\omega \wedge \omega = 0$ for any form of odd degree. We can, however, wedge even degree forms with themselves. (Which you know already for 0-forms where the wedge product is just multiplication of functions.)

As a more specific example, consider $M = \mathbf{R}^3$ and $\omega = \omega_\mu dx^\mu$ and $\eta = \eta_\mu dx^\mu$. We then have

$$\begin{aligned}\omega \wedge \eta &= (\omega_1 dx^1 + \omega_2 dx^2 + \omega_3 dx^3) \wedge (\eta_1 dx^1 + \eta_2 dx^2 + \eta_3 dx^3) \\ &= (\omega_1\eta_2 - \omega_2\eta_1) dx^1 \wedge dx^2 + (\omega_2\eta_3 - \omega_3\eta_2) dx^2 \wedge dx^3 + (\omega_3\eta_1 - \omega_1\eta_3) dx^3 \wedge dx^1\end{aligned}$$

Notice that the components that arise are precisely those of the cross-product acting on vectors in $\mathbf{R}^3$. This is no coincidence: what we usually think of as the cross-product between vectors is really a wedge product between forms. We'll have to wait to Section 3 to understand how to map from one to the other.

It can also be shown that the wedge product is associative, meaning

$$\omega \wedge (\eta \wedge \lambda) = (\omega \wedge \eta) \wedge \lambda$$

We can then drop the brackets in any such product.

Given a basis $\{f^\mu\}$ of $\Lambda^1(M)$, a basis of $\Lambda^p(M)$ can be constructed by wedge products $\{f^{\mu_1} \wedge \ldots \wedge f^{\mu_p}\}$. We will usually work with the coordinate basis $\{dx^\mu\}$. This means that any $p$-form $\omega$ can be written locally as

$$\omega = \frac{1}{p!}\, \omega_{\mu_1 \ldots \mu_p}\, dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_p} \tag{2.25}$$

Although locally any $p$-form can be written as (2.25), this may not be true globally. This, and related issues, will become of some interest in Section 2.4.3.

### 2.4.1 The Exterior Derivative

We learned in Section 2.3.1 how to construct a one-form $df$ from a function $f$. In a coordinate basis, this one-form has components (2.17),

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu$$

We can extend this definition to higher forms. The *exterior derivative* is a map

$$d : \Lambda^p(M) \to \Lambda^{p+1}(M)$$

In local coordinates (2.25), the exterior derivative acts as

$$(d\omega) = \frac{1}{p!} \frac{\partial \omega_{\mu_1 \ldots \mu_p}}{\partial x^\nu} dx^\nu \wedge dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_p} \tag{2.26}$$

Equivalently we have

$$(d\omega)_{\mu_1\ldots\mu_{p+1}} = (p+1)\,\partial_{[\mu_1}\omega_{\mu_2\ldots\mu_{p+1}]} \tag{2.27}$$

Importantly, if we subsequently act with the exterior derivative again, we get

$$d(d\omega) = 0$$

because the derivatives are anti-symmetrised and hence vanish. This holds true for any $p$-form, a fact which is sometimes expressed as

$$d^2 = 0$$

It can be shown that the exterior derivative satisfies a number of further properties,

- $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^p \omega \wedge d\eta$, where $\omega \in \Lambda^p(M)$.
- $d(\varphi^*\omega) = \varphi^*(d\omega)$ where $\varphi^*$ is the pull-back associated to the map between manifolds, $\varphi : M \to N$
- Because the exterior derivative commutes with the pull-back, it also commutes with the Lie derivative. This ensures that we have $d(\mathcal{L}_X\omega) = \mathcal{L}_X(d\omega)$.

A $p$-form $\omega$ is said to be *closed* if $d\omega = 0$ everywhere. It is *exact* if $\omega = d\eta$ everywhere for some $\eta$. Because $d^2 = 0$, an exact form is necessary closed. The question of when the converse is true is interesting: we'll discuss this more in Section 2.4.3.

**Examples**

Suppose that we have a one-form $\omega = \omega_\mu dx^\mu$, the exterior derivative gives a 2-form

$$(d\omega)_{\mu\nu} = \partial_\mu\omega_\nu - \partial_\nu\omega_\mu \quad \Rightarrow \quad d\omega = \frac{1}{2}(\partial_\mu\omega_\nu - \partial_\nu\omega_\mu)dx^\mu \wedge dx^\nu$$

As a specific example of this example, suppose that we take the one-form to live on $\mathbf{R}^3$, with

$$\omega = \omega_1 dx^1 + \omega_2 dx^2 + \omega_3 dx^3$$

Since this is a field, each of the components $\omega_\mu$ is a function of $x^1$, $x^2$ and $x^3$. The exterior derivative is given by

$$\begin{aligned} d\omega &= \partial_2\omega_1\, dx^2 \wedge dx^1 + \partial_3\omega_1\, dx^3 \wedge dx^1 + \partial_1\omega_2\, dx^1 \wedge dx^2 \\ &\qquad + \partial_3\omega_2\, dx^3 \wedge dx^2 + \partial_1\omega_3\, dx^1 \wedge dx^3 + \partial_2\omega_3\, dx^2 \wedge dx^3 \\ &= (\partial_1\omega_2 - \partial_2\omega_1)dx^1 \wedge dx^2 + (\partial_2\omega_3 - \partial_3\omega_2)dx^2 \wedge dx^3 + (\partial_3\omega_1 - \partial_1\omega_3)dx^3 \wedge dx^1 \end{aligned} \tag{2.28}$$

Notice that there's no term like $\partial_1\omega_1$ because this would come with a $dx^1 \wedge dx^1 = 0$.

In the olden days (before this course), we used to write vector fields in $\mathbf{R}^3$ as $\boldsymbol{\omega} = (\omega^1, \omega^2, \omega^3)$ and compute the curl $\nabla \times \boldsymbol{\omega}$. But the components of the curl are precisely the components that appear in $d\omega$. In fact, our "vector" $\boldsymbol{\omega}$ was really a one-form and the curl turned it into a two-form. It's a happy fact that in $\mathbf{R}^3$, vectors, one-forms and two-forms all have three components, which allowed us to conflate them in our earlier courses. (In fact, there is a natural map between them that we will meet in Section 3.)

Suppose instead that we start with a 2-form $B$ in $\mathbf{R}^3$, which we write as

$$B = B_1\, dx^2 \wedge dx^3 + B_2\, dx^3 \wedge dx^1 + B_3\, dx^1 \wedge dx^2$$

Taking the exterior derivative now gives

$$\begin{aligned} dB &= \partial_1 B_1\, dx^1 \wedge dx^2 \wedge dx^3 + \partial_2 B_2\, dx^2 \wedge dx^3 \wedge dx^1 + \partial_3 B_3\, dx^3 \wedge dx^1 \wedge dx^2 \\ &= (\partial_1 B_1 + \partial_2 B_2 + \partial_3 B_3)\, dx^1 \wedge dx^2 \wedge dx^3 \end{aligned} \tag{2.29}$$

This time there is just a single component, but again it's something familiar. Had we written the original three components of the two-form in old school vector notation $\mathbf{B} = (B_1, B_2, B_3)$, then the single component of $dB$ is what we previously called $\nabla \cdot \mathbf{B}$.

**The Lie Derivative Yet Again**

There is yet another operation that we can construct on $p$-forms. Given a vector field $X \in \mathfrak{X}(M)$, we can construct the *interior product*, a map $\iota_X : \Lambda^p(M) \to \Lambda^{p-1}(M)$. If $\omega \in \Lambda^p(M)$, we define a $\iota_X \omega \in \Lambda^{p-1}(M)$ by

$$\iota_X \omega(Y_1, \ldots, Y_{p-1}) = \omega(X, Y_1, \ldots, Y_{p-1}) \tag{2.30}$$

In other words, we just put $X$ in the first argument of $\omega$. Acting on functions $f$, we simply define $\iota_X f = 0$.

The anti-symmetry of forms means that $\iota_X \iota_Y = -\iota_Y \iota_X$. Moreover, you can check that

$$\iota_X(\omega \wedge \eta) = \iota_X \omega \wedge \eta + (-1)^p \omega \wedge \iota_X \eta$$

where $\omega \in \Lambda^p(M)$.

Consider a 1-form $\omega$. There are two different ways to act with $\iota_X$ and $d$ to give us back a one-form. These are

$$\iota_X d\omega = \iota_X \frac{1}{2}(\partial_\mu \omega_\nu - \partial_\nu \omega \mu) dx^\mu \wedge dx^\nu = X^\mu \partial_\mu \omega_\nu dx^\nu - X^\nu \partial_\mu \omega_\nu dx^\mu$$

and

$$d\iota_X\omega = d(\omega_\mu X^\mu) = X^\mu\partial_\nu\omega_\mu dx^\nu + \omega_\mu\partial_\nu X^\mu dx^\nu$$

Adding the two together gives

$$(d\iota_X + \iota_X d)\omega = (X^\mu\partial_\mu\omega_\nu + \omega_\mu\partial_\nu X^\mu)dx^\nu$$

But this is exactly the same expression we saw previously when computing the Lie derivative (2.22) of a one-form. We learn that

$$\mathcal{L}_X\omega = (d\iota_X + \iota_X d)\omega \tag{2.31}$$

This expression is sometimes referred to as *Cartan's magic formula*. A similar calculation shows that (2.31) holds for any $p$-form $\omega$.

### 2.4.2 Forms You Know and Love

There are a number of examples of differential forms that you've met already, but likely never called them by name.

#### The Electromagnetic Field

The electromagnetic gauge field $A_\mu = (\phi, \mathbf{A})$ should really be thought of as the components of a one-form on spacetime $\mathbf{R}^4$. (Here I've set $c = 1$.) We write

$$A = A_\mu(x)dx^\mu$$

Taking the exterior derivative yields a 2-form $F = dA$, given by

$$F = \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu = \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)dx^\mu \wedge dx^\nu$$

But this is precisely the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ that we met in our lectures on Electromagnetism. The components are the electric and magnetic fields, arranged as

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{2.32}$$

By construction, we also have $dF = d^2A = 0$. In this context, this is sometimes called the *Bianchi identity*; it yields two of the four Maxwell equations. In old school vector calculus notation, these are $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E} + \partial\mathbf{B}/\partial t = 0$. We need a little more structure to get the other two as we will see later in this chapter.

The gauge field $A$ is not unique. Given any function $\alpha$, we can always shift it by a gauge transformation

$$A \to A + d\alpha \quad \Rightarrow \quad A_\mu \to A_\mu + \partial_\mu \alpha$$

This leaves the field strength invariant because $F \to F + d(d\alpha) = F$.

**Phase Space and Hamilton's Equations**

In classical mechanics, the phase space is a manifold $M$ parameterised by coordinates $(q^i, p_j)$ where $q^i$ are the positions of particles and $p_j$ the momenta. Recall from our lectures on Classical Dynamics that the Hamiltonian $H(q,p)$ is a function on $M$, and Hamilton's equations are

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} \tag{2.33}$$

Phase space also comes with a structure called a *Poisson bracket*, defined on a pair of functions $f$ and $g$ as

$$\{f, g\} = \frac{\partial f}{\partial q^j}\frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q^j}$$

Then the time evolution of any function $f$ can be written as

$$\dot{f} = \{f, H\}$$

which reproduces Hamilton's equations if we take $f = q^i$ or $f = p_i$.

Underlying this story is the mathematical structure of forms. The key idea is that we have a manifold $M$ and a function $H$ on $M$. We want a machinery that turns the function $H$ into a vector field $X_H$. Particles then follow trajectories in phase space that are integral curves generated by $X_H$.

To achieve this, we introduce a *symplectic two-form* $\omega$ on an even-dimensional manifold $M$. This two form must be closed, $d\omega = 0$, and non-degenerate, which means that the top form $\omega \wedge \ldots \wedge \omega \neq 0$. We'll see why we need these requirements as we go along. A manifold $M$ equipped with a symplectic two-form is called a *symplectic manifold*.

Any 2-form provides a map $\omega : T_p(M) \to T^*_p(M)$. Given a vector field $X \in \mathfrak{X}(M)$, we can simply take the interior product with $\omega$ to get a one-form $\iota_X\omega$. However, we want to go the other way: given a function $H$, we can always construct a one-form $dH$, and we'd like to exchange this for a vector field. We can do this if the map

$\omega : T_p(M) \to T_p^*(M)$ is actually an isomorphism, so the inverse exists. This turns out to be true provided that $\omega$ is non-degenerate. In this case, we can define the vector field $X_H$ by solving the equation

$$\iota_{X_H}\omega = -dH \tag{2.34}$$

If we introduce coordinates $x^\mu$ on the manifold, then the component form of this equation is

$$X^\mu_H \omega_{\mu\nu} = -\partial_\nu H$$

We denote the inverse as $\omega^{\mu\nu} = -\omega^{\nu\mu}$ such that $\omega^{\mu\nu}\omega_{\nu\rho} = \delta^\mu_\rho$. The components of the vector field are then

$$X^\mu_H = -\omega^{\nu\mu}\partial_\nu H = \omega^{\mu\nu}\partial_\nu H$$

The integral curves generated by $X_H$ obey the differential equation (2.9)

$$\frac{dx^\mu}{dt} = X^\mu_H = \omega^{\mu\nu}\partial_\nu H$$

These are the general form of Hamilton's equations. They reduce to our earlier form (2.33) if we write $x^\mu = (q^i, p_j)$ are choose the symplectic form to have block diagonal form $\omega^{\mu\nu} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

To define the Poisson structure, we first note that we can repeat the map (2.34) to turn any function $f$ into a vector field $X_f$ obeying $\iota_{X_f}\omega = -df$. But we can then feed these vector fields back into the original 2-form $\omega$. This gives us a Poisson bracket,

$$\{f, g\} = \omega(X_g, X_f) = -\omega(X_f, X_g)$$

Or, in components,

$$\{f, g\} = \omega^{\mu\nu}\partial_\mu f\,\partial_\nu g$$

There are many other ways to write this Poisson bracket structure in invariant form. For example, backtracking through various definitions we find

$$\{f, g\} = -\iota_{X_f}\omega(X_g) = df(X_g) = X_g(f)$$

The equation of motion in Poisson bracket structure is then

$$\dot{f} = \{f, H\} = X_H(f) = \mathcal{L}_{X_H} f$$

which tells us that the Lie derivative along $X_H$ generates time evolution.

We haven't yet explained why the symplectic two-form must be closed, $d\omega = 0$. You can check that this is needed so that the Poisson bracket obeys the Jacobi identity. Alternatively, it ensures that the symplectic form itself is invariant under Hamiltonian flow, in the sense that $\mathcal{L}_{X_H}\omega = 0$. To see this, we use (2.31)

$$\mathcal{L}_{X_H}\omega = (d\iota_{X_H} + \iota_{X_H}d)\omega = \iota_{X_H}d\omega$$

The second equality follows from the fact that $d\iota_{X_H}\omega = -d(dH) = 0$. If we insist that $d\omega = 0$ then we find $\mathcal{L}_{X_H}\omega = 0$ as promised.

### 2.4.3 A Sniff of de Rham Cohomology

The exterior derivative is a map which squares to zero, $d^2 = 0$. It turns out that one can have a lot of fun with such maps. We will now explore a little bit of this fun.

First a repeat of definitions we met already: a $p$-form is *closed* if $d\omega = 0$ everywhere. A $p$-form is *exact* is $\omega = d\eta$ everywhere for some $\eta$. Because $d^2 = 0$, exact implies closed. However, the converse is not necessarily true. It turns out that the way in which closed forms fail to be exact captures interesting facts about the topology of the underlying manifold.

We've met this kind of question before. In electromagnetism, we have a magnetic field $\mathbf{B}$ which obeys $\nabla \cdot \mathbf{B} = 0$. We then argue that this means we can write the magnetic field as $\mathbf{B} = \nabla \times \mathbf{A}$. This is more properly expressed the language of forms. We we saw in the previous section, the magnetic field is really a 2-form

$$B = B_1\, dx^2 \wedge dx^3 + B_2\, dx^3 \wedge dx^1 + B_3\, dx^1 \wedge dx^2$$

We computed the exterior derivative in (2.29); it is

$$dB = (\partial_1 B_1 + \partial_2 B_2 + \partial_3 B_3)\, dx^1 \wedge dx^2 \wedge dx^3$$

We see that the Maxwell equation $\nabla \cdot \mathbf{B} = 0$ is really the statement that $B$ is a closed two-form, obeying $dB = 0$. We also saw in (2.28) if we write $B = dA$ for some one-form $A$, then the components are given by $\mathbf{B} = \nabla \times \mathbf{A}$. Clearly writing $B = dA$ ensures that $dB = 0$. But when is the converse true? We have the following statement (which we leave unproven)

**Theorem (The Poincaré Lemma)**: On $M = \mathbf{R}^n$, closed implies exact.

Since we've spent a lot of time mapping manifolds to $\mathbf{R}^n$, this also has consequence for a general manifold $M$. It means that if $\omega$ is a closed $p$-form, then in any neighbourhood $\mathcal{O} \subset M$ it is always possible to find a $\eta \in \Lambda^{p-1}(M)$ such that $\omega = d\eta$ on $\mathcal{O}$. The catch is that it may not be possible to find such an $\eta$ everywhere on the manifold.

**An Example**

Consider the one-dimensional manifold $M = \mathbf{R}$. We can take a one-form $\omega = f(x)dx$. This is always closed because it is a top form. It is also exact. We introduce the function

$$g(x) = \int_0^x dx'\ f(x')$$

Then $\omega = dg$.

Now consider the topologically more interesting one-dimensional manifold $\mathbf{S}^1$, which we can view as the phase $e^{i\theta} \in \mathbf{C}$. We can introduce the form $\omega = d\theta$ on $\mathbf{S}^1$. The way its written makes it look like its an exact form, but this is an illusion because, as we stressed in Section 2.1, $\theta$ is not a good coordinate everywhere on $\mathbf{S}^1$ because it's not single valued. Indeed, it's simple to see that there is no single-valued function $g(\theta)$ on $\mathbf{S}^1$ such that $\omega = dg$. So on $\mathbf{S}^1$, we can construct a form which, locally, can be written as $d\theta$ but globally cannot be written as $d$(something). So we have a form that is closed but not exact.

**Another Example**

On $M = \mathbf{R}^2$, the Poincaré lemma ensures that all closed forms are exact. However, things change if we remove a single point and consider $\mathbf{R}^2 - \{0, 0\}$. Consider the one-form defined by

$$\omega = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$$

This is not a smooth one-form on $\mathbf{R}^2$ because of the divergence at the origin. But removing that point means that $\omega$ becomes acceptable. We can check that $\omega$ is closed,

$$d\omega = -\frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right) dy \wedge dx + \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) dx \wedge dy = 0$$

where the $= 0$ follows from a little bit of algebra. $\omega$ is exact if we can find a function $f$, defined everywhere on $\mathbf{R}^2 - \{0, 0\}$ such that $\omega = df$, which means

$$\omega = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \quad \Rightarrow \quad \frac{\partial f}{\partial x} = -\frac{y}{x^2+y^2} \ \text{ and } \ \frac{\partial f}{\partial y} = \frac{x}{x^2+y^2}$$

We can certainly integrate these equations; the result is

$$f(x, y) = \tan^{-1}\left(\frac{y}{x}\right) + \text{constant}$$

But this is not a smooth function everywhere on $\mathbf{R}^2 - \{0, 0\}$. This means that we can't, in fact, write $\omega = df$ for a well defined function on $\mathbf{R}^2 - \{0, 0\}$. We learn that removing a point makes a big difference: now closed no longer implies exact.

There is a similar story for $\mathbf{R}^3$. Indeed, this is how magnetic monopoles sneak back into physics. You can learn more about this in the lectures on Gauge Theory.

### Betti Numbers

We denote the set of all closed $p$-forms on a manifold $M$ as $Z^p(M)$. Equivalently, $Z^p(M)$ is the kernel of the map $d : \Lambda^p(M) \to \Lambda^{p+1}(M)$.

We denote the set of all exact $p$-forms on a manifold $M$ as $B^p(M)$. Equivalently, $B^p(M)$ is the range of $d : \Lambda^{p-1}(M) \to \Lambda^p(M)$.

The $p^{\text{th}}$ *de Rham cohomology group* is defined to be

$$H^p(M) = Z^p(M)/B^p(M)$$

The quotient here is an equivalence class. Two closed forms $\omega, \omega' \in Z^p(M)$ are said to be equivalent if $\omega = \omega' + \eta$ for some $\eta \in B^p(M)$. We say that $\omega$ and $\omega'$ sit in the same equivalence class $[\omega]$. The cohomology group $H^p(M)$ is the set of equivalence classes; in other words, it consists of closed forms mod exact forms.

The *Betti numbers* $B_p$ of a manifold $M$ are defined as

$$B_p = \dim H^p(M)$$

It turns out that these are always finite. The Betti number $B_0 = 1$ for any connected manifold. This can be traced to the existence of constant functions which are clearly closed but, because there are no $p = -1$ forms, are not exact. The higher Betti numbers are non-zero only if the manifold has some interesting topology. Finally, the *Euler character* is defined as the alternating sum of Betti numbers,

$$\chi(M) = \sum_p (-1)^p B_p \tag{2.35}$$

Here are some simple examples. We've already seen that the circle $\mathbf{S}^1$ has a closed, non-exact one-form. This means that $B_1 = 1$ and $\chi = 0$. The sphere $\mathbf{S}^n$ has only $B_n = 1$ and $\chi = 1 + (-1)^n$. The torus $\mathbf{T}^n$ has $B_p = \binom{n}{p}$ and $\chi = 0$.

### 2.4.4 Integration

We have learned how to differentiate on manifolds by using a vector field $X$. Now it is time to learn how to integrate. It turns out that the things that we integrate on manifolds are forms.

### Integrating over Manifolds

To start, we need to orient ourselves. An *volume form*, or *orientation* on a manifold of dimension $\dim(M) = n$ is a nowhere-vanishing top form $v$. Any top form has just a single component and can be locally written as

$$v = v(x)\, dx^1 \wedge \ldots \wedge dx^n$$

where we require $v(x) \neq 0$. If such a top form exists everywhere on the manifold, then $M$ is said to be *orientable*.

The orientation is called right-handed if $v(x) > 0$ everywhere, and left-handed if $v(x) < 0$ everywhere. Given one volume form $v$, we can always construct another by multiplying by a function, giving $\tilde{v} = fv$ where $f(x) > 0$ everywhere or $f(x) < 0$ everywhere.

It's not enough to just write down a volume form with $v(x) \neq 0$ locally. We must also ensure that we can patch these volume forms together over the manifold, without the handedness changing. Suppose that we have two sets of coordinates, $x^\mu$ and $\tilde{x}^\mu$ that overlap on some region. In the new coordinates, the volume form is given by

$$v = v(x)\frac{\partial x^1}{\partial \tilde{x}^{\mu_1}} d\tilde{x}^{\mu_1} \wedge \ldots \wedge \frac{\partial x^n}{\partial \tilde{x}^{\mu_n}} d\tilde{x}^{\mu_n} = v(x) \det\left(\frac{\partial x^\mu}{\partial \tilde{x}^\nu}\right) d\tilde{x}^1 \wedge \ldots \wedge d\tilde{x}^n$$

which has the same orientation provided

$$\det\left(\frac{\partial x^\mu}{\partial \tilde{x}^\nu}\right) > 0 \tag{2.36}$$

Non-orientable manifolds cannot be covered by overlapping charts such that (2.36) holds. Examples include the Möbius strip and real projective space $\mathbf{RP}^n$ for $n$ even. (In contrast $\mathbf{RP}^n$ is orientable for $n$ odd, and $\mathbf{CP}^n$ is orientable for all $n$.) In these lectures, we deal only with orientable manifolds.

Given a volume form $v$ on $M$, we can integrate any function $f : M \to \mathbf{R}$ over the manifold. In a chart $\phi : \mathcal{O} \to U$, with coordinates $x^\mu$, we have

$$\int_{\mathcal{O}} fv = \int_U dx_1 \ldots dx_n \; f(x)v(x)$$

On the right-hand-side, we're just doing normal integration over some part of $\mathbf{R}^n$. The volume form is playing the role of a measure, telling us how much to weight various parts of the integral. To integrate over the entire manifold, we divide the manifold up into different regions, each covered by a single chart. We then perform the integral over each region and sum the results.

**Integrating over Submanifolds**

We don't have to integrate over the full manifold $M$. We can integrate over some lower dimensional submanifold.

A manifold $\Sigma$ with dimension $k < n$ is a *submanifold* of $M$ if we can find a map $\phi : \Sigma \rightarrow M$ which is one-to-one (which ensures that $\Sigma$ doesn't intersect itself in $M$) and $\phi_* : T_p(\Sigma) \rightarrow T_{\phi(p)}(M)$ is one-to-one.

We can then integrate a $k$-form $\omega$ on $M$ over a $k$-dimensional submanifold $\Sigma$. We do this by pulling back the $k$-form to $\Sigma$ and writing

$$\int_{\phi(\Sigma)} \omega = \int_{\Sigma} \phi^* \omega$$

For example, suppose that we have a one-form $A$ living over $M$. If $C$ is a one-dimensional manifold, the we can introduce a map $\sigma : C \rightarrow M$ which defines non-intersecting, one-dimensional curve $\sigma(C)$ which is a submanifold of $M$. We can then pull-back $A$ onto this curve and integrate to get

$$\int_{\sigma(C)} A = \int_C \sigma^* A$$

This probably looks more familiar in coordinates. If the curve traces out a path $x^\mu(\tau)$ in $M$, we have

$$\int_C \sigma^* A = \int d\tau \ A_\mu(x) \frac{dx^\mu}{d\tau}$$

But this is precisely the way the worldline of a particle couples to the electromagnetic field, as we previously saw in (1.20).

### 2.4.5 Stokes' Theorem

Until now, we have considered only smooth manifolds. There is a slight generalisation that will be useful. We define a *manifold with boundary* in the same way as a manifold, except the charts map $\phi : \mathcal{O} \rightarrow U$ where $U$ is an open subset of $\mathbf{R}^{n+} = \{\{x^1, \ldots, x^n\}$ such that $x^n \geq 0\}$. The boundary has co-dimension 1 and is denoted $\partial M$: it is the submanifold with coordinates $x^n = 0$.

Consider a manifold $M$ with boundary $\partial M$. If the dimension of the manifold is $\dim(M) = n$ then for any $(n-1)$-form $\omega$, we have the following wonderfully simple result

$$\int_M d\omega = \int_{\partial M} \omega \qquad (2.37)$$

This is *Stokes' theorem*. We do not prove it here, but instead give a few examples.

First, consider $n = 1$ with $M$ the interval $I$. We introduce coordinates $x \in [a, b]$ on the interval. The 0-form $\omega = \omega(x)$ is simply a function and $d\omega = (d\omega/dx)dx$. In this case, the two sides of Stokes' theorem can be evaluated to give

$$\int_M d\omega = \int_a^b \frac{d\omega}{dx} dx \quad \text{and} \quad \int_{\partial M} \omega = \omega(b) - \omega(a)$$

Equating the two, we see that Stokes' theorem is simply a restatement of the fundamental theorem of calculus.

Next, we take $M \subset \mathbf{R}^2$ to be a manifold with boundary. We introduce a one-form with coordinates

$$\omega = \omega_1 dx^1 + \omega_2 dx^2 \quad \Rightarrow \quad d\omega = \left( \frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2} \right) dx^1 \wedge dx^2$$

In this case, the ingredients in Stokes' theorem are

$$\int_M dw = \int_M \left( \frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2} \right) dx^1 dx^2 \quad \text{and} \quad \int_{\partial M} \omega = \int_{\partial M} \omega_1 dx^1 + \omega_2 dx^2$$

Equating the two gives the result usually referred to as Green's theorem in the plane.

Finally, consider $M \subset \mathbf{R}^3$ to be a manifold with boundary, with a 2-form

$$\omega = \omega_1 \, dx^2 \wedge dx^3 + \omega_2 \, dx^3 \wedge dx^1 + \omega_3 \, dx^1 \wedge dx^2$$

The right-hand-side of Stokes theorem is

$$\int_{\partial M} \omega_1 \, dx^2 dx^3 + \omega_2 \, dx^3 dx^1 + \omega_3 \, dx^1 dx^2$$

Meanwhile, we computed the exterior derivative of a 2-form in (2.29). The left-hand-side of Stokes' theorem then gives

$$\int_M d\omega = \int_M (\partial_1 \omega_1 + \partial_2 \omega_2 + \partial_3 \omega_3) \, dx^1 dx^2 dx^3$$

This time, equating the two gives us Gauss' divergence theorem.

We see that Stokes' theorem, as written in (2.37), is the mother of all integral theorems, packaging many famous results in a single formula. We'll revisit this in Section 3.2.4 where we relate Stokes' theorem to a more explicit form of the divergence theorem.

# 3. Introducing Riemannian Geometry

We have yet to meet the star of the show. There is one object that we can place on a manifold whose importance dwarfs all others, at least when it comes to understanding gravity. This is the metric.

The existence of a metric brings a whole host of new concepts to the table which, collectively, are called *Riemannian geometry*. In fact, strictly speaking we will need a slightly different kind of metric for our study of gravity, one which, like the Minkowski metric, has some strange minus signs. This is referred to as *Lorentzian Geometry* and a slightly better name for this section would be "Introducing Riemannian and Lorentzian Geometry". However, for the purposes of this section the differences are minor. The novelties of Lorentzian geometry will become more pronounced later in the course when we explore some of the physical consequences such as horizons.

## 3.1 The Metric

In Section 1, we informally introduced the metric as a way to measure distances between points. It does, indeed, provide this service but it is not its initial purpose. Instead, the metric is an inner product on each vector space $T_p(M)$.

**Definition**: A *metric* $g$ is a $(0,2)$ tensor field that is:

- Symmetric: $g(X,Y) = g(Y,X)$.
- Non-Degenerate: If, for any $p \in M$, $g(X,Y)\big|_p = 0$ for all $Y \in T_p(M)$ then $X_p = 0$.

With a choice of coordinates, we can write the metric as

$$g = g_{\mu\nu}(x)\, dx^\mu \otimes dx^\nu$$

The object $g$ is often written as a line element $ds^2$ and this expression is abbreviated as

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu dx^\nu$$

This is the form that we saw previously in (1.4). The metric components can extracted by evaluating the metric on a pair of basis elements,

$$g_{\mu\nu}(x) = g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right)$$

The metric $g_{\mu\nu}$ is a symmetric matrix. We can always pick a basis $e_\mu$ of each $T_p(M)$ so that this matrix is diagonal. The non-degeneracy condition above ensures that none of these diagonal elements vanish. Some are positive, some are negative. Sylvester's law of inertia is a theorem in algebra which states that the number of positive and negative entries is independent of the choice of basis. (This theorem has nothing to do with inertia. But Sylvester thought that if Newton could have a law of inertia, there should be no reason he couldn't.) The number of negative entries is called the *signature* of the metric.

### 3.1.1 Riemannian Manifolds

For most applications of differential geometry, we are interested in manifolds in which all diagonal entries of the metric are positive. A manifold equipped with such a metric is called a *Riemannian manifold*. The simplest example is *Euclidean space* $\mathbf{R}^n$ which, in Cartesian coordinates, is equipped with the metric

$$g = dx^1 \otimes dx^1 + \ldots + dx^n \otimes dx^n$$

The components of this metric are simply $g_{\mu\nu} = \delta_{\mu\nu}$.

A general Riemannian metric gives us a way to measure the length of a vector $X$ at each point,

$$|X| = \sqrt{g(X,X)}$$

It also allows us to measure the angle between any two vectors $X$ and $Y$ at each point, using

$$g(X,Y) = |X||Y|\cos\theta$$

The metric also gives us a way to measure the distance between two points $p$ and $q$ along a curve in $M$. The curve is parameterised by $\sigma : [a,b] \to M$, with $\sigma(a) = p$ and $\sigma(b) = q$. The distance is then

$$\text{distance} = \int_a^b dt\ \sqrt{g(X,X)\big|_{\sigma(t)}}$$

where $X$ is a vector field that is tangent to the curve. If the curve has coordinates $x^\mu(t)$, the tangent vector is $X^\mu = dx^\mu/dt$, and the distance is

$$\text{distance} = \int_a^b dt \sqrt{g_{\mu\nu}(x)\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}$$

Importantly, this distance does not depend on the choice of parameterisation of the curve; this is essentially the same calculation that we did in Section 1.2 when showing the reparameterisation invariance of the action for a particle.

### 3.1.2 Lorentzian Manifolds

For the purposes of general relativity, we will be working with a manifold in which one of the diagonal entries of the metric is negative. A manifold equipped with such a metric is called *Lorentzian*.

The simplest example of a Lorentzian metric is *Minkowski space*. This is $\mathbf{R}^n$ equipped with the metric

$$\eta = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + \ldots + dx^{n-1} \otimes dx^{n-1}$$

The components of the Minkowski metric are $\eta_{\mu\nu} = \text{diag}(-1, +1, \ldots, +1)$. As this example shows, on a Lorentzian manifold we usually take the coordinate index $x^\mu$ to run from $0, 1, \ldots, n-1$.

At any point $p$ on a general Lorentzian manifold, it is always possible to find an orthonormal basis $\{e_\mu\}$ of $T_p(M)$ such that, locally, the metric looks like the Minkowski metric

$$g_{\mu\nu}\Big|_p = \eta_{\mu\nu} \tag{3.1}$$

This fact is closely related to the equivalence principle; we'll describe the coordinates that allow us to do this in Section 3.3.2.

In fact, if we find one set of coordinates in which the metric looks like Minkowski space at $p$, it is simple to exhibit other coordinates. Consider a different basis of vector fields related by

$$\tilde{e}_\mu = \Lambda^\nu{}_\mu e_\nu$$

Then, in this basis the components of the metric are

$$\tilde{g}_{\mu\nu} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu g_{\rho\sigma}$$

This leaves the metric in Minkowski form at $p$ if

$$\eta_{\mu\nu} = \Lambda^\rho{}_\mu(p) \Lambda^\sigma{}_\nu(p) \eta_{\rho\sigma} \tag{3.2}$$

This is the defining equation for a Lorentz transformation that we saw previously in (1.14). We see that viewed locally – which here means at a point $p$ – recover some basic features of special relativity. Note, however, that if we choose coordinates so that the metric takes the form (3.1) at some point $p$, it will likely differ from the Minkowski metric as we move away from $p$.

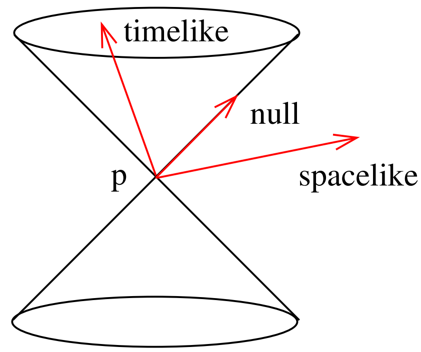



**Figure 21:** The lightcone at a point $p$, with three different types of tangent vectors.

The fact that, locally, the metric looks like the Minkowski metric means that we can import some ideas from special relativity. At any point $p$, a vector $X_p \in T_p(M)$ is said to be *timelike* if $g(X_p, X_p) < 0$, *null* if $g(X_p, X_p) = 0$, and *spacelike* if $g(X_p, X_p) > 0$.

At each point on $M$, we can then draw *lightcones*, which are the null tangent vectors at that point. There are both past-directed and future-directed lightcones at each point, as shown in Figure 21. The novelty is that the directions of these lightcones can vary smoothly as we move around the manifold. This specifies the causal structure of spacetime, which determines who can be friends with whom. We'll see more of this later in the lectures.

We can again use the metric to determine the length of curves. The nature of a curve at a point is inherited by the nature of its tangent vector. A curve is called *timelike* if its tangent vector is everywhere timelike. In this case, we can again use the metric to measure the distance along the curve between two points $p$ and $q$. Given a parametrisation $x^\mu(t)$, this distance is,

$$\tau = \int_a^b dt\, \sqrt{-g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}$$

This is called the *proper time*. It is, in fact, something we've met before: it is precisely the action (1.27) for a point particle moving in the spacetime with metric $g_{\mu\nu}$.

### 3.1.3 The Joys of a Metric

Whether we're on a Riemannian or Lorentzian manifold, there are a number of bounties that the metric brings.

### The Metric as an Isomophism

First, the metric gives us a natural isomorphism between vectors and covectors: $g : T_p(M) \to T^*_p(M)$ for each $p$, with the one-form constructed from the contraction of $g$ and a vector field $X$.

In a coordinate basis, we write $X = X^\mu \partial_\mu$. This is mapped to a one-form which, because this is a natural isomorphism, we also call $X$. This notation is less annoying than you might think; in components the one-form is written is $X = X_\mu dx^\mu$. The components are then related by

$$X_\mu = g_{\mu\nu} X^\nu$$

Physicists usually say that we use the metric to lower the index from $X^\mu$ to $X_\mu$. But in their heart, they mean "the metric provides a natural isomorphism between a vector space and its dual".

Because $g$ is non-degenerate, the matrix $g_{\mu\nu}$ is invertible. We denote the inverse $g^{\mu\nu}$, with $g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho$. These are the components of a symmetric $(2,0)$ tensor $\hat{g} = g^{\mu\nu} \partial_\mu \otimes \partial_\nu$. More importantly, the inverse metric allows us to raise the index on a one-form to give us back the original tangent vector,

$$X^\mu = g^{\mu\nu} X_\nu$$

In Euclidean space, with Cartesian coordinates, the metric is simply $g_{\mu\nu} = \delta_{\mu\nu}$ which is so simple it hides the distinction between vectors and one-forms. This is the reason we didn't notice the the difference between these spaces when we were five.

### The Volume Form

The metric also gives us a natural volume form on the manifold $M$. On a Riemannian manifold, this is defined as

$$v = \sqrt{\det g_{\mu\nu}}\, dx^1 \wedge \ldots \wedge dx^n$$

The determinant is usually simply written as $\sqrt{g} = \sqrt{\det g_{\mu\nu}}$. On a Lorentzian manifold, the determinant is negative and we instead have

$$v = \sqrt{-g}\, dx^0 \wedge \ldots \wedge dx^{n-1} \tag{3.3}$$

As defined, the volume form looks coordinate dependent. Importantly, it is not. To see this, introduce some rival coordinates $\tilde{x}^\mu$, with

$$dx^\mu = A^\mu{}_\nu d\tilde{x}^\nu \quad \text{where} \quad A^\mu{}_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}$$

In the new coordinates, the wedgey part of the volume form becomes

$$dx^1 \wedge \ldots \wedge dx^n = A^1{}_{\mu_1} \ldots A^n{}_{\mu_n} d\tilde{x}^{\mu_1} \wedge \ldots \wedge d\tilde{x}^{\mu_n}$$

We can rearrange the one-forms into the order $d\tilde{x}^1 \wedge \ldots \wedge d\tilde{x}^n$. We pay a price of $+$ or $-1$ depending on whether $\{\mu_1, \ldots, \mu_n\}$ is an even or odd permutation of $\{1, \ldots, n\}$. Since we're summing over all indices, this is the same as summing over all permutations $\pi$ of $\{1, \ldots, n\}$, and we have

$$\begin{aligned} dx^1 \wedge \ldots \wedge dx^n &= \sum_{\text{perms } \pi} \text{sign}(\pi)\, A^1_{\pi(1)} \ldots A^n_{\pi(n)} d\tilde{x}^1 \wedge \ldots \wedge d\tilde{x}^n \\ &= \det(A)\, d\tilde{x}^1 \wedge \ldots \wedge d\tilde{x}^n \end{aligned}$$

where $\det(A) > 0$ if the change of coordinates preserves the orientation. This factor of $\det(A)$ is the usual Jacobian factor that one finds when changing the measure in an integral.

Meanwhile, the metric components transform as

$$g_{\mu\nu} = \frac{\partial \tilde{x}^\rho}{\partial x^\mu} \frac{\partial \tilde{x}^\sigma}{\partial x^\nu} \tilde{g}_{\rho\sigma} = (A^{-1})^\rho{}_\mu (A^{-1})^\sigma{}_\nu \tilde{g}_{\rho\sigma}$$

and so the determinant becomes

$$\det g_{\mu\nu} = (\det\ A^{-1})^2 \det \tilde{g}_{\mu\nu} = \frac{\det \tilde{g}_{\mu\nu}}{(\det A)^2}$$

We see that the factors of $\det A$ cancel, and we can equally write the volume form as

$$v = \sqrt{|\tilde{g}|}\, d\tilde{x}^1 \wedge \ldots \wedge d\tilde{x}^n$$

The volume form (3.3) may look more familiar if we write it as

$$v = \frac{1}{n!} v_{\mu_1 \ldots \mu_n} dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_n}$$

Here the components $v_{\mu_1 \ldots \mu_n}$ are given in terms of the totally anti-symmetric object $\epsilon_{\mu_1 \ldots \mu_n}$ with $\epsilon_{1 \ldots n} = +1$ and other components determined by the sign of the permutation,

$$v_{\mu_1 \ldots \mu_n} = \sqrt{|g|}\, \epsilon_{\mu_1 \ldots \mu_n} \tag{3.4}$$

Note that $v_{\mu_1 \ldots \mu_n}$ is a tensor, which means that $\epsilon_{\mu_1 \ldots \mu_n}$ can't quite be a tensor: instead, it is a tensor divided by $\sqrt{|g|}$. It is sometimes said to be a *tensor density*. The anti-symmetric tensor density arises in many places in physics. In all cases, it should be viewed as a volume form on the manifold. (In nearly all cases, this volume form arises from a metric as here.)

As with other tensors, we can use the metric to raise the indices and construct the a volume form with all indices up

$$v^{\mu_1\ldots\mu_n} = g^{\mu_1\nu_1}\ldots g^{\mu_n\nu_n} v_{\nu_1\ldots\nu_n} = \pm\frac{1}{\sqrt{|g|}}\epsilon^{\mu_1\ldots\mu_n}$$

where we get a $+$ sign for a right-handed basis and a $-$ sign for a left-handed basis. Here $\epsilon^{\mu_1\ldots\mu_n}$ is again a totally anti-symmetric tensor density with $\epsilon^{1\ldots n} = +1$. Note, however, that while we raise the indices on $v_{\mu_1\ldots\mu_n}$ using the metric, this statement doesn't quite hold for $\epsilon_{\mu_1\ldots\mu_n}$ which takes values 1 or 0 regardless of whether the indices are all down or all up. This reflects the fact that it is a tensor density, rather than a genuine tensor.

The existence of a natural volume form means that, given a metric, we can integrate any function $f$ over the manifold. We will sometimes write this as

$$\int_M f v = \int_M d^n x\ \sqrt{\pm g} f$$

The metric $\sqrt{\pm g}$ provides a measure on the manifold that tells us what regions of the manifold are weighted more strongly than the others in the integral.

**The Hodge Dual**

On an oriented manifold $M$, we can use the totally anti-symmetric tensor $\epsilon_{\mu_1,\ldots,\mu_n}$ to define a map which takes a $p$-form $\omega \in \Lambda^p(M)$ to an $(n-p)$-form, denoted $(\star\,\omega) \in \Lambda^{n-p}(M)$, defined by

$$(\star\,\omega)_{\mu_1\ldots\mu_{n-p}} = \frac{1}{p!}\sqrt{|g|}\,\epsilon_{\mu_1\ldots\mu_{n-p}\nu_1\ldots\nu_p}\omega^{\nu_1\ldots\nu_p} \tag{3.5}$$

This map is called the *Hodge dual.* It is independent of the choice of coordinates.

It's not hard to check that,

$$\star\,(\star\,\omega) = \pm(-1)^{p(n-p)}\omega \tag{3.6}$$

where the $+$ sign holds for Riemannian manifolds and the $-$ sign for Lorentzian manifolds. (To prove this, it's useful to first show that $v^{\mu_1\ldots\mu_p\rho_1\ldots\rho_{n-p}}v_{\nu_1\ldots\nu_p\rho_1\ldots\rho_{n-p}} = \pm p!(n-p)!\delta^{\mu_1}_{[\nu_1}\ldots\delta^{\mu_p}_{\nu_p]}$, again with the $\pm$ sign for Riemannian/Lorentzian manifolds.)

It's worth returning to some high school physics and viewing it through the lens of our new tools. We are very used to taking two vectors in $\mathbf{R}^3$, say $\mathbf{a}$ and $\mathbf{b}$, and taking the cross-product to find a third vector

$$\mathbf{a} \times \mathbf{b} = \mathbf{c}$$

In fact, we really have objects that live in three different spaces here, related by the Euclidean metric $\delta_{\mu\nu}$. First we use this metric to relate the vectors to one-forms. The cross-product is then really a wedge product which gives us back a 2-form. We then use the metric twice more, once to turn to the two-form back into a one-form using the Hodge dual, and again to turn the one-form into a vector. Of course, none of these subtleties bothered us when we were 15. But when we start thinking about curved manifolds, with a non-trivial metric, these distinctions become important.

The Hodge dual allows us to define an inner product on each $\Lambda^p(M)$. If $\omega, \eta \in \Lambda^p(M)$, we define

$$\langle \eta, \omega \rangle = \int_M \eta \wedge \star\omega$$

which makes sense because $\star\omega \in \Lambda^{n-p}(M)$ and so $\eta \wedge \star\omega$ is a top form that can be integrated over the manifold.

With such an inner product in place, we can also start to play the kind of games that are familiar from quantum mechanics and look at operators on $\Lambda^p(M)$ and their adjoints. The one operator that we have introduced on the space of forms is the exterior derivative, defined in Section 2.4.1. Its adjoint is defined by the following result:

**Claim**: For $\omega \in \Lambda^p(M)$ and $\alpha \in \Lambda^{p-1}(M)$,

$$\langle d\alpha, \omega \rangle = \langle \alpha, d^\dagger \omega \rangle \tag{3.7}$$

where the adjoint operator $d^\dagger : \Lambda^p(M) \to \Lambda^{p-1}(M)$ is given by

$$d^\dagger = \pm(-1)^{np+n-1} \star d \star$$

with, again, the $\pm$ sign for Riemannian/Lorentzian manifolds respectively.

**Proof**: This is simply the statement of integration-by-parts for forms. On a closed manifold $M$, Stokes' theorem tells us that

$$0 = \int_M d(\alpha \wedge \star\omega) = \int_M d\alpha \wedge \star\omega + (-1)^{p-1} \alpha \wedge d \star \omega$$

The first term is simply $\langle d\alpha, \omega\rangle$. The second term also takes the form of an inner product which, up to a sign, is proportional to $\langle \alpha, \star d \star \omega\rangle$. To determine the sign, note that $d\star\omega \in \Lambda^{n-p+1}(M)$ so, using (3.6), we have $\star\star\, d\star\omega = \pm(-1)^{(n-p+1)(p-1)} d\star\omega$. Putting this together gives

$$\langle d\alpha, \omega\rangle = \pm(-1)^{np+n-1}\langle \alpha, \star d \star \omega\rangle$$

as promised. $\square$

### 3.1.4 A Sniff of Hodge Theory

We can combine $d$ and $d^\dagger$ to construct the *Laplacian*, $\triangle : \Lambda^p(M) \to \Lambda^p(M)$, defined as

$$\triangle = (d + d^\dagger)^2 = dd^\dagger + d^\dagger d$$

where the second equality follows because $d^2 = d^{\dagger\,2} = 0$. The Laplacian can be defined on both Riemannian manifolds, where it is positive definite, and Lorentzian manifolds. Here we restrict our discussion to Riemannian manifolds.

Acting on functions $f$, we have $d^\dagger f = 0$ (because $\star f$ is a top form so $d\star f = 0$). That leaves us with,

$$\begin{aligned}
\triangle(f) &= -\star\, d \star (\partial_\mu f\, dx^\mu) \\
&= -\frac{1}{(n-1)!}\star d\left((\partial_\mu f) g^{\mu\nu}\sqrt{|g|}\,\epsilon_{\nu\rho_1\ldots\rho_{n-1}} dx^{\rho_1}\wedge\ldots\wedge dx^{\rho_{n-1}}\right) \\
&= -\frac{1}{(n-1)!}\star \partial_\sigma\left(\sqrt{|g|} g^{\mu\nu}\partial_\mu f\right)\epsilon_{\nu\rho_1\ldots\rho_{n-1}} dx^\sigma \wedge dx^{\rho_1}\wedge\ldots\wedge dx^{\rho_{n-1}} \\
&= -\star\, \partial_\nu\left(\sqrt{|g|} g^{\mu\nu}\partial_\mu f\right) dx^1\wedge\ldots\wedge dx^n \\
&= -\frac{1}{\sqrt{|g|}}\partial_\nu\left(\sqrt{|g|} g^{\mu\nu}\partial_\mu f\right)
\end{aligned}$$

This form of the Laplacian, acting on functions, appears fairly often in applications of differential geometry.

There is a particularly nice story involving $p$-forms $\gamma$ that obey

$$\triangle\gamma = 0$$

Such forms are said to be *harmonic*. An harmonic form is necessarily closed, meaning $d\gamma = 0$, and *co-closed*, meaning $d^\dagger\gamma = 0$. This follows by writing

$$\langle \gamma, \triangle\gamma\rangle = \langle d\gamma, d\gamma\rangle + \langle d^\dagger\gamma, d^\dagger\gamma\rangle = 0$$

and noting that the inner product is positive-definite.

There are some rather pretty facts that relate the existence of harmonic forms to de Rham cohomology. The space of harmonic $p$-forms on a manifold $M$ is denoted $\text{Harm}^p(M)$. First, the *Hodge decomposition theorem*, which we state without proof: any $p$-form $\omega$ on a compact, Riemannian manifold can be uniquely decomposed as

$$\omega = d\alpha + d^\dagger\beta + \gamma$$

where $\alpha \in \Lambda^{p-1}(M)$ and $\beta \in \Lambda^{p+1}(M)$ and $\gamma \in \text{Harm}^p(M)$. This result can then be used to prove:

**Hodge's Theorem**: There is an isomorphism

$$\text{Harm}^p(M) \cong H^p(M)$$

where $H^p(M)$ is the de Rham cohomology group introduced in Section 2.4.3. In particular, the Betti numbers can be computed by counting the number of linearly independent harmonic forms,

$$B_p = \dim\, \text{Harm}^p(M)$$

**Proof**: First, let's show that any harmonic form $\gamma$ provides a representative of $H^p(M)$. As we saw above, any harmonic $p$-form is closed, $d\gamma = 0$, so $\gamma \in Z^p(M)$. But the unique nature of the Hodge decomposition tells us that $\gamma \neq d\beta$ for some $\beta$.

Next, we need to show that any equivalence class $[\omega] \in H^p(M)$ can be represented by a harmonic form. We decompose $\omega = d\alpha + d^\dagger\beta + \gamma$. By definition $[\omega] \in H^p(M)$ means that $d\omega = 0$ so we have

$$0 = \langle d\omega, \beta\rangle = \langle \omega, d^\dagger\beta\rangle = \langle d\alpha + d^\dagger\beta + \gamma, d^\dagger\beta\rangle = \langle d^\dagger\beta, d^\dagger\beta\rangle$$

where, in the final step, we "integrated by parts" and used the fact that $dd\alpha = d\gamma = 0$. Because the inner product is positive definite, we must have $d^\dagger\beta = 0$ and, hence, $\omega = \gamma + d\alpha$. Any other representative $\tilde{\omega} \sim \omega$ of $[\omega] \in H^p(M)$ differs by $\tilde{\omega} = \omega + d\eta$ and so, by the Hodge decomposition, is associated to the same harmonic form $\gamma$. $\square$

### 3.2 Connections and Curvature

We've already met one version of differentiation in these lectures. A vector field $X$ is, at heart, a differential operator and provides a way to differentiate a function $f$. We write this simply as $X(f)$.

As we saw previously, differentiating higher tensor fields is a little more tricky because it requires us to subtract tensor fields at different points. Yet tensors evaluated at different points live in different vector spaces, and it only makes sense to subtract these objects if we can first find a way to map one vector space into the other. In Section 2.2.4, we used the flow generated by $X$ as a way to perform this mapping, resulting in the idea of the Lie derivative $\mathcal{L}_X$.

There is, however, a different way to take derivatives, one which ultimately will prove more useful. The derivative is again associated to a vector field $X$. However, this time we introduce a different object, known as a *connection* to map the vector spaces at one point to the vector spaces at another. The result is an object, distinct from the Lie derivative, called the *covariant derivative*.

### 3.2.1 The Covariant Derivative

A *connection* is a map $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$. We usually write this as $\nabla(X, Y) = \nabla_X Y$ and the object $\nabla_X$ is called the *covariant derivative*. It satisfies the following properties for all vector fields $X$, $Y$ and $Z$,

- $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$
- $\nabla_{(fX+gY)} Z = f\nabla_X Z + g\nabla_Y Z$ for all functions $f, g$.
- $\nabla_X(fY) = f\nabla_X Y + (\nabla_X f)Y$ where we define $\nabla_X f = X(f)$

The covariant derivative endows the manifold with more structure. To elucidate this, we can evaluate the connection in a basis $\{e_\mu\}$ of $\mathfrak{X}(M)$. We can always express this as

$$\nabla_{e_\rho} e_\nu = \Gamma^\mu_{\rho\nu} e_\mu \tag{3.8}$$

with $\Gamma^\mu_{\rho\nu}$ the components of the connection. It is no coincidence that these are denoted by the same greek letter that we used for the Christoffel symbols in Section 1. However, for now, you should not conflate the two; we'll see the relationship between them in Section 3.2.3.

The name "connection" suggests that $\nabla$, or its components $\Gamma^\mu_{\nu\rho}$, connect things. Indeed they do. We will show in Section 3.3 that the connection provides a map from the tangent space $T_p(M)$ to the tangent space at any other point $T_q(M)$. This is what allows the connection to act as a derivative.

We will use the notation

$$\nabla_\mu = \nabla_{e_\mu}$$

This makes the covariant derivative $\nabla_\mu$ look similar to a partial derivative. Using the properties of the connection, we can write a general covariant derivative of a vector field as

$$\begin{aligned}\nabla_X Y &= \nabla_X(Y^\mu e_\mu)\\ &= X(Y^\mu)e_\mu + Y^\mu \nabla_X e_\mu\\ &= X^\nu e_\nu(Y^\mu)e_\mu + X^\nu Y^\mu \nabla_\nu e_\mu\\ &= X^\nu \left(e_\nu(Y^\mu) + \Gamma^\mu_{\nu\rho}Y^\rho\right) e_\mu\end{aligned}$$

The fact that we can strip of the overall factor of $X^\nu$ means that it makes sense to write the components of the covariant derivative as

$$\nabla_\nu Y = (e_\nu(Y^\mu) + \Gamma^\mu_{\nu\rho}Y^\rho)e_\mu$$

Or, in components,

$$(\nabla_\nu Y)^\mu = e_\nu(Y^\mu) + \Gamma^\mu_{\nu\rho}Y^\rho \tag{3.9}$$

Note that the covariant derivative coincides with the Lie derivative on functions, $\nabla_X f = \mathcal{L}_X f = X(f)$. It also coincides with the old-fashioned partial derivative: $\nabla_\mu f = \partial_\mu f$. However, its action on vector fields differs. In particular, the Lie derivative $\mathcal{L}_X Y = [X, Y]$ depends on both $X$ and the first derivative of $X$ while, as we have seen above, the covariant derivative depends only on $X$. This is the property that allows us to write $\nabla_X = X^\nu \nabla_\nu$ and think of $\nabla_\mu$ as an operator in its own right. In contrast, there is no way to write "$\mathcal{L}_X = X^\mu \mathcal{L}_\mu$". While the Lie derivative has its uses, the ability to define $\nabla_\mu$ means that this is best viewed as the natural generalisation of the partial derivative to curved space.

**Differentiation as Punctuation**

In a coordinate basis, in which $e_\mu = \partial_\mu$, the covariant derivative (3.9) becomes

$$(\nabla_\nu Y)^\mu = \partial_\nu Y^\mu + \Gamma^\mu_{\nu\rho}Y^\rho \tag{3.10}$$

We will differentiate often. To save ink, we use the sloppy, and sometimes confusing, notation

$$(\nabla_\nu Y)^\mu = \nabla_\nu Y^\mu$$

This means, in particular, that $\nabla_\nu Y^\mu$ is the $\mu^{\rm th}$ component of $\nabla_\nu Y$, rather than the differentiation of the function $Y^\mu$. Furthermore, we will sometimes denote covariant differentiation using a semi-colon

$$\nabla_\nu Y^\mu = Y^\mu{}_{;\nu}$$

Meanwhile, the partial derivative is denoted using a mere comma, $\partial_\mu Y^\nu = Y^\nu{}_{,\mu}$. The expression (3.10) then reads

$$Y^\mu{}_{;\nu} = Y^\mu{}_{,\nu} + \Gamma^\mu_{\nu\rho} Y^\rho$$

Note that the $Y^\mu_{;\nu}$ are components of a bona fide tensor. In contrast, the $Y^\mu_{,\nu}$ are not components of a tensor. And, as we now show, neither are the $\Gamma^\mu_{\rho\nu}$.

**The Connection is a Not a Tensor**

The connection is not a tensor. We can see this immediately from the definition $\nabla(X, fY) = \nabla_X(fY) = f\nabla_X Y + (X(f))Y$. This is not linear in the second argument, which is one of the requirements of a tensor.

To illustrate this, we can ask what the connection looks like in a different basis,

$$\tilde{e}_\nu = A^\mu{}_\nu e_\mu \tag{3.11}$$

for some invertible matrix $A$. If $e_\mu$ and $\tilde{e}_\mu$ are both coordinate bases, then

$$A^\mu{}_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}$$

We know from (2.23) that the components of a $(1,2)$ tensor transform as

$$\tilde{T}^\mu{}_{\nu\rho} = (A^{-1})^\mu{}_\tau A^\lambda{}_\nu A^\sigma{}_\rho T^\tau{}_{\lambda\sigma} \tag{3.12}$$

We can now compare this to the transformation of the connection components $\Gamma^\mu_{\rho\nu}$. In the basis $\tilde{e}_\mu$, we have

$$\nabla_{\tilde{e}_\rho} \tilde{e}_\nu = \tilde{\Gamma}^\mu_{\rho\nu} \tilde{e}_\mu$$

Substituting in the transformation (3.11), we have

$$\tilde{\Gamma}^\mu_{\rho\nu}\tilde{e}_\mu = \nabla_{(A^\sigma{}_\rho e_\sigma)}(A^\lambda{}_\nu e_\lambda) = A^\sigma{}_\rho \nabla_{e_\sigma}(A^\lambda{}_\nu e_\lambda) = A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} e_\tau + A^\sigma{}_\rho e_\lambda \partial_\sigma A^\lambda{}_\nu$$

We can write this as

$$\begin{aligned}
\tilde{\Gamma}^\mu_{\rho\nu}\tilde{e}_\mu &= \left(A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} + A^\sigma{}_\rho \partial_\sigma A^\tau{}_\nu\right) e_\tau \\
&= \left(A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} + A^\sigma{}_\rho \partial_\sigma A^\tau{}_\nu\right) (A^{-1})^\mu{}_\tau \tilde{e}_\mu
\end{aligned}$$

Stripping off the basis vectors $\tilde{e}_\mu$, we see that the components of the connection transform as

$$\tilde{\Gamma}^\mu_{\rho\nu} = (A^{-1})^\mu{}_\tau A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} + (A^{-1})^\mu{}_\tau A^\sigma{}_\rho \partial_\sigma A^\tau{}_\nu \tag{3.13}$$

The first term coincides with the transformation of a tensor (3.12). But the second term, which is independent of $\Gamma$, but instead depends on $\partial A$, is novel. This is the characteristic transformation property of a connection.

**Differentiating Other Tensors**

We can use the Leibnizarity of the covariant derivative to extend its action to any tensor field. It's best to illustrate this with an example.

Consider a one-form $\omega$. If we differentiate $\omega$, we will get another one-form $\nabla_X\omega$ which, like any one-form, is defined by its action on vector fields $Y \in \mathfrak{X}(M)$. To construct this, we will insist that the connection obeys the Leibnizarity in the modified sense that

$$\nabla_X(\omega(Y)) = (\nabla_X\omega)(Y) + \omega(\nabla_X Y)$$

But $\omega(Y)$ is simply a function, which means that we can also write this as

$$\nabla_X(\omega(Y)) = X(\omega(Y))$$

Putting these together gives

$$(\nabla_X\omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y)$$

In coordinates, we have

$$\begin{aligned} X^\mu(\nabla_\mu\omega)_\nu Y^\nu &= X^\mu\partial_\mu(\omega_\nu Y^\nu) - \omega_\nu X^\mu(\partial_\mu Y^\nu + \Gamma^\nu_{\mu\rho}Y^\rho) \\ &= X^\mu(\partial_\mu\omega_\rho - \Gamma^\nu_{\mu\rho}\omega_\nu)Y^\rho \end{aligned}$$

where, crucially, the $\partial Y$ terms cancel in going from the first to the second line. This means that the overall result is linear in $Y$ and we may define $\nabla_X\omega$ without reference to the vector field $Y$ on which is acts. In components, we have

$$(\nabla_\mu\omega)_\rho = \partial_\mu\omega_\rho - \Gamma^\nu_{\mu\rho}\omega_\nu$$

As for vector fields, we also write this as

$$(\nabla_\mu\omega)_\rho \equiv \nabla_\mu\omega_\rho \equiv \omega_{\rho;\mu} = \omega_{\rho,\mu} - \Gamma^\nu_{\mu\rho}\omega_\nu$$

This kind of argument can be extended to a general tensor field of rank $(p,q)$, where the covariant derivative is defined by,

$$T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q;\rho} = T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q,\rho} + \Gamma^{\mu_1}_{\rho\sigma}T^{\sigma\mu_2\ldots\mu_p}{}_{\nu_1\ldots\nu_q} + \ldots + \Gamma^{\mu_p}_{\rho\sigma}T^{\mu_1\ldots\mu_{p-1}\sigma}{}_{\nu_1\ldots\nu_q}$$
$$-\Gamma^{\sigma}_{\rho\nu_1}T^{\mu_1\ldots\mu_p}{}_{\sigma\nu_2\ldots\nu_q} - \ldots - \Gamma^{\sigma}_{\rho\nu_q}T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_{q-1}\sigma}$$

The pattern is clear: for every upper index $\mu$ we get a $+\Gamma T$ term, while for every lower index we get a $-\Gamma T$ term.

Now that we can differentiate tensors, we will also need to extend our punctuation notation slightly. If more than two subscripts follow a semi-colon (or, indeed, a comma) then we differentiate respect to both, doing the one on the left first. So, for example, $X^{\mu}{}_{;\nu\rho} = \nabla_\rho\nabla_\nu X^\mu$.

### 3.2.2 Torsion and Curvature

Even though the connection is not a tensor, we can use it to construct two tensors. The first is a rank $(1,2)$ tensor $T$ known as *torsion*. It is defined to act on $X, Y \in \mathfrak{X}(M)$ and $\omega \in \Lambda^1(M)$ by

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

The other is a rank $(1,3)$ tensor $R$, known as *curvature*. It acts on $X, Y, Z \in \mathfrak{X}(M)$ and $\omega \in \Lambda^1(M)$ by

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

The curvature tensor is also called the *Riemann tensor*.

Alternatively, we could think of torsion as a map $T : \mathfrak{X}(M)\times\mathfrak{X}(M) \to \mathfrak{X}(M)$, defined by

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

Similarly, the curvature $R$ can be viewed as a map from $\mathfrak{X}(M)\times\mathfrak{X}(M)$ to a differential operator acting on $\mathfrak{X}(M)$,

$$R(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]} \tag{3.14}$$

### Checking Linearity

To demonstrate that $T$ and $R$ are indeed tensors, we need to show that they are linear in all arguments. Linearity in $\omega$ is straightforward. For the others, there are some small calculations to do. For example, we must show that $T(\omega; fX, Y) = fT(\omega; X, Y)$. To see this, we just run through the definitions of the various objects,

$$T(\omega; fX, Y) = \omega(\nabla_{fX}Y - \nabla_Y(fX) - [fX, Y])$$

We then use $\nabla_{fX}Y = f\nabla_X Y$ and $\nabla_Y(fX) = f\nabla_Y X + Y(f)\, X$ and $[fX, Y] = f[X, Y] - Y(f)X$. The two $Y(f)X$ terms cancel, leaving us with

$$\begin{aligned} T(\omega; fX, Y) &= f\omega(\nabla_X Y - \nabla_Y X - [X, Y]) \\ &= fT(\omega; X, Y) \end{aligned}$$

Similarly, for the curvature tensor we have

$$\begin{aligned} R(\omega; fX, Y, Z) &= \omega(\nabla_{fX}\nabla_Y Z - \nabla_Y\nabla_{fX}Z - \nabla_{[fX,Y]}Z \\ &= \omega(f\nabla_X\nabla_Y Z - \nabla_Y(f\nabla_X Z) - \nabla_{(f[X,Y]-Y(f)X)}Z) \\ &= \omega(f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - Y(f)\nabla_X Z - \nabla_{f[X,Y]}Z + \nabla_{Y(f)X}Z) \\ &= \omega(f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - Y(f)\nabla_X Z - f\nabla_{[X,Y]}Z + Y(f)\nabla_X Z) \\ &= f\omega(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z) \\ &= fR(\omega; X, Y, Z) \end{aligned}$$

Linearity in $Y$ follows from linearity in $X$. But we still need to check linearity in $Z$,

$$\begin{aligned} R(\omega; X, Y, fZ) &= \omega(\nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ)) \\ &= \omega(\nabla_X(f\nabla_Y Z + Y(f)Z) - \nabla_Y(f\nabla_X Z + X(f)Z) \\ &\qquad -f\nabla_{[X,Y]}Z - [X, Y](f)Z) \\ &= \omega(f\nabla_X\nabla_Y + X(f)\nabla_Y Z + Y(f)\nabla_X Z + X(Y(f))Z \\ &\qquad -f\nabla_Y\nabla_X Z - Y(f)\nabla_X Z - X(f)\nabla_Y Z - Y(X(f))Z \\ &\qquad -f\nabla_{[X,Y]}Z - [X, Y](f)Z) \\ &= fR(\omega; X, Y, Z) \end{aligned}$$

Thus, both torsion and curvature define new tensors on our manifold.

### Components

We can evaluate these tensors in a coordinate basis $\{e_\mu\} = \{\partial_\mu\}$, with the dual basis $\{f^\mu\} = \{dx^\mu\}$. The components of the torsion are

$$\begin{aligned} T^\rho{}_{\mu\nu} &= T(f^\rho; e_\mu, e_\nu) \\ &= f^\rho(\nabla_\mu e_\nu - \nabla_\nu e_\mu - [e_\mu, e_\nu]) \\ &= f^\rho(\Gamma^\sigma_{\mu\nu} e_\sigma - \Gamma^\sigma_{\nu\mu} e_\sigma) \\ &= \Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu} \end{aligned}$$

where we've used the fact that, in a coordinate basis, $[e_\mu, e_\nu] = [\partial_\mu, \partial_\nu] = 0$. We learn that, even though $\Gamma^\rho_{\mu\nu}$ is not a tensor, the anti-symmetric part $\Gamma^\rho_{[\mu\nu]}$ does form a tensor. Clearly the torsion tensor is anti-symmetric in the lower two indices

$$T^\rho{}_{\mu\nu} = -T^\rho{}_{\nu\mu}$$

Connections which are symmetric in the lower indices, so $\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu}$ have $T^\rho{}_{\mu\nu} = 0$. Such connections are said to be *torsion-free*.

The components of the curvature tensor are given by

$$R^\sigma{}_{\rho\mu\nu} = R(f^\sigma; e_\mu, e_\nu, e_\rho)$$

Note the slightly counterintuitive, but standard ordering of the indices; the indices $\mu$ and $\nu$ that are associated to covariant derivatives $\nabla_\mu$ and $\nabla_\nu$ go at the end. We have

$$\begin{aligned} R^\sigma{}_{\rho\mu\nu} &= f^\sigma(\nabla_\mu \nabla_\nu e_\rho - \nabla_\nu \nabla_\mu e_\rho - \nabla_{[e_\mu, e_\nu]} e_\rho) \\ &= f^\sigma(\nabla_\mu \nabla_\nu e_\rho - \nabla_\nu \nabla_\mu e_\rho) \\ &= f^\sigma(\nabla_\mu(\Gamma^\lambda_{\nu\rho} e_\lambda) - \nabla_\nu(\Gamma^\lambda_{\mu\rho} e_\lambda)) \\ &= f^\sigma((\partial_\mu \Gamma^\lambda_{\nu\rho}) e_\lambda + \Gamma^\lambda_{\nu\rho} \Gamma^\tau_{\mu\lambda} e_\tau - (\partial_\nu \Gamma^\lambda_{\mu\rho}) e_\lambda - \Gamma^\lambda_{\mu\rho} \Gamma^\tau_{\nu\lambda} e_\tau) \\ &= \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\lambda_{\nu\rho} \Gamma^\sigma_{\mu\lambda} - \Gamma^\lambda_{\mu\rho} \Gamma^\sigma_{\nu\lambda} \end{aligned} \tag{3.15}$$

Clearly the Riemann tensor is anti-symmetric in its last two indices

$$R^\sigma{}_{\rho\mu\nu} = -R^\sigma{}_{\rho\nu\mu}$$

Equivalently, $R^\sigma{}_{\rho\mu\nu} = R^\sigma{}_{\rho[\mu\nu]}$. There are a number of further identities of the Riemann tensor of this kind. We postpone this discussion to Section 3.4.

#### The Ricci Identity

There is a closely related calculation in which both the torsion and Riemann tensors appears. We look at the commutator of covariant derivatives acting on vector fields. Written in an orgy of anti-symmetrised notation, this calculation gives

$$\begin{aligned}\nabla_{[\mu}\nabla_{\nu]}Z^\sigma &= \partial_{[\mu}(\nabla_{\nu]}Z^\sigma) + \Gamma^\sigma_{[\mu|\lambda|}\nabla_{\nu]}Z^\lambda - \Gamma^\rho_{[\mu\nu]}\nabla_\rho Z^\sigma \\ &= \partial_{[\mu}\partial_{\nu]}Z^\sigma + (\partial_{[\mu}\Gamma^\sigma_{\nu]\rho})Z^\rho + (\partial_{[\mu}Z^\rho)\Gamma^\sigma_{\nu]\rho} + \Gamma^\sigma_{[\mu|\lambda|}\partial_{\nu]}Z^\lambda \\ &\quad + \Gamma^\sigma_{[\mu|\lambda|}\Gamma^\lambda_{\nu]\rho}Z^\rho - \Gamma^\rho_{[\mu\nu]}\nabla_\rho Z^\sigma\end{aligned}$$

The first term vanishes, while the third and fourth terms cancel against each other. We're left with

$$2\nabla_{[\mu}\nabla_{\nu]}Z^\sigma = R^\sigma{}_{\rho\mu\nu}Z^\rho - T^\rho{}_{\mu\nu}\nabla_\rho Z^\sigma \tag{3.16}$$

where the torsion tensor is $T^\rho{}_{\mu\nu} = \Gamma^\rho_{[\mu\nu]}$ and the Riemann tensor appears as

$$R^\sigma{}_{\rho\mu\nu} = 2\partial_{[\mu}\Gamma^\sigma_{\nu]\rho} + 2\Gamma^\sigma_{[\mu|\lambda|}\Gamma^\lambda_{\nu]\rho}$$

which coincides with (3.15). The expression (3.16) is known as the *Ricci identity*.

### 3.2.3 The Levi-Civita Connection

So far, our discussion of the connection $\nabla$ has been entirely independent of the metric. However, something nice happens if we have both a connection and a metric. This something nice is called the *fundamental theorem of Riemannian geometry*. (Happily, it's also true for Lorentzian geometries.)

**Theorem**: There exists a unique, torsion free, connection that is compatible with a metric $g$, in the sense that

$$\nabla_X g = 0$$

for all vector fields $X$.

**Proof**: We start by showing uniqueness. Suppose that such a connection exists. Then, by Leibniz

$$X(g(Y,Z)) = \nabla_X(g(Y,Z)) = (\nabla_X g)(Y,Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

Since $\nabla_X g = 0$, this becomes

$$X(g(Y,Z)) = g(\nabla_X Y, Z) + g(\nabla_X Z, Y)$$

By cyclic permutation of $X$, $Y$ and $Z$, we also have

$$\begin{aligned}
Y(g(Z,X)) &= g(\nabla_Y Z, X) + g(\nabla_Y X, Z)\\
Z(g(X,Y)) &= g(\nabla_Z X, Y) + g(\nabla_Z Y, X)
\end{aligned}$$

Since the torsion vanishes, we have

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

We can use this to write the cyclically permuted equations as

$$\begin{aligned}
X(g(Y,Z)) &= g(\nabla_Y X, Z) + g(\nabla_X Z, Y) + g([X,Y],Z)\\
Y(g(Z,X)) &= g(\nabla_Z Y, X) + g(\nabla_Y X, Z) + g([Y,Z],X)\\
Z(g(X,Y)) &= g(\nabla_X Z, Y) + g(\nabla_Z Y, X) + g([Z,X],Y)
\end{aligned}$$

Add the first two of these equations, and subtract the third. We find

$$\begin{aligned}
g(\nabla_Y X, Z) = \frac{1}{2}\Big[&X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y))\\
&- g([X,Y],Z) - g([Y,Z],X) + g([Z,X],Y)\Big]
\end{aligned} \tag{3.17}$$

But with a non-degenerate metric, this specifies the connection uniquely. We'll give an expression in terms of components in (3.18) below.

It remains to show that the object $\nabla$ defined this way does indeed satisfy the properties expected of a connection. The tricky one turns out to be the requirement that $\nabla_{fX} Y = f\nabla_X Y$. We can see that this is indeed the case as follows:

$$\begin{aligned}
g(\nabla_{fY} X, Z) &= \frac{1}{2}\Big[X(g(fY,Z)) + fY(g(Z,X)) - Z(g(X,fY))\\
&\qquad - g([X,fY],Z) - g([fY,Z],X) + g([Z,X],fY)\Big]\\
&= \frac{1}{2}\Big[fX(g(Y,Z)) + X(f)g(Y,Z) + fY(g(Z,X)) - fZ(g(X,Y))\\
&\qquad - Z(f)g(X,Y) - fg([X,Y],Z) - X(f)g(Y,Z) - fg([Y,Z],X)\\
&\qquad + Z(f)g(Y,X) + fg([Z,X],Y)\Big]\\
&= g(f\nabla_Y X, Z)
\end{aligned}$$

The other properties of the connection follow similarly. □

The connection (3.17), compatible with the metric, is called the *Levi-Civita* connection. We can compute its components in a coordinate basis $\{e_\mu\} = \{\partial_\mu\}$. This is particularly simple because $[\partial_\mu, \partial_\nu] = 0$, leaving us with

$$g(\nabla_\nu e_\mu, e_\rho) = \Gamma^\lambda_{\nu\mu} g_{\lambda\rho} = \frac{1}{2}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

Multiplying by the inverse metric gives

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) \tag{3.18}$$

The components of the Levi-Civita connection are called the *Christoffel symbols*. They are the objects (1.31) we met already in Section 1 when discussing geodesics in spacetime. For the rest of these lectures, when discussing a connection we will always mean the Levi-Civita connection.

**An Example: Flat Space**

In flat space $\mathbf{R}^d$, endowed with either Euclidean or Minkowski metric, we can always pick Cartesian coordinates, in which case the Christoffel symbols vanish. However, in other coordinates this need not be the case. For example, in Section 1.1.1, we computed the flat space Christoffel symbols in polar coordinates (1.10). They don't vanish. But because the Riemann tensor is a genuine tensor, if it vanishes in one coordinate system then it must vanishes in all of them. Given some horrible coordinate system, with $\Gamma^\rho_{\mu\nu} \neq 0$, we can always compute the corresponding Riemann tensor to see if the space is actually flat after all.

**Another Example: The Sphere $\mathbf{S}^2$**

Consider $\mathbf{S}^2$ with radius $r$ and the round metric

$$ds^2 = r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

We can extract the Christoffel symbols from those of flat space in polar coordinates (1.10). The non-zero components are

$$\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta \quad , \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \frac{\cos\theta}{\sin\theta} \tag{3.19}$$

From these, it is straightforward to compute the components of the Riemann tensor. They are most simply expressed as $R_{\sigma\rho\mu\nu} = g_{\sigma\lambda} R^\lambda{}_{\rho\mu\nu}$ and are given by

$$R_{\theta\phi\theta\phi} = R_{\phi\theta\phi\theta} = -R_{\theta\phi\phi\theta} = -R_{\phi\theta\theta\phi} = r^2\sin^2\theta \tag{3.20}$$

with the other components vanishing.

### 3.2.4 The Divergence Theorem

Gauss' Theorem, also known as the divergence theorem, states that if you integrate a total derivative, you get a boundary term. There is a particular version of this theorem in curved space that we will need for later applications.

As a warm-up, we have the following result:

**Lemma:** The contraction of the Christoffel symbols can be written as

$$\Gamma^\mu_{\mu\nu} = \frac{1}{\sqrt{g}}\partial_\nu \sqrt{g} \tag{3.21}$$

On Lorentzian manifolds, we should replace $\sqrt{g}$ with $\sqrt{|g|}$.

**Proof:** From (3.18), we have

$$\Gamma^\mu_{\mu\nu} = \frac{1}{2}g^{\mu\rho}\partial_\nu g_{\mu\rho} = \frac{1}{2}\mathrm{tr}(g^{-1}\partial_\nu g) = \frac{1}{2}\mathrm{tr}(\partial_\nu \log g)$$

However, there's a useful identity for the log of any diagonalisable matrix: they obey

$$\mathrm{tr}\log A = \log\det A$$

This is clearly true for a diagonal matrix, since the determinant is the product of eigenvalues while the trace is the sum. But both trace and determinant are invariant under conjugation, so this is also true for diagonalisable matrices. Applying it to our metric formula above, we have

$$\Gamma^\mu_{\mu\nu} = \frac{1}{2}\mathrm{tr}(\partial_\nu \log g) = \frac{1}{2}\partial_\nu \log\det g = \frac{1}{2}\frac{1}{\det g}\partial_\nu \det g = \frac{1}{\sqrt{\det g}}\partial_\nu\sqrt{\det g}$$

which is the claimed result. $\square$

With this in hand, we can now prove the following:

**Divergence Theorem:** Consider a region of a manifold $M$ with boundary $\partial M$. Let $n^\mu$ be an outward-pointing, unit vector orthogonal to $\partial M$. Then, for any vector field $X^\mu$ on $M$, we have

$$\int_M d^n x\ \sqrt{g}\,\nabla_\mu X^\mu = \int_{\partial M} d^{n-1}x\ \sqrt{\gamma}\,n_\mu X^\mu$$

where $\gamma_{ij}$ is the pull-back of the metric to $\partial M$, and $\gamma = \det\gamma_{ij}$. On a Lorentzian manifold, a version of this formula holds only if $\partial M$ is purely timelike or purely spacelike, which ensures that $\gamma \neq 0$ at any point.

**Proof:** Using the lemma above, the integrand is

$$\sqrt{g}\,\nabla_\mu X^\mu = \sqrt{g}\left(\partial_\mu X^\mu + \Gamma^\mu_{\mu\nu}X^\nu\right) = \sqrt{g}\left(\partial_\mu X^\mu + X^\nu \frac{1}{\sqrt{g}}\partial_\nu\sqrt{g}\right) = \partial_\mu\left(\sqrt{g}X^\mu\right)$$

The integral is then

$$\int_M d^nx\ \sqrt{g}\,\nabla_\mu X^\mu = \int_M d^nx\ \partial_\mu\left(\sqrt{g}X^\mu\right)$$

which now is an integral of an ordinary partial derivative, so we can apply the usual divergence theorem that we are familiar with. It remains only to evaluate what's happening at the boundary $\partial M$. For this, it is useful to pick coordinates so that the boundary $\partial M$ is a surface of constant $x^n$. Furthermore, we will restrict to metrics of the form

$$g_{\mu\nu} = \begin{pmatrix} \gamma_{ij} & 0 \\ 0 & N^2 \end{pmatrix}$$

Then by our usual rules of integration, we have

$$\int_M d^nx\ \partial_\mu\left(\sqrt{g}X^\mu\right) = \int_{\partial M} d^{n-1}x\ \sqrt{\gamma N^2}X^n$$

The unit normal vector $n^\mu$ is given by $n^\mu = (0, 0, \ldots, 1/N)$, which satisfies $g_{\mu\nu}n^\mu n^\nu = 1$ as it should. We then have $n_\mu = g_{\mu\nu}n^\nu = (0, 0, \ldots, N)$, so we can write

$$\int_M d^nx\ \sqrt{g}\,\nabla_\mu X^\mu = \int_{\partial M} d^{n-1}x\ \sqrt{\gamma}\,n_\mu X^\mu$$

which is the result we need. As the final expression is a covariant quantity, it is true in general. □

In Section 2.4.5, we advertised Stokes' theorem as the mother of all integral theorems. It's perhaps not surprising to hear that the divergence theorem is a special case of Stokes' theorem. To see this, here's an alternative proof that uses the languages of forms.

**Another Proof:** Given the volume form $v$ on $M$, and a vector field $X$, we can contract the two to define an $n-1$ form $\omega = \iota_X v$. (This is the interior product that we previously met in (2.30).) It has components

$$\omega_{\mu_1\ldots\mu_{n-1}} = \sqrt{g}\,\epsilon_{\mu_1\ldots\mu_n}X^{\mu_n}$$

If we now take the exterior derivative, $d\omega$, we have a top-form. Since the top form is unique up to multiplication, $d\omega$ must be proportional to the volume form. Indeed, it's not hard to show that

$$(d\omega)_{\mu_1\ldots\mu_n} = \sqrt{g}\,\epsilon_{\mu_1\ldots\mu_n}\nabla_\nu X^\nu$$

This means that, in form language, the integral over $M$ that we wish to consider can be written as

$$\int_M d^nx\ \sqrt{g}\,\nabla_\mu X^\mu = \int_M d\omega$$

Now we invoke Stokes' theorem, to write

$$\int_M d\omega = \int_{\partial M}\omega$$

We now need to massage $\omega$ into the form needed. First, we introduce a volume form $\hat{v}$ on $\partial M$, with components

$$\hat{v}_{\mu_1\ldots\mu_{n-1}} = \sqrt{\gamma}\epsilon_{\mu_1\ldots\mu_{n-1}}$$

This is related to the volume form on $M$ by

$$\frac{1}{n}v_{\mu_1\ldots\mu_{n-1}\nu} = \hat{v}_{[\mu_1\ldots\mu_{n-1}}n_{\nu]}$$

where $n^\mu$ is the orthonormal vector that we introduced previously. We then have

$$\omega_{\mu_1\ldots\mu_{n-1}} = \sqrt{\gamma}\,(n_\nu X^\nu)\tilde{\epsilon}_{\mu_1\ldots\mu_{n-1}}$$

The divergence theorem then follows from Stokes' theorem. □

### 3.2.5 The Maxwell Action

Let's briefly turn to some physics. We take the manifold $M$ to be spacetime. In classical field theory, the dynamical degrees of freedom are objects that take values at each point in $M$. We call these objects *fields*. The simplest such object is just a function which, in physics, we call a scalar field.

As we described in Section 2.4.2, the theory of electromagnetism is described by a one-form field $A$. In fact, there is a little more structure because we ask that the theory is invariant under gauge transformations

$$A \to A + d\alpha$$

To achieve this, we construct a field strength $F = dA$ which is indeed invariant under gauge transformations. The next question to ask is: what are the dynamics of these fields?

The most elegant and powerful way to describe the dynamics of classical fields is provided by the action principle. The action is a functional of the fields, constructed by integrating over the manifold. The differential geometric language that we've developed in these lectures tells us that there are, in fact, very few actions one can write down.

To see this, suppose that our manifold has only the 2-form $F$ but is not equipped with a metric. If spacetime has dimension $\dim(M) = 4$ (it does!) then we need to construct a 4-form to integrate over $M$. There is only one of these at our disposal, suggesting the action

$$S_{\rm top} = -\frac{1}{2}\int F\wedge F$$

If we expand this out in the electric and magnetic fields using (2.32), we find

$$S_{\rm top} = \int dx^0 dx^1 dx^2 dx^3 \; \mathbf{E}\cdot\mathbf{B}$$

Actions of this kind, which are independent of the metric, are called *topological*. They are typically unimportant in classical physics. Indeed, we can locally write $F\wedge F = d(A\wedge F)$, so the action is a total derivative and does not affect the classical equations of motion. Nonetheless, topological actions often play a subtle and interesting roles in quantum physics. For example, the action $S_{\rm top}$ underlies the theory of topological insulators. You can read more about this in Section 1 of the lectures on Gauge Theory.

To construct an action that gives rise to interesting classical dynamics, we need to introduce a metric. The existence of a metric allows us to introduce a second two-form, $\star F$, and construct the action

$$S_{\rm Maxwell} = -\frac{1}{2}\int F\wedge\star F = -\frac{1}{4}\int d^4x\sqrt{-g}g^{\mu\nu}g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma} = -\frac{1}{4}\int d^4x\sqrt{-g}\; F^{\mu\nu}F_{\mu\nu}$$

This is the Maxwell action, now generalised to a curved spacetime. If we restrict to flat Minkowski space, the components are $F^{\mu\nu}F_{\mu\nu} = 2(\mathbf{B}^2 - \mathbf{E}^2)$. As we saw in our lectures on Electromagnetism, varying this action gives the remaining two Maxwell equations. In the elegant language of differential geometry, these take the simple form

$$d\star F = 0$$

We can also couple the gauge field to an electric current. This is described by a one-form $J$, and we write the action

$$S = \int -\frac{1}{2}F\wedge\star F + A\wedge\star J$$

We require that this action is invariant under gauge transformations $A \to A + d\alpha$. The action transforms as

$$S \to S + \int d\alpha \wedge \star J$$

After an integration by parts, the second term vanishes provided that

$$d \star J = 0$$

which is the requirement of current conservation expressed in the language of forms. The Maxwell equations now have a source term, and read

$$d \star F = \star J \tag{3.22}$$

We see that the rigid structure of differential geometry leads us by the hand to the theories that govern our world. We'll see this again in Section 4 when we discuss gravity.

**Electric and Magnetic Charges**

To define electric and magnetic charges, we integrate over submanifolds. For example, consider a three-dimensional spatial submanifold $\Sigma$. The electric charge in $\Sigma$ defined to be

$$Q_e = \int_\Sigma \star J$$

It's simple to check that this agrees with our usual definition $Q_e = \int d^3x \ J^0$ in flat Minkowski space. Using the equation of motion (3.22), we can translate this into an integral of the field strength

$$Q_e = \int_\Sigma d \star F = \int_{\partial\Sigma} \star F \tag{3.23}$$

where we have used Stokes' theorem to write this as an integral over the boundary $\partial\Sigma$. The result is the general form of Gauss' law, relating the electric charge in a region to the electric field piercing the boundary of the region. Similarly, we can define the magnetic charge

$$Q_m = \int_{\partial\Sigma} F$$

When we first meet Maxwell theory, we learn that magnetic charges do not exist, courtesy of the identity $dF = 0$. However, this can be evaded in topologically more interesting spaces.

The statement of current conservation $d \star J = 0$ means that the electric charge $Q_e$ in a region cannot change unless current flows in or out of that region. This fact, familiar from Electromagnetism, also has a nice expression in terms of forms. Consider a cylindrical region of spacetime $V$, ending on two spatial hypersurfaces $\Sigma_1$ and $\Sigma_2$ as shown in the figure. The boundary of $V$ is then

$$\partial V = \Sigma_1 \cup \Sigma_2 \cup B$$

where $B$ is the cylindrical timelike hypersurface.

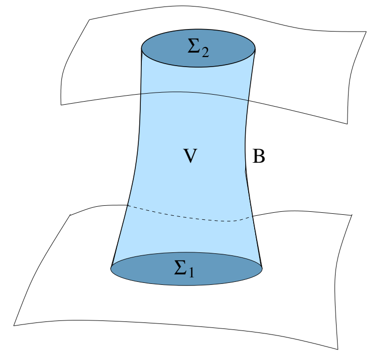



**Figure 22:**

We require that $J = 0$ on $B$, which is the statement that no current flows in or out of the region. Then we have

$$Q_e(\Sigma_1) - Q_e(\Sigma_2) = \int_{\Sigma_1} \star J - \int_{\Sigma_2} \star J = \int_{\partial V} \star J = \int_V d \star J = 0$$

which tells us that the electric charge remains constant in time.

**Maxwell Equations Using Connections**

The form of the Maxwell equations given above makes no reference to a connection. It does, however, use the metric, buried in the definition of the Hodge $\star$.

There is an equivalent formulation of the Maxwell equation using the covariant derivative. This will also serve to highlight the relationship between the covariant and exterior derivatives. First note that, given a one-form $A \in \Lambda^1(M)$, we can define the field strength as

$$F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu$$

where the Christoffel symbols have cancelled out by virtue of the anti-symmetry. This is what allowed us to define the exterior derivative without the need for a connection.

Next, consider the current one-form $J$. We can recast the statement of current conservation as follows:

**Claim:**

$$d \star J = 0 \quad \Leftrightarrow \quad \nabla_\mu J^\mu = 0$$

**Proof:** We have

$$\nabla_\mu J^\mu = \partial_\mu J^\mu + \Gamma^\mu_{\mu\rho} J^\rho = \frac{1}{\sqrt{-g}}\partial_\mu \left(\sqrt{-g}J^\mu\right)$$

where, in the second equality, we have used our previous result (3.21): $\Gamma^\mu_{\mu\nu} = \partial_\nu \log\sqrt{|g|}$. But this final form is proportional to $d\star J$, with the Hodge dual defined in (3.5). □

As an aside, in Riemannian signature the formula

$$\nabla_\mu J^\mu = \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}J^\mu)$$

provides a quick way of computing the divergence in different coordinate systems (if you don't have the inside cover of Jackson to hand). For example, in spherical polar coordinates on $\mathbf{R}^3$, we have $g = r^4\sin^2\theta$. Plug this into the expression above to immediately find

$$\nabla\cdot\mathbf{J} = \frac{1}{r^2}\partial_r(r^2 J^r) + \frac{1}{\sin\theta}\,\partial_\theta(\sin\theta\, J^\theta) + \partial_\phi J^\phi$$

The Maxwell equation (3.22) can also be written in terms of the covariant derivative

**Claim:**

$$d\star F = \star J \quad\Leftrightarrow\quad \nabla_\mu F^{\mu\nu} = J^\nu \tag{3.24}$$

**Proof:** We have

$$\begin{aligned}\nabla_\mu F^{\mu\nu} &= \partial_\mu F^{\mu\nu} + \Gamma^\mu_{\mu\rho}F^{\rho\nu} + \Gamma^\nu_{\mu\rho}F^{\mu\rho} \\ &= \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}F^{\mu\nu}\right) + \Gamma^\nu_{\mu\rho}F^{\mu\rho} = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}F^{\mu\nu}\right)\end{aligned}$$

where, in the second equality, we've again used (3.21) and in the final equality we've used the fact that $\Gamma^\nu_{\mu\rho}$ is symmetric while $F^{\nu\rho}$ is anti-symmetric. To complete the proof, you need to chase down the definitions of the Hodge dual (3.5) and the exterior derivative (2.26). (If you're struggling to match factors of $\sqrt{-g}$, then remember that the volume form $v = \sqrt{-g}\epsilon$ is a tensor, while the epsilon symbol $\epsilon_{\mu_1\ldots\mu_4}$ is a tensor density.) □

## 3.3 Parallel Transport

Although we have now met a number of properties of the connection, we have not yet explained its name. What does it connect?

The answer is that the connection connects tangent spaces, or more generally any tensor vector space, at different points of the manifold. This map is called *parallel transport*. As we stressed earlier, such a map is necessary to define differentiation.

Take a vector field $X$ and consider some associated integral curve $C$, with coordinates $x^\mu(\tau)$, such that

$$X^\mu\Big|_C = \frac{dx^\mu(\tau)}{d\tau} \tag{3.25}$$

We say that a tensor field $T$ is *parallely transported* along $C$ if

$$\nabla_X T = 0 \tag{3.26}$$

Suppose that the curve $C$ connects two points, $p \in M$ and $q \in M$. The requirement (3.26) provides a map from the vector space defined at $p$ to the vector space defined at $q$.

To illustrate this, consider the parallel transport of a second vector field $Y$. In components, the condition (3.26) reads

$$X^\nu \left(\partial_\nu Y^\mu + \Gamma^\mu_{\nu\rho} Y^\rho\right) = 0$$

If we now evaluate this on the curve $C$, we can think of $Y^\mu = Y^\mu(x(\tau))$, which obeys

$$\frac{dY^\mu}{d\tau} + X^\nu \Gamma^\mu_{\nu\rho} Y^\rho = 0 \tag{3.27}$$

These are a set of coupled, ordinary differential equations. Given an initial condition at, say $\tau = 0$, corresponding to point $p$, these equations can be solved to find a unique vector at each point along the curve.

Parallel transport is path dependent. It depends on both the connection, and the underlying path which, in this case, is characterised by the vector field $X$.

This is the second time we've used a vector field $X$ to construct maps between tensors at different points in the manifold. In Section 2.2.2, we used $X$ to generate a flow $\sigma_t : M \to M$, which we could then use to pull-back or push-forward tensors from one point to another. This was the basis of the Lie derivative. This is not the same as the present map. Here, we're using $X$ only to define the curve, while the connection does the work of telling relating tensor spaces along the curve.

### 3.3.1 Geodesics Revisited

A *geodesic* is a curve tangent to a vector field $X$ that obeys

$$\nabla_X X = 0 \tag{3.28}$$

Along the curve $C$, we can substitute the expression (3.25) into (3.27) to find

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu{}_{\rho\nu}\frac{dx^\rho}{d\tau}\frac{dx^\nu}{d\tau} = 0 \tag{3.29}$$

This is precisely the geodesic equation (1.30) that we derived in Section 1 by considering the action for a particle moving in spacetime. In fact, we find that the condition (3.28) results in geodesics with affine parameterisation.

For the Levi-Civita connection, we have $\nabla_X g = 0$. This ensures that for any vector field $Y$ parallely transported along a geodesic $X$, so $\nabla_X Y = \nabla_X X = 0$, we have

$$\frac{d}{d\tau} g(X, Y) = 0$$

This tells us that the vector field $Y$ makes the same angle with the tangent vector along each point of the geodesic.

### 3.3.2 Normal Coordinates

Geodesics lend themselves to the construction of a particularly useful coordinate system. On a Riemannian manifold, in the neighbourhood of a point $p \in M$, we can always find coordinates such that

$$g_{\mu\nu}(p) = \delta_{\mu\nu} \quad \text{and} \quad g_{\mu\nu,\rho}(p) = 0 \tag{3.30}$$

The same holds for Lorentzian manifolds, now with $g_{\mu\nu}(p) = \eta_{\mu\nu}$. These are referred to as *normal coordinates*. Because the first derivative of the metric vanishes, normal coordinates have the property that, at the point $p$, the Christoffel symbols vanish: $\Gamma^\mu_{\nu\rho}(p) = 0$. Generally, away from $p$ we will have $\Gamma^\mu_{\nu\rho} \neq 0$. Note, however, that it is not generally possible to ensure that the second derivatives of the metric also vanish. This, in turn, means that it's not possible to pick coordinates such that the Riemann tensor vanishes at a given point.

There are a number of ways to demonstrate the existence of coordinates (3.30). The brute force way is to start with some metric $\tilde{g}_{\mu\nu}$ in coordinates $\tilde{x}^\mu$ and try to find a change of coordinates to $x^\mu(\tilde{x})$ which does the trick. In the new coordinates,

$$\frac{\partial \tilde{x}^\rho}{\partial x^\mu}\frac{\partial \tilde{x}^\sigma}{\partial x^\nu}\tilde{g}_{\rho\sigma} = g_{\mu\nu} \tag{3.31}$$

We'll take the point $p$ to be the origin in both sets of coordinates. Then we can Taylor expand

$$\tilde{x}^\rho = \left.\frac{\partial \tilde{x}^\rho}{\partial x^\mu}\right|_{x=0} x^\mu + \frac{1}{2} \left.\frac{\partial^2 \tilde{x}^\rho}{\partial x^\mu \partial x^\nu}\right|_{x=0} x^\mu x^\nu + \ldots$$

We insert this into (3.31), together with a Taylor expansion of $\tilde{g}_{\rho\sigma}$, and try to solve the resulting partial differential equations to find the coefficients $\partial\tilde{x}/\partial x$ and $\partial^2\tilde{x}/\partial x^2$ that do the job. For example, the first requirement is

$$\left.\frac{\partial \tilde{x}^\rho}{\partial x^\mu}\right|_{x=0} \left.\frac{\partial \tilde{x}^\sigma}{\partial x^\nu}\right|_{x=0} \tilde{g}_{\rho\sigma}(p) = \delta_{\mu\nu}$$

Given any $\tilde{g}_{\rho\sigma}(p)$, it's always possible to find $\partial\tilde{x}/\partial x$ so that this is satisfied. In fact, a little counting shows that there are many such choices. If $\dim M = n$, then there are $n^2$ independent coefficients in the matrix $\partial\tilde{x}/\partial x$. The equation above puts $\frac{1}{2}n(n+1)$ conditions on these. That still leaves $\frac{1}{2}n(n-1)$ parameters unaccounted for. But this is to be expected: this is precisely the dimension of the rotational group $SO(n)$ (or the Lorentz group $SO(1, n-1)$ that leaves the flat metric unchanged.

We can do a similar counting at the next order. There are $\frac{1}{2}n^2(n+1)$ independent elements in the coefficients $\partial^2\tilde{x}^\rho/\partial x^\mu\partial x^\nu$. This is exactly the same number of conditions in the requirement $g_{\mu\nu,\rho}(p) = 0$.

We can also see why we shouldn't expect to set the second derivative of the metric to zero. Requiring $g_{\mu\nu,\rho\sigma} = 0$ is $\frac{1}{4}n^2(n+1)^2$ constraints. Meanwhile, the next term in the Taylor expansion is $\partial^3\tilde{x}^\rho/\partial x^\mu\partial x^\nu\partial x^\lambda$ which has $\frac{1}{6}n^2(n+1)(n+2)$ independent coefficients. We see that the numbers no longer match. This time we fall short, leaving

$$\frac{1}{4}n^2(n+1)^2 - \frac{1}{6}n^2(n+1)(n+2) = \frac{1}{12}n^2(n^2-1)$$

unaccounted for. This, therefore, is the number of ways to characterise the second derivative of the metric in a manner that cannot be undone by coordinate transformations. Indeed, it is not hard to show that this is precisely the number of independent coefficients in the Riemann tensor. (For $n = 4$, there are 20 coefficients of the Riemann tensor.)

#### The Exponential Map

There is a rather pretty, direct way to construct the coordinates (3.30). This uses geodesics. The rough idea is that, given a tangent vector $X_p \in T_p(M)$, there is a unique affinely parameterised geodesic through $p$ with tangent vector $X_p$ at $p$. We then

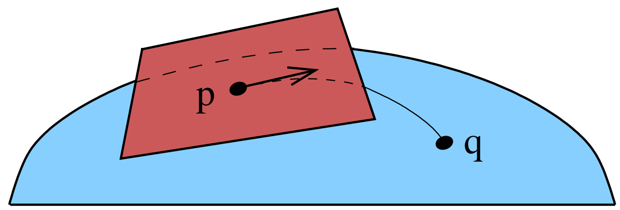



**Figure 23:** Start with a tangent vector, and follow the resulting geodesic to get the exponential map.

label any point $q$ in the neighbourhood of $p$ by the coordinates of the geodesic that take us to $q$ in some fixed amount of time. It's like throwing a ball in all possible directions, and labelling points by the initial velocity needed for the ball to reach that point in, say, 1 second.

Let's put some flesh on this. We introduce any coordinate system (not necessarily normal coordinates) $\tilde{x}^\mu$ in the neighbourhood of $p$. Then the geodesic we want solves the equation (3.29) subject to the requirements

$$\left.\frac{d\tilde{x}^\mu}{d\tau}\right|_{\tau=0} = \tilde{X}^\mu_p \quad \text{with} \quad \tilde{x}^\mu(\tau=0) = 0$$

There is a unique solution.

This observation means that we can define a map,

$$\text{Exp} : T_p(M) \to M$$

Given $X_p \in T_p(M)$, construct the appropriate geodesic and the follow it for some affine distance which we take to be $\tau = 1$. This gives a point $q \in M$. This is known as the *exponential map* and is illustrated in the Figure 23.

There is no reason that the exponential map covers all of the manifold $M$. It could well be that there are points which cannot be reached from $p$ by geodesics. Moreover, it may be that there are tangent vectors $X_p$ for which the exponential map is ill-defined. In general relativity, this occurs if the spacetime has singularities. Neither of these issues are relevant for our current purpose.

Now pick a basis $\{e_\mu\}$ of $T_p(M)$. The exponential map means that tangent vector $X_p = X^\mu e_\mu$ defines a point $q$ in the neighbourhood of $p$. We simply assign this point coordinates

$$x^\mu(q) = X^\mu$$

These are the normal coordinates.

If we pick the initial basis $\{e_\mu\}$ to be orthonormal, then the geodesics will point in orthogonal directions which ensures that the metric takes the form $g_{\mu\nu}(p) = \delta_{ab}$.

To see that the first derivative of the metric also vanishes, we first fix a point $q$ associated to a given tangent vector $X \in T_p(M)$. The tells us that the point $q$ sits a distance $\tau = 1$ along the geodesic. We can now ask: what tangent vector will take us a different distance along this same geodesic? Because the geodesic equation (3.29) is homogeneous in $\tau$, if we halve the length of $X$ then we will travel only half the distance along the geodesic, i.e. to $\tau = 1/2$. In general, the tangent vector $\tau X$ will take us a distance $\tau$ along the geodesic

$$\text{Exp} : \tau X_p \to x^\mu(\tau) = \tau X^\mu$$

This means that the geodesics in these coordinates take the particularly simply form

$$x^\mu(\tau) = \tau X^\mu$$

Since these are geodesics, they must solve the geodesic equation (3.29). But, for trajectories that vary linearly in time, this is just

$$\Gamma^\mu_{\rho\nu}(x(\tau))\, X^\rho X^\nu = 0$$

This holds at any point along the geodesic. At most points $x(\tau)$, this equation only holds for those choices of $X^\rho$ which take us along the geodesic in the first place. However, at $x(\tau) = 0$, corresponding to the point $p$ of interest, this equation must hold for any tangent vector $X^\mu$. This means that $\Gamma^\mu_{(\rho\nu)}(p) = 0$ which, for a torsion free connection, ensures that $\Gamma^\mu_{\rho\nu}(p) = 0$.

Vanishing Christoffel symbols means that the derivative of the metric vanishes. This follows for the Levi-Civita connection by writing $2g_{\mu\sigma}\Gamma^\sigma_{\rho\nu} = g_{\mu\rho,\nu} + g_{\mu\nu,\rho} - g_{\rho\nu,\mu}$. Symmetrising over $(\mu\rho)$ means that the last two terms cancel, leaving us with $g_{\mu\rho,\nu} = 0$ when evaluated at $p$.

**The Equivalence Principle**

Normal coordinates play an important conceptual role in general relativity. Any observer at point $p$ who parameterises her immediate surroundings using coordinate constructed by geodesics will experience a locally flat metric, in the sense of (3.30).

This is the mathematics underlying the Einstein equivalence principle. This principle states that any freely falling observer, performing local experiments, will not experience a gravitational field. Here "freely falling" means the observer follows geodesics, as we saw in Section 1 and will naturally use normal coordinates. In this context, the coordinates are called a *local inertial frame*. The lack of gravitational field is the statement that $g_{\mu\nu}(p) = \eta_{\mu\nu}$.

Key to the understanding the meaning and limitations of the equivalence principle is the the word "local". There is a way to distinguish whether there is a gravitational field at $p$: we compute the Riemann tensor. This depends on the second derivative of the metric and, in general, will be non-vanishing. However, to measure the effects of the Riemann tensor, one typically has to compare the result of an experiment at $p$ with an experiment at a nearby point $q$: this is considered a "non-local" observation as far as the equivalence principle goes. In the next two subsections, we give examples of physics that depends on the Riemann tensor.

### 3.3.3 Path Dependence: Curvature and Torsion

Take a tangent vector $Z_p \in T_p(M)$, and parallel transport it along a curve $C$ to some point $r \in M$. Now parallel transport it along a different curve $C'$ to the same point $r$. How do the resulting vectors differ?

To answer this, we construct each of our curves $C$ and $C'$ from two segments, generated by linearly independent vector fields, $X$ and $Y$ satisfying $[X, Y] = 0$ as shown in Figure 24. To make life easy, we'll the point $r$ to be close to the original point $p$.

We pick normal coordinates $x^\mu = (\tau, \lambda, 0, \ldots)$ so that the starting point is at $x^\mu(p) = 0$ while the tangent vectors are aligned along the coordinates, $X = \partial/\partial\tau$ and $Y = \partial/\partial\sigma$. Then the other stopped points along the curve are $x^\mu(q) = (\delta\tau, 0, 0, \ldots)$, $x^\mu(r) = (\delta\tau, \delta\sigma, 0, \ldots)$ and $x^\mu(s) = (0, \delta\sigma, 0, \ldots)$ where $\delta\tau$ and $\delta\sigma$ are taken to be small. This set-up is shown in Figure 24.

First we parallel transport $Z_p$ along $X$ to $Z_q$. Along the curve, $Z^\mu$ solves (3.27)

$$\frac{dZ^\mu}{d\tau} + X^\nu \Gamma^\mu_{\rho\nu} Z^\rho = 0 \tag{3.32}$$

We Taylor expand the solution as

$$Z^\mu_q = Z^\mu_p + \left.\frac{dZ^\mu}{d\tau}\right|_{\tau=0} \delta\tau + \frac{1}{2} \left.\frac{d^2 Z^\mu}{d\tau^2}\right|_{\tau=0} \delta\tau^2 + \mathcal{O}(\delta\tau^3)$$

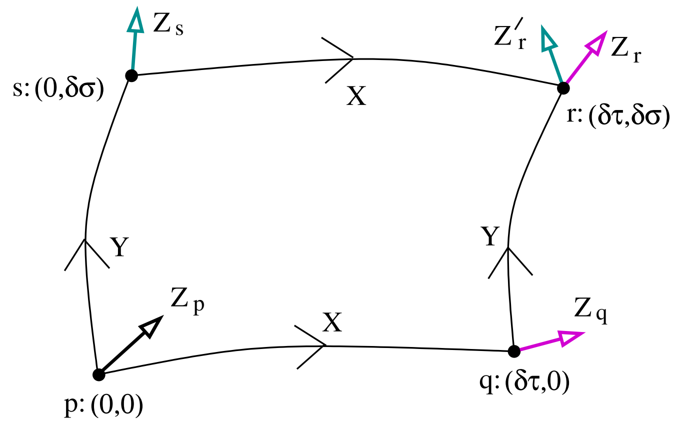



**Figure 24:** Parallel transporting a vector $Z_p$ along two different paths does not give the same answer.

From (3.32), we have $dZ^\mu/d\tau\big|_0 = 0$ because, in normal coordinates, $\Gamma^\mu{}_{\rho\nu}(p) = 0$. We can calculate the second derivative by differentiating (3.32) to find

$$\begin{aligned} \left.\frac{d^2 Z^\mu}{d\tau^2}\right|_{\tau=0} &= -\left.\left(X^\nu Z^\rho \frac{d\Gamma^\mu_{\rho\nu}}{d\tau} + \frac{dX^\nu}{d\tau} Z^\rho \Gamma^\mu_{\rho\nu} + X^\nu \frac{dZ^\rho}{d\tau}\Gamma^\mu_{\rho\nu}\right)\right|_p \\ &= -\left. X^\nu Z^\rho \frac{d\Gamma^\mu_{\rho\nu}}{d\tau}\right|_p \\ &= -(X^\nu X^\sigma Z^\rho \Gamma^\mu_{\rho\nu,\sigma})_p \end{aligned} \tag{3.33}$$

Here the second line follows because we're working in normal coordinates at $p$, and the final line because $\tau$ is the parameter along the integral curve of $X$, so $d/d\tau = X^\sigma \partial_\sigma$. We therefore have

$$Z^\mu_q = Z^\mu_p - \frac{1}{2}(X^\nu X^\sigma Z^\rho \Gamma^\mu_{\rho\nu,\sigma})_p\, \delta\tau^2 + \ldots \tag{3.34}$$

Now we parallel transport once more, this time along $Y$ to $Z^\mu_r$. The Taylor expansion now takes the form

$$Z^\mu_q = Z^\mu_p + \left.\frac{dZ^\mu}{d\sigma}\right|_q \delta\sigma + \frac{1}{2}\left.\frac{d^2 Z^\mu}{d\sigma^2}\right|_q \delta\sigma^2 + \mathcal{O}(\delta\sigma^3) \tag{3.35}$$

We can again evaluate the first derivative $dZ^\mu/d\sigma|_q$ using the analog of the parallel transport equation (3.32),

$$\left.\frac{dZ^\mu}{d\sigma}\right|_q = -(Y^\nu Z^\rho \Gamma^\mu_{\rho\nu})_q$$

Since we're working in normal coordinates about $p$ and not $q$, we no longer get to argue that this term vanishes. Instead we Taylor expand about $p$ to get

$$(Y^\nu Z^\rho \Gamma^\mu_{\rho\nu})_q = (Y^\nu Z^\rho X^\sigma \Gamma^\mu_{\rho\nu,\sigma})_p \, \delta\tau + \ldots$$

Note that in principle we should also Taylor expand $Y^\nu$ and $Z^\rho$ but, at leading order, these will multiply $\Gamma^\mu_{\rho\nu}(p) = 0$, so they only contribute at next order. The second order term in the Taylor expansion (3.35) involves $d^2Z^\mu/d\sigma^2|_q$ and there is an expression similar to (3.33). To leading order the $dX^\nu/d\sigma$ and $dZ^\rho/d\sigma$ terms are again absent because they are multiplied by $\Gamma^\mu_{\rho\nu}(q) = d\Gamma^\mu_{\rho\nu}/d\tau|_p \, \delta\tau$. We therefore have

$$\begin{aligned} \left.\frac{d^2 Z^\mu}{d\sigma^2}\right|_q &= -(Y^\nu Y^\sigma Z^\rho \Gamma^\mu_{\rho\nu,\sigma})_q + \ldots \\ &= -(Y^\nu Y^\sigma Z^\rho \Gamma^\mu_{\rho\nu,\sigma})_p + \ldots \end{aligned}$$

where we replaced the point $q$ with point $p$ because they differ only subleading terms proportional to $\delta\tau$. The upshot is that this time the difference between $Z^\mu_r$ and $Z^\mu_q$ involves two terms,

$$Z^\mu_r = Z^\mu_q - (Y^\nu Z^\rho X^\sigma \Gamma^\mu_{\rho\nu,\sigma})_p \, \delta\tau\delta\sigma - \frac{1}{2}(Y^\nu Y^\sigma Z^\rho \Gamma^\mu_{\rho\nu,\sigma})_p \, \delta\sigma^2 + \ldots$$

Finally, we can relate $Z^\mu_q$ to $Z^\mu_p$ using the expression (3.34) that we derived previously. We end up with

$$Z^\mu_r = Z^\mu_p - \frac{1}{2}(\Gamma^\mu_{\rho\nu,\sigma})_p \left[X^\nu X^\sigma Z^\rho \, \delta\tau^2 + 2Y^\nu Z^\rho X^\sigma \, \delta\sigma\delta\tau + Y^\nu Y^\sigma Z^\rho \, \delta\sigma^2\right]_p + \ldots$$

where $\ldots$ denotes any terms cubic or higher in small quantities.

Now suppose we go along the path $C'$, first visiting point $s$ and then making our way to $r$. We can read the answer off directly from the result above, simply by swapping $X$ and $Y$ and $\sigma$ and $\tau$; only the middle term changes,

$$Z'^\mu_r = Z^\mu_p - \frac{1}{2}(\Gamma^\mu_{\rho\nu,\sigma})_p \left[X^\nu X^\sigma Z^\rho \, \delta\tau^2 + 2X^\nu Z^\rho Y^\sigma \, \delta\sigma\delta\tau + Y^\nu Y^\sigma Z^\rho \, \delta\sigma^2\right]_p + \ldots$$

We find that

$$\begin{aligned} \Delta Z^\mu_r = Z^\mu_r - Z'^\mu_r &= -(\Gamma^\mu_{\rho\nu,\sigma} - \Gamma^\mu_{\rho\sigma,\nu})_p (Y^\nu Z^\rho X^\sigma)_p \, \delta\sigma\delta\tau + \ldots \\ &= (R^\mu{}_{\rho\sigma\nu} Y^\nu Z^\rho X^\sigma)_p \, \delta\sigma\delta\tau + \ldots \end{aligned}$$

where, in the final equality, we've used the expression for the Riemann tensor in components (3.15), which simplifies in normal coordinates as $\Gamma^\mu_{\rho\sigma}(p) = 0$. Note that, to the order we're working, we could equally as well evaluate $R^\mu{}_{\rho\sigma\nu} X^\nu Z^\rho Y^\sigma$ at the point $r$; the two differ only by higher order terms.

Although our calculation was performed with a particular choice of coordinates, the end result is written as an equality between tensors and must, therefore, hold in any coordinate system. This is a trick that we will use frequently throughout these lectures: calculations are considerably easier in normal coordinates. But if the resulting expression relate tensors then the final result must be true in any coordinate system.

We have discovered a rather nice interpretation of the Riemann tensor: it tells us the path dependence of parallel transport. The calculation above is closely related to the idea of *holonomy*. Here, one transports a vector around a closed curve $C$ and asks how the resulting vector compares to the original. This too is captured by the Riemann tensor. A particularly simple example of non-trivial holonomy comes from parallel transport of a vector on a sphere: the direction that you end up pointing in depends on the path you take.

**The Meaning of Torsion**

We discarded torsion almost as soon as we met it, choosing to work with the Levi-Civita connection which has vanishing torsion, $\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu}$. Moreover, as we will see in Section 4, torsion plays no role in the theory of general relativity which makes use of the Levi-Civita connection. Nonetheless, it is natural to ask: what is the geometric meaning of torsion? There is an answer to this that makes use of the kind of parallel transport arguments we used above.

This time, we start with two vectors $X, Y \in T_p(M)$. We pick coordinates $x^\mu$ and write these vectors as $X = X^\mu \partial_\mu$ and $Y = Y^\mu \partial_\mu$. Starting from $p \in M$, we can use these two vectors to construct two points infinitesimally close to $p$. We call these points $r$ and $s$ respectively: they have coordinates

$$r : x^\mu + X^\mu \epsilon \quad \text{and} \quad s : x^\mu + Y^\mu \epsilon$$

where $\epsilon$ is some infinitesimal parameter.

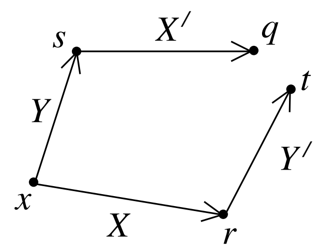



**Figure 25:**

We now parallel transport the vector $X \in T_p(M)$ along the direction of $Y$ to give a new vector $X' \in T_s(M)$. Similarly, we parallel transport $Y$ along the direction of $X$ to get a new vector $Y' \in T_r(M)$. These new vectors have components

$$X' = (X^\mu - \epsilon \Gamma^\mu_{\nu\rho} Y^\nu X^\rho)\partial_\mu \quad \text{and} \quad Y' = (Y^\mu - \epsilon \Gamma^\mu_{\nu\rho} X^\nu Y^\rho)\partial_\mu$$

Each of these tangent vectors now defines a new point. Starting from point $s$, and moving in the direction of $X'$, we see that we get a new point $q$ with coordinates

$$q : x^\mu + (X^\mu + Y^\mu)\epsilon - \epsilon^2 \Gamma^\mu_{\nu\rho} Y^\nu X^\rho$$

Meanwhile, if we sit at point $r$ and move in the direction of $Y'$, we get to a typically different point, $t$, with coordinates

$$t : x^\mu + (X^\mu + Y^\mu)\epsilon - \epsilon^2 \Gamma^\mu_{\nu\rho} X^\nu Y^\rho$$

We see that if the connection has torsion, so $\Gamma^\mu_{\nu\rho} \neq \Gamma^\mu_{\rho\nu}$, then the two points $q$ and $t$ do not coincide. In other words, torsion measures the failure of the parallelogram shown in figure to close.

### 3.3.4 Geodesic Deviation

Consider now a one-parameter family of geodesics, with coordinates $x^\mu(\tau; s)$. Here $\tau$ is the affine parameter along the geodesics, all of which are tangent to the vector field $X$ so that, along the surface spanned by $x^\mu(\tau, s)$, we have

$$X^\mu = \left.\frac{\partial x^\mu}{\partial \tau}\right|_s$$

Meanwhile, $s$ labels the different geodesics, as shown in Figure 26. We take the tangent vector in the $s$ direction to be generated by a second vector field $S$ so that,

$$S^\mu = \left.\frac{\partial x^\mu}{\partial s}\right|_\tau$$

The tangent vector $S^\mu$ is sometimes called the *deviation vector*; it takes us from one geodesic to a nearby geodesic with the same affine parameter $\tau$.

The family of geodesics sweeps out a surface embedded in the manifold. This gives us some freedom in the way we assign coordinates $s$ and $\tau$. In fact, we can always pick coordinates $s$ and $t$ on the surface such that $S = \partial/\partial s$ and $T = \partial/\partial t$, ensuring that

$$[S, X] = 0$$

Roughly speaking, we can do this if we use $\tau$ and $s$ as coordinates on some submanifold of $M$. Then the vector fields can be written simply as $X = \partial/\partial\tau$ and $S = \partial/\partial s$ and $[X, S] = 0$.

We can ask how neighbouring geodesics behave. Do they converge? Or do they move further apart? Now consider a connection $\Gamma$ with vanishing torsion, so that $\nabla_X S - \nabla_S X = [X, S]$. Since $[X, S] = 0$, we have

$$\nabla_X \nabla_X S = \nabla_X \nabla_S X = \nabla_S \nabla_X X + R(X, S)X$$

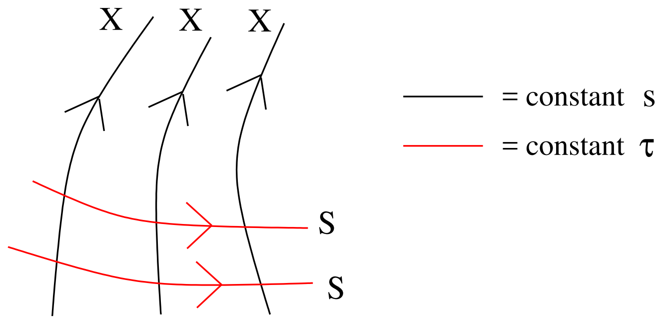



**Figure 26:** The black lines are geodesics generated by $X$. The red lines label constant $\tau$ and are generated by $S$, with $[X,S]=0$.

where, in the second equality, we've used the expression (3.14) for the Riemann tensor as a differential operator. But $\nabla_X X = 0$ because $X$ is tangent to geodesics, and we have

$$\nabla_X \nabla_X S = R(X,S)X$$

In index notation, this is

$$X^\nu \nabla_\nu (X^\rho \nabla_\rho S^\mu) = R^\mu{}_{\nu\rho\sigma} X^\nu X^\rho S^\sigma$$

If we further restrict to an integral curve $C$ associated to the vector field $X$, as in (3.25), this equation becomes

$$\frac{d^2 S^\mu}{d\tau^2} = R^\mu{}_{\nu\rho\sigma} X^\nu X^\rho S^\sigma \tag{3.36}$$

The left-hand-side tells us how the deviation vector $S^\mu$ changes as we move along the geodesic. In other words, it is the relative acceleration of neighbouring geodesics. We learn that this relative acceleration is controlled by the Riemann tensor.

Experimentally, such geodesic deviations are called *tidal forces*. We met a simple example in Section 1.2.4.

### An Example: the Sphere $\mathbf{S}^2$ Again

It is simple to determine the geodesics on the sphere $\mathbf{S}^2$ of radius $r$. Using the Christoffel symbols (3.19), the geodesic equations are

$$\frac{d^2\theta}{d\tau^2} = \sin\theta\cos\theta\left(\frac{d\phi}{d\tau}\right)^2 \quad \text{and} \quad \frac{d^2\phi}{d\tau^2} = -2\frac{\cos\theta}{\sin\theta}\frac{d\phi}{d\tau}\frac{d\theta}{d\tau}$$

The solutions are great circles. The general solution is a little awkward in these coordinates, but there are two simple solutions.

- We can set $\theta = \pi/2$ with $\dot{\theta} = 0$ and $\dot{\phi} =$ constant. This is a solution in which the particle moves around the equator. Note that this solution doesn't work for other values of $\theta$.
- We can set $\dot{\phi} = 0$ and $\dot{\theta} =$ constant. These are paths of constant longitude and are geodesics for any constant value of $\phi$. Note, however, that our coordinates go a little screwy at the poles $\theta = 0$ and $\theta = \pi$.

To illustrate geodesic deviation, we'll look at the second class of solutions; the particle moves along $\theta = v\tau$, with the angle $\phi$ specifying the geodesic. This set-up is simple enough that we don't need to use any fancy Riemann tensor techniques: we can just understand the geodesic deviation using simple geometry. The distance between the geodesic at $\phi = 0$ and the geodesic at some other longitude $\phi$ is

$$s(\tau) = r\phi \sin\theta = r\phi \sin(v\tau) \tag{3.37}$$

Now let's re-derive this result using out fancy technology. The geodesics are generated by the vector field $X^\theta = v$. Meanwhile, the separation between geodesics at a fixed $\tau$ is $S^\phi = s(\tau)$. The geodesic deviation equation in the form (3.36) is

$$\frac{d^2 s}{d\tau^2} = v^2 R^\phi{}_{\theta\theta\phi}\, s(\tau)$$

We computed the Riemann tensor for $\mathbf{S}^2$ in (3.20); the relevant component is

$$R_{\phi\theta\theta\phi} = -r^2 \sin^2\theta \quad \Rightarrow \quad R^\phi{}_{\theta\theta\phi} = g^{\phi\phi} R_{\phi\theta\theta\phi} = -1 \tag{3.38}$$

and the geodesic deviation equation becomes simply

$$\frac{d^2 s}{d\tau^2} = -v^2 s$$

which is indeed solved by (3.37).

### 3.4 More on the Riemann Tensor and its Friends

Recall that the components of the Riemann tensor are given by (3.15),

$$R^\sigma{}_{\rho\mu\nu} = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\lambda_{\nu\rho}\Gamma^\sigma_{\mu\lambda} - \Gamma^\lambda_{\mu\rho}\Gamma^\sigma_{\nu\lambda} \tag{3.39}$$

We can immediately see that the Riemann tensor is anti-symmetric in the final two indices

$$R^\sigma{}_{\rho\mu\nu} = -R^\sigma{}_{\rho\nu\mu}$$

However, there are also a number of more subtle symmetric properties satisfied by the Riemann tensor when we use the Levi-Civita connection. Logically, we could have discussed this back in Section 3.2. However, it turns out that a number of statements are substantially simpler to prove using normal coordinates introduced in Section 3.3.2.

**Claim**: If we lower an index on the Riemann tensor, and write $R_{\sigma\rho\mu\nu} = g_{\sigma\lambda}R^{\lambda}{}_{\rho\mu\nu}$ then the resulting object also obeys the following identities

- $R_{\sigma\rho\mu\nu} = -R_{\sigma\rho\nu\mu}$.
- $R_{\sigma\rho\mu\nu} = -R_{\rho\sigma\mu\nu}$.
- $R_{\sigma\rho\mu\nu} = R_{\mu\nu\sigma\rho}$.
- $R_{\sigma[\rho\mu\nu]} = 0$.

**Proof**: We work in normal coordinates, with $\Gamma^{\lambda}{}_{\mu\nu} = 0$ at a point. The Riemann tensor can then be written as

$$\begin{aligned} R_{\sigma\rho\mu\nu} &= g_{\sigma\lambda}\left(\partial_\mu\Gamma^{\lambda}{}_{\nu\rho} - \partial_\nu\Gamma^{\lambda}{}_{\mu\rho}\right) \\ &= \frac{1}{2}\left(\partial_\mu(\partial_\nu g_{\sigma\rho} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\nu\rho}) - \partial_\nu(\partial_\mu g_{\sigma\rho} + \partial_\rho g_{\mu\sigma} - \partial_\sigma g_{\mu\rho})\right) \\ &= \frac{1}{2}\left(\partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\mu\partial_\sigma g_{\nu\rho} - \partial_\nu\partial_\rho g_{\mu\sigma} + \partial_\nu\partial_\sigma g_{\mu\rho}\right) \end{aligned}$$

where, in going to the second line, we used the fact that $\partial_\mu g^{\lambda\sigma} = 0$ in normal coordinates. The first three symmetries are manifest; the final one follows from a little playing. (It is perhaps quicker to see the final symmetry if we return to the Christoffel symbols where, in normal coordinates, we have $R^{\sigma}{}_{\rho\mu\nu} = \partial_\mu\Gamma^{\sigma}{}_{\rho\nu} - \partial_\nu\Gamma^{\sigma}{}_{\rho\mu}$.) But since the symmetry equations are tensor equations, they must hold in all coordinate systems. □

**Claim**: The Riemann tensor also obeys the *Bianchi identity*

$$\nabla_{[\lambda}R_{\sigma\rho]\mu\nu} = 0 \tag{3.40}$$

Alternatively, we can anti-symmetrise on the final two indices, in which case this can be written as $R^{\sigma}{}_{\rho[\mu\nu;\lambda]} = 0$.

**Proof**: We again use normal coordinates, where $\nabla_\lambda R_{\sigma\rho\mu\nu} = \partial_\lambda R_{\sigma\rho\mu\nu}$ at the point $p$. Schematically, we have $R = \partial\Gamma + \Gamma\Gamma$, so $\partial R = \partial^2\Gamma + \Gamma\partial\Gamma$ and the final $\Gamma\partial\Gamma$ term is absent in normal coordinates. This means that we just have $R = \partial^2\Gamma$ which, in its full coordinated glory, is

$$\partial_\lambda R_{\sigma\rho\mu\nu} = \frac{1}{2}\partial_\lambda\left(\partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\mu\partial_\sigma g_{\nu\rho} - \partial_\nu\partial_\rho g_{\mu\sigma} + \partial_\nu\partial_\sigma g_{\mu\rho}\right)$$

Now anti-symmetrise on the three appropriate indices to get the result. $\square$

For completeness, we should mention that the identities $R_{\sigma[\rho\mu\nu]} = 0$ and $\nabla_{[\lambda}R_{\sigma\rho]\mu\nu} = 0$ (sometimes called the first and second Bianchi identities respectively) are more general, in the sense that they hold for arbitrary torsion free connection. In contrast, the other two identities, $R_{\sigma\rho\mu\nu} = -R_{\rho\sigma\mu\nu}$ and $R_{\sigma\rho\mu\nu} = R_{\mu\nu\sigma\rho}$ hold only for the Levi-Civita connection.

### 3.4.1 The Ricci and Einstein Tensors

There are a number of further tensors that we can build from the Riemann tensor.

First, given a rank $(1,3)$ tensor, we can always construct a rank $(0,2)$ tensor by contraction. If we start with the Riemann tensor, the resulting object is called the *Ricci tensor*. It is defined by

$$R_{\mu\nu} = R^\rho{}_{\mu\rho\nu}$$

The Ricci tensor inherits its anti-symmetry from the Riemann tensor. We write $R_{\mu\nu} = g^{\sigma\rho}R_{\sigma\mu\rho\nu} = g^{\rho\sigma}R_{\rho\nu\sigma\mu}$, giving us

$$R_{\mu\nu} = R_{\nu\mu}$$

We can go one step further and create a function $R$ over the manifold. This is the *Ricci scalar*,

$$R = g^{\mu\nu}R_{\mu\nu}$$

The Bianchi identity (3.40) has a nice implication for the Ricci tensor. If we write the Bianchi identity out in full, we have

$$\begin{aligned} & & \nabla_\lambda R_{\sigma\rho\mu\nu} + \nabla_\sigma R_{\rho\lambda\mu\nu} + \nabla_\rho R_{\lambda\sigma\mu\nu} = 0 \\ \times\ g^{\mu\lambda}g^{\rho\nu} & \quad\Rightarrow\quad & \nabla^\mu R_{\mu\sigma} - \nabla_\sigma R + \nabla^\nu R_{\nu\sigma} = 0 \end{aligned}$$

which means that

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R$$

This motivates us to introduce the *Einstein tensor*,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$$

which has the property that it is covariantly constant, meaning

$$\nabla^\mu G_{\mu\nu} = 0 \tag{3.41}$$

We'll be seeing much more of the Ricci and Einstein tensors in the next section.

### 3.4.2 Connection 1-forms and Curvature 2-forms

Calculating the components of the Riemann tensor is straightforward but extremely tedious. It turns out that here is a slightly different way of repackaging the connection and the torsion and curvature tensors using the language of forms. This not only provides a simple way to actually compute the the Riemann tensor, but also offers some useful conceptual insight.

#### Vielbeins

Until now, we have typically worked with a coordinate basis $\{e_\mu\} = \{\partial_\mu\}$. However, we could always pick a basis of vector fields that has no such interpretation. For example, a linear combination of a coordinate basis, say

$$\hat{e}_a = e_a{}^\mu\,\partial_\mu$$

will not, in general, be a coordinate basis itself.

Given a metric, there is a non-coordinate basis that will prove particularly useful for computing the curvature tensor. This is the basis such that, on a Riemannian manifold,

$$g(\hat{e}_a, \hat{e}_b) = g_{\mu\nu} e_a{}^\mu e_b{}^\nu = \delta_{ab}$$

Alternatively, on a Lorentzian manifold we take

$$g(\hat{e}_a, \hat{e}_b) = g_{\mu\nu} e_a{}^\mu e_b{}^\nu = \eta_{ab} \tag{3.42}$$

The components $e_a{}^\mu$ are called *vielbeins* or *tetrads*. (On an $n$-dimensional manifold, these objects are usually called "German word for $n$"-beins. For example, one-dimensional manifolds have einbeins; four-dimensional manifolds have vierbeins.)

The is reminiscent of our discussion in Section 3.1.2 where we mentioned that we can always find coordinates so that any metric will look flat at a point. In (3.42), we've succeeded in making the manifold look flat everywhere (at least in a patch covered by a chart). There are no coordinates that do this, but there's nothing to stop us picking a basis of vector fields that does the job. In what follows, $\mu\nu$ indices are raised/lowered with the metric $g_{\mu\nu}$ while $a, b$ indices are raised/lowered with the flat metric $\delta_{ab}$ or $\eta_{ab}$. We will phrase our discussion in the context of Lorentzian manifolds, with an eye to later applications to general relativity.

The vielbeins aren't unique. Given a set of vielbeins, we can always find another set related by

$$\tilde{e}_a{}^\mu = e_b{}^\mu (\Lambda^{-1})^b{}_a \quad \text{with} \quad \Lambda_a{}^c \Lambda_b{}^d \eta_{cd} = \eta_{ab} \tag{3.43}$$

These are Lorentz transformations. However now they are *local Lorentz transformation*, because $\Lambda$ can vary over the manifold. These local Lorentz transformations are a redundancy in the definition of the vielbeins in (3.42).

The dual basis of one-forms $\{\hat{\theta}^a\}$ is defined by $\hat{\theta}^a(\hat{e}_b) = \delta^a_b$. They are related to the coordinate basis by

$$\hat{\theta}^a = e^a{}_\mu dx^\mu$$

Note the different placement of indices: $e^a{}_\mu$ is the inverse of $e_a{}^\mu$, meaning it satisfies $e^a{}_\mu e_b{}^\mu = \delta^a_b$ and $e^a{}_\mu e_a{}^\nu = \delta^\nu_\mu$. In the non-coordinate basis, the metric on a Lorentzian manifold takes the form

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu = \eta_{ab} \hat{\theta}^a \otimes \hat{\theta}^b \quad \Rightarrow \quad g_{\mu\nu} = e^a{}_\mu e^b{}_\nu \eta_{ab}$$

For Riemannian manifolds, we replace $\eta_{ab}$ with $\delta_{ab}$.

### The Connection One-Form

Given a non-coordinate basis $\{\hat{e}_a\}$, we can define the components of a connection in the usual way (3.8)

$$\nabla_{\hat{e}_c} \hat{e}_b = \Gamma^a_{cb}\, \hat{e}_a$$

Note that, annoyingly, these are not the same functions as $\Gamma^\mu_{\rho\nu}$, which are the components of the connection computed in the coordinate basis! You need to pay attention to whether the components are Greek $\mu, \nu$ etc which tells you that we're in the coordinate basis, or Roman $a, b$ etc which tells you we're in the vielbein basis.

We then define matrix-valued *connection one-form* as

$$\omega^a{}_b = \Gamma^a_{cb}\,\hat{\theta}^c \tag{3.44}$$

This is sometimes referred to as the *spin connection* because of the role it plays in defining spinors on curved spacetime. We'll describe this in Section 4.5.6.

The connection one-forms don't transform covariantly under local Lorentz transformations (3.43). Instead, in the new basis, the components of the connection one-form are defined as $\nabla_{\tilde{\hat{e}}_b}\tilde{\hat{e}}_c = \tilde{\Gamma}^a_{bc}\tilde{\hat{e}}_a$. You can check that the connection one-form transforms as

$$\tilde{\omega}^a{}_b = \Lambda^a{}_c\,\omega^c{}_d(\Lambda^{-1})^d{}_b + \Lambda^a{}_c(d\Lambda^{-1})^c{}_b \tag{3.45}$$

The second term reflects the fact that the original connection components $\Gamma^\mu_{\nu\rho}$ do not transform as a tensor, but with an extra term involving to the derivative of the coordinate transformation (3.13). This now shows up as an extra term involving the derivative of the local Lorentz transformation.

There is a rather simple way to compute the connection one-forms, at least for a torsion free connection. This follows from the first of two *Cartan structure relations*:

**Claim**: For a torsion free connection,

$$d\hat{\theta}^a + \omega^a{}_b \wedge \hat{\theta}^b = 0 \tag{3.46}$$

**Proof:** We first look at the second term,

$$\omega^a{}_b \wedge \hat{\theta}^b = \Gamma^a_{cb}\,(e^c{}_\mu dx^\mu) \wedge (e^b{}_\nu dx^\nu)$$

The components $\Gamma^a{}_{cb}$ are related to the coordinate basis components by

$$\Gamma^a_{cb} = e^a{}_\rho e_c{}^\mu\left(\partial_\mu e_b{}^\rho + e_b{}^\nu \Gamma^\rho_{\mu\nu}\right) = e^a{}_\rho e_c{}^\mu \nabla_\mu e_b{}^\rho \tag{3.47}$$

So

$$\begin{aligned}\omega^a{}_b \wedge \hat{\theta}^b &= e^a{}_\rho e_c{}^\lambda e^c{}_\mu e^b{}_\nu\,(\partial_\lambda e_b{}^\rho + e_b{}^\sigma \Gamma^\rho_{\lambda\sigma})\,dx^\mu \wedge dx^\nu \\ &= e^a{}_\rho e^b{}_\nu \partial_\mu e_b{}^\rho\,dx^\mu \wedge dx^\nu\end{aligned}$$

where, in the second line we've used $e_c{}^\lambda e^c{}_\mu = \delta^\lambda_\mu$ and the fact that the connection is torsion free so $\Gamma^\rho_{[\mu\nu]} = 0$. Now we use the fact that $e^b{}_\nu e_b{}^\rho = \delta^\rho_\nu$, so $e^b{}_\nu \partial_\mu e_b{}^\rho = -e_b{}^\rho \partial_\mu e^b{}_\nu$. We have

$$\begin{aligned}\omega^a{}_b \wedge \hat{\theta}^b &= -e^a{}_\rho e_b{}^\rho \partial_\mu e^b{}_\nu\,dx^\mu \wedge dx^\nu \\ &= -\partial_\mu e^a{}_\nu\,dx^\mu \wedge dx^\nu = -d\hat{\theta}^a\end{aligned}$$

which completes the proof. □

The discussion above was for a general connection. For the Levi-Civita connection, we have a stronger result

**Claim**: For the Levi-Civita connection, the connection one-form is anti-symmetric

$$\omega_{ab} = -\omega_{ba} \tag{3.48}$$

**Proof**: This follows from the explicit expression (3.47) for the components $\Gamma^a_{bc}$. Lowering an index, we have

$$\Gamma_{abc} = \eta_{ad}\, e^d{}_\rho e_b{}^\mu \nabla_\mu e_c{}^\rho = -\eta_{ad}\, e_c{}^\rho e_b{}^\mu \nabla_\mu e^d{}_\rho = -\eta_{cf} e^f{}_\sigma e_b{}^\mu \nabla_\mu (\eta_{ad} g^{\rho\sigma} e^d{}_\rho)$$

where, in the final equality, we've used the fact that the connection is compatible with the metric to raise the indices of $e^d{}_\rho$ inside the covariant derivative. Finishing off the derivation, we then have

$$\Gamma_{abc} = -\eta_{cf}\, e^f{}_\rho e_b{}^\mu \nabla_\mu e_a{}^\rho = -\Gamma_{cba}$$

The result then follows from the definition $\omega_{ab} = \Gamma_{acb}\hat{\theta}^c$. □

The Cartan structure equation (3.46), together with the anti-symmetry condition (3.48), gives a quick way to compute the spin connection. It's instructive to do some counting to see how these two equations uniquely define $\omega^a{}_b$. In particular, since $\omega_{ab}$ is anti-symmetric, one might think that it has $\frac{1}{2}n(n-1)$ independent components, and these can't possibly be fixed by the $n$ Cartan structure equations (3.46). But this is missing the fact that $\omega_{ab}$ are not numbers, but are one-forms. So the true number of components in $\omega_{ab}$ is $n \times \frac{1}{2}n(n-1)$. Furthermore, the Cartan structure equation is an equation relating 2-forms, each of which has $\frac{1}{2}n(n-1)$ components. This means that it's really $n \times \frac{1}{2}n(n-1)$ equations. We see that the counting does work, and the two fix the spin connection uniquely.

### The Curvature Two-Form

We can compute the components of the Riemann tensor in our non-coordinate basis,

$$R^a{}_{bcd} = R(\hat{\theta}^a; \hat{e}_c, \hat{e}_d, \hat{e}_b)$$

The anti-symmetry of the last two indices, $R^a{}_{bcd} = -R^a{}_{bdc}$, makes this ripe for turning into a matrix of two-forms,

$$\mathcal{R}^a{}_b = \frac{1}{2} R^a{}_{bcd}\, \hat{\theta}^c \wedge \hat{\theta}^d \tag{3.49}$$

The second of the two Cartan structure relations states that this can be written in terms of the curvature one-form as

$$\mathcal{R}^a{}_b = d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b \tag{3.50}$$

The proof of this is mechanical and somewhat tedious. It's helpful to define the quantities $[\hat{e}_a, \hat{e}_b] = f_{ab}{}^c\, \hat{e}_c$ along the way, since they appear on both left and right-hand sides.

### 3.4.3 An Example: the Schwarzchild Metric

The connection one-form and curvature two-form provide a slick way to compute the curvature tensor associated a metric. The reason for this is that computing exterior derivatives takes significantly less effort than computing covariant derivatives. We will illustrate this for metrics of the form,

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{3.51}$$

For later applications, it will prove useful to compute the Riemann tensor for this metric with general $f(r)$. However, if we want to restrict to the Schwarzchild metric we can take

$$f(r) = \sqrt{1 - \frac{2GM}{r}} \tag{3.52}$$

The basis of non-coordinate one-forms is

$$\hat{\theta}^0 = f\, dt \quad , \quad \hat{\theta}^1 = f^{-1}\, dr \quad , \quad \hat{\theta}^2 = r\, d\theta \quad , \quad \hat{\theta}^3 = r \sin\theta\, d\phi \tag{3.53}$$

Note that the one-forms $\hat{\theta}$ should not be confused with the angular coordinate $\theta$! In this basis, the metric takes the simple form

$$ds^2 = \eta_{ab} \hat{\theta}^a \otimes \hat{\theta}^b$$

We now compute $d\hat{\theta}^a$. Caclulationally, this is straightforward. In particular, it's substantially easier than computing the covariant derivative because there's no messy connection to worry about. The exterior derivatives are simply

$$d\hat{\theta}^0 = f'\, dr \wedge dt \quad , \quad d\hat{\theta}^1 = 0 \quad , \quad d\hat{\theta}^2 = dr \wedge d\theta \quad , \quad d\hat{\theta}^3 = \sin\theta\, dr \wedge d\phi + r\cos\theta\, d\theta \wedge d\phi$$

The first Cartan structure relation, $d\hat{\theta}^a = -\omega^a{}_b \wedge \hat{\theta}^b$, can then be used to read off the connection one-form. The first equation tells us that $\omega^0{}_1 = f' f dt = f'\, \hat{\theta}^0$. We then use the anti-symmetry (3.48), together with raising and lowering by the Minkowski metric $\eta = \text{diag}(-1, +1, +1, +1)$ to get $\omega^1{}_0 = \omega_{10} = -\omega_{01} = \omega^0{}_1$. The Cartan structure equation then gives $d\hat{\theta}^1 = -\omega^1{}_0 \wedge \hat{\theta}^0 + \ldots$ and the $\omega^1{}_0 \wedge \hat{\theta}^0$ contribution happily vanishes because it is proportional to $\hat{\theta}^0 \wedge \hat{\theta}^0 = 0$.

Next, we take $\omega^2{}_1 = f d\theta = (f/r)\hat{\theta}^2$ to solve the $d\hat{\theta}^2$ structure equation. The anti-symmetry (3.48) gives $\omega^1{}_2 = -\omega^2{}_1 = -(f/r)\hat{\theta}^2$ and this again gives a vanishing contribution to the $d\hat{\theta}^1$ structure equation.

Finally, the $d\hat{\theta}^3$ equation suggests that we take $\omega^3{}_1 = f\sin\theta d\phi = (f/r)\hat{\theta}^3$ and $\omega^3{}_2 = \cos\theta d\phi = (1/r)\cot\theta\,\hat{\theta}^3$. These anti-symmetric partners $\omega^1{}_3 = -\omega^3{}_1$ and $\omega^2{}_3 = -\omega^3{}_2$ do nothing to spoil the $d\hat{\theta}^1$ and $d\hat{\theta}^2$ structure equations, so we're home dry. The final result is

$$\begin{aligned} \omega^0{}_1 = \omega^1{}_0 = f'\,\hat{\theta}^0 \quad &, \quad \omega^2{}_1 = -\omega^1{}_2 = \frac{f}{r}\hat{\theta}^2 \\ \omega^3{}_1 = -\omega^1{}_3 = \frac{f}{r}\hat{\theta}^3 \quad &, \quad \omega^3{}_2 = -\omega^2{}_3 = \frac{\cot\theta}{r}\hat{\theta}^3 \end{aligned}$$

Now we can use this to compute the curvature two-form. We will focus on

$$\mathcal{R}^0{}_1 = d\omega^0{}_1 + \omega^0{}_c \wedge \omega^c{}_1$$

We have

$$d\omega^0{}_1 = f' d\hat{\theta}^0 + f'' dr \wedge \hat{\theta}^0 = \Big((f')^2 + f'' f\Big) dr \wedge dt$$

The second term in the curvature 2-form is $\omega^0{}_c \wedge \omega^c{}_1 = \omega^0{}_1 \wedge \omega^1{}_1 = 0$. So we're left with

$$\mathcal{R}^0{}_1 = \Big((f')^2 + f'' f\Big) dr \wedge dt = \Big((f')^2 + f'' f\Big)\hat{\theta}^1 \wedge \hat{\theta}^0$$

The other curvature 2-forms can be computed in a similar fashion. We can now read off the components of the Riemann tensor in the non-coordinate basis using (3.49). (We should remember that we get a contribution from both $R^0{}_{101}$ and $R^0{}_{110} = -R^0{}_{101}$, which cancels the factor of 1/2 in (3.49).) After lowering an index, we find that the non-vanishing components of the Riemann tensor are

$$\begin{aligned} R_{0101} &= ff'' + (f')^2 \\ R_{0202} &= \frac{ff'}{r} \\ R_{0303} &= \frac{ff'}{r} \\ R_{1212} &= -\frac{ff'}{r} \\ R_{1313} &= -\frac{ff'}{r} \\ R_{2323} &= \frac{1-f^2}{r^2} \end{aligned}$$

We can also convert this back to the coordinates $x^\mu = (t, r, \theta, \phi)$ using

$$R_{\mu\nu\rho\sigma} = e^a{}_\mu e^b{}_\nu e^c{}_\rho e^d{}_\sigma R_{abcd}$$

This is particularly easy in this case because the matrices $e_a{}^\mu$ defining the one-forms (3.53) are diagonal. We then have

$$\begin{aligned}
R_{trtr} &= ff'' + (f')^2 \\
R_{t\theta t\theta} &= f^3 f' r \\
R_{t\phi t\phi} &= f^3 f' r \sin^2\theta \\
R_{r\theta r\theta} &= -\frac{f' r}{f} \\
R_{r\phi r\phi} &= -\frac{f' r}{f}\sin^2\theta \\
R_{\theta\phi\theta\phi} &= (1-f^2) r^2 \sin^2\theta
\end{aligned} \tag{3.54}$$

Finally, if we want to specialise to the Schwarzchild metric with $f(r)$ given by (3.52), we have

$$\begin{aligned}
R_{trtr} &= -\frac{2GM}{r^3} \\
R_{t\theta t\theta} &= \frac{GM(r-2GM)}{r^2} \\
R_{t\phi t\phi} &= \frac{GM(r-2GM)}{r^2}\sin^2\theta \\
R_{r\theta r\theta} &= -\frac{GM}{r-2GM} \\
R_{r\phi r\phi} &= -\frac{GM\sin^2\theta}{r-2GM} \\
R_{\theta\phi\theta\phi} &= 2GMr\sin^2\theta
\end{aligned}$$

Although the calculation is a little lengthy, it turns out to be considerably quicker than first computing the Levi-Civita connection and subsequently motoring through to get the Riemann tensor components.

### 3.4.4 The Relation to Yang-Mills Theory

It is no secret that the force of gravity is geometrical. However, the other forces are equally as geometrical. The underlying geometry is something called a fibre bundle, rather than the geometry of spacetime.

We won't describe fibre bundles in this course, but we can exhibit a clear similarity between the structures that arise in general relativity and the structures that arise in the other forces, which are described by Maxwell theory and its generalisation to Yang-Mills theory.

Yang-Mills theory is based on a simple Lie group $G$ which, for this discussion, we will take to be $SU(N)$ or $U(N)$. If we take $G = U(1)$, then Yang-Mills theory reduces to Maxwell theory. The theory is described in terms of an object that physicists call a *gauge potential*. This is a spacetime "vector" $A_\mu$ which lives in the Lie algebra of $G$. In more down to earth terms, each components is an anti-Hermitian $N \times N$ matrix, $(A_\mu)^a{}_b$, with $a, b = 1, \ldots, N$. In fact, as we saw above, this "vector" is really a one-form. The novelty is that it's a Lie algebra-valued one-form.

Mathematicians don't refer to $A_\mu$ as a gauge potential. Instead, they call it a connection (on a fibre bundle). This relationship becomes clearer if we look at how $A_\mu$ changes under a gauge transformation

$$\tilde{A}_\mu = \Omega A_\mu \Omega^{-1} + \Omega \partial_\mu \Omega^{-1}$$

where $\Omega(x) \in G$. This is identical to the transformation property (3.45) of the one-form connection under local Lorentz transformations.

In Yang-Mills, as in Maxwell theory, we construct a field strength. In components, this is given by

$$(F_{\mu\nu})^a{}_b = \partial_\mu (A_\nu)^a{}_b - \partial_\nu (A_\mu)^a{}_b + [A_\mu, A_\nu]^a{}_b$$

Alternatively, in the language of forms, the field strength becomes

$$F^a_{\ b} = dA^a{}_b + A^a{}_c \wedge A^c{}_b$$

Again, there is an obvious similarity with the curvature 2-form introduced in (3.50). Mathematicians refer the Yang-Mills field strength the "curvature".

A particularly quick way to construct the Yang-Mills field strength is to take the commutator of two covariant derivatives. It is simple to check that

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = F_{\mu\nu}$$

where I've suppressed the $a, b$ indices on both sides. This is the gauge theory version of the Ricci identity (3.16): for a torsion free connection,

$$[\nabla_\mu, \nabla_\nu] Z^\sigma = R^\sigma{}_{\rho\mu\nu} Z^\rho$$

# 4. The Einstein Equations

It is now time to do some physics. The force of gravity is mediated by a gravitational field. The glory of general relativity is that this field is identified with a metric $g_{\mu\nu}(x)$ on a 4d Lorentzian manifold that we call spacetime.

This metric is not something fixed; it is, like all other fields in Nature, a dynamical object. This means that there are rules which govern how this field evolves in time. The purpose of this section is to explore these rules and some of their consequences.

We will start by understanding the dynamics of the gravitational field in the absence of any matter. We will then turn to understand how the gravitational field responds matter – or, more precisely, to energy and momentum – in Section 4.5.

## 4.1 The Einstein-Hilbert Action

All our fundamental theories of physics are described by action principles. Gravity is no different. Furthermore, the straight-jacket of differential geometry places enormous restrictions on the kind of actions that we can write down. These restrictions ensure that the action is something intrinsic to the metric itself, rather than depending on our choice of coordinates.

Spacetime is a manifold $M$, equipped with a metric of Lorentzian signature. An action is an integral over $M$. We know from Section 2.4.4 that we need a volume-form to integrate over a manifold. Happily, as we have seen, the metric provides a canonical volume form, which we can then multiply by any scalar function. Given that we only have the metric to play with, the simplest such (non-trivial) function is the Ricci scalar $R$. This motivates us to consider the wonderfully concise action

$$S = \int d^4x\, \sqrt{-g} R \tag{4.1}$$

This is the *Einstein-Hilbert action*. Note that the minus sign under the square-root arises because we are in a Lorentzian spacetime: the metric has a single negative eigenvalue and so its determinant, $g = \det g_{\mu\nu}$, is negative.

As a quick sanity check, recall that the Ricci tensor takes the schematic form (3.39) $R \sim \partial\Gamma + \Gamma\Gamma$ while the Levi-Civita connection itself is $\Gamma \sim \partial g$. This means that the Einstein-Hilbert action is second order in derivatives, just like most other actions we consider in physics.

### Varying the Einstein-Hilbert Action

We would like to determine the Euler-Lagrange equations arising from the action (4.1). We do this in the usual way, by starting with some fixed metric $g_{\mu\nu}(x)$ and seeing how the action changes when we shift

$$g_{\mu\nu}(x) \to g_{\mu\nu}(x) + \delta g_{\mu\nu}(x)$$

Writing the Ricci scalar as $R = g^{\mu\nu} R_{\mu\nu}$, the Einstein-Hilbert action clearly changes as

$$\delta S = \int d^4x \, \Big( (\delta\sqrt{-g}) g^{\mu\nu} R_{\mu\nu} + \sqrt{-g}(\delta g^{\mu\nu}) R_{\mu\nu} + \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} \Big) \tag{4.2}$$

It turns out that it's slightly easier to think of the variation in terms of the inverse metric $\delta g^{\mu\nu}$. This is equivalent to the variation of the metric $\delta g_{\mu\nu}$; the two are related by

$$g_{\rho\mu} g^{\mu\nu} = \delta^\nu_\rho \quad \Rightarrow \quad (\delta g_{\rho\mu}) g^{\mu\nu} + g_{\rho\mu} \delta g^{\mu\nu} = 0 \quad \Rightarrow \quad \delta g^{\mu\nu} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\rho\sigma}$$

The middle term in (4.2) is already proportional to $\delta g^{\mu\nu}$. We now deal with the first and third terms in turn. We will need the following result:

**Claim:** The variation of $\sqrt{-g}$ is given by

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\, g_{\mu\nu}\, \delta g^{\mu\nu}$$

**Proof:** We use the fact that any diagonalisable matrix $A$ obeys the identity

$$\log\det A = \operatorname{tr}\log A$$

This is obviously true for diagonal matrices. (The determinant is the product of eigenvalues while the trace is the sum of eigenvalues.) But because both the determinant and the trace are invariant under conjugation, it is also true for a diagonalisable matrix. Using this, we have,

$$\frac{1}{\det A}\,\delta(\det A) = \operatorname{tr}(A^{-1}\delta A)$$

Applying this to the metric, we have

$$\delta\sqrt{-g} = \frac{1}{2}\frac{1}{\sqrt{-g}}\,(-g)\, g^{\mu\nu}\, \delta g_{\mu\nu} = \frac{1}{2}\sqrt{-g}\, g^{\mu\nu}\, \delta g_{\mu\nu}$$

Using $g^{\mu\nu}\delta g_{\mu\nu} = -g_{\mu\nu}\delta g^{\mu\nu}$ then gives the result. □

So far, we have managed to write the variation of the action (4.2) as

$$\delta S=\int d^4x\ \sqrt{-g}\ \left(R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}\right)\delta g^{\mu\nu}+\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu}$$

We now need only worry about the final term. For this, we use:

**Claim:** The variation of the Ricci tensor is a total derivative

$$\delta R_{\mu\nu}=\nabla_\rho\,\delta\Gamma^\rho_{\mu\nu}-\nabla_\nu\,\delta\Gamma^\rho_{\mu\rho}$$

where

$$\delta\Gamma^\rho_{\mu\nu}=\frac{1}{2}g^{\rho\sigma}\left(\nabla_\mu\delta g_{\sigma\nu}+\nabla_\nu\delta g_{\sigma\mu}-\nabla_\sigma\delta g_{\mu\nu}\right)$$

**Proof:** We start by looking at the variation of the Christoffel symbols, $\Gamma^\rho_{\mu\nu}$. First note that, although the Christoffel symbol itself is not a tensor, the variation $\delta\Gamma^\rho_{\mu\nu}$ is a tensor. This is because it is the difference of Christoffel symbols, one computed using $g_{\mu\nu}$ and the other using $g_{\mu\nu}+\delta g_{\mu\nu}$. But the extra derivative term in the transformation of $\Gamma^\rho_{\mu\nu}$ is independent of the metric and so cancels out when we take the difference, leaving us with an object which transforms nicely as a tensor.

This is a useful observation. At any point $p\in M$ we can choose to work in normal coordinates such that $\partial_\rho g_{\mu\nu}=0$ and, correspondingly, $\Gamma^\rho_{\mu\nu}=0$. Then, to linear order in the variation, the change in the Christoffel symbol evaluated at $p$ is

$$\begin{aligned}\delta\Gamma^\rho_{\mu\nu}&=\frac{1}{2}g^{\rho\sigma}\left(\partial_\mu\delta g_{\sigma\nu}+\partial_\nu\delta g_{\sigma\mu}-\partial_\sigma\delta g_{\mu\nu}\right)\\&=\frac{1}{2}g^{\rho\sigma}\left(\nabla_\mu\delta g_{\sigma\nu}+\nabla_\nu\delta g_{\sigma\mu}-\nabla_\sigma\delta g_{\mu\nu}\right)\end{aligned}$$

where we're at liberty to replace the partial derivatives with covariant derivatives because they differ only by the Christoffel symbols $\Gamma^\rho_{\mu\nu}$ which, in normal coordinates, vanish at $p$. However, both the left and right-hand sides of this equation are tensors which means that although we derived this expression using normal coordinates, it must hold in any coordinate system. Moreover, the point $p$ was arbitrary so the final expression holds generally.

Next we look at the variation of the Riemann tensor. In normal coordinates, the expression (3.39) becomes

$$R^\sigma{}_{\rho\mu\nu}=\partial_\mu\Gamma^\sigma_{\nu\rho}-\partial_\nu\Gamma^\sigma_{\mu\rho}$$

and the variation is

$$\delta R^\sigma{}_{\rho\mu\nu} = \partial_\mu \, \delta\Gamma^\sigma_{\nu\rho} - \partial_\nu \, \delta\Gamma^\sigma_{\mu\rho} = \nabla_\mu \, \delta\Gamma^\sigma_{\nu\rho} - \nabla_\nu \, \delta\Gamma^\sigma_{\mu\rho}$$

where, as before, we replace partial derivatives with covariant derivatives as we are working in normal coordinates where the Christoffel symbols vanish. Once again, our final expression relates two tensors and must, therefore, hold in any coordinate system. Contracting indices (and working to leading order), we have

$$\delta R_{\rho\nu} = \nabla_\mu \, \delta\Gamma^\mu_{\nu\rho} - \nabla_\nu \, \delta\Gamma^\mu_{\rho\mu}$$

as claimed. □

The upshot of these calculations is that

$$g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\mu X^\mu \quad \text{with} \quad X^\mu = g^{\rho\nu} \, \delta\Gamma^\mu_{\rho\nu} - g^{\mu\nu} \, \delta\Gamma^\rho_{\nu\rho}$$

The variation of the action (4.2) can then be written as

$$\delta S = \int d^4x \, \sqrt{-g} \left[ \left( R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \delta g^{\mu\nu} + \nabla_\mu X^\mu \right) \tag{4.3}$$

This final term is a total derivative and, by the divergence theorem of Section 3.2.4, we ignore it. Requiring that the action is extremised, so $\delta S = 0$, we have the equations of motion

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0 \tag{4.4}$$

where $G_{\mu\nu}$ is the Einstein tensor defined in Section 3.4.1. These are the *Einstein field equations* in the absence of any matter. In fact they simplify somewhat: if we contract (4.4) with $g^{\mu\nu}$, we find that we must have $R = 0$. Substituting this back in, the vacuum Einstein equations are simply the requirement that the metric is Ricci flat,

$$R_{\mu\nu} = 0 \tag{4.5}$$

These deceptively simple equations hold a myriad of surprises. We will meet some of the solutions as we go along, notably gravitational waves in Section 5.2 and black holes in Section **??**.

Before we proceed, a small comment. We happily discarded the boundary term in (4.3), a standard practice whenever we invoke the variational principle. It turns out that there are some situations in general relativity where we should not be quite so cavalier. In such circumstances, one can be more careful by invoking the so-called Gibbons-Hawking boundary term.

### 4.1.1 An Aside on Dimensional Analysis

As it stands, there's something a little fishy about the action (4.1): it doesn't have the right dimensions. This isn't such an issue since we have just a single term in the action and multiplying the action by a constant doesn't change the classical equations of motion. Nonetheless, it will prove useful to get it right at this stage.

If we take the coordinates $x^\mu$ to have dimension of length, then the metric $g_{\mu\nu}$ is necessarily dimensionless. The Ricci scalar involves two spatial derivatives so has dimension $[R] = L^{-2}$. Including the integration measure, the action (4.1) then has dimensions $[S] = L^2$. However, actions should have dimensions of energy $\times$ time (it's the same dimensions as $\hbar$), or $[S] = ML^2T^{-1}$. This means that the Einstein-Hilbert action should be multiplied by a constant with the appropriate dimensions. We take

$$S = \frac{c^3}{16\pi G}\int d^4x\ \sqrt{-g}R$$

where $c$ is the speed of light and $G$ is Newton's constant,

$$G \approx 6.67 \times 10^{-11}\ \mathrm{m}^3\,\mathrm{kg}^{-1}s^{-2}$$

This factor doesn't change the equation of motion in vacuum, but we will see in Section 4.5 that it determines the strength of the coupling between the gravitational field and matter, as we might expect.

It's no fun carrying around a morass of fundamental constants in all our equations. For this reason, we often work in "natural units" in which various constants are set equal to 1. From now on, we will set $c = 1$. (Any other choice of $c$, including $3 \times 10^8$, is simply dumb.) This means that units of length and time are equated.

However, different communities have different conventions when it comes to $G$. Relativists will typically set $G = 1$. Since we have already set $c = 1$, we have $[G] = LM^{-1}$. Setting $G = 1$ then equates mass with length. This is useful when discussing gravitational phenomenon where the mass is often directly related to the length. For example, the Schwarzchild radius of black hole is $R_s = 2GM/c^2$ which becomes simply $R_s = 2M$ once we set $G = c = 1$.

However, if you're interested in phenomena other than gravity, then it's no more sensible to set $G = 1$ than to set, say, the Fermi coupling for the weak force $G_F = 1$. Instead, it is more useful to choose the convention where $\hbar = 1$, a choice which equates energy with inverse time (also known as frequency). With this convention, Newton's constant has dimension $[G] = M^{-2}$. The corresponding energy scale is known as the (reduced) Planck mass; it is given by

$$M_{pl}^2 = \frac{\hbar c}{8\pi G}$$

It is around $10^{18}$ GeV. This is a very high energy scale, way beyond anything we have probed in experiment. This can be traced to the weakness of the gravitational force. With $c = \hbar = 1$, we can equally well write the Einstein-Hilbert action as

$$S = \frac{1}{2} M_{pl}^2 \int d^4x \, \sqrt{-g} R$$

You might be tempted to set $c = \hbar = G = 1$. This leaves us with no remaining dimensional quantities. It is typically a bad idea, not least because dimensional analysis is a powerful tool and one we do not want to lose. In these lectures, we will focus only on gravitational physics. Nonetheless, we will retain $G$ in all equations.

### 4.1.2 The Cosmological Constant

We motivated the Einstein-Hilbert action as the simplest term we could write down. While it's true that it's the simplest term that results in interesting dynamics for the gravitational field, there is in fact a simpler term which we could add to the action. This comes from multiplying the volume form by a constant. The resulting action is

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g}\, (R - 2\Lambda)$$

Here $\Lambda$ is referred to as the *cosmological constant*. It has dimension $[\Lambda] = L^{-2}$. The minus sign in the action comes from thinking of the Lagrangian as "$T - V$": the cosmological constant is like the potential energy $V$.

Varying the action as before now yields the Einstein equations,

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = -\Lambda g_{\mu\nu}$$

This time, if we contract with $g^{\mu\nu}$, we get $R = 4\Lambda$. Substituting this back in, the vacuum Einstein equations in the presence of a cosmological constant become

$$R_{\mu\nu} = \Lambda g_{\mu\nu}$$

We will solve these shortly in Section 4.2.

#### Higher Derivative Terms

The Einstein-Hilbert action (with cosmological constant) is the simplest thing we can write down but it is not the only possibility, at least if we allow for higher derivative terms. For example, there are three terms that contain four derivatives of the metric,

$$S_{4-\text{deriv}} = \int d^4x\ \sqrt{-g}\left(c_1 R^2 + c_2 R_{\mu\nu}R^{\mu\nu} + c_3 R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}\right)$$

with $c_1$, $c_2$ and $c_3$ dimensionless constants. General choices of these constants will result in higher order equations of motion which do not have a well-defined initial value problem. Nonetheless, it turns out that one can find certain combinations of these terms which conspire to keep the equations of motion second order. This is known as *Lovelock's theorem*. In $d = 4$ dimensions, this combination has a rather special topological property: a generalisation of the Gauss-Bonnet theorem states that

$$\frac{1}{8\pi^2}\int_M d^4x\ \sqrt{g}\left(R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}\right) = \chi(M)$$

where $\chi(M) \in \mathbf{Z}$ is the Euler character of $M$ that we previously defined in (2.35). In Lorentzian signature, this combination of curvature terms is also a total derivative and does not affect the classical equations of motion.

As in any field theory, higher derivative terms in the action only become relevant for fast varying fields. In General Relativity, they are unimportant for all observed physical phenomena and we will not discuss them further in this course.

### 4.1.3 Diffeomorphisms Revisited

Here's a simple question: how many degrees of freedom are there in the metric $g_{\mu\nu}$? Since this is a symmetric $4 \times 4$ matrix, our first guess is $\frac{1}{2} \times 4 \times 5 = 10$.

However, not all the components of the metric $g_{\mu\nu}$ are physical. Two metrics which are related by a change of coordinates, $x^\mu \rightarrow \tilde{x}^\mu(x)$ describe the same physical spacetime. This means that there is a redundancy in any given representation of the metric, which removes precisely 4 of the 10 degrees of freedom, leaving just 6 behind.

Mathematically, this redundancy is implement by diffeomorphisms. (We defined diffeomorphisms in Section 2.1.3.) Given a diffeomorphism, $\phi : M \rightarrow M$, we can use this to map all fields, including the metric, on $M$ to a new set of fields on $M$. The end result is physically indistinguishable from where we started: it describes the same system, but in different coordinates. Such diffeomorphisms are analogous to the gauge symmetries that are familiar in Maxwell and Yang-Mills theory.

Let's look more closely at the implication of these diffeomorphisms for the path integral. We'll consider a diffeomorphism that takes a point with coordinates $x^\mu$ to a nearby point with coordinates

$$x^\mu \to \tilde{x}^\mu(x) = x^\mu + \delta x^\mu$$

We could view this either as an "active change", in which one point with coordinates $x^\mu$ is mapped another point with coordinates $x^\mu + \delta x^\mu$, or as a "passive" change, in which we use two different coordinate charts to label the same point. Ultimately, the two views lead to the same place. We'll adopt the passive perspective here, simply because we have a lot of experience of changing coordinates. Later we'll revert to the active picture.

We can think of the change of coordinates as generated by an infinitesimal vector field $X$,

$$\delta x^\mu = -X^\mu(x)$$

The metric transforms as

$$g_{\mu\nu}(x) \to \tilde{g}_{\mu\nu}(\tilde{x}) = \frac{\partial x^\rho}{\partial \tilde{x}^\mu}\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma}(x)$$

With our change of coordinate $\tilde{x}^\mu = x^\mu - X^\mu(x)$, with infinitesimal $X^\mu$, we can invert the Jacobian matrix to get

$$\frac{\partial \tilde{x}^\mu}{\partial x^\rho} = \delta^\mu_\rho - \partial_\rho X^\mu \quad \Rightarrow \quad \frac{\partial x^\rho}{\partial \tilde{x}^\mu} = \delta^\rho_\mu + \partial_\mu X^\rho$$

where the inverse holds to leading order in the variation $X$. Continuing to work infinitesimally, we then have

$$\begin{aligned} \tilde{g}_{\mu\nu}(\tilde{x}) &= \left(\delta^\rho_\mu + \partial_\mu X^\rho\right)\left(\delta^\sigma_\nu + \partial_\nu X^\sigma\right) g_{\rho\sigma}(x) \\ &= g_{\mu\nu}(x) + g_{\mu\rho}(x)\partial_\nu X^\rho + g_{\nu\rho}(x)\partial_\mu X^\rho \end{aligned}$$

Meanwhile, we can Taylor expand the left-hand side

$$\tilde{g}_{\mu\nu}(\tilde{x}) = \tilde{g}_{\mu\nu}(x + \delta x) = \tilde{g}_{\mu\nu}(x) - X^\lambda \partial_\lambda g_{\mu\nu}(x)$$

Comparing the the different metrics at the same point $x$, we find that the metric undergoes the infinitesimal change

$$\delta g_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) = X^\lambda \partial_\lambda g_{\mu\nu} + g_{\mu\rho}\partial_\nu X^\rho + g_{\nu\rho}\partial_\mu X^\rho \tag{4.6}$$

But this is something we've seen before: it is the Lie derivative of the metric. In other words, if we act with an infinitesimal diffeomorphism along $X$, the metric changes as

$$\delta g_{\mu\nu} = (\mathcal{L}_X g)_{\mu\nu}$$

This makes sense: it's like the leading term in a Taylor expansion along $X$.

In fact, we can also massage (4.6) into a slightly different form. We lower the index on $X^\rho$ in the last two $\partial X^\rho$ terms by taking the metric inside the derivative. This results in two further terms in which the derivative hits the metric, and these must be cancelled off. We're left with

$$\delta g_{\mu\nu} = \partial_\mu X_\nu + \partial_\nu X_\mu + X^\rho \left(\partial_\rho g_{\mu\nu} - \partial_\mu g_{\rho\nu} - \partial_\nu g_{\mu\rho}\right)$$

But the terms in the brackets are the Christoffel symbols, $2g_{\rho\sigma}\Gamma^\sigma_{\mu\nu}$. We learn that the infinitesimal change in the metric can be written as

$$\delta g_{\mu\nu} = \nabla_\mu X_\nu + \nabla_\nu X_\mu \tag{4.7}$$

Let's now see what this means for the path integral. Under a general change of the metric, the Einstein-Hilbert action changes as (4.3)

$$\delta S = \int d^4x\, \sqrt{-g}\, G^{\mu\nu}\, \delta g_{\mu\nu}$$

where we have discarded the boundary term. Insisting that $\delta S = 0$ for *any* variation $\delta g_{\mu\nu}$ gives the equation of motion $G^{\mu\nu} = 0$. In contrast, symmetries of the action are those variations $\delta g_{\mu\nu}$ for which $\delta S = 0$ for any choice of metric. Since diffeomorphisms are (gauge) symmetries, we know that the action is invariant under changes of the form (4.7). Using the fact that $G_{\mu\nu}$ is symmetric, we must have

$$\delta S = 2\int d^4x \sqrt{-g}\, G^{\mu\nu} \nabla_\mu X_\nu = 0 \quad \text{for all } X_\mu(x)$$

After integrating by parts, we find that this results in something familiar: the Bianchi identity

$$\nabla_\mu G^{\mu\nu} = 0$$

We already know that the Bianchi identity holds from our work in Section 3.4, but the derivation there was a little fiddly. Here we learn that, from the path integral perspective, the Bianchi identity is a result of diffeomorphism invariance.

In fact it makes sense that the two are connected. Naively, the Einstein equation $G_{\mu\nu} = 0$ comprises ten independent equations. But, as we've seen, diffeomorphism invariance means that there aren't ten independent components of the metric, so one might worry that the Einstein equations are overdetermined. Happily, diffeomorphisms also ensure that not all the Einstein equations are independent either; they are related by the four Bianchi constraints. We see that, in fact, the Einstein equations give only six independent conditions on the six independent degrees of freedom in the metric.

## 4.2 Some Simple Solutions

We will now look for some simple solutions to the Einstein equations

$$R_{\mu\nu} = \Lambda g_{\mu\nu}$$

As we will see, the solutions take a very different form depending on whether $\Lambda$ is zero, positive or negative.

### Minkowski Space

Let's start with $\Lambda = 0$. Here the vacuum Einstein equations reduce to $R_{\mu\nu} = 0$. If we're looking for the simplest solution to this equation, it's tempting to suggest $g_{\mu\nu} = 0$. Needless to say, this isn't allowed! The tensor field $g_{\mu\nu}$ is a metric and, as defined in Section 3, must be non-degenerate. Indeed, the existence of the inverse $g^{\mu\nu}$ was assumed in the derivation of the Einstein equations from the action.

While this restriction is natural geometrically, it is rather unusual from the perspective of a physical theory. It is not a holonomic constraint on the physical degrees of freedom: instead it is an inequality $\det g_{\mu\nu} < 0$ (together with the requirement that $g_{\mu\nu}$ has one, rather than three, negative eigenvalues). Other fields in the Standard Model don't come with such restrictions. Instead, it is reminiscent of fluid mechanics where one has to insist that matter density obeys $\rho(\mathbf{x}, t) > 0$. Ultimately, it seems likely that this restriction is telling us that the gravitational field is not fundamental and should be replaced by something else in regimes where $\det g_{\mu\nu}$ is getting small.

The restriction that $\det g_{\mu\nu} \neq 0$ means that the simplest Ricci flat metric is Minkowski space, with

$$ds^2 = -dt^2 + d\mathbf{x}^2$$

Of course, this is far from the only metric obeying $R_{\mu\nu} = 0$. Another example is provided by the Schwarzchild metric,

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \qquad (4.8)$$

which we will discuss further in Section **??**. We will meet more solutions as the course progresses.

### 4.2.1 de Sitter Space

We now turn to the Einstein equations with $\Lambda > 0$. Once again, there are many solutions. Since it's a pain to solve the Einstein equations, let's work with an ansatz that we've already seen. Suppose that we look for solutions of the form

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{4.9}$$

We already computed the components of the Riemann tensor for such a metric in Section 3.4.3 using the technology of curvature 2-forms. From the result, given in (3.54), we can easily check that the Ricci tensor is diagonal with components

$$R_{tt} = -f^4 R_{rr} = f^3\left(f'' + \frac{2f'}{r} + \frac{f'^2}{f}\right)$$

and

$$R_{\phi\phi} = \sin^2\theta R_{\theta\theta} = \left(1 - f^2 - 2ff'r\right)\sin^2\theta$$

The resulting Ricci tensor can indeed be made to be proportional to the metric, with $R_{\mu\nu} = \Lambda g_{\mu\nu}$. Comparing to (4.9), we see that the function $f(r)$ must satisfy two constraints. The first comes from the $tt$ and $rr$ components,

$$f'' + \frac{2f'}{r} + \frac{f'^2}{f} = -\frac{\Lambda}{f} \tag{4.10}$$

The second comes from the $\theta\theta$ and $\phi\phi$ components,

$$1 - 2ff'r - f^2 = \Lambda r^2 \tag{4.11}$$

It's simple to see that both conditions are satisfied by the choice

$$f(r) = \sqrt{1 - \frac{r^2}{R^2}} \quad \text{with} \quad R^2 = \frac{3}{\Lambda}$$

The resulting metric takes the form

$$ds^2 = -\left(1 - \frac{r^2}{R^2}\right) dt^2 + \left(1 - \frac{r^2}{R^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{4.12}$$

This is *de Sitter space*. Or, more precisely, it is the *static patch* of de Sitter space; we'll see what this latter statement means shortly.

### Geodesics in de Sitter

To interpret this metric, it's useful to understand the behaviour of geodesics. We can see immediately that the presence of the non-trivial $g_{tt}(r)$ term means that a particle won't sit still at constant $r \neq 0$; instead it is pushed to smaller values of $g_{tt}(r)$, or larger values of $r$.

We can put some more flesh on this. Because the metric (4.12) has the a similar form to the Schwarzchild metric, we simply need to follow the steps that we already took in Section 1.3. First we write down the action for a particle in de Sitter space. We denote the proper time of the particle as $\sigma$. (In Section 1.3, we used $\tau$ to denote proper time, but we'll need this for a different time coordinate defined below.) Working with the more general metric (4.9), the action is

$$S_{dS} = \int d\sigma \Big[ - f(r)^2 \dot{t}^2 + f(r)^{-2} \dot{r}^2 + r^2 (\dot{\theta}^2 + \sin^2\theta\, \dot{\phi}^2) \Big] \tag{4.13}$$

where $\dot{x}^\mu = dx^\mu/d\sigma$.

Any degree of freedom which appears only with time derivatives in the Lagrangian is called *ignorable*. They lead to conserved quantities. The Lagrangian above has two ignorable degrees of freedom: $\phi(\sigma)$ and $t(\sigma)$. The first leads to the conserved quantity that we call angular momentum,

$$l = \frac{1}{2}\frac{dL}{d\dot{\phi}} = r^2 \sin^2\theta\, \dot{\phi}$$

where the factor of $1/2$ in front of $dL/d\dot{\phi}$ arises because the kinetic terms in (4.13) don't come with the usual factor of $1/2$. Meanwhile, the conserved quantity associated to $t(\sigma)$ is usually referred to as the energy

$$E = -\frac{1}{2}\frac{dL}{d\dot{t}} = f(r)^2 \dot{t} \tag{4.14}$$

The equations of motion arising from the action (4.13) should be supplemented with the constraint that tells us whether we're dealing with a massive or massless particle. For a massive particle, the constraint ensures that the trajectory is timelike,

$$-f(r)^2\dot{t}^2 + f(r)^{-2}\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2) = -1$$

Without loss of generality, we can restrict to geodesics that lie in the $\theta = \pi/2$ plane, so $\dot{\theta} = 0$ and $\sin^2\theta = 1$. Replacing $\dot{t}$ and $\dot{\phi}$ with $E$ and $l$ respectively, the constraint becomes

$$\dot{r}^2 + V_{\rm eff}(r) = E^2$$

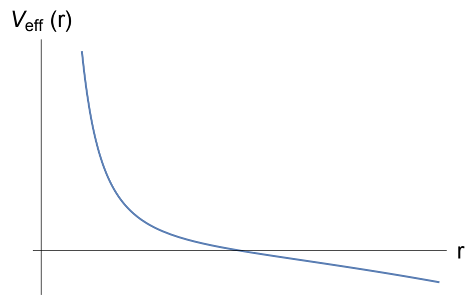



**Figure 27:** The effective potential for a massive particle in de Sitter with angular momentum..

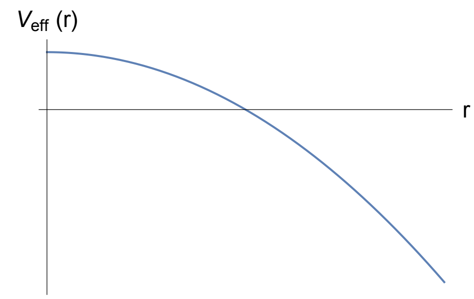



**Figure 28:** ...and with no angular momentum.

where the effective potential is given by

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right) f(r)^2$$

For geodesics in de Sitter, we therefore have

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right)\left(1 - \frac{r^2}{R^2}\right)$$

This is shown in the figures for $l \neq 0$ and for $l = 0$. We can immediately see the key physics: the potential pushes the particle out to larger values of $r$.

We focus on geodesics with vanishing angular momentum, $l = 0$. In this case, the potential is an inverted harmonic oscillator. A particle sitting stationary at $r = 0$ is a geodesic, but it is unstable: if it has some initial velocity then it will move away from the origin, following the trajectory

$$r(\sigma) = R\sqrt{E^2 - 1}\sinh\left(\frac{\sigma}{R}\right) \tag{4.15}$$

The metric (4.12) is singular at $r = R$, which might make us suspect that something fishy is going on there. But whatever this fishiness is, it's not visible in the solution (4.15) which shows that any observer reaches $r = R$ in finite proper time $\sigma$.

The fishiness reveals itself if we look at the coordinate time $t$. This also has the interpretation of the time experienced by someone sitting at the point $r = 0$. Using (4.14), the trajectory (4.15) evolves as

$$\frac{dt}{d\sigma} = E\left(1 - \frac{r^2}{R^2}\right)^{-1}$$

It is simple to check that $t(\sigma) \to \infty$ as $r(\sigma) \to R$. (For example, suppose that we have $r(\sigma) = R$ at some value $\sigma = \sigma_\star$ of proper time. Then look at what happens just before this time by expanding $\sigma = \sigma_\star - \epsilon$ with $\epsilon$ small. The equation above becomes $dt/d\epsilon = -\alpha/\epsilon$ for some constant $\alpha$, telling us that $t(\epsilon) \sim -\log(\epsilon/R)$ and we indeed find that $t \to \infty$ and $\epsilon \to 0$.) This means that while a guy on the trajectory (4.15) sails right through the point $r = R$ in finite proper time, according to his companion waiting at $r = 0$ this will take infinite time.

This strange behaviour is, it turns out, similar to what happens at the horizon of a black hole, which is the surface $r = 2GM$ in the metric (4.8). (We will look at more closely at this in Section **??**.) However, the Schwarzchild metic also has a singularity at $r = 0$, whereas the de Sitter metric looks just like flat space at $r = 0$. (To see this, simply Taylor expand the coefficients of the metric around $r = 0$.) Instead, de Sitter space seems like an inverted black hole in which particles are pushed outwards to $r = R$. But how should we interpret this radius? We will get more intuition for this as we proceed.

**de Sitter Embeddings**

We will have to wait until Section 4.4.2 to get a full understanding of the physics behind this. But we can make some progress by writing the de Sitter metric in different coordinates. In fact, it turns out that there's a rather nice way of embedding de Sitter space as a sub-manifold of $\mathbf{R}^{1,4}$, with metric

$$ds^2 = -(dX^0)^2 + \sum_{i=1}^{4}(dX^i)^2 \tag{4.16}$$

We will now show that the de Sitter space metric (4.12) is a metric on the sub-manifold in $\mathbf{R}^{1,4}$ defined by the timelike hyperboloid

$$-(X^0)^2 + \sum_{i=1}^{4}(X^i)^2 = R^2 \tag{4.17}$$

There are a number of different ways to parameterise solutions to this constraint. Suppose that we choose to treat $X^4$ as a special coordinate. We define the sum of the first three spatial coordinates to be

$$r^2 = (X^1)^2 + (X^2)^2 + (X^3)^2 \tag{4.18}$$

so the constraint (4.17) becomes

$$R^2 - r^2 = -(X^0)^2 + (X^4)^2$$

We can parameterise solutions to this equation as

$$X^0 = \sqrt{R^2 - r^2}\sinh(t/R) \quad \text{and} \quad X^4 = \sqrt{R^2 - r^2}\cosh(t/R) \tag{4.19}$$

The variation is then

$$\begin{aligned} dX^0 &= \sqrt{1 - \frac{r^2}{R^2}}\cosh(t/R)dt - \frac{r}{\sqrt{R^2 - r^2}}\sinh(t/R)dr \\ dX^4 &= \sqrt{1 - \frac{r^2}{R^2}}\sinh(t/R)dt - \frac{r}{\sqrt{R^2 - r^2}}\cosh(t/R)dr \end{aligned}$$

Meanwhile the variation of $X^i$, with $i = 1, 2, 3$, is just the familiar line element for $\mathbf{R}^3$: $\sum_{i=1}^3 (dX^i)^2 = dr^2 + r^2 d\Omega_2^2$ where $d\Omega_2^2$ is the metric on the unit 2-sphere. A two line calculation then shows that the pull-back of the 5d Minkowski metric (4.16) onto the hypersurface (4.17) gives the de Sitter metric in the static patch coordinates (4.12).

The choice of coordinates (4.18) and (4.19) are not the most intuitive. First, they single out $X^4$ as special, when the constraint (4.17) does no such thing. This hides some of the symmetry of de Sitter space. Moreover, the coordinates do not cover the whole of the hyperboloid, since they restrict only to $X^4 \geq 0$.

We can do better. Consider instead the solution to the constraint (4.17)

$$X^0 = R\sinh(\tau/R) \quad \text{and} \quad X^i = R\cosh(\tau/R)y^i \tag{4.20}$$

where the $y^i$, with $i = 1, 2, 3, 4$, obey $\sum_i (y^i)^2 = 1$ and so parameterise a unit 3-sphere. These coordinates have the advantage that they retain (more of) the symmetry of de Sitter space, and cover the whole space. Substituting this into the 5d Minkowski metric (4.16) gives a rather different metric on de Sitter space,

$$ds^2 = -d\tau^2 + R^2\cosh^2(\tau/R)\, d\Omega_3^2 \tag{4.21}$$

where $d\Omega_3^2$ denotes the metric on the unit 3-sphere. These are known as *global coordinates*, since they cover the whole space. (Admittedly, any choice of coordinates on $\mathbf{S}^3$ will suffer from the familiar problem of coordinate singularities at the poles.) Since this metric is related to (4.12) by a change of coordinates, it too must obey the Einstein equation. (We'll check this explicitly in Section 4.6 where we discuss a class of metrics of this form.)

These coordinates provide a much clearer intuition for the physics of de Sitter space: it is a time-dependent solution in which a spatial $\mathbf{S}^3$ first shrinks to a minimal radius $R$, and subsequently expands. This is shown in the figure. The expansionary phase is a fairly good approximation to our current universe on large scales; you can learn more about this in the lectures on Cosmology.

The cosmological interpretation of an expanding universe is much harder to glean from the static patch coordinates (4.12) in which the space appears to be unchanging in time. Indeed, de Sitter himself originally discovered the metric in the static patch coordinates. He noticed that light is redshifted in this metric, which then caused all sorts of confusion when trying to understand whether the redshift of galaxies (then known as the de Sitter effect!) should be viewed as evidence for an expanding universe. There is a lesson here: it can be difficult to stare at a metric and get a sense for what you're looking at.

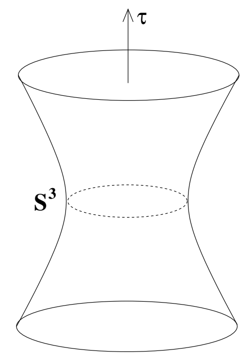



**Figure 29:**

The global coordinates clearly show that there is nothing fishy happening when $X^4 = 0$, the surface which corresponds to $r = R$ in (4.12). This is telling us that this is nothing but a coordinate singularity. (As, indeed, is the $r = 2GM$ singularity in the Schwarzchild metric.) Nonetheless, there is still some physics lurking in this coordinate singularity, which we will extract over the next few sections.

### 4.2.2 Anti-de Sitter Space

We again look for solutions to the Einstein equations,

$$R_{\mu\nu} = \Lambda g_{\mu\nu}$$

now with a negative cosmological constant $\Lambda < 0$. We can again use the ansatz (4.9) and again find the constraints (4.10) and (4.11). The fact that $\Lambda$ is now negative means that our previous version of $f(r)$ no longer works, but it's not hard to find the tweak: the resulting metric takes the form

$$ds^2 = -\left(1 + \frac{r^2}{R^2}\right) dt^2 + \left(1 + \frac{r^2}{R^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{4.22}$$

with $R^2 = -3/\Lambda$. This is the metric of *anti-de Sitter space*, also known simply as AdS.

Sometimes this metric is written by introducing the coordinate $r = R\sinh\rho$, after which it takes the form

$$ds^2 = -\cosh^2\!\rho\, dt^2 + R^2 d\rho^2 + R^2\sinh^2\rho\left(d\theta^2 + \sin^2\theta d\phi^2\right) \tag{4.23}$$

Now there's no mysterious coordinate singularity in the $r$ direction and, indeed, we will see shortly that these coordinates now cover the entire space.

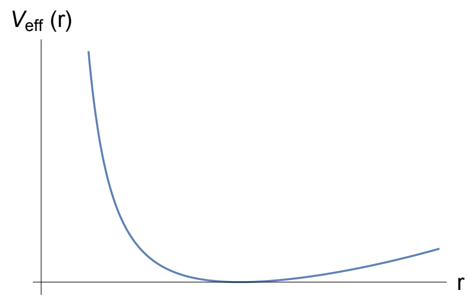



**Figure 30:** The effective potential for a massive particle in anti-de Sitter with angular momentum..

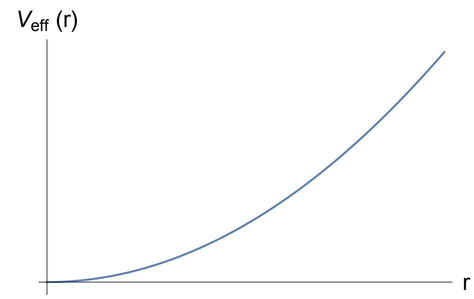



**Figure 31:** ...and with no angular momentum.

### Geodesics in Anti-de Sitter

Because the anti-de Sitter metric (4.22) falls in the general class (4.9), we can import the geodesic equations that we derived for de Sitter space. The radial trajectory of a massive particle moving in the $\theta = \pi/2$ plane is again governed by

$$\dot{r}^2 + V_{\text{eff}}(r) = E^2 \tag{4.24}$$

but this time with the effective potential

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right)\left(1 + \frac{r^2}{R^2}\right)$$

Again, $l = r^2\dot{\phi}$ is the angular momentum of the particle. This potential is shown in the figures for $l \neq 0$ and $l = 0$. From this, we can immediately see how geodesics behave. If there is no angular momentum, so $l = 0$, anti-de Sitter space acts like a harmonic potential, pushing the particle towards the origin $r = 0$. Geodesics oscillate backwards and forwards around $r = 0$.

In contrast, if the particle also has angular momentum then the potential has a minimum at $r_\star^2 = Rl$. This geodesic is like a motorcycle wall-of-death trick, with the angular momentum keeping the particle pinned up the potential. Other geodsics spin in the same fashion, while oscillating about $r_\star$. Importantly, particles with finite energy $E$ cannot escape to $r \to \infty$: they are trapped by the spacetime to live within some finite distance of the origin.

The picture that emerges from this analysis is that AdS is like a harmonic trap, pushing particles to the origin. This comes with something of a puzzle however because, as we will see below (and more in Section 4.3), AdS is a homogeneous space which, roughly speaking, means that all points are the same. How is it possible that AdS acts acts like a harmonic trap, pushing particles to $r = 0$, yet is also a homogeneous space?!

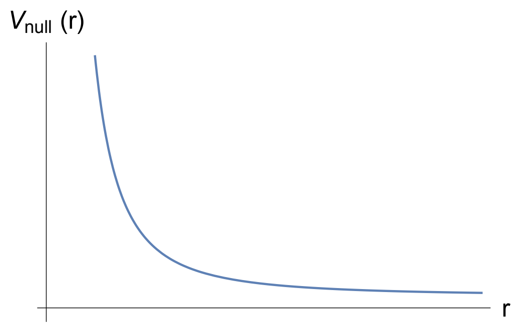



**Figure 32:** The potential experienced by massless particles in AdS.

To answer this question, consider a guy sitting stationary at the origin $r = 0$. This is a geodesic. From his perspective, intrepid AdS explorers on other geodesics (with, say, $l = 0$) will oscillate backwards and forwards about the origin $r = 0$, just like a particle in a harmonic trap. However, since these explorers will themselves be travelling on a geodesic, they are perfectly entitled to view themselves as sedentary, stay-at-home types, sitting perfectly still at their 'origin', watching the other folk fly around them. In this way, just as everyone in de Sitter can view themselves at the centre of the universe, with other observers moving away from them, everyone in anti-de Sitter can view themselves in the centre of the universe, with other observers flying around them.

We can also look at the fate of massless particles. This time the action is supplemented by the constraint

$$-f(r)^2\dot{t}^2 + f(r)^{-2}\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) = 0$$

This tells us that the particle follows a null geodesic.. The equation (4.24) gets replaced by

$$\dot{r}^2 + V_{\rm null}(r) = E^2$$

with the effective potential now given by

$$V_{\rm null}(r) = \frac{l^2}{r^2}\left(1 + \frac{r^2}{R^2}\right)$$

This potential is again shown in Figure 32. This time the potential is finite as $r \to \infty$, which tells us that there is no obstacle to light travelling as far as it likes: it suffers only the usual gravitational redshift. We learn that AdS spacetime confines massive particles, but not massless ones.

To solve the equations for a massless particle, it's simplest to work in the coordinates $r = R\sinh\rho$ that we introduced in (4.23). If we restrict to vanishing angular momentum, $l = 0$, the equation above becomes

$$R\dot{\rho} = \pm\frac{E}{\cosh\rho} \quad \Rightarrow \quad R\sinh\rho = E(\sigma - \sigma_0)$$

where $\sigma$ is the affine geodesic parameter. We see that $\rho \to \infty$ only in infinite affine time, $\sigma \to \infty$. However, it's more interesting to see what happens in coordinate time. This follows by recalling the definition of $E$ in (4.14),

$$E = \cosh^2\!\rho\,\dot{t}$$

(Equivalently, you can see this by dint of the fact that we have a null geodesic, with $\cosh\!\rho\,\dot{t} = \pm R\dot{\rho}$.) We then find

$$R\tan(t/R) = E(\sigma - \sigma_0)$$

So as $\sigma \to \infty$, the coordinate time tends to $t \to \pi R/2$. We learn that not only do light rays escape to $\rho = \infty$, but they do so in a finite coordinate time $t$. This means that to make sense of dynamics in AdS, we must specify some boundary conditions at infinity to dictate what happens to massless particles or fields when they reach it.

Anti-de Sitter space does not appear to have any cosmological applications. However, it turns out to be the place where we best understand quantum gravity, and so has been the object of a great deal of study.

**Anti-de Sitter Embeddings**

Like its $\Lambda > 0$ cousin, anti-de Sitter spacetime also has a natural embedding in a 5d spacetime. This time, it sits within $\mathbf{R}^{2,3}$, with metric

$$ds^2 = -(dX^0)^2 - (dX^4)^2 + \sum_{i=1}^{3}(dX^i)^2 \tag{4.25}$$

where it lives as the hyperboloid,

$$-(X^0)^2 - (X^4)^2 + \sum_{i=1}^{3}(X^i)^2 = -R^2 \tag{4.26}$$

We can solve this constraint by

$$X^0 = R\cosh\rho\,\sin(t/R) \quad , \quad X^4 = R\cosh\rho\,\cos(t/R) \quad , \quad X^i = Ry^i\sinh\rho \tag{4.27}$$

where the $y^i$, with $i = 1, 2, 3$, obey $\sum_i (y^i)^2 = 1$ and so parameterise a unit 2-sphere. Substituting this into the metric (4.25) gives the anti-de Sitter metric in the coordinates (4.23).

In fact there is one small subtlety: the embedding hyperboloid has topology $\mathbf{S}^1 \times \mathbf{R}^3$, with $\mathbf{S}^1$ corresponding to a compact time direction. This can be seen in the parameterisation (4.27), where the time coordinate takes values $t \in [0, 2\pi R)$. However, the AdS metrics (4.22) or (4.23) have no such restriction, with $t \in (-\infty, +\infty)$. They are the universal covering of the hyperboloid (4.26).

There is another parameterisation of the hyperboloid that is also useful. It takes the rather convoluted form

$$X^i = \frac{r}{R} x^i \text{ for } i = 0, 1, 2 \quad , \quad X^4 - X^3 = r \quad , \quad X^4 + X^3 = \frac{R^2}{r} + \frac{r}{R^2} \eta_{ij} x^i x^j$$

with $r \in [0, \infty)$. Although the change of coordinates is tricky, the metric is very straightforward, taking the form

$$ds^2 = R^2 \frac{dr^2}{r^2} + \frac{r^2}{R^2} \eta_{ij} dx^i dx^j \tag{4.28}$$

These coordinates don't cover the whole of AdS; instead they cover only one-half of the hyperboloid, restricted to $X^4 - X^3 > 0$. This is known as the *Poincaré patch* of AdS. Moreover, the time coordinate, which already extends over the full range $x^0 \in (-\infty, +\infty)$, cannot be further extended. This means that as $x^0$ goes from $-\infty$ to $+\infty$ in (4.28), in global coordinates (4.22), the time coordinate $t$ goes only from 0 to $2\pi R$.

Two other choices of coordinates are also commonly used to describe the Poincaré patch. If we set $z = R^2/r$, then we have

$$ds^2 = \frac{R^2}{z^2} \left( dz^2 + \eta_{ij} dx^i dx^j \right)$$

Alternatively, if we set $r = Re^\rho$, we have

$$ds^2 = R^2 d\rho^2 + e^{2\rho} \eta_{ij} dx^i dx^j$$

In each case, massive particles fall towards $r = 0$, or $z = \infty$, or $\rho = -\infty$.

## 4.3 Symmetries

We introduced the three spacetimes – Minkowski, de Sitter and anti-de Sitter – as simple examples of solutions to the Einstein equations. In fact, what makes them special are their symmetries.

The symmetries of Minkowski space are very familiar: they consist of translations and rotations in space and time, the latter splitting into genuine rotations and Lorentz boosts. It's hard to overstate the importance of these symmetries: on a fixed Minkowski background they are responsible for the existence of energy, momentum and angular momentum.

The purpose of this section is to find a way to characterise the symmetries of a general metric.

### 4.3.1 Isometries

Intuitively, the notion of symmetry is clear. If you hold up a round sphere, it looks the same no matter what way you rotate it. In contrast, if the sphere has dimples and bumps, then the rotational symmetry is broken. The distinction between these two should be captured by the metric. Roughly speaking, the metric on a round sphere looks the same at all points, while the metric on a dimpled sphere will depend on where you sit. We want a way to state this mathematically.

To do this, we need the concept of a *flow* that we introduced in Section 2.2.3. Recall that a flow on a manifold $M$ is a one-parameter family of diffeomorphisms $\sigma_t : M \to M$. A flow can be identified with a vector field $K \in \mathfrak{X}(M)$ which, at each point in $M$, points along tangent vectors to the flow

$$K^\mu = \frac{dx^\mu}{dt}$$

This flow is said to be an *isometry*, if the metric looks the same at each point along a given flow line. Mathematically, this means that an isometry satisfies

$$\mathcal{L}_K g = 0 \quad \Leftrightarrow \quad \nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \tag{4.29}$$

where the equivalence of the two expressions was shown in Section 4.1.3. This is the *Killing equation* and any $K$ satisfying this equation is known as a *Killing vector field*.

Sometimes it is possible to stare at a metric and immediately write down a Killing vector. Suppose that the metric components $g_{\mu\nu}(x)$ do not depend on one particular coordinate, say $y \equiv x^1$. Then the vector field $X = \partial/\partial y$ is a Killing vector, since

$$(\mathcal{L}_{\partial_y} g)_{\mu\nu} = \frac{\partial g_{\mu\nu}}{\partial y} = 0$$

However, we have met coordinates like $y$ before: they become the ignorable coordinates in the Lagrangian for a particle moving in the metric $g_{\mu\nu}$, resulting in conserved quantities. We once again see the familiar link between symmetries and conserved quantities. We'll explore this more in Section 4.3.2 and again later in the lectures.

There is a group structure underlying these symmetries. Or, more precisely, a Lie algebra structure. This follows from the result (2.15)

$$\mathcal{L}_X\mathcal{L}_Y - \mathcal{L}_Y\mathcal{L}_X = \mathcal{L}_{[X,Y]}$$

(Strictly speaking, we showed this in (2.15) only for Lie derivatives acting on vector fields, but it can be checked that it holds on arbitrary tensor fields.) This means that Killing vectors too form a Lie algebra. This is the Lie algebra of the isometry group of the manifold.

**An Example: Minkowski Space**

As a particularly simple example, consider Minkowski spacetime with $g_{\mu\nu} = \eta_{\mu\nu}$. The Killing equation is

$$\partial_\mu K_\nu + \partial_\nu K_\mu = 0$$

There are two forms of solutions. We can take

$$K_\mu = c_\mu$$

for any constant vector $c_\mu$. These generate translations in Minkowski space. Alternatively, we can take

$$K_\mu = \omega_{\mu\nu}x^\nu$$

with $\omega_{\mu\nu} = -\omega_{\nu\mu}$. These generate rotations and Lorentz boosts in Minkowski space.

We can see the emergence of the algebra structure more clearly. We define Killing vectors

$$P_\mu = \partial_\mu \quad \text{and} \quad M_{\mu\nu} = \eta_{\mu\rho}x^\rho\partial_\nu - \eta_{\nu\rho}x^\rho\partial_\mu \tag{4.30}$$

There are 10 such Killing vectors in total; 4 from translations and six from rotations and boosts. A short calculation shows that they obey

$$[P_\mu, P_\nu] = 0 \quad , \quad [M_{\mu\nu}, P_\sigma] = -\eta_{\mu\sigma}P_\nu + \eta_{\sigma\nu}P_\nu$$
$$[M_{\mu\nu}, M_{\rho\sigma}] = \eta_{\mu\sigma}M_{\nu\rho} + \eta_{\nu\rho}M_{\mu\sigma} - \eta_{\mu\rho}M_{\nu\sigma} - \eta_{\nu\sigma}M_{\mu\rho}$$

which we recognise as the Lie algebra of the Poincaré group $\mathbf{R}^4 \times SO(1,3)$.

#### More Examples: de Sitter and anti-de Sitter

The isometries of de Sitter and anti-de Sitter are simplest to see from their embeddings. The constraint (4.17) that defines de Sitter space is invariant under the rotations of $\mathbf{R}^{1,4}$, and so de Sitter inherits the $SO(1,4)$ isometry group. Similarly, the constraint (4.26) that defines anti-de Sitter is invariant under the rotations of $\mathbf{R}^{2,3}$. Correspondingly, AdS has the isometry group $SO(2,3)$. Note that both of these groups are 10 dimensional: in terms of counting, these spaces are just as symmetric as Minkowski space.

It is simple to write down the 10 Killing spinors in the parent 5d spacetime: they are

$$M_{AB} = \eta_{AC} X^C \partial_B - \eta_{BC} X^C \partial_A$$

where $X^A$, $A = 0,1,2,3,4$ are coordinates in 5d and $\eta_{AB}$ is the appropriate Minkowski metric, with signature $(-++++)$ for de Sitter and $(--+++)$ for anti-de Sitter. In either case, the Lie algebra is that of the appropriate Lorentz group,

$$[M_{AB}, M_{CD}] = \eta_{AD} M_{BC} + \eta_{BC} M_{AD} - \eta_{AC} M_{BD} - \eta_{BD} M_{AC}$$

Importantly, the embedding hyperbolae (4.17) and (4.26) are both invariant under these Killing vectors, in the sense that the flows generated by $M_{AB}$ keep take us from one point on the hyperbolae to another. This means that the Killing vectors are inherited by de Sitter and anti-de Sitter spaces respectively.

For example, we can consider de Sitter space in the static patch with $r^2 = (X^1)^2 + (X^2)^2 + (X^3)^2$ and (4.19)

$$X^0 = \sqrt{R^2 - r^2}\sinh(t/R) \quad \text{and} \quad X^4 = \sqrt{R^2 - r^2}\cosh(t/R)$$

We know that the metric in the static patch (4.12) is independent of time. This means that $K = \partial_t$ is Killing vector. Pushed forwards to the 5d space, this becomes

$$\frac{\partial}{\partial t} = \frac{\partial X^A}{\partial t}\frac{\partial}{\partial X^A} = \frac{1}{R}\left(X^4 \frac{\partial}{\partial X^0} + X^0 \frac{\partial}{\partial X^4}\right) \tag{4.31}$$

In fact, this Killing vector highlights a rather important subtlety with de Sitter space. As we go on, we will see that timelike Killing vectors – those obeying $g_{\mu\nu}K^\mu K^\nu < 0$ everywhere – play a rather special role because we can use them to define energy. (We'll describe this for particles in the next section.)

In anti-de Sitter space, there is no problem in finding a timelike Killing vector. Indeed, we can see it by eye in the global coordinates (4.22), where it is simply $K = \partial_t$. But de Sitter is another story.

If we work in the static patch (4.12), then the Killling vector (4.31) is a timelike Killing vector. Indeed, we used this to derive the conserved energy $E$ when discussing geodesics in de Sitter in Section 4.2.1. But we know that the static patch does not cover all of de Sitter spacetime.

Indeed, we extend the Killing vector (4.31) over the entire space, it is not timelike everywhere. To see this, note that when $X^4 > 0$ and $X^0 = 0$, the Killing vector pushes us forwards in the $X^0$ direction, but when $X_4 < 0$ and $X^0 = 0$ it pushes us backwards in the $X^0$ direction. This means that the Killing vector field points to the future in some parts of space and to the past in others! Correspondingly, if we try to define an energy using this Killing vector it will be positive in some parts of the space and negative in others. Relatedly, in parts of the space where $X^4 = 0$ and $X^0 \neq 0$, the Killing vector pushes us in the $X^4$ direction, and so is spacelike.

The upshot of this discussion is an important feature of de Sitter space: there is no global, positive conserved energy. This tallies with our metric in global coordinates (4.21) which is time dependent and so does not obviously have a timelike Killing vector. The lack of a globally defined energy is one of several puzzling aspects of de Sitter space: we'll meet more as we proceed.

### 4.3.2 A First Look at Conserved Quantities

Emmy Noether taught us that symmetries are closely related to conserved quantities. In the present context, this means that any dynamics taking place in a spacetime with an isometry will have a conserved quantity.

Suppose, for example, that there is a massive particle moving in a spacetime with metric $g$. The particle will follow some trajectory $x^\mu(\tau)$, with $\tau$ the proper time. If the spacetime admits a Killing vector $K$, then we can construct the quantity that is conserved along the geodesic,

$$Q = K_\mu \frac{dx^\mu}{d\tau} \tag{4.32}$$

To see that $Q$ is indeed unchanging, we compute

$$\begin{aligned}
\frac{dQ}{d\tau} &= \partial_\nu K_\mu \frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau} + K_\mu \frac{d^2x^\mu}{d\tau^2} \\
&= \partial_\nu K_\mu \frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau} - K_\mu \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} \\
&= \nabla_\nu K_\mu \frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau} = 0
\end{aligned}$$

where in the second line we've used the geodesic equation and, in the final equality, we've used the symmetry of the Killing equation.

The derivation above looks rather different from our usual formulation of Noether's theorem. For this reason, it's useful re-derive the Killing equation and corresponding conserved charge by playing the usual Noether games. We can do this by looking at the action for a massive particle (in the form (1.32))

$$S = \int d\tau \; g_{\mu\nu}(x)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}$$

Now we can play the usual Noether games. Consider the infinitesimal transformation

$$\delta x^\mu = K^\mu(x)$$

The action transforms as

$$\begin{aligned}
\delta S &= \int d\tau \; \partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}K^\rho + 2g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dK^\nu}{d\tau} \\
&= \int d\tau \; \partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}K^\rho + 2\frac{dx^\mu}{d\tau}\left(\frac{dK_\mu}{d\tau} - K^\nu\frac{dg_{\mu\nu}}{d\tau}\right) \\
&= \int d\tau \; \left(\partial_\rho g_{\mu\nu}K^\rho - 2K^\rho\partial_\nu g_{\mu\rho} + 2\partial_\nu K_\mu\right)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \\
&= \int d\tau \; 2\nabla_\nu K_\mu \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}
\end{aligned}$$

The transformation is a symmetry of the action if $\delta S = 0$. From the symmetry of the $\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}$ terms, this is true provided that $K_\mu$ obeys the Killing equation

$$\nabla_{(\nu}K_{\mu)} = 0$$

Noether's theorem then identifies the charge $Q$ defined in (4.32) as the conserved quantity arising from this symmetry.

We met examples of these conserved quantities in Section 4.2 when discussing geodesics in de Sitter and anti-de Sitter spacetimes. (And, indeed, in Section 1.3 when discussing the geodesic orbits around a black hole). Both the energy $E$ and the angular momentum $l$ are Noether charges of this form.

Killing vectors have further roles to play in identifying conserved quantities. In Section 4.5.5, we'll describe how we can use Killing vectors to define energy and momentum of fields in a background spacetime. Later, in Section **??**, we will see how Killing vectors can be used to identify conserved quantities, known as *Komar integrals*, associated to spacetime itself.

## 4.4 Asymptotics of Spacetime

The three solutions – Minkowski, de Sitter, and anti-de Sitter – have different spacetime curvature and differ in their symmetries. But there is a more fundamental distinction: they have different behaviour at infinity.

This is important because we will ultimately want to look at more complicated solutions. These may have reduced symmetries, or no symmetries at all. But, providing fields are suitably localised, they will asymptote to one of the three symmetric spaces described above. This gives us a way to characterise whether physics is happening "in Minkowski spacetime", "in de Sitter", or "in anti-de Sitter".

It turns out that "inifinity" in Lorentzian spacetimes is more interesting than you might have thought. One can go to infinity along spatial, timelike or null directions, and each of these may have a different structure. It will be useful to introduce a tool to visualise infinity of spacetime.

### 4.4.1 Conformal Transformations

Given a spacetime $M$ with metric $g_{\mu\nu}$, we may construct a new metric $\tilde{g}_{\mu\nu}$ by a *conformal transformation*,

$$\tilde{g}_{\mu\nu}(x) = \Omega^2(x) g_{\mu\nu}(x) \tag{4.33}$$

with $\Omega(x)$ a smooth, non-vanishing function.

Typically $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ describe very different spacetimes, with distances in the two considerably warped. However, the conformal transformation preserves angles. In particular, in a Lorentzian spacetime, this means that two metrics related by a conformal transformation have the same causal structure. A vector field $X$ which is everywhere-null with respect to the metric $g_{\mu\nu}$ will also be everywhere-null with respect to $\tilde{g}_{\mu\nu}$,

$$g_{\mu\nu} X^\mu X^\nu = 0 \quad \Leftrightarrow \quad \tilde{g}_{\mu\nu} X^\mu X^\nu = 0$$

Similarly, vectors that are timelike/spacelike with respect to $g_{\mu\nu}$ will continue to be timelike/spacelike separated with respect to $\tilde{g}_{\mu\nu}$.

A conformal transformation of the metric does not change the causal structure. However, any other change of the metric does. This fact is sometimes summarised in the slogan "the causal structure is $9/10^{\text{th}}$ of the metric". Although, taking into account diffeomorphism invariance, a better slogan would be "the causal structure is around $5/6^{\text{th}}$ of the metric".

#### Conformal Transformations and Geodesics

A particle trajectory which is timelike with respect to $g_{\mu\nu}$ will necessarily also be timelike with respect to $\tilde{g}_{\mu\nu}$. But because distances get screwed up under a conformal transformation, there is no reason to expect that a timelike geodesic will map to a timelike geodesic. However, it turns out that null geodesics do map to null geodesics, although the affine parameterisation gets messed up along the way.

To see this, we first compute the Christoffel symbols in the new metric. They are

$$\begin{aligned}\Gamma^\mu_{\rho\sigma}[\tilde{g}] &= \frac{1}{2}\tilde{g}^{\mu\nu}\left(\partial_\rho\tilde{g}_{\nu\sigma} + \partial_\sigma\tilde{g}_{\rho\nu} - \partial_\nu\tilde{g}_{\rho\sigma}\right) \\ &= \frac{1}{2}\Omega^{-2}g^{\mu\nu}\left(\partial_\rho(\Omega^2 g_{\nu\sigma}) + \partial_\sigma(\Omega^2 g_{\rho\nu}) - \partial_\nu(\Omega^2 g_{\rho\sigma})\right) \\ &= \Gamma^\mu_{\rho\sigma}[g] + \Omega^{-1}\left(\delta^\mu_\sigma\nabla_\rho\Omega + \delta^\mu_\rho\nabla_\sigma\Omega - g_{\rho\sigma}\nabla^\mu\Omega\right)\end{aligned}$$

where, in the final line, we've replaced $\partial$ with $\nabla$ on the grounds that the derivatives are hitting a scalar function $\Omega(x)$ so it makes no difference.

If we have an affinely parameterised geodesic in the metric $g$

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}[g]\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0$$

then in the metric $\tilde{g}$ we have

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}[\tilde{g}]\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = \Omega^{-1}\left(\delta^\mu_\sigma\nabla_\rho\Omega + \delta^\mu_\rho\nabla_\sigma\Omega - g_{\rho\sigma}\nabla^\nu\Omega\right)\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau}$$

The right-hand side looks like a mess. And for timelike or spacelike geodesics, it is. But for null geodesics we have

$$g_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0$$

so at least one term on the right-hand side vanishes. The others can be written as

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}[\tilde{g}]\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 2\frac{dx^\mu}{d\tau}\frac{1}{\Omega}\frac{d\Omega}{d\tau}$$

But this is the equation for a geodesic that is not affinely parameterised, as in (1.28). So a conformal transformation does map null geodesics to null geodesics as claimed.

### The Weyl Tensor

Our favourite curvature tensors are not invariant under conformal transformations. However, it turns out that there is a combination of curvature tensors does not change under conformal transformations. This is the *Weyl tensor*. In a manifold of dimension $n$, it is defined as

$$C_{\mu\nu\rho\sigma} = R_{\mu\nu\rho\sigma} - \frac{2}{n-2}\left(g_{\mu[\rho}R_{\sigma]\nu} - g_{\nu[\rho}R_{\sigma]\mu}\right) + \frac{2}{(n-1)(n-2)} R g_{\mu[\rho}g_{\sigma]\nu}$$

The Weyl tensor has all the symmetry properties of the Riemann tensor, but with the additional property that if you contract any pair of indices with a metric then it vanishes. In this sense, it can be viewed as the "trace-free" part of the Riemann tensor.

## 4.4.2 Penrose Diagrams

There are a number of interesting and deep stories associated to conformal transformations (4.33). For example, there are a class of theories that are invariant under conformal transformations of Minkowski space; these so-called conformal field theories describe physics at a second order phase transition. But here we want to use conformal transformations to understand what happens at infinity of spacetime.

The main idea is to perform a conformal transformation that pulls infinity to some more manageable, finite distance. Obviously this transformation will mangle distances, but it will retain the causal structure of the original spacetime. We can then draw this causal structure on a very finite piece of paper (e.g. A4). The resulting picture is called a *Penrose diagram*, named after its discoverers, Roger Penrose and Brandon Carter. We will illustrate this with a series of examples.

### Minkowski Space

We start with Minkowski space. It turns out that, even here, infinity is rather subtle.

It will be simplest if we first work in $d = 1 + 1$ dimensions, where the Minkowski metric takes the form

$$ds^2 = -dt^2 + dx^2 \tag{4.34}$$

The first thing we do is introduce light-cone coordinates,

$$u = t - x \quad \text{and} \quad v = t + x$$

In these coordinates, the Minkowski metric is even simpler

$$ds^2 = -du\, dv$$

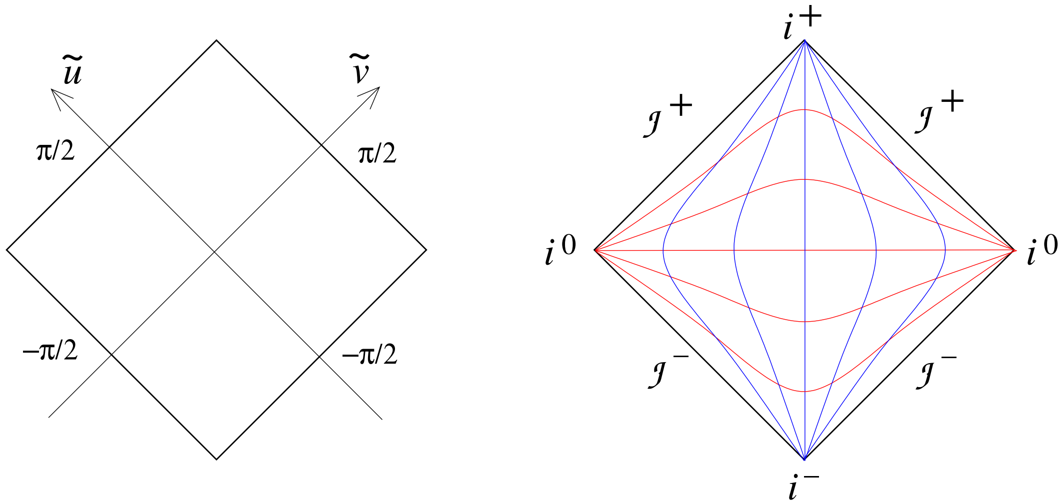



**Figure 33:** The Penrose diagram for $d = 1 + 1$ Minkowski space.

Both of these light-cone coordinates take values over the full range of $\mathbf{R}$: $u, v \in (-\infty, \infty)$. In an attempt to make things more finite, we will introduce another coordinate that traverses the full range of $u$ and $v$ over a finite interval. A convenient choice is

$$u = \tan\tilde{u} \quad \text{and} \quad v = \tan\tilde{v} \tag{4.35}$$

where we now cover the whole of Minkowski space as $\tilde{u}, \tilde{v} \in (-\pi/2, +\pi/2)$. Note that, strictly speaking, we shouldn't include the points $\tilde{u}, \tilde{v} = \pm\pi/2$ since these correspond to $u, v = \pm\infty$.

In the new coordinates, the metric takes the form

$$ds^2 = -\frac{1}{\cos^2\tilde{u}\,\cos^2\tilde{v}}d\tilde{u}\,d\tilde{v}$$

Notice that the metric diverges as we approach the boundary of Minkowski space, where $\tilde{u}$ or $\tilde{v} \to \pm\pi/2$. However, we can now do our conformal transformation. We define the new metric

$$d\tilde{s}^2 = (\cos^2\tilde{u}\,\cos^2\tilde{v})\,ds^2 = -d\tilde{u}\,d\tilde{v}$$

After the conformal map, nothing bad happens as we approach $\tilde{u}, \tilde{v} \to \pm\pi/2$. It is customary to now add in the "points at infinity", $\tilde{u} = \pm\pi/2$ and $\tilde{v} = \pm\pi/2$, an operation that goes by the name of *conformal compactification*.

The Penrose diagram is a pictorial representation of this space. As in other relativistic diagrams, we insist that light-rays go at 45°. We take time to be in the vertical direction, and space in the horizontal. This means that we draw the lightcone $\tilde{u}$ and $\tilde{v}$ coordinates at 45°. The resulting diagram is shown with the $\tilde{u}$ and $\tilde{v}$ axes on the left-hand side of Figure 33.

We can also dress our Penrose diagram in various ways. For example, we could draw geodesics with respect to the original metric (4.34). These are shown in the right-hand side of Figure 33; the verticalish blue lines are timelike geodesics of constant $x$; the horizontalish red lines are spacelike geodesics of constant $t$. We have also listed the different kinds of "infinity" for Minkowski space. They are

- All timelike geodesics start at the point labelled $i^-$, with $(\tilde{u}, \tilde{v}) = (-\pi/2, -\pi/2)$ and end at the point labelled $i^+$ with $(\tilde{u}, \tilde{v}) = (+\pi/2, +\pi/2)$. In other words, this is the origin and fate of all massive particles. These points are referred to as *past* and *future timelike infinity* respectively.

- All spacelike geodesics begin or end at one of the two points labelled $i^0$, either $(\tilde{u}, \tilde{v}) = (-\pi/2, +\pi/2)$ or $(\tilde{u}, \tilde{v}) = (+\pi/2, -\pi/2)$. These points are *spacelike infinity*.

- All null curves start on the boundary labelled $\mathcal{I}^-$, with $\tilde{u} = -\pi/2$ and arbitrary $\tilde{v}$, or $\tilde{v} = -\pi/2$ and arbitrary $\tilde{u}$. This boundary is pronounced "scri-minus" and known, more formally, as *past null infinity*. Such null curves end on the boundary labelled $\mathcal{I}^+$, with $\tilde{u} = +\pi/2$ and arbitrary $\tilde{v}$, or $\tilde{v} = +\pi/2$ and arbitrary $\tilde{u}$. This is pronounced "scri-plus" and known as *future null infinity*.

  We see from the picture that there are more ways to "go to infinity" in a null direction than in a timelike or spacelike direction. This is one of the characteristic features of Minkowski space.

The Penrose diagram allows us to immediately visualise the causal structure of Minkowski space. For example, as timelike curves approach $i^+$, their past lightcone encompasses more and more of the spacetime, as shown in the left-hand side of Figure 34. This means that an observer in Minkowski space can see everything (in principle) as long as they wait long enough. Relatedly, give any two points in Minkowski space, they are causally connected in both the past and future, meaning that their past and future lightcones necessarily intersect, as shown in the Mondrian painting on the right-hand side of Figure 34. This means that there was always an event in the past that could influence both points, and always an event in the future that can be influenced

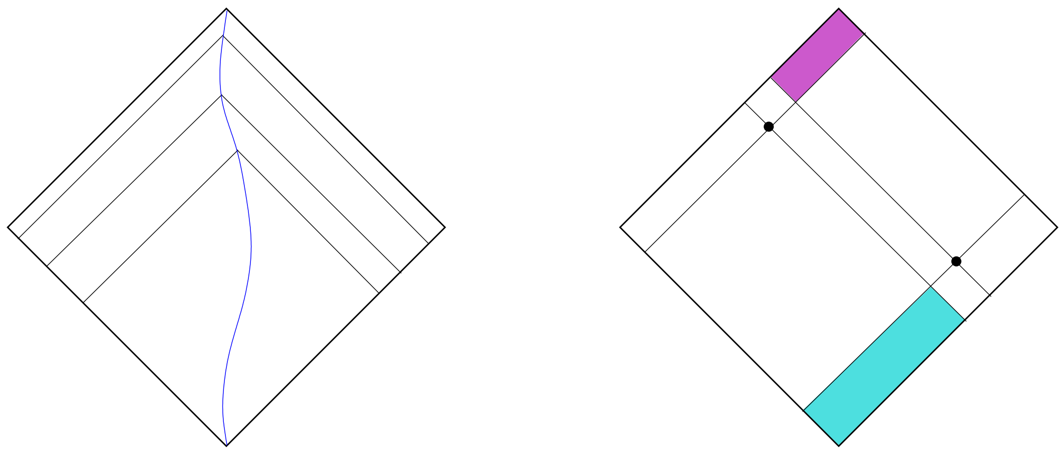


**Figure 34:** On the left: you will eventually see everything. On the right: every two points share some of their causal past and some of their causal future.

by both. These comments may seem trivial, but we will soon see that they don't hold in other spacetimes, including the one we call home.

Let's now repeat the analysis for Minkowski space in $d = 3+1$ dimensions, with the metric

$$ds^2 = -dt^2 + dr^2 + r^2\, d\Omega_2^2$$

where $d\Omega_2^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the round metric on $\mathbf{S}^2$ (and is not to be confused with the conformal factor $\Omega(x)$ that we introduced earlier). Again we introduce lightcone coordinates

$$u = t - r \quad \text{and} \quad v = t + r$$

and write this metric as

$$ds^2 = -du\, dv + \frac{1}{4}(u-v)^2 d\Omega_2^2$$

In the finite-range coordinates (4.35), the metric becomes

$$ds^2 = \frac{1}{4\cos^2\tilde{u}\,\cos^2\tilde{v}}\left(-4d\tilde{u}\, d\tilde{v} + \sin^2(\tilde{u}-\tilde{v})d\Omega_2^2\right)$$

Finally, we do the conformal transformation to the new metric

$$d\tilde{s}^2 = -4d\tilde{u}\, d\tilde{v} + \sin^2(\tilde{u}-\tilde{v})d\Omega_2^2 \tag{4.36}$$

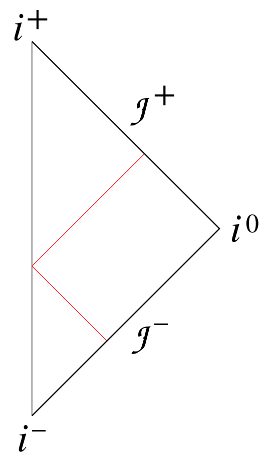



**Figure 35:** The Penrose diagram for 4d Minkowski spacetime.

There is an additional requirement that didn't arise in 2d: we must insist that $v \geq u$ so that $r \geq 0$, as befits a radial coordinate. This means that, after a conformal compactification, $\tilde{u}$ and $\tilde{v}$ take values in

$$-\frac{\pi}{2} \leq \tilde{u} \leq \tilde{v} \leq \frac{\pi}{2}$$

To draw a diagram corresponding to the spacetime (4.36), we're going to have to ditch some dimensions. We chose not to depict the $\mathbf{S}^2$, and only show the $\tilde{u}$ and $\tilde{v}$ directions. The resulting Penrose diagram for $d = 3 + 1$ dimensional Minkowski space is shown in Figure 35.

Every point in the diagram corresponds to an $\mathbf{S}^2$ of radius $\sin(\tilde{u} - \tilde{v})$, except for the left-hand line which sits at $\tilde{u} = \tilde{v}$ where this $\mathbf{S}^2$ shrinks to a point. This is not a boundary of Minkowski space; it is simply the origin $r = 0$. To illustrate this, we've drawn a null geodesic in red in the figure; it starts at $\mathcal{I}^-$ and and when it hits the $r = 0$ vertical line, it simply bounces off and ends up at $\mathcal{I}^+$.

The need to draw a 4d space on a 2d piece of paper is something of a limitation of Penrose diagrams. It means that they're only really useful for spacetimes that have an obvious $\mathbf{S}^2$ sitting inside them that we can drop. Or, to state it more precisely, spacetimes that have an $SO(3)$ isometry. But these spacetimes are the simplest and tend to be the most important.

We have seen that Minkowski space has a null boundary, together with a couple of points at spatial and temporal infinity. This naturally lends itself to asking questions about scattering of massless fields: we set up some initial data on $\mathcal{I}^-$, let it evolve, and read off its fate at $\mathcal{I}^+$. In quantum field theory, this is closely related to the object we call the S-matrix.

**de Sitter Space**

The global coordinates for de Sitter space are (4.21),

$$ds^2 = -d\tau^2 + R^2 \cosh^2(\tau/R)\, d\Omega_3^2$$

To construct the Penrose diagram we work with *conformal time*, defined by

$$\frac{d\eta}{d\tau} = \frac{1}{R\cosh(\tau/R)}$$

The solution is

$$\cos\eta = \frac{1}{\cosh(\tau/R)} \tag{4.37}$$

with $\eta \in (-\pi/2, +\pi/2)$ as $\tau \in (-\infty, +\infty)$. In conformal time, de Sitter space has the metric

$$ds^2 = \frac{R^2}{\cos^2\eta}\left(-d\eta^2 + d\Omega_3^2\right)$$

We write the metric on the $\mathbf{S}^3$ as

$$d\Omega_3^2 = d\chi^2 + \sin^2\chi d\Omega_2^2 \tag{4.38}$$

with $\chi \in [0, \pi]$. The de Sitter metric is conformally equivalent to

$$d\tilde{s}^2 = -d\eta^2 + d\chi^2 + \sin^2\chi d\Omega_2^2$$

After a conformal compactification, $\eta \in [-\pi/2, +\pi/2]$ and $\chi \in [0, \pi]$. The Penrose diagram is shown in Figure 36.

The two vertical lines are not boundaries of the spacetime; they are simply the north and south poles of the $\mathbf{S}^3$. The boundaries are the horizontal lines at the top and bottom: they are labelled both as $i^\pm$ and $\mathcal{I}^\pm$, reflecting the fact that they are where both timelike and null geodesics originate and terminate.

We learn that de Sitter spacetime has a *spacelike* $\mathbf{S}^3$ boundary. (The normal to this boundary is timelike.)

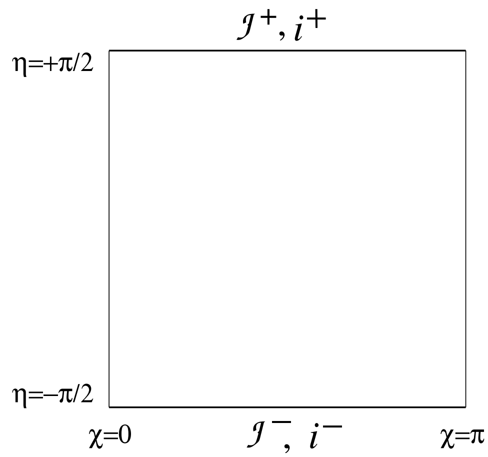



**Figure 36:** The Penrose diagram for de Sitter.

The causal structure of de Sitter spacetime is very different from Minkowski. It is not true that if an observer waits long enough then she will be able to see everything that's happening. For example, an observer who sits at the north pole (the left-hand side of the figure) will ultimately be able to see exactly half the spacetime, as shown in the left-hand side of Figure 37. The boundary of this space (the diagonal line in the figure) is her *event horizon*. It is similar to the event horizon of a black hole in the sense that signals from beyond the horizon cannot reach her. However, as is clear from the picture, it is an observer-dependent horizon: someone else will have an entirely different event horizon. In this context, these are sometimes referred to as *cosmological horizons*.

Furthermore, the observer at the north pole will only be able to communicate with another half of the spacetime, as shown in the middle of Figure 37. The boundary of the region of influence is known as the *particle horizon*. You should think of it as the furthest distance light can travel since the beginning of time. The intersection of these two regimes is called the (northern) *causal diamond* and is shown as the red triangle in the right-hand figure. An observer sitting at the southern pole also has a causal diamond, shown in blue in the right-hand side of Figure 37. It is causally disconnected from the northern diamond.

This state of affairs was nicely summarised by Schrödinger who, in 1956, wrote

> "It does seem rather odd that two or more observers, even such as sat on the same school bench in the remote past, should in future, when they have followed different paths in life, experience different worlds, so that

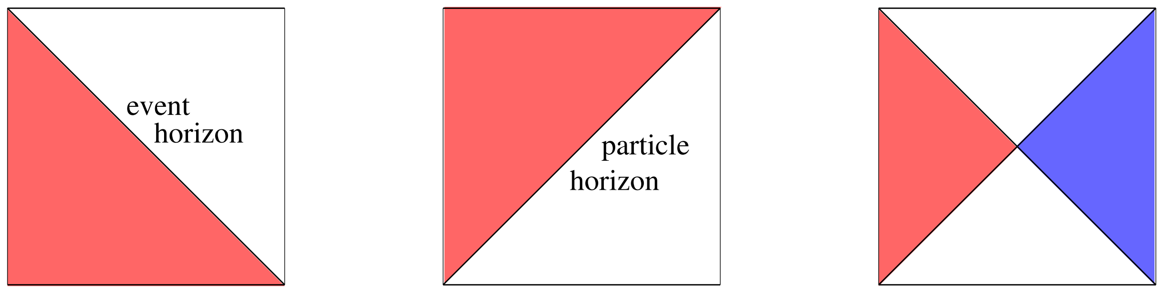



**Figure 37:** On the left: an observer at the north pole does not see everything. She has an event horizon. In the middle: Nor can she influence everything: she has a particle horizon. On the right: the causal diamond for an observer at the north pole (in red) and at the south pole (in blue).

> eventually certain parts of the experienced world of one of them should remain by principle inaccessible to the other and vice versa."

In asymptotically de Sitter spacetimes, it would appear that the natural questions involve setting some initial conditions on spacelike $\mathcal{I}^-$, letting it evolve, and reading off the data on $\mathcal{I}^+$. One of the lessons of the development of quantum mechanics is that we shouldn't talk about things that cannot, even in principle, be measured. Yet in de Sitter space we see that no single observer has an overview of the whole space. This causes a number of headaches and, as yet, unresolved conceptual issues when we try to discuss quantum gravity in de Sitter space.

Finally, we can use the Penrose diagram to answer a lingering puzzle about the static patch of de Sitter, in which the metric takes the form (4.12)

$$ds^2 = -\left(1 - \frac{r^2}{R^2}\right) dt^2 + \left(1 - \frac{r^2}{R^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{4.39}$$

The question is: how should we interpret the divergence at $r = R$?

To answer this, we will look at where the surface $r = R$ sits in the Penrose diagram. First, we look at the embedding of the static patch in $\mathbf{R}^{1,4}$, given in (4.19)

$$X^0 = \sqrt{R^2 - r^2}\sinh(t/R) \quad \text{and} \quad X^4 = \sqrt{R^2 - r^2}\cosh(t/R)$$

Naively the surface $r = R$ corresponds to $X^0 = X^4 = 0$. But that's a little too quick. To see this, we consider what happens as we approach $r \to R$ by writing $r = R(1-\epsilon^2/2)$, with $\epsilon \ll 1$. We then have

$$X^0 \approx R\epsilon\sinh(t/R) \quad \text{and} \quad X^4 \approx R\epsilon\cosh(t/R)$$

We can now send $\epsilon \to 0$, keeping $X^0$ and $X^4$ finite provided that we also send $t \to \pm\infty$. To do this, we must ensure that we keep the combination $\epsilon\, e^{\pm t/R}$ finite. This means that we can identify the surface $r = R$ with the lines $X^0 = \pm X^4$.

Now we translate this into global coordinates. These were given in (4.20),

$$X^0 = R\sinh(\tau/R) \quad \text{and} \quad X^4 = R\cosh(\tau/R)\cos\chi$$

where $\chi$ is the polar angle on $\mathbf{S}^3$ that we introduced in (4.38). After one further map to conformal time (4.37), we find that the lines $X^0 = \pm X^4$ become

$$\sin\eta = \pm\cos\chi \quad \Rightarrow \quad \chi = \pm(\eta - \pi/2)$$

But these are precisely the diagonal lines in the Penrose diagram that appear as horizons for people living on the poles.

It's also simple to check that the point $r = 0$ in the static patch corresponds to the north pole $\chi = 0$ in global coordinates and, furthermore, $t = \tau$ along this line.

The upshot is that the static patch of de Sitter (4.39) provides coordinates that cover only the northern causal diamond of de Sitter, with the coordinate singularity at $r = R$ coinciding with the past and future observer-dependent horizons.

One advantage of the static patch coordinates is that they clearly exhibit a timelike Killing vector, $K = \partial_t$. This moves us from a surface of constant $t$ to another surface of constant $t$. But we argued in Section 4.3.1 that there was no global timelike Killing vector field in de Sitter since, in $\mathbf{R}^{1,4}$, the Killing vector is given by (**??**). The Penrose diagram makes this simpler to visualise. If we extend the Killing vector beyond the static patch, it acts as shown in the figure. It is timelike and future pointing only in the northern causal diamond. It is also timelike in the southern causal diamond, but points towards the past. Meanwhile it is a spacelike Killing vector in both the upper and lower quadrants.

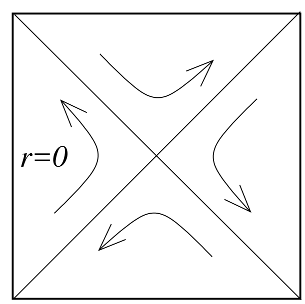



**Figure 38:** Killing vectors in de Sitter

### Anti-de Sitter Space

The global coordinates for anti-de Sitter space are (4.23),

$$ds^2 = -\cosh^2\!\rho\, dt^2 + R^2 d\rho^2 + R^2\sinh^2\!\rho\, d\Omega_2^2$$

with $\rho \in [0, +\infty)$. To construct the Penrose diagram, this time we introduce a "conformal radial coordinate" $\psi$, defined by

$$\frac{d\psi}{d\rho} = \frac{1}{\cosh \rho}$$

This is very similar to the conformal map of de Sitter space, but with time replaced by space. The solution is

$$\cos \psi = \frac{1}{\cosh \rho}$$

One difference from the de Sitter analysis is that since $\rho \in [0, \infty)$, the conformal coordinate lives in $\psi \in [0, \pi/2)$. The metric on anti-de Sitter becomes

$$ds^2 = \frac{R^2}{\cos^2 \psi} \left( -d\tilde{t}^{\,2} + d\psi^2 + \sin^2 \psi \, d\Omega_2^2 \right) = \frac{R^2}{\cos^2 \psi} \left( -d\tilde{t}^{\,2} + d\Omega_3^2 \right)$$

where we've introduced the dimensionless time coordinates $\tilde{t} = t/R$. We learn that the anti-de Sitter metric is conformally equivalent to

$$d\tilde{s}^2 = -d\tilde{t}^{\,2} + d\psi^2 + \sin^2 \psi d\,\Omega_2^2$$

where, after a conformal compactification, $\tilde{t} \in (-\infty, +\infty)$ and $\psi \in [0, \pi/2]$. The Penrose diagram is shown in figure. It is an infinite strip. The left-hand edge at $\psi = 0$ is not a boundary: it is the spatial origin where the $\mathbf{S}^2$ shrinks to zero size. In contrast, the right-hand edge at $\psi = \pi/2$ is the boundary of spacetime.

The boundary is labelled $\mathcal{I}$. In terms of our previous notation, it should be viewed as a combination of $\mathcal{I}^-$, $\mathcal{I}^+$ and $i^0$, since null paths begin and end here, as do spacelike paths. The boundary is now timelike (it has spacelike normal vector), and has topology

$$\mathcal{I} = \mathbf{R} \times \mathbf{S}^2$$

with $\mathbf{R}$ the time factor.

The Penrose diagram allows us to immediately see that light rays hit the boundary in finite conformal time, confirming the calculation that we did in Section 4.2.2. If we want to specify physics in AdS, we need to say something about what happens at the boundary. For example, in the figure we have shown a light ray simply emerging from the boundary at one time and absorbed at some later time. Another choice would be to impose reflecting boundary conditions, so that the light ray bounces back and forth for ever. In this way, anti-de Sitter space is very much like a box, with massive particles trapped in the interior and massless particles able to bounce off the boundary.

In field theoretic language, we could start with initial data on some $d = 3$ dimensional spacelike hypersurface $\Sigma$ and try to evolve it in time. This is what we usually do in physics. But in AdS, this information is not sufficient. This is because we can find points to the future of $\Sigma$ which are in causal contact with the boundary. This means that what happens there depends on the choices we make on the boundary. In fancy language, we say that AdS is not *globally hyperbolic*: there exists no Cauchy surface on which we can specify initial data .

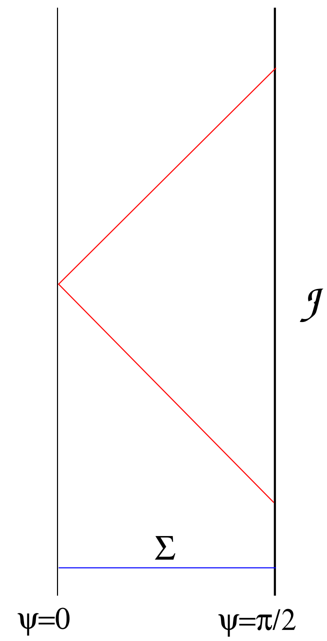



**Figure 39:** The Penrose diagram for AdS

AdS is the setting for our best-understood theories of quantum gravity. It turns out that gravitational dynamics in asymptotically AdS spacetimes is entirely equivalent to a quantum field theory living on the boundary $\mathcal{I}$. This idea goes by many names, including the *AdS/CFT correspondence*, *gauge-gravity duality*, or simply *holography*.

## 4.5 Coupling Matter

Until now, we've only discussed the dynamics of vacuum spacetime, with matter consigned to test particles moving on geodesics. But matter is not merely an actor on the spacetime stage: instead it backreacts, and affects the dynamics of spacetime itself.

### 4.5.1 Field Theories in Curved Spacetime

The first question we should ask is: how does matter couple to the spacetime metric? This is simplest to describe when matter takes the form of fields which themselves are governed by a Lagrangian. (We will look at what happens when matter is made of particles, albeit ones that form fluids, in Section 4.5.4.)

#### Scalar Fields

As a simple example, consider a scalar field $\phi(x)$. In flat space, the action takes the form

$$S_{\text{scalar}} = \int d^4x \, \left( -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\,\partial_\nu(\phi) - V(\phi) \right) \tag{4.40}$$

with $\eta^{\mu\nu}$ the Minkowski metric. The minus sign in front of the derivative terms follows from the choice of signature $(-+++)$. This differs from, say, the lectures on quantum field theory, but ensures that the action takes the form "kinetic energy" - "potential energy".

It is straightforward to generalise this to describe a field moving in curved spacetime: we simply need to replace the Minkowski metric with the curved metric, and ensure that we're integrating over a multiple of the volume form. In practice, this means that we have

$$S_{\text{scalar}} = \int d^4x\, \sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\phi\,\nabla_\nu\phi - V(\phi)\right) \tag{4.41}$$

Note that we've upgraded the derivatives from $\partial_\mu$ to $\nabla_\mu$, although in this case it's redundant because, on a scalar field, $\nabla_\mu\phi = \partial_\mu\phi$. Nonetheless, it will prove useful shortly.

Note, however, that curved spacetime also introduces new possibilities for us to add to the action. For example, we could equally well consider the theory

$$S_{\text{scalar}} = \int d^4x\, \sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\phi\,\nabla_\nu\phi - V(\phi) - \frac{1}{2}\xi R\phi^2\right) \tag{4.42}$$

for some constant $\xi$. This reduces to the flat space action (4.40) when we take $g_{\mu\nu} = \eta_{\mu\nu}$ since the Ricci scalar is then $R = 0$, but it gives different dynamics for each choice of $\xi$. To derive the equation of motion for $\phi$, we vary the action (4.42) with respect to $\phi$, keeping $g_{\mu\nu}$ fixed for now

$$\begin{aligned}\delta S_{\text{scalar}} &= \int d^4x\, \sqrt{-g}\left(-g^{\mu\nu}\nabla_\mu\delta\phi\,\nabla_\nu\phi - \frac{\partial V}{\partial\phi}\delta\phi - \xi R\phi\delta\phi\right)\\ &= \int d^4x\, \sqrt{-g}\left[\left(g^{\mu\nu}\nabla_\mu\nabla_\nu\phi - \frac{\partial V}{\partial\phi} - \xi R\phi\right)\delta\phi - \nabla_\mu\left(\delta\phi\nabla^\mu\phi\right)\right]\end{aligned}$$

Notice that although the covariant derivatives $\nabla_\mu$ could be replaced by $\partial_\mu$ on the first line, they're crucially important on the second where we needed the fact that $\nabla_\mu g_{\rho\sigma} = 0$ to do the integration by parts. The final term is a boundary term (using the divergence theorem proven in Section 3.2.4) and can be discarded. This leaves us with the equation of motion for a scalar field in curved spacetime,

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi - \frac{\partial V}{\partial\phi} - \xi R\phi = 0$$

Again, the covariant derivatives are needed here: we could write $\nabla_\mu\nabla_\nu\phi = \nabla_\mu\partial_\nu\phi$ except it looks stupid. But $\nabla_\mu\nabla_\nu\phi \neq \partial_\mu\partial_\nu\phi$.

**Maxwell Theory**

We already met the action for Maxwell theory in Section 3.2.5 as an example of integrating forms over manifolds. It is given by

$$S_{\text{Maxwell}} = -\frac{1}{2}\int F\wedge\star F = -\frac{1}{4}\int d^4x\, \sqrt{-g}\, g^{\mu\rho}g^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma} \tag{4.43}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu$. (The equivalence of these two expressions follows because of anti-symmetry, with the Levi-Civita connections in the final term cancelling.) This time, the equations of motion are

$$\nabla_\mu F^{\mu\nu} = 0$$

Indeed, this is the only covariant tensor that we can write down that generalises the flat space result $\partial_\mu F^{\mu\nu} = 0$.

### 4.5.2 The Einstein Equations with Matter

To understand how fields backreact on spacetime, we just need to consider the combined action

$$S = \frac{1}{16\pi G}\int d^4x\,\sqrt{-g}(R - 2\Lambda) + S_M$$

where $S_M$ is the action for matter fields which, as we have seen above, depends on both the matter fields and the metric. We know what happens when we vary the Einstein-Hilbert action with respect to the metric. Now we care about $S_M$. We define the *energy-momentum tensor* to be

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}} \tag{4.44}$$

Notice that $T_{\mu\nu}$ is symmetric, a property that it inherits from the metric $g_{\mu\nu}$. If we vary the full action with respect to the metric, we have

$$\delta S = \frac{1}{16\pi G}\int d^4x\,\sqrt{-g}\,(G_{\mu\nu} + \Lambda g_{\mu\nu})\,\delta g^{\mu\nu} - \frac{1}{2}\int d^4x\,\sqrt{-g}\,T_{\mu\nu}\,\delta g^{\mu\nu}$$

From this we can read off the equations of motion,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G\,T_{\mu\nu} \tag{4.45}$$

These are the full Einstein equations, describing gravity coupled to matter.

There are a number of different ways of writing this. First, the cosmological constant is sometimes absorbed as just another component of the energy-momentum tensor,

$$(T_{\mu\nu})_\Lambda = -\frac{\Lambda}{8\pi G}g_{\mu\nu} \tag{4.46}$$

One reason for this is that the matter fields can often mimic a cosmological constant and it makes sense to bundle all such terms together. (For example, a scalar field sitting at an extremal point of a potential is indistinguishable from a cosmological constant.) In this case, we just have

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}$$

where $T_{\mu\nu}$ now includes the cosmological term.

Taking the trace (i.e. contracting with $g^{\mu\nu}$) then gives

$$-R = 8\pi G T$$

with $T = g^{\mu\nu}T_{\mu\nu}$. We can use this to directly relate the Ricci tensor to the energy momentum

$$R_{\mu\nu} = 8\pi G \left( T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right) \tag{4.47}$$

This form will also have its uses in what follows.

### 4.5.3 The Energy-Momentum Tensor

The action $S_M$ is constructed to be diffeomorphism invariant. This means that we can replay the argument of Section 2.1.3 that led us to the Bianchi identity: if we vary the metric by a diffeomorphism $\delta g_{\mu\nu} = (\mathcal{L}_X g)_{\mu\nu} = 2\nabla_{(\mu}X_{\nu)}$, then we have

$$\delta S_M = -2 \int d^4x \, \sqrt{-g} T_{\mu\nu} \nabla^\mu X^\nu = 0 \quad \text{for all } X^\mu$$

This tells us that the energy momentum tensor is necessarily covariantly conserved,

$$\nabla_\mu T^{\mu\nu} = 0 \tag{4.48}$$

Of course, this was necessary to make the Einstein equation (4.45) consistent, since we know that $\nabla_\mu G^{\mu\nu} = 0$. Indeed, viewed from the action principle, both the Bianchi identity and $\nabla_\mu T^{\mu\nu} = 0$ have the same origin.

Although we've introduced the energy-momentum tensor as something arising from curved spacetime, it is also an important object in theories in flat space that have nothing to do with gravity. In that setting, the energy-momentum tensor arises as the Noether currents associated to translational invariance in space and time.

A hint of this is already apparent in (4.48) which, restricted to flat space, gives the expected conservation law enjoyed by Noether currents, $\partial_\mu T^{\mu\nu} = 0$. However, there is a rather slick argument that makes the link to Noether's theorem tighter.

In flat space, the energy-momentum tensor comes from invariance under translations $x^\mu \to x^\mu + X^\mu$, with constant $X^\mu$. There's a standard trick to compute the Noether current associated to any symmetry which involves promoting the symmetry parameters to be functions of the spacetime coordinates, so $\delta x^\mu = X^\mu(x)$. The action restricted to flat space is not invariant under such a shift. But it's simple to construct an action that is invariant: we simply couple the fields to a background metric and allow that to also vary. This is precisely the kinds of action we've been considering in this section. The change of the action in flat space where we don't let the metric vary must be equal and opposite to the change of the action where we let the metric vary but don't change $x^\mu$ (because the combination of the two vanishes). We must have

$$\delta S_{\text{flat}} = -\int d^4x \;\left.\frac{\delta S_M}{\delta g^{\mu\nu}}\right|_{g_{\mu\nu}=\eta_{\mu\nu}} \delta g^{\mu\nu}$$

But the variation of the metric *without* changing the point $x^\mu$ is $\delta g_{\mu\nu} = \partial_\mu X_\mu + \partial_\nu X_\mu$. (The Christoffel symbols in the more familiar expression with $\nabla_\mu$ come from the $\partial g_{\mu\nu}$ term in (4.7), and this is precisely the term we neglect.) We have

$$\delta S_{\text{flat}} = -2\int d^4x \;\left.\frac{\delta S_M}{\delta g^{\mu\nu}}\right|_{g_{\mu\nu}=\eta_{\mu\nu}} \partial^\mu X^\nu = -2\int d^4x \;\partial^\mu\left(\left.\frac{\delta S_M}{\delta g^{\mu\nu}}\right|_{g_{\mu\nu}=\eta_{\mu\nu}}\right) X^\nu$$

But we know that $\delta S_{\text{flat}} = 0$ whenever $X^\mu =$ constant, since this is precisely what it means for the theory to be translationally invariant. We learn that the conserved Noether current in flat space is

$$\left.T_{\mu\nu}\right|_{\text{flat}} = -2\left.\frac{\delta S_M}{\delta g^{\mu\nu}}\right|_{g_{\mu\nu}=\eta_{\mu\nu}}$$

which is the flat space version of (4.44).

### Examples of the Energy-Momentum Tensor

It is straightforward to compute the energy-momentum tensor for a scalar field. We take the action (4.41) and vary with respect to the metric. We will need the result $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\, g_{\mu\nu}\,\delta g^{\mu\nu}$ from Section 4.1. We then find

$$\delta S_{\text{scalar}} = \int d^4x\;\sqrt{-g}\left(\frac{1}{4}g_{\mu\nu}\nabla^\rho\phi\,\nabla_\rho\phi + \frac{1}{2}g_{\mu\nu}V(\phi) - \frac{1}{2}\nabla_\mu\phi\nabla_\nu\phi\right)\delta g^{\mu\nu}$$

where the first two terms come from varying $\sqrt{-g}$ and the third comes from varying the metic in the gradient term. This gives us the energy momentum tensor

$$T_{\mu\nu} = \nabla_\mu\phi\nabla_\nu\phi - g_{\mu\nu}\left(\frac{1}{2}\nabla^\rho\phi\,\nabla_\rho\phi + V(\phi)\right) \tag{4.49}$$

If we now restrict to flat space, with $g_{\mu\nu} = \eta_{\mu\nu}$, we find, for example,

$$T_{00} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi)$$

where $\nabla$ is the usual 3d spatial derivative. We recognise this as the energy density of a scalar field.

We can play the same game with the Maxwell action (4.43). Varying with respect to the metric, we have

$$\delta S_{\text{Maxwell}} = -\frac{1}{4}\int d^4x\,\sqrt{-g}\left(-\frac{1}{2}g_{\mu\nu}F^{\rho\sigma}F_{\rho\sigma} + 2g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma}\right)\delta g^{\mu\nu}$$

So the energy momentum tensor is given by

$$T_{\mu\nu} = g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma} - \frac{1}{4}g_{\mu\nu}F^{\rho\sigma}F_{\rho\sigma}$$

In flat space, with $g_{\mu\nu} = \eta_{\mu\nu}$,

$$T_{00} = \frac{1}{2}\mathbf{E}^2 + \frac{1}{2}\mathbf{B}^2$$

with $F_{0i} = -E_i$, the electric field, and $F_{ij} = \epsilon_{ijk}B_k$ the magnetic field. Again, we recognise this as the energy density in the electric and magnetic fields. You can read more about the properties of the Maxwell energy-momentum tensor in the lecture on Electromagnetism.

### 4.5.4 Perfect Fluids

Take any kind of object in the universe. Throw a bunch of them together, heat them up, and gently splash. The resulting physics will be described by the equations of fluid dynamics.

A perfect fluid is described by its *energy density* $\rho(\mathbf{x},t)$, pressure $P(\mathbf{x},t)$ and a velocity 4-vector $u^\mu(\mathbf{x},t)$ such that $u^\mu u_\mu = -1$. The pressure and energy density are not unrelated: there is an identity between them that is usually called the *equation of state*,

$$P = P(\rho)$$

Common examples include *dust*, which consists of massive particles floating around, moving very slowly so that the pressure is $P = 0$, and *radiation*, which is a fluid made of many photons for which $P = \rho/3$.

The energy-momentum tensor for a perfect fluid is given by

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + Pg^{\mu\nu} \tag{4.50}$$

If we are in Minkowski space, so $g_{\mu\nu} = \eta_{\mu\nu}$ and the fluid is at rest, so $u^\mu = (1,0,0,0)$, then the energy momentum tensor is

$$T^{\mu\nu} = \text{diag}(\rho, P, P, P)$$

We see that $T^{00} = \rho$, as expected for the energy density. More generally, for a moving fluid we have $T_{\mu\nu}u^\mu u^\nu = \rho$, which means that $\rho$ is the energy density measured by an observer co-moving with the fluid.

The energy-momentum tensor must obey

$$\nabla_\mu T^{\mu\nu} = 0$$

A short calculation shows that this is equivalent to two relations between the fluid variables. The first is

$$u^\mu \nabla_\mu \rho + (\rho + P)\nabla_\mu u^\mu = 0 \tag{4.51}$$

This is the relativistic generalisation of mass conservation for a fluid. Here "mass" has been replaced by energy density $\rho$. The first term, $u^\mu \nabla_\mu \rho$ calculates how fast the energy density is changing as we move along $u^\mu$. The second term tells us the answer: it depends on $\nabla_\mu u^\mu$, the rate of flow of fluid out of a region.

The second constraint is

$$(\rho + P)u^\mu \nabla_\mu u^\nu = -(g^{\mu\nu} + u^\mu u^\nu)\nabla_\mu P \tag{4.52}$$

This is the relativistic generalisation of the Euler equation, the fluid version of Newton's second law "F=ma". The left-hand side of the equation should be viewed as "mass $\times$ acceleration", the right-hand side is the force which, in a fluid, is due to pressure differences. You can learn more about these equations in the relativistic context and their solutions in chapter 3 of the lectures on Cosmology.

### 4.5.5 The Slippery Business of Energy Conservation

In flat space, the existence of an energy-momentum tensor ensures that we can define the conserved quantities, energy and momentum. In curved spacetime, things are significantly more subtle.

To see this, it's useful to compare the energy-momentum tensor $T^{\mu\nu}$ with a current $J^\mu$ which arises from a global symmetry (such as, for example, the phase rotation of a complex scalar field). In flat space, both obey current conservation

$$\partial_\mu J^\mu = 0 \quad \text{and} \quad \partial_\mu T^{\mu\nu} = 0 \tag{4.53}$$

From these, we can construct conserved charge by integrating over a spatial volume $\Sigma$

$$Q(\Sigma) = \int_\Sigma d^3x\ J^0 \quad \text{and} \quad P^\mu(\Sigma) = \int_\Sigma d^3x\ T^{0\mu}$$

To see that these are indeed conserved, we simply need to integrate over a spacetime volume $V$, bounded by $\Sigma$ at past and future times $t_1$ and $t_2$. We then have

$$0 = \int_V d^4x\ \partial_\mu J^\mu = \Delta Q(\Sigma) + \int_B d^3x\ n_i J^i$$

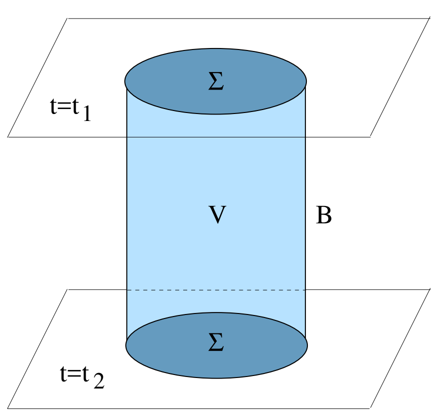



**Figure 40:** Charge conservation in flat spacetime...

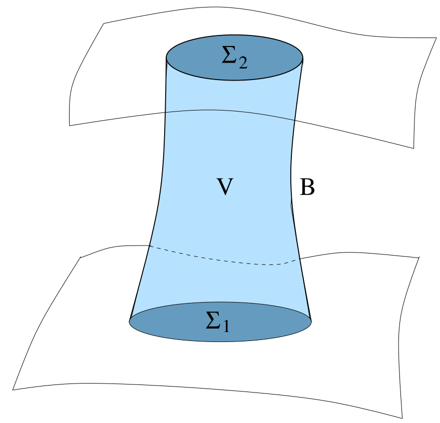



**Figure 41:** ...and in curved spacetime.

where $\Delta Q(\Sigma) = [Q(\Sigma)](t_2) - [Q(\Sigma)](t_1)$ and $n^i$ is the outward-pointing normal to $B$, the timelike boundary of $V$, as shown in the left-hand figure. Provided that no current flows out of the region, meaning $J_i = 0$ when evaluated on $B$, we have $\Delta Q(\Sigma) = 0$. (Often, we take $\Sigma = \mathbf{R}^3$ so that $B = \mathbf{S}^2_\infty \times I$ with $I$ and interval, and we only have to require that there are no currents at infinity.) This is the statement of charge conservation.

In Minkowski space, this same argument works just as well for $P^\mu(\Sigma)$, meaning that we are able to assign conserved energy and momentum to fields in some region, providing that no currents leak out through the boundary.

Let's now contrast this to the situation in curved spacetime. The conservation laws (4.53) are replaced by their covariant versions

$$\nabla_\mu J^\mu = 0 \quad \text{and} \quad \nabla_\mu T^{\mu\nu} = 0$$

We can replay the argument above, now invoking the divergence theorem from Section 3.2.4

$$0 = \int_V d^4x \, \sqrt{-g} \, \nabla_\mu J^\mu = \int_{\partial V} d^3x \, \sqrt{|\gamma|} \, n_\mu J^\mu$$

with $\gamma_{ij}$ the pull-back of the metric to $\partial V$ and $n^\mu$ the normal vector. We consider a spacetime volume $V$ with boundary

$$\partial V = \Sigma_1 \cup \Sigma_2 \cup B$$

Here $\Sigma_1$ and $\Sigma_2$ are past and future spacelike boundaries, while $B$ is the timelike boundary as shown in the right-hand figure. If we again insist that no current flows out of the region by requiring that $J^\mu n_\mu = 0$ when evaluated on $B$, then the expression above becomes

$$Q(\Sigma_2) = Q(\Sigma_1)$$

where the charge $Q(\Sigma)$ evaluated on a spacelike hypersurface $\Sigma$ is defined by

$$Q(\Sigma) = \int_\Sigma d^3x \, \sqrt{\gamma}\, n_\mu J^\mu$$

This means that, for a vector field, covariant conservation is the same thing as actual conservation. The story above is a repeat of the one we told using differential forms in Section 3.2.5.

Now let's try to tell a similar story for the energy-momentum tensor. In analogy to the derivation above, it's clear that we should try to manipulate the integral

$$0 = \int_V d^4x \, \sqrt{-g}\, \nabla_\mu T^{\mu\nu}$$

The problem is that we don't have a divergence theorem for integrals of this kind because of the hanging $\nu$ index on the energy momentum tensor. The key to deriving the divergence theorem for $J^\mu$ was the expression

$$\nabla_\mu J^\mu = \partial_\mu J^\mu + \Gamma^\mu_{\mu\rho} J^\rho = \frac{1}{\sqrt{-g}} \partial_\mu \left( \sqrt{-g} J^\mu \right)$$

This allows us to turn a covariant derivative into a normal derivative which gives a boundary term in the integral. However, the same expression for the energy-momentum tensor reads

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\mu\rho} T^{\rho\nu} + \Gamma^\nu_{\mu\rho} T^{\mu\rho} = \frac{1}{\sqrt{-g}} \partial_\mu \left( \sqrt{-g} T^{\mu\nu} \right) + \Gamma^\nu_{\mu\rho} T^{\mu\rho}$$

That extra term involving the Christoffel symbol stops us converting the integral of $\nabla_\mu T^{\mu\nu}$ into a boundary term. Indeed, we can rewrite $\nabla_\mu T^{\mu\nu} = 0$ as

$$\partial_\mu \left( \sqrt{-g} T^{\mu\nu} \right) = -\sqrt{-g}\, \Gamma^\nu_{\mu\rho} T^{\mu\rho} \tag{4.54}$$

then the right-hand side looks like a driving force which destroys conservation of energy and momentum. We learn that, for a higher tensor like $T^{\mu\nu}$, covariant conservation is not the same thing as actual conservation!

**Conserved Energy from a Killing Vector**

We can make progress by introducing one further ingredient. If our spacetime has a Killing vector field $K$, we can construct a current from the energy-momentum tensor by writing

$$J_T^\nu = K_\mu T^{\mu\nu}$$

Taking the covariant divergence of the current gives

$$\nabla_\nu J_T^\nu = (\nabla_\nu K_\mu) T^{\mu\nu} + K_\mu \nabla_\nu T^{\mu\nu} = 0$$

where the first term vanishes by virtue of the Killing equation, with $T^{\mu\nu}$ imprinting its symmetric indices on $\nabla_\nu K_\mu$, and the second term vanishes by (4.48).

Now we're in business. We can construct conserved charges from the current $J_T^\mu$ as explained above,

$$Q_T(\Sigma) = \int_\Sigma d^3x\ \sqrt{\gamma}\ n_\mu J^\mu$$

The interpretation of these charges depends on the properties of the Killing vector. If $K^\mu$ is everywhere timelike, meaning $g_{\mu\nu} K^\mu K^\nu < 0$ at all points, then the charge can be identified with the energy of the matter

$$E = Q_T(\Sigma)$$

If the Killing vector is everywhere spacelike, meaning $g_{\mu\nu} K^\mu K^\nu > 0$ at all points, then the charge can be identified with the momentum of the matter.

**Conserved Energy Without a Killing Vector?**

There are situations where spacetime does not have a Killing vector yet, intuitively, we would still like to associate something analogous to energy. This is where things start to get subtle.

A simple situation where this arises is two orbiting stars. It turns out that the resulting spacetime does not admit a timelike Killing vector. As we will describe in some detail in Section 5.3, as the stars orbit they emit gravitational waves, losing energy, which causes them to slowly spiral towards each other. This fits nicely with the lack of a timelike Killing vector, since we wouldn't expect to define a conserved energy for the stars.

However, it certainly feels like we should be able to define a conserved energy for the total system which, in this case, means stars together with the gravitational waves. In particular, we would like to quantify the amount of energy lost by the stars and carried away by the gravitational waves. But this requires us to define something new, namely the energy density in the gravitational field. And this is where the trouble starts!

There's an obvious way to proceed, one that starts by returning to our original definition of the energy-momentum tensor (4.44)

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}}$$

A naive guess would be to include the action for both matter and gravity in this definition, giving an energy momentum tensor which includes both matter and gravity

$$T^{\rm total}_{\mu\nu} = -\frac{2}{\sqrt{-g}}\left(\frac{1}{16\pi G_N}\frac{\delta S_{EH}}{\delta g^{\mu\nu}} + \frac{\delta S_M}{\delta g^{\mu\nu}}\right)$$

But this gives

$$T^{\rm total}_{\mu\nu} = -\frac{1}{8\pi G_N}G_{\mu\nu} + T_{\mu\nu} = 0$$

which vanishes by the Einstein field equations. The idea that the total energy of the universe vanishes, with negative gravitational energy cancelling the positive energy from matter sounds like it might be something important. It turns out to be very good for selling pseudo-scientific books designed to make the world think you're having deep thoughts. It's not, however, particularly good for anything to do with physics. For example, it's not as if electrons and positrons are suddenly materialising everywhere in space, their mass energy cancelled by the gravitational energy. That's not the way the universe works. Instead the right way to think about this equation is to simply appreciate that energy in general relativity is subtle.

Clearly, we should try to do better to understand the energy carried in the gravitational field. Unfortunately, it turn out that doing better is challenging. There are compelling arguments that show that there is no tensor that can be thought of as the capturing the local energy density of the gravitational field. Roughly speaking, these arguments start from the observation that the energy in the Newtonian gravitational field is proportional to $(\nabla\Phi)^2$. We should therefore expect that the relativistic version of energy density is proportional to the first derivative of the metric. Yet the equivalence principle tells us that we can always find coordinates – those experienced by a freely falling observer – which ensure that the first derivative of the metric vanish at any given point. But a tensor that vanishes in one coordinate system also vanishes in another.

We'll confront these issues again in Section 5.3 when we try to answer the question of how much energy emitted in gravitational waves by a binary star system. We'll see there that, for this case, there are simplifications that mean we can converge on a sensible answer.

### 4.5.6 Spinors

In flat space, fermions transform in the spinor representation of the $SO(1,3)$ Lorentz group. Recall from the lectures on Quantum Field Theory that we first introduce gamma matrices obeying the Clifford algebra

$$\{\gamma_a, \gamma_b\} = 2\eta_{ab} \tag{4.55}$$

Notice that we've put indices $a, b = 0, 1, 2, 3$ on the gamma matrices, rather than the more familiar $\mu, \nu$. This is deliberate. In $d = 3+1$ dimensions, each of the $\gamma_a$ is a $4 \times 4$ matrix.

We can use these gamma matrices to construct generators of the Lorentz group,

$$S_{ab} = \frac{1}{4}[\gamma_a, \gamma_b]$$

These define the spinor representation of the Lorentz group. We write a Lorentz transformation $\Lambda$ as

$$\Lambda = \exp\left(\frac{1}{2}\lambda^{ab} M_{ab}\right)$$

with $M_{ab}$ the usual Lorentz generators (defined, for example, in (4.30)) and $\lambda^{ab}$ a choice of 6 numbers that specify the particular Lorentz transformation. Then the corresponding transformation in the spinor representation is given by

$$S[\Lambda] = \exp\left(\frac{1}{2}\lambda^{ab} S_{ab}\right)$$

A Dirac spinor field $\psi_\alpha(x)$ is then a 4-component complex vector that, under a Lorentz transformation, changes as

$$\psi(x) \to S[\Lambda]\, \psi(\Lambda^{-1} x) \tag{4.56}$$

In Minkowski space, the action for the spinor is

$$S_{\text{Dirac}} = \int d^4x \; i\left(\bar{\psi}\gamma^a \partial_a \psi + m\bar{\psi}\psi\right)$$

with $\bar{\psi} = \psi^\dagger \gamma^0$. The magic of gamma matrices ensures that this action is invariant under Lorentz transformations, despite having just a single derivative. Our task in this section is to generalise this action to curved spacetime.

We can already see some obstacles. The gamma matrices (4.55) are defined in Minkowski space and it's not clear that they would retain their magic if generalised to curved space. Furthermore, what should we do with the derivative? We might suspect that it gets replaced by a a covariant derivative, but what connection do we choose?

To answer these questions, we will need to invoke the vierbeins and connection one-form that we met in Section 3.4.2. Recall that the vierbeins $\hat{e}_a = e_a{}^\mu \partial_\mu$ are a collection of 4 vector fields, that allow us "diagonalise" the metric. We define $e^a{}_\mu$ to be the inverse of $e_a{}^\mu$, meaning it satisfies $e^a{}_\mu e_b{}^\mu = \delta^a_b$ and $e^a{}_\mu e_a{}^\nu = \delta^\nu_\mu$. The metric can then be written

$$g_{\mu\nu} = e^a{}_\mu e^b{}_\nu \eta_{ab}$$

These formula are really telling us that we should raise/lower the $\mu, \nu$ indices using the metric $g_{\mu\nu}$ but raise/lower the $a, b$ indices using the Minkowski metric $\eta_{ab}$.

The formalism of vierbeins allowed us to introduce a the idea of a *local Lorentz transformation* $\Lambda(x)$, defined in (3.43), which acts on the vierbeins as

$$e_a{}^\nu \to \tilde{e}_a{}^\nu = e_b{}^\nu (\Lambda^{-1})^b{}_a \quad \text{with} \quad \Lambda_a{}^c \Lambda_b{}^d \eta_{cd} = \eta_{ab}$$

This local Lorentz transformation can now be promoted to act on a spinor field as (4.56), again with $S[\Lambda]$ depending on the coordinate $x$.

We want our action to be invariant under these local Lorentz transformations. In particular, we might expect to run into difficulties with the derivative which, after a Lorentz transformation, now hits $\Lambda(x)$ as well as $\psi(x)$. But this is exactly the kind of problem that we've met before when writing down actions for gauge theories, and we know very well how to solve it: we simply need to include a connection in the action that transforms accordingly. To this end, we construct the covariant derivative acting on a spinor field

$$\nabla_\mu \psi = \partial_\mu \psi + \frac{1}{2} \omega_\mu^{ab} S_{ab} \psi$$

with $\omega_\mu^{ab}$ the appropriate connection. But what is it?

The right choice is the connection one-form, also known as the spin connection, that we met in Section 3.4.2. From (3.44) and (3.47), we have

$$(\omega^a{}_b)_\mu = \Gamma^a_{cb} e^c{}_\mu = e^a{}_\rho \nabla_\mu e_b{}^\rho$$

This does the trick because of its inhomogeneous transformation (3.45)

$$(\omega^a{}_b)_\mu \to \Lambda^a{}_c (\omega^c{}_d)_\mu (\Lambda^{-1})^d{}_b + \Lambda^a{}_c (\partial_\mu \Lambda^{-1})^c{}_b$$

This cancels the contribution from the derivative in the same way as the covariant derivative in a non-Abelian gauge theory. The generalisation of the Dirac action to curved space is then simply

$$S_{\rm Dirac} = \int d^4x \, \sqrt{-g}i \left( \bar{\psi}\gamma^a e_a{}^\mu \nabla_\mu \psi + m\bar{\psi}\psi \right)$$

There are a number of reasons to be interested in coupling fermions to gravity. First, and most obviously, both are constituents of our universe and its important to understand how they fit together. Second, they are important for more formal aspects of mathematical physics: they are the key component in Witten's simple proof of the positive mass theorem, and there are reasons to suspect that the quantisation of gravity ultimately requires supersymmetry at a high energy scale.

However, there is one thing that you probably shouldn't do with them, which is put them on the right-hand side of the Einstein equation and solve them. This is because fermions are quantum fields and do not have a macroscopic, classical analog.

Of course, all fields are, at heart, quantum. But for bosonic fields, it makes sense to think of them classically where they can be viewed as the quantum fields with high occupation number. The familiar in electromagnetism, where the classical electric and magnetic field can be thought of containing many photons. This means that it makes sense to find spacetimes which solve the Einstein equations $G_{\mu\nu} = 8\pi G T_{\mu\nu}$ where the curvature is supported by a profile for scalar fields on the right-hand-side.

In contrast, there is no classical limit of fermionic fields. This is because the Pauli exclusion principle prohibits a large occupation number. If you therefore attempt to find a spacetime supported by a fermionic $T_{\mu\nu}$, you are really looking for a gravitational solution sourced by precisely one quantum excitation. Given the feebleness of gravity on the microscopic scale, this is unlikely to be interesting.

This is not to say that fermions don't affect gravity. Important examples for gravitating fermionic systems include white dwarfs and neutron stars. But in each of these cases there is a separation of scales where one can first neglect gravity and find an effective equation of state for the fermions, and subsequently understand how this backreacts on spacetime. If you want to understand the spacetime directly from the Dirac equation than you have a complicated many-body problem on your hands.

### 4.5.7 Energy Conditions

If we know the kind of matter that fills spacetime, then we can just go ahead and solve the Einstein equations. However, we will often want to make more general statements about the allowed properties of spacetime without reference to any specific matter content. In this case, it is useful to place certain restrictions on the kinds of energy-momentum tensor that we consider physical.

These restrictions, known as *energy conditions*, capture the rough idea that energy should be positive. A number of classic results in general relativity, such as the singularity theorems, rely on these energy conditions as assumptions.

There are a bewildering number of these energy conditions. Moreover, it is not difficult to find examples of matter which violate most of them! We now describe a number of the most important energy conditions, together with their limitations.

- **Weak Energy Condition:** This states that, for any timelike vector field $X$,

  $$T_{\mu\nu}X^\mu X^\nu \geq 0 \quad \text{for all } X \text{ with } X_\mu X^\mu < 0$$

  The idea is that this quantity is the energy seen by an observer moving along the timelike integral curves of $X$, and this should be positive. A timelike curve can get arbitrarily close to a null curve so, by continuity, the weak energy condition can be extended to timelike and null curves

  $$T_{\mu\nu}X^\mu X^\nu \geq 0 \quad \text{for all } X \text{ with } X_\mu X^\mu \leq 0 \tag{4.57}$$

  To get a feel for this requirement, let's first impose it on the energy-momentum tensor for a perfect fluid (4.50). We will consider timelike vectors $X$ normalised to $X \cdot X = -1$. We then have

  $$T_{\mu\nu}X^\mu X^\nu = (\rho + P)(u \cdot X)^2 - P \geq 0$$

  We work in the rest frame of the fluid, so $u^\mu = (1, 0, 0, 0)$ and consider constant timelike vector fields, $X^\mu = (\cosh\varphi, \sinh\varphi, 0, 0)$. These describe the worldlines of observers boosted with rapidity $\varphi$ with respect to the fluid. The weak energy condition then gives us

  $$(\rho + P)\cosh^2\varphi - P \geq 0 \quad \Rightarrow \quad \begin{cases} \rho \geq 0 & \varphi = 0 \\ P \geq -\rho & \varphi \to \infty \end{cases}$$

  The first condition $\rho \geq 0$ is what we expect from the weak energy condition: it ensures that the energy density is positive. The second condition $P \geq -\rho$ says that negative pressure is acceptable, just as long as it's not too negative.

There are, however, situations in which negative energy density makes physical sense. Indeed, we've met one already: if we view the cosmological constant as part of the energy momentum tensor, as in (4.46), then any $\Lambda < 0$ violates the weak energy condition. Viewed this way, anti-de Sitter spacetime violates the weak energy condition.

We can also look at how this condition fares for scalar fields. From the energy-momentum tensor (4.49), we have

$$(X^\mu \partial_\mu \phi)^2 + \frac{1}{2}\partial_\mu \phi \partial^\mu \phi + V(\phi) \geq 0 \tag{4.58}$$

The first term is positive, but the second term can have either sign. In fact, it turns out that the first and second term combined are always positive. To see this, define the vector $Y$ orthogonal to $X$

$$Y_\mu = \partial_\mu \phi + X_\mu (X^\nu \partial_\nu \phi)$$

This satisfies $X_\mu Y^\mu = 0$: it is the projection of $\partial_\mu \phi$ onto directions orthogonal to $X$. Because $X$ is timelike, $Y$ must be spacelike (or null) and so obeys $Y_\mu Y^\mu \geq 0$. The weak energy condition (4.58) can be rewritten as

$$\frac{1}{2}(X^\mu \partial_\mu \phi)^2 + \frac{1}{2} Y_\mu Y^\mu + V(\phi) \geq 0$$

Now the first two terms are positive. We see that the weak energy condition is satisfied provided that $V(\phi) \geq 0$. However, it is violated in any classical theory with $V(\phi) \leq 0$ and there's no reason to forbid such negative potentials for a scalar field.

- **Strong Energy Condition:** There is a different, less immediately intuitive, energy condition. This is the requirement that, for any timelike vector field $X$,

$$R_{\mu\nu} X^\mu X^\nu \geq 0$$

This is the *strong energy condition*. It is poorly named. The strong energy condition is neither stronger nor weaker than the weak energy condition: it is simply different. We'll learn the motivation for the strong energy condition in Section **??** when we come to discuss the Raychaudhuri equation. It ensures that timelike geodesics converge, which can be viewed as the statement that gravity is attractive.

Using the form of the Einstein equations (4.47), the strong energy condition requires

$$\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right)X^\mu X^\nu \geq 0$$

for all timelike vector fields $X$. As before, continuity ensures that we can extend this to timelike and null vector fields, $X \cdot X \leq 0$.

If we take $X \cdot X = -1$ then, applied to a perfect fluid (4.50), the strong energy condition requires

$$(\rho + P)(u \cdot X)^2 - P + \frac{1}{2}(3P - \rho) \geq 0$$

As before, we consider the fluid in its rest frame with $u^\mu = (1, 0, 0, 0)$ and look at this condition for boosted observers with $X^\mu = (\cosh\varphi, \sinh\varphi, 0, 0)$. We have

$$(\rho + P)\cosh^2\varphi + \frac{1}{2}(P - \rho) \geq 0 \quad \Rightarrow \quad \begin{cases} P \geq -\rho/3 & \varphi = 0 \\ P \geq -\rho & \varphi \to \infty \end{cases}$$

Once again, it is not difficult to find situations where the strong energy condition is violated. Most strikingly, a cosmological constant $\Lambda > 0$ is not compatible with the strong energy condition. In fact, we may have suspected this because neighbouring, timelike geodesics in de Sitter space are pulled apart by the expansion of space. In fact, the strong energy condition forbids any FRW universe with $\ddot{a} > 0$, but there are at least two periods when our own universe underwent accelerated expansion: during inflation, and now.

Finally, it's not hard to show that any classical scalar field with a positive potential energy will violate the strong energy condition.

- **Null Energy Condition:** The null energy condition

$$T_{\mu\nu}X^\mu X^\nu \geq 0 \quad \text{for all } X \text{ with } X \cdot X = 0$$

This is implied by both weak and strong energy conditions, but the converse is not true: the null energy condition is strictly weaker than both the weak and strong conditions. This, of course, means that it is less powerful if we wield it to prove various statements. However, the null energy condition has the advantage that it is satisfied by any sensible classical field theory and any perfect fluid that obeys $\rho + P \geq 0$.

- **Dominant Energy Condition:** There is also an energy condition which is stronger than the weak condition. For any future-directed timelike vector $X$, we can define the current

$$J^\mu = -T^{\mu\nu} X_\nu \tag{4.59}$$

This is energy density current as seen by an observer following the lines of $X$. The dominant energy condition requires that, in addition to the weak energy condition (4.57), the current is either timelike or null, so

$$J_\mu J^\mu \leq 0$$

This is the reasonable statement that energy doesn't flow faster than light.

One can check that the extra condition (4.59) is satisfied for a scalar field. For a perfect fluid we have

$$J^\mu = -(\rho + P)(u \cdot X) u^\mu - P X^\mu$$

It's simple to check that the requirement $J_\mu J^\mu \leq 0$ is simply $\rho^2 \geq P^2$.

The validity of the various energy conditions becomes murkier still in the quantum world. We can consider quantum matter coupled to a classical yet dynamical spacetime through the equation

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G \langle T_{\mu\nu} \rangle$$

where $\langle T_{\mu\nu} \rangle$ is the expectation value of the energy-momentum tensor. Each of the energy conditions listed above is violated by fairly standard quantum field theories. There is, however, a somewhat weaker statement that holds true in general. This is the *averaged null energy condition*. It can be proven that, along an infinite, *achronal* null geodesic, any reasonable quantum field theory obeys

$$\int_{-\infty}^{+\infty} d\lambda \, \langle T_{\mu\nu} \rangle X^\mu X^\nu \geq 0$$

Here $\lambda$ is an affine parameter along the null geodesic and the vector $X^\mu$ points along the geodesic and is normalised to $X^\mu \partial_\mu \lambda = 1$. Here the word "achronal" means that no two points on the geodesic can be connected by a timelike curve. (As a counterexample, consider an infinite null ray on $M = \mathbf{R} \times \mathbf{S}^1$ which continually orbits the spatial circle. This geodesic is not achronal and the averaged null energy condition is not, in general, obeyed along this geodesic.)

## 4.6 A Taste of Cosmology

There are surprisingly few phenomena in Nature where we need to solve the Einstein equations sourced by matter,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

However there is one situation where the role of $T_{\mu\nu}$ on the right-hand side is crucial: this is cosmology, the study of the universe as a whole.

### 4.6.1 The FRW Metric

The key assumption of cosmology is that the universe is spatially homogeneous and isotropic. This means restricts the our choices of spatial geometry to one of three: these are

- Euclidean Space $\mathbf{R}^3$: This space has vanishing curvature and the familiar metric

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

- Sphere $\mathbf{S}^3$: This space has uniform positive curvature and metric

$$ds^2 = \frac{1}{1-r^2}dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

  With this choice of coordinates, we have implicitly set the radius of the sphere to 1.

- Hyperboloid $\mathbf{H}^3$: This space has uniform negative curvature and metric

$$ds^2 = \frac{1}{1+r^2}dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

The existence of three symmetric spaces is entirely analogous to the the three different solutions we discussed Section 4.2. de Sitter and anti-de Sitter are have constant *spacetime* curvature, supplied by the cosmological constant. The metrics above have constant *spatial* curvature. Note, however, that the metric on $\mathbf{S}^3$ coincides with the spatial part of the de Sitter metric in coordinates (4.12), while the metric on $\mathbf{H}^3$ coincides with the spatial part of the anti-de Sitter metric in coordinates (4.22)

We write these spatial metrics in unified form,

$$ds^2 = \gamma_{ij}dx^i dx^j = \frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \quad \text{with } k = \begin{cases} +1 & \mathbf{S}^3 \\ 0 & \mathbf{R}^3 \\ -1 & \mathbf{H}^3 \end{cases}$$

In cosmology, we wish to describe a spacetime in which space expands as the universe evolves. We do this with metrics of the form

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j \tag{4.60}$$

This is the Friedmann-Robertson-Walker, or FRW metric. (It is also known as the FLRW metric, with Lemaître's name a worthy addition to the list.) The dimensionless scale factor $a(t)$ should be viewed as the "size" of the spatial dimensions (a name which makes more sense for the compact $\mathbf{S}^3$ than the non-compact $\mathbf{R}^3$, but is mathematically sensible for both.) Note that de Sitter space in global coordinates (4.21) is an example of an FRW metric with $k = +1$.

### Curvature Tensors

We wish to solve the Einstein equations for metrics that take the FRW form. Our first task is to compute the Ricci tensor. We start with the Christoffel symbols: it is straightforward to find $\Gamma^\mu_{00} = \Gamma^0_{i0} = 0$ and

$$\Gamma^0_{ij} = a\dot{a}\gamma_{ij} \quad , \quad \Gamma^i_{0j} = \frac{\dot{a}}{a}\delta^i_j \quad , \quad \Gamma^i_{jk} = \frac{1}{2}\gamma^{il}\left(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}\right)$$

To compute the Ricci tensor, we use the expression

$$R_{\mu\nu} = \partial_\rho\Gamma^\rho_{\nu\mu} - \partial_\nu\Gamma^\rho_{\rho\mu} + \Gamma^\lambda_{\nu\mu}\Gamma^\rho_{\rho\lambda} - \Gamma^\lambda_{\rho\mu}\Gamma^\rho_{\nu\lambda} \tag{4.61}$$

which we get from contracting indices on the similar expression (3.39) for the Riemann tensor

It's not hard to see that $R_{0i} = 0$. The quick argument is that there's no covariant 3-vector that could possibly sit on the right-hand side. The other components need a little more work.

**Claim:**

$$R_{00} = -3\frac{\ddot{a}}{a}$$

**Proof:** Using the non-vanishing Christoffel symbols listed above, we have

$$R_{00} = -\partial_0 \Gamma^i_{i0} - \Gamma^j_{i0}\Gamma^i_{j0} = -3\frac{d}{dt}\left(\frac{\dot{a}}{a}\right) - 3\left(\frac{\dot{a}}{a}^2\right)$$

which gives the claimed result $\square$

**Claim:**

$$R_{ij} = \left(\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2}\right) g_{ij}$$

**Proof:** This is straightforward to show for $k = 0$ FRW metrics where the spatial metric is flat. It's a little more annoying for the $k = \pm 1$ metrics. A trick that simplifies life is to compute the components of $R_{ij}$ at the spatial origin $\mathbf{x} = 0$ where the spatial metric is $\gamma_{ij} = \delta_{ij}$, and then use covariance to argue that the right result must have $R_{ij} \sim \gamma_{ij}$. In doing this, we just have to remember not to set $\mathbf{x} = 0$ too soon, since we will first need to differentiate the Christoffel symbols and then evaluate them at $\mathbf{x} = 0$.

We start by writing the spatial metric in Cartesian coordinates, on the grounds that it's easier to differentiate in this form

$$\gamma_{ij} = \delta_{ij} + \frac{kx_i x_j}{1 - k\mathbf{x}\cdot\mathbf{x}}$$

The Christoffel symbols depend on $\partial\gamma_{ij}$ and the Ricci tensor on $\partial^2\gamma_{ij}$. This means that if we want to evaluate the Ricci tensor at the origin $\mathbf{x} = 0$, we only need to work with the metric to quadratic order in $x$. This simplifies things tremendously since

$$\gamma_{ij} = \delta_{ij} + kx_i x_j + \mathcal{O}(x^4)$$

Similarly, we have

$$\gamma^{ij} = \delta^{ij} - kx^i x^j + \mathcal{O}(x^4)$$

where $i, j$ indices are raised and lowered using $\delta^{ij}$. Plugging these forms into the expression for the Christoffel symbols gives

$$\Gamma^i_{jk} = kx^i\delta_{jk} + \mathcal{O}(x^3)$$

With this in hand, we can compute the Ricci tensor

$$\begin{aligned} R_{ij} &= \partial_\rho \Gamma^\rho_{ij} - \partial_j \Gamma^\rho_{\rho i} + \Gamma^\lambda_{ij}\Gamma^\rho_{\rho\lambda} - \Gamma^\lambda_{\rho i}\Gamma^\rho_{j\lambda} \\ &= (\partial_0 \Gamma^0_{ij} + \partial_k \Gamma^k_{ij}) - \partial_j \Gamma^k_{ki} + (\Gamma^0_{ij}\Gamma^k_{k0} + \Gamma^k_{ij}\Gamma^l_{lk}) - (\Gamma^0_{ki}\Gamma^k_{j0} + \Gamma^k_{0i}\Gamma^0_{jk} + \Gamma^k_{li}\Gamma^l_{jk}) \end{aligned}$$

We can drop the $\Gamma^k_{ij}\Gamma^l_{lk}$ term since it vanishes at $\mathbf{x} = 0$. Furthermore, we can now safely replace any undifferentiated $\gamma_{ij}$ in the Christoffel symbols with $\delta_{ij}$. What's left gives

$$\begin{aligned} R_{ij} &= \left(\partial_0(a\dot{a}) + 3k - k + 3\dot{a}^2 - \dot{a}^2 - \dot{a}^2\right)\delta_{ij} + \mathcal{O}(x^2) \\ &= \left(a\ddot{a} + 2\dot{a}^2 + 2k\right)\delta_{ij} + \mathcal{O}(x^2) \end{aligned}$$

We now invoke the covariance argument to write

$$R_{ij} = \left(a\ddot{a} + 2\dot{a}^2 + 2k\right)\gamma_{ij} = \frac{1}{a^2}\left(a\ddot{a} + 2\dot{a}^2 + 2k\right)g_{ij}$$

as promised $\square$

With these results, we can now compute the Ricci scalar: it is

$$R = 6\left(\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right)$$

Finally, the Einstein tensor has components

$$G_{00} = 3\left(\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right) \quad \text{and} \quad G_{ij} = -\left(2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right)g_{ij}$$

Our next task is to understand the matter content in the universe.

### 4.6.2 The Friedmann Equations

We take the universe to be filled with perfect fluids of the kind that we introduced in Section 4.5.4. The energy momentum tensor is

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + P g^{\mu\nu}$$

But we assume that the fluid is at rest in the preferred frame of the universe, meaning that $u^\mu = (1, 0, 0, 0)$ in the FRW coordinates (4.60). As we saw in Section 4.5.4, the constraint $\nabla_\mu T^{\mu\nu} = 0$ gives the condition (4.51)

$$u^\mu \nabla_\mu \rho + (\rho + P)\nabla_\mu u^\mu = 0 \quad \Rightarrow \quad \dot{\rho} + \frac{3\dot{a}}{a}(\rho + P) = 0$$

where we've used $\nabla_\mu u^\mu = \partial_\mu u^\mu + \Gamma^\mu_{\mu\rho}u^\rho = \Gamma^i_{i0}u^0$, and the expression (4.61) for the Christoffel symbols. This is known as the *continuity equation*: it expresses the conservation of energy in an expanding universe. You can check that the second constraint (4.52) is trivial when applied to homogeneous and isotropic fluids. (It plays a role when we consider the propagation of sound waves in the universe.)

To make progress, we also need the equation of state. The fluids of interest have rather simple equations of state, taking the form

$$P = w\rho$$

with constant $w$. Of particular interest are the cases $w = 0$, corresponding to pressureless dust, and $w = 1/3$ corresponding to radiation.

For a given equation of state, the continuity equation becomes

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a}$$

So we learn that the energy density $\rho$ dilutes as the universe expands, with

$$\rho = \frac{\rho_0}{a^{3(1+w)}} \tag{4.62}$$

with $\rho_0$ an integration constant. For pressureless dust, we have $\rho \sim 1/a^3$ which is the expected scaling of energy density with volume. For radiation we have $\rho \sim 1/a^4$, which is the due to the scaling with volume together with an extra factor from redshift.

Now we can look at the Einstein equations. The temporal component is

$$G_{00} + \Lambda g_{00} = 8\pi G T_{00} \quad \Rightarrow \quad \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \tag{4.63}$$

This is the *Friedmann equation.* In conjunction with (4.62), it tells us how the universe expands.

We also have the spatial components of the Einstein equation,

$$\begin{aligned} G_{ij} + \Lambda g_{ij} = 8\pi G T_{ij} \quad &\Rightarrow \quad 2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} - \Lambda = -8\pi G P \\ &\Rightarrow \quad \frac{\ddot{a}}{a} - \frac{\Lambda}{3} = -\frac{4\pi G}{3}(\rho + 3P) \end{aligned}$$

This is the *Raychaudhuri equation.* It is not independent of the Friedmann equation; if you differentiate (4.63) with respect to time, you can derive Raychaudhuri.

There is plenty of physics hiding in these equations. Some particularly simple solutions can be found by setting $k = \Lambda = 0$ and looking at a universe dominated by a single fluid with energy density scaling as (4.62). The Friedmann equation becomes

$$\left(\frac{\dot{a}}{a}\right)^2 \sim \frac{1}{a^{3(1+w)}} \quad \Rightarrow \quad a(t) = \left(\frac{t}{t_0}\right)^{2/(3+3w)}$$

Picking $w = 1/3$ we have $a(t) \sim t^{1/2}$ which describes the expansion of our universe when it was dominated by radiation (roughly the first 50,000 years). Picking $w = 0$ we have $a(t) \sim t^{2/3}$ which describes the expansion of our universe when it was dominated by matter (roughly the following 10 billion years). You can find many more solutions of the Friedmann equations and a discussion of the relevant physics in the lectures on Cosmology.

# 5. When Gravity is Weak

The elegance of the Einstein field equations ensures that they hold a special place in the hearts of many physicists. However, any fondness you may feel for these equations will be severely tested if you ever try to solve them. The Einstein equations comprise ten, coupled partial differential equations. While a number of important solutions which exhibit large amounts symmetry are known, the general solution remains a formidable challenge.

We can make progress by considering situations in which the metric is almost flat. We work with $\Lambda = 0$ and consider metrics which, in so-called *almost-inertial coordinates* $x^\mu$, takes the form

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{5.1}$$

Here $\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$ is the Minkowski metric. The components $h_{\mu\nu}$ are assumed to be small perturbation of this metric: $h_{\mu\nu} \ll 1$.

Our strategy is to expand the Einstein equations to linear order in the small perturbation $h_{\mu\nu}$. At this order, we can think of gravity as a symmetric "spin 2" field $h_{\mu\nu}$ propagating in flat Minkowski space $\eta_{\mu\nu}$. To this end, all indices will now be raised and lowered with $\eta_{\mu\nu}$ rather than $g_{\mu\nu}$. For example, we have

$$h^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}$$

Our theory will exhibit a Lorentz invariance, under which $x^\mu \to \Lambda^\mu{}_\nu x^\nu$ and the gravitational field transforms as

$$h^{\mu\nu}(x) \to \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \, h^{\rho\sigma}(\Lambda^{-1}x)$$

In this way, we construct a theory around flat space that starts to look very much like the other field theories that we meet in physics.

## 5.1 Linearised Theory

To proceed, we need to construct the various curvature tensors from the metric (5.1). For each, we work at linear order in $h$. To leading order, the inverse metric is

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$$

The Christoffel symbols are then

$$\Gamma^\sigma_{\nu\rho} = \frac{1}{2}\eta^{\sigma\lambda}\left(\partial_\nu h_{\lambda\rho} + \partial_\rho h_{\nu\lambda} - \partial_\lambda h_{\nu\rho}\right) \tag{5.2}$$

The Riemann tensor is

$$R^\sigma{}_{\rho\mu\nu} = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\lambda_{\nu\rho}\Gamma^\sigma_{\mu\lambda} - \Gamma^\lambda_{\mu\rho}\Gamma^\sigma_{\nu\lambda}$$

The $\Gamma\Gamma$ terms are second order in $h$, so to linear order we have

$$\begin{aligned} R^\sigma{}_{\rho\mu\nu} &= \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} \\ &= \frac{1}{2}\eta^{\sigma\lambda}\left(\partial_\mu\partial_\rho h_{\nu\lambda} - \partial_\mu\partial_\lambda h_{\nu\rho} - \partial_\nu\partial_\rho h_{\mu\lambda} + \partial_\nu\partial_\lambda h_{\mu\rho}\right) \end{aligned} \tag{5.3}$$

The Ricci tensor is then

$$R_{\mu\nu} = \frac{1}{2}\left(\partial^\rho\partial_\mu h_{\nu\rho} + \partial^\rho\partial_\nu h_{\mu\rho} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h\right)$$

with $h = h^\mu{}_\mu$ the trace of $h_{\mu\nu}$ and $\Box = \partial^\mu\partial_\mu$. The Ricci scalar is

$$R = \partial^\mu\partial^\nu h_{\mu\nu} - \Box h \tag{5.4}$$

By the time we get to the Einstein tensor, we've amassed quite a collection of terms

$$G_{\mu\nu} = \frac{1}{2}\Big[\partial^\rho\partial_\mu h_{\nu\rho} + \partial^\rho\partial_\nu h_{\mu\rho} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h - \left(\partial^\rho\partial^\sigma h_{\rho\sigma} - \Box h\right)\eta_{\mu\nu}\Big] \tag{5.5}$$

The Bianchi identity for the full Einstein tensor is $\nabla^\mu G_{\mu\nu} = 0$. For the linearised Einstein tensor, this reduces to

$$\partial^\mu G_{\mu\nu} = 0 \tag{5.6}$$

It's simple to check explicitly that this is indeed obeyed by (5.5).

The Einstein equations then become the linear, but somewhat complicated, set of partial differential equations

$$\partial^\rho\partial_\mu h_{\nu\rho} + \partial^\rho\partial_\nu h_{\mu\rho} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h - \left(\partial^\rho\partial^\sigma h_{\rho\sigma} - \Box h\right)\eta_{\mu\nu} = 16\pi G T_{\mu\nu} \tag{5.7}$$

where, for consistency, the source $T_{\mu\nu}$ must also be suitably small. The left-hand side of this equation should be viewed as a second order, linear differential operator acting on $h_{\mu\nu}$. This is known as the *Lichnerowicz operator*.

**The Fierz-Pauli Action**

The linearised equations of motion can be derived from an action principle, first written down by Fierz and Pauli,

$$S_{FP} = \frac{1}{8\pi G}\int d^4x\ \left[-\frac{1}{4}\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + \frac{1}{2}\partial_\rho h_{\mu\nu}\partial^\nu h^{\rho\mu} + \frac{1}{4}\partial_\mu h\partial^\mu h - \frac{1}{2}\partial_\nu h^{\mu\nu}\partial_\mu h\right] \tag{5.8}$$

This is the expansion of the Einstein-Hilbert action to quadratic order in $h$ (after some integration by parts). (At linear order, the expansion of the Lagrangian is equal to the linearised Ricci scalar (5.4) which is a total derivative.)

Varying the Fierz-Pauli action, and performing some integration by parts, we have

$$\begin{aligned}\delta S_{FP} &= \frac{1}{8\pi G}\int d^4x\ \left[\frac{1}{2}\partial_\rho\partial^\rho h_{\mu\nu} - \partial^\rho\partial_\nu h_{\rho\mu} - \frac{1}{2}\partial^\rho\partial_\rho h\eta_{\mu\nu} + \frac{1}{2}\partial_\nu\partial_\mu h + \frac{1}{2}\partial_\rho\partial_\sigma h^{\rho\sigma}\eta_{\mu\nu}\right]\delta h^{\mu\nu} \\ &= \frac{1}{8\pi G}\int d^4x\ \Big[-G_{\mu\nu}\,\delta h^{\mu\nu}\Big] \end{aligned} \tag{5.9}$$

We see that the Fierz-Pauli action does indeed give the vacuum Einstein equations $G_{\mu\nu} = 0$. We can then couple matter by adding $T_{\mu\nu}h^{\mu\nu}$ to the action.

### 5.1.1 Gauge Symmetry

Linearised gravity has a rather pretty gauge symmetry. This is inherited from the diffeomorphisms of the full theory. To see this, we repeat our consideration of infinitesimal diffeomorphisms from Section 4.1.3. Under an infinitesimal change of coordinates

$$x^\mu \to x^\mu - \xi^\mu(x)$$

with $\xi$ assumed to be small. The metric changes by (4.6)

$$\delta g_{\mu\nu} = (\mathcal{L}_\xi g)_{\mu\nu} = \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu$$

When the metric takes the form (5.1), this can be viewed as a transformation of the linearised field $h_{\mu\nu}$. Because both $\xi$ and $h$ are small, the covariant derivative should be taken using the vanishing connection of Minkowski space. We then have

$$h_{\mu\nu} \to h_{\mu\nu} + (\mathcal{L}_\xi\eta)_{\mu\nu} = h_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu \tag{5.10}$$

This looks very similar to the gauge transformation of Maxwell theory, where the gauge potential shifts as $A_\mu \to A_\mu + \partial_\mu\alpha$. Just as the electromagnetic field strength $F_{\mu\nu} = 2\partial_{[\mu}A_{\nu]}$ is gauge invariant, so is the linearised Riemann tensor $R^\sigma{}_{\rho\mu\nu}$.

We can quickly check that the Fierz-Pauli action is invariant under the gauge symmetry (5.10). From (5.9), we have

$$\delta S_{FP} = -\frac{1}{8\pi G}\int d^4x\ 2G_{\mu\nu}\partial^\mu\xi^\nu = +\frac{1}{8\pi G}\int d^4x\ 2(\partial^\mu G_{\mu\nu})\xi^\nu = 0$$

where, in the second equality, we've integrated by parts (and discarded the boundary term) and in the third equality we've invoked the linearised Bianchi identity (5.6). In fact, this is just the same argument that we used to derive the Bianchi identity in Section 4.1.3, now played backwards.

When doing calculations in electromagnetism, it's often useful to pick a gauge. One of the most commonly used is Lorentz gauge,

$$\partial^\mu A_\mu = 0$$

Once we impose this condition, the Maxwell equations $\partial^\mu F_{\mu\nu} = j_\nu$ reduce to the wave equations

$$\Box A_\nu = j_\nu$$

We solved these equations in detail in the lectures on Electromagnetism.

We can impose a similar gauge fixing condition in linearised gravity. In this case, the analog of Lorentz gauge is called *de Donder gauge*

$$\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h = 0 \tag{5.11}$$

To see that this is always possible, suppose that you are handed a metric that doesn't obey the de Donder condition but instead satisfies $\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h = f_\nu$ for some functions $f_\nu$. Then do a gauge transformation (5.10). Your new gauge potential will satisfy $\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h + \Box\xi_\nu = f_\nu$. So if you pick a gauge transformation $\xi_\mu$ that obeys $\Box\xi_\mu = f_\mu$ then your new metric will be in de Donder gauge.

There is a version of de Donder gauge condition (5.11) that we can write down in the full non-linear theory. We won't need it in this course, but it's useful to know it exists. It is

$$g^{\mu\nu}\Gamma^\rho_{\mu\nu} = 0 \tag{5.12}$$

This isn't a tensor equation because the connection $\Gamma^\rho_{\mu\nu}$ is not a tensor. Indeed, if a tensor vanishes in one choice of coordinates then it vanishes for all choices while the whole point of a gauge fixing condition is to pick out a preferred choice of coordinates. If we substitute in the linearised Christoffel symbols (5.2), this reduces to the de Donder gauge condition.

The non-linear gauge condition (5.12) has a number of nice features. For example, in general the wave operator (or, on a Riemannian manifold, the Laplacian $\triangle$) is $\Box = \nabla^\mu\nabla_\mu = g^{\mu\nu}(\partial_\nu\partial_\mu - \Gamma^\rho_{\nu\mu}\partial_\rho)$. If we fix the gauge (5.12), the annoying connection term vanishes and we simply have $\Box = g^{\mu\nu}\partial_\mu\partial_\nu$. A similar simplification happens if we compute the covariant divergence of a one-form in this gauge: $\nabla^\mu\omega_\mu = g^{\mu\nu}\nabla_\mu\omega_\nu = g^{\mu\nu}(\partial_\mu\omega_\nu - \Gamma^\rho_{\mu\nu}\omega_\rho) = \partial^\mu\omega_\mu$.

Back in our linearised world, de Donder gauge greatly simplifies the Einstein equation (5.7), which now become

$$\Box h_{\mu\nu} - \frac{1}{2}\Box h \eta_{\mu\nu} = -16\pi G T_{\mu\nu} \tag{5.13}$$

It is useful to define

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu}$$

Taking the trace of both sides gives $\bar{h} = \eta^{\mu\nu}\bar{h}_{\mu\nu} = -h$ so, given $\bar{h}_{\mu\nu}$ we can trivially reconstruct $h_{\mu\nu}$ as

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\bar{h}\eta_{\mu\nu} \tag{5.14}$$

Written in terms of $\bar{h}_{\mu\nu}$, the linearised Einstein equations in de Donder gauge (5.13) then reduce once again to a bunch of wave equations

$$\Box \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu} \tag{5.15}$$

and we can simply import the solutions from electromagnetism to learn something about gravity. We'll look at some examples shortly.

### 5.1.2 The Newtonian Limit

Under certain circumstances, the linearised equations of general relativity reduce to the familiar Newtonian theory of gravity. These circumstances occur when we have a low-density, slowly moving distribution of matter.

For simplicity, we'll look at a stationary matter configuration. This means that we take

$$T_{00} = \rho(\mathbf{x})$$

with the other components vanishing. Since nothing depends on time, we can replace the wave operator is just the Laplacian in $\mathbf{R}^3$: $\Box = -\partial_t^2 + \nabla^2 = \nabla^2$. The Einstein equations are then simply

$$\nabla^2 \bar{h}_{00} = -16\pi G \rho(\mathbf{x}) \quad \text{and} \quad \nabla^2 \bar{h}_{0i} = \nabla^2 \bar{h}_{ij} = 0$$

With suitable boundary conditions, the solutions to these equations are

$$\bar{h}_{00} = -4\Phi(\mathbf{x}) \quad \text{and} \quad \bar{h}_{0i} = \bar{h}_{ij} = 0 \tag{5.16}$$

where the field $\Phi$ is identified with the Newtonian gravitational potential, obeying (0.1)

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

Translating this back to $h_{\mu\nu}$ using (5.14), we use $\bar{h} = +4\Phi$ to find

$$h_{00} = -2\Phi \;, \quad h_{ij} = -2\Phi\delta_{ij} \;, \quad h_{0i} = 0$$

Putting this back into the full metric $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, we have

$$ds^2 = -(1+2\Phi)dt^2 + (1-2\Phi)d\mathbf{x}\cdot d\mathbf{x}$$

If we take a $\Phi = -GM/r$ as expected for a point mass, we find that this coincides with the leading expansion of the Schwarzchild metric (4.8). (The $g_{00}$ term turns out to be exact; the $g_{ij}$ term is the leading order Taylor expansion of $(1+2\Phi)^{-1}$.)

Way back in Section 1.2, we gave a naive, intuitive discussion of curved spacetime. There we already anticipated that the Newtonian potential $\Phi$ would appear in the $g_{00}$ component of the metric (1.26). However, in solving the Einstein equations, we learn that this is necessarily accompanied by an appearance of $\Phi$ in the $g_{ij}$ component. Ultimately, this is the reason for the factor of 2 discrepancy between the Newtonian and relativistic predictions for light bending that we met in Section 1.3

## 5.2 Gravitational Waves

A long time ago, in a galaxy far far away, two black holes collided. Here a "long time ago" means 1.3 billion years ago. And "far far away" means a distance of about 1.3 billion light years.

To say that this was a violent event is something of an understatement. One of the black holes was roughly 35 times heavier than the Sun, the other about 30 times heavier. When they collided they merged to form a black hole whose mass was about 62 times heavier than the Sun. Now $30 + 35 \neq 62$. This means that some mass, or equivalently energy, went missing during the collision. In a tiny fraction of a second, this pair of black holes emitted an energy equivalent to three times the mass of the Sun.

That, it turns out, is quite a lot of energy. For example, nuclear bombs convert the mass of a handful of atoms into energy. But here we're talking about solar masses, not atomic masses. In fact, for that tiny fraction of a second, these colliding black holes released more energy than all the stars in all the galaxies in the visible universe put together.

But the most astonishing part of the story is how we know this collision happened. It's because, on September 14th, 2015, at 9.30 in the morning UK time, we felt it. The collision of the black holes was so violent that it caused an enormous perturbation of spacetime. Like dropping a stone in a pond, these ripples propagated outwards as gravitational waves. These ripples started 1.3 billion years ago, roughly at the time that multi-cellular life was forming here on Earth. They then travelled through the cosmos at the speed of light. The ripples hit the outer edge of our galaxy about 50,000 years ago, at a time when humans were hanging out with neanderthals. The intervening 50,000 years gave us just enough time to band together into hunter-gatherer tribes, develop cohesive societies bound by false religions, invent sophisticated language and writing, discover mathematics, understand the theory that governs the spacetime continuum and, finally, build a machine that is capable of detecting the ripples, turning it on just in time for the gravitational wave to hit the south pole and pass, up through the Earth, triggering the detector.

The purpose of this section is to tell the story above in equations.

### 5.2.1 Solving the Wave Equation

Gravitational waves propagate in vacuum, in the absence of any sources. This means that we need to solve the linearised equation

$$\Box \bar{h}_{\mu\nu} = 0 \tag{5.17}$$

One solution is provided by the gravitational wave

$$\bar{h}_{\mu\nu} = \mathrm{Re}\left(H_{\mu\nu}\, e^{ik_\rho x^\rho}\right) \tag{5.18}$$

Here $H_{\mu\nu}$ a complex, symmetric polarisation matrix and the wavevector $k^\mu$ is a real 4-vector. Usually when writing these solutions we are lazy and drop the Re on the right-hand side, leaving it implicit that one takes the real part. This plane wave ansatz solves the linearised Einstein equation (5.17) provided that the wavevector is null,

$$k_\mu k^\mu = 0$$

This tells us that gravitational waves, like light waves, travel at the speed of light. If we write the wavevector as $k^\mu = (\omega, \mathbf{k})$, with $\omega$ the frequency, then this condition becomes $\omega = \pm|\mathbf{k}|$.

Because the wave equation is linear, we may superpose as many different waves of the form (5.18) as we wish. In this way, we build up the most general solution to the wave equation.

Naively, the polarisation matrix $H_{\mu\nu}$ has 10 components. But we still have to worry about gauge issues. The ansatz (5.18) satisfies the de Donder gauge condition $\partial^\mu \bar{h}_{\mu\nu} = 0$ only if

$$k^\mu H_{\mu\nu} = 0 \tag{5.19}$$

This tells us that the polarisation is transverse to the direction of propagation. Furthermore, the choice of de Donder gauge does not exhaust our ability to make gauge transformations. If we make a further gauge transformation $h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$, then

$$\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \partial^\rho \xi_\rho \eta_{\mu\nu}$$

This transformation leaves the solution in de Donder gauge $\partial^\mu \bar{h}_{\mu\nu} = 0$ provided that

$$\Box \xi_\nu = 0$$

In particular, we can take

$$\xi_\mu = \lambda_\mu \, e^{ik_\rho x^\rho}$$

which obeys $\Box \xi_\mu = 0$ because $k_\rho k^\rho = 0$. A gauge transformation of this type shifts the polarisation matrix to

$$H_{\mu\nu} \to H_{\mu\nu} + i \left( k_\mu \lambda_\nu + k_\nu \lambda_\mu - k^\rho \lambda_\rho \eta_{\mu\nu} \right) \tag{5.20}$$

Polarisation matrices that differ in this way describe the same gravitational wave. We now choose the gauge transformation $\lambda_\mu$ in order to further set

$$H_{0\mu} = 0 \quad \text{and} \quad H^\mu{}_\mu = 0 \tag{5.21}$$

These conditions, in conjunction with (5.19), are known as *transverse traceless gauge*. Because $H$ is traceless, this choice of gauge has the advantage that $h_{\mu\nu} = \bar{h}_{\mu\nu}$.

At this stage we can do some counting. The polarisation matrix $H_{\mu\nu}$ has 10 components. The de Donder condition (5.19) gives 4 constraints, and there are 4 residual gauge transformations (5.20). The upshot is that there are just $10 - 4 - 4 = 2$ independent polarisations in $H_{\mu\nu}$.

(There is a similar counting in Maxwell theory. The polarisation of $A_\mu$ seemingly has 4 components. The Lorentz gauge $\partial^\mu A_\mu = 0$ kills one of them, and a residual gauge transformation kills another, leaving the 2 familiar polarisation states of light.)

#### An Example

Consider a wave propagating in the $z$ direction. The wavevector is

$$k_\mu = (\omega, 0, 0, \omega)$$

The condition (5.19) sets $H_{0\nu} + H_{3\nu} = 0$. The additional constraint (5.21) restricts the polarisation matrix to be

$$H_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & H_+ & H_X & 0 \\ 0 & H_X & -H_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.22}$$

Both $H_+$ and $H_X$ can be complex; we take the real part when computing the metric in (5.18). Here we see explicitly the two polarisation states $H_+$ and $H_X$. We'll see below how to interpret these two polarisations.

### 5.2.2 Bobbing on the Waves

What do you feel if a gravitational wave passes you by? Well, if you're happy to be modelled as a pointlike particle, moving along a geodesic, then the answer is simple: you feel nothing at all. This follows from the equivalence principle. Instead, it's all about your standing relative to your neighbours.

This relative physics is captured by the geodesic deviation equation that we met in Section 3.3.4. Consider a family of geodesics $x^\mu(\tau, s)$, with $s$ labelling the different geodesics, and $\tau$ the affine parameter along any geodesic. The vector field tangent to these geodesics is the velocity 4-vector

$$u^\mu = \left.\frac{\partial x^\mu}{\partial \tau}\right|_s$$

Meanwhile, the displacement vector $S^\mu$ takes us between neighbouring geodesics,

$$S^\mu = \left.\frac{\partial x^\mu}{\partial s}\right|_\tau$$

We previously derived the geodesic deviation equation (3.36)

$$\frac{d^2 S^\mu}{d\tau^2} = R^\mu{}_{\rho\sigma\nu} u^\rho u^\sigma S^\nu$$

We'll consider the situation where, in the absence of the gravitational wave, our family of geodesics are sitting happily in a rest frame, with $u^\mu = (1,0,0,0)$. As the gravitational wave passes, the geodesics will change as

$$u^\mu = (1,0,0,0) + \mathcal{O}(h)$$

Fortunately, we won't need to compute the details of this. We will compute the deviation to leading order in the metric perturbation $h$, but the Riemann tensor is already $\mathcal{O}(h)$, which means that we can neglect the corrections in the other terms. Similarly, we can replace the proper time $\tau$ for the coordinate time $t$. We then have

$$\frac{d^2 S^\mu}{dt^2} = R^\mu{}_{00\nu} S^\nu$$

The Riemann tensor in the linearised regime was previously computed in (5.3)

$$R^\mu{}_{\rho\sigma\nu} = \frac{1}{2} g^{\mu\lambda} \left( \partial_\sigma \partial_\rho h_{\nu\lambda} - \partial_\sigma \partial_\lambda h_{\nu\rho} - \partial_\nu \partial_\rho h_{\sigma\lambda} + \partial_\nu \partial_\lambda h_{\sigma\rho} \right)$$

Using $h_{\mu 0} = 0$, the component we need is simply

$$R^\mu{}_{00\nu} = \frac{1}{2} \partial_0^2 h^\mu{}_\nu$$

Our geodesic deviation equation is then

$$\frac{d^2 S^\mu}{dt^2} = \frac{1}{2} \frac{d^2 h^\mu{}_\nu}{dt^2} S^\nu \tag{5.23}$$

We see that the gravitational wave propagating, say, the $z$ direction with polarisation vector (5.22) affects neither $S^0$ nor $S^3$. The only effect on the geodesics is in the $(x,y)$-plane, transverse to the direction of propagation. For simplicity, we will solve this equation in the $z = 0$ plane.

<u>$H_+$ Polarisation:</u> If we set $H_X = 0$ in (5.22), then the geodesic deviation equation (5.23) becomes

$$\frac{d^2 S^1}{dt^2} = -\frac{\omega^2}{2} H_+ e^{i\omega t} S^1 \quad \text{and} \quad \frac{d^2 S^2}{dt^2} = +\frac{\omega^2}{2} H_+ e^{i\omega t} S^2$$

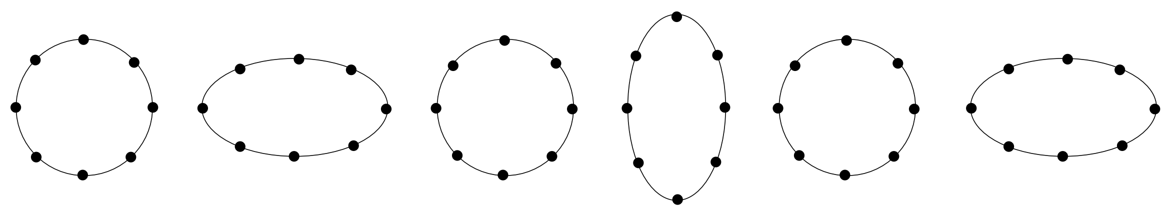
We can solve these perturbatively in $H_+$. Keeping terms of order $\mathcal{O}(h)$ only, we have

$$S^1(t) = S^1(0)\left(1 + \frac{1}{2} H_+ e^{i\omega t} + \ldots\right) \quad \text{and} \quad S^2(t) = S^2(0)\left(1 - \frac{1}{2} H_+ e^{i\omega t} + \ldots\right) \tag{5.24}$$

where, as we mentioned previously, we should take the real part of the right-hand-side. (Recall that $H_+$ can also be complex.)

From these solutions, we can determine the way in which geodesics are affected by a passing wave. Think of the displacement vector $S^\mu$ as the distance from the origin to a neighbouring geodesic. We will consider a family of neighbouring geodesics corresponding to a collection of particles which, at time $t=0$, are arranged around a circle of radius $R$. This means that we have initial conditions $S^a(t=0)$ satisfying $S^1(0)^2 + S^2(0)^2 = R^2$.

The solutions (5.24) tell us how these geodesics evolve. The relative minus sign between the two equations means that when geodesics move outwards in, say, the $x^1 = x$ direction, they move inwards in the $x^2 = y$ direction, and vice-versa. The net result is that, as time goes on, these particles will evolve from a circle to an ellipse and back again, displaced like this:



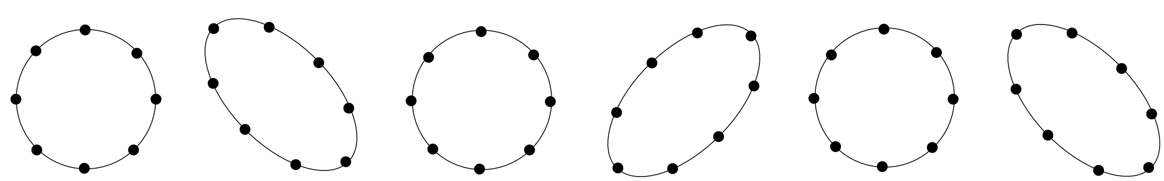
<u>$H_X$ Polarisation:</u> If we set $H_+ = 0$ in (5.22), then the geodesic deviation equation (5.23) becomes

$$\frac{d^2S^1}{dt^2} = -\frac{\omega^2}{2} H_X e^{i\omega t} S^2 \quad \text{and} \quad \frac{d^2S^2}{dt^2} = -\frac{\omega^2}{2} H_X e^{i\omega t} S^1$$

Again, we solve these perturbatively in $H_X$. We have

$$S^1(t) = S^1(0) + \frac{1}{2} S^2(0) H_X e^{i\omega t} + \ldots \quad \text{and} \quad S^2(t) = S^2(0) + \frac{1}{2} S^1(0) H_X e^{i\omega t} + \ldots$$

The displacement is the same as previously, but rotated by 45°. (To see this, note that the displacements $S^1(t) \pm S^2(t)$ have the same functional form as (5.24).) This means that this time the displacement of geodesics looks like this:



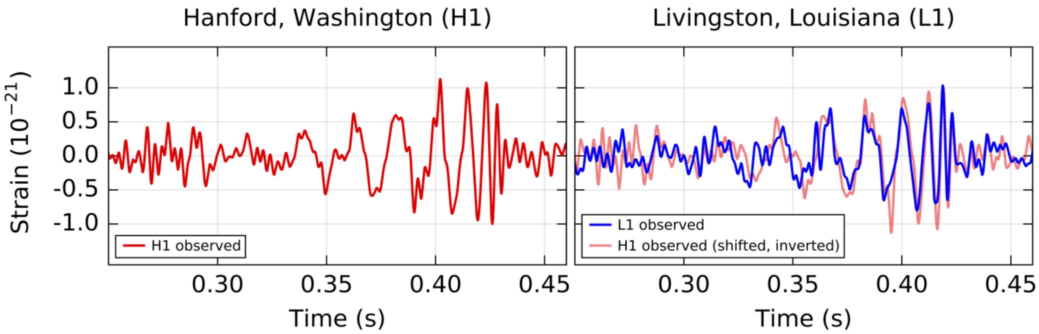



**Figure 42:** The discovery of gravitational waves by the LIGO detectors.

We can also take linear combinations of the polarisation states. Adding the two polarisations above give an elliptic displacement whose axis rotates. This is analogous to the circular polarisation of light.

The displacements due to gravitational waves are invariant under rotations by $\pi$. This contrasts with polarisation of light which is described by a vector, and so is only invariant under $2\pi$ rotations. This reflects the fact that graviton has spin 2, while the photon has spin 1.

### Gravitational Wave Detectors

Gravitational wave detectors are interferometers. They bounce light back and forth between two arms, with the mirrors at either end playing the role of test masses.

If the gravitational wave travels perpendicular to the plane of the detector, it will shorten one arm and lengthen the other. With the arms aligned along the $x$ and $y$ axes, the maximum change in length can be read from (5.24),

$$L' = L\left(1 \pm \frac{H_+}{2}\right) \quad \Rightarrow \quad \frac{\delta L}{L} = \frac{H_+}{2}$$

To get a ballpark figure for this, we need to understand how large we expect $H_+$ to be from any plausible astrophysical source. We'll do this in Section 5.3.2. It turns out it's not really very large at all: typical sources have $H_+ \sim 10^{-21}$. The lengths of each arm in the LIGO detectors is around $L \sim 3$ km, meaning that we have to detect a change in length of $\delta L \sim 10^{-18}$ m. This seems like a crazy small number: it's smaller than the radius of a proton, and around $10^{12}$ times smaller than the wavelength of the light used in the interferometer. Nonetheless, the sensitivity of the detectors is up to the task and the LIGO observatories detected gravitational waves for the first time in 2015. For this, three members of the collaboration were awarded the 2017 Nobel prize. Subsequently, the LIGO and VIRGO detectors have observed a large number of mergers involving black holes and neutron stars.

### 5.2.3 Exact Solutions

We have found a wave-like solution to the linearised Einstein equations. The metric for a wave moving in, say, the positive $z$ direction takes the form

$$ds^2 = -dt^2 + (\delta_{ab} + h_{ab}(z-t))dx^a dx^b + dz^2 \tag{5.25}$$

where the $a, b = 1, 2$ indices run over the spatial directions transverse to the direction of the wave. Because the wave equation is linear, any function $h_{ab}(z-t)$ is a solution to the linearised Einstein equations; the form that we gave in (5.18) is simply the Fourier decomposition of the general solution.

Because gravitational waves are so weak, the linearised metric is entirely adequate for any properties of gravitational waves that we wish to calcuate. Nonetheless, it's natural to ask if this solution has an extension to the full non-linear Einstein equations. Rather surprisingly, it turns out that there is.

For a wave propagating in the positive $z$ direction, we first introduce lightcone coordinates

$$u = t - z \quad , \quad v = t + z$$

Then we consider the *plane wave* ansatz, sometimes called the *Brinkmann metric*

$$ds^2 = -du dv + dx^a dx^a + H_{ab}(u) x^a x^b du^2$$

Note that our linearised gravitational wave (5.25) is not of this form; there is some (slightly fiddly) change of coordinates that takes us between the two metrics. One can show that the Brinkmann metric is Ricci flat, and hence solves the vacuum Einstein equations, for *any* traceless metric $H_{ab}$

$$R_{\mu\nu} = 0 \quad \Leftrightarrow \quad H^a{}_a(u) = 0$$

The general metric again has two independent polarisation states,

$$H_{ab}(u) = \begin{pmatrix} H_{11}(u) & H_{12}(u) \\ H_{12}(u) & -H_{11}(u) \end{pmatrix}$$

It is unusual to find solutions on non-linear PDEs which depend on arbitrary functions, like $H_{11}(u)$ and $H_{12}(u)$. The Brinkmann metrics are a rather special exception.

## 5.3 Making Waves

The gravitational wave solutions described in the previous section are plane waves. They come in from infinity, and go out to infinity. In reality however, gravitational waves start at some point and radiate out.

As we will see, the story is entirely analogous to what we saw in our earlier course on Electromagnetism. There, you generate electromagnetic waves by shaking electric charges. Similarly, we generate gravitational waves by shaking masses. The purpose of this section is to make this precise.

### 5.3.1 The Green's Function for the Wave Equation

Our starting point is the linearised Einstein equation (5.15),

$$\Box\bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu} \tag{5.26}$$

which assumes that both the source, in the guise of the energy momentum tensor $T_{\mu\nu}$, and the perturbed metric $\bar{h}_{\mu\nu}$ are small. This is simply a bunch of decoupled wave equations. We already solved these in Section 6 of the lectures on Electromagnetism, and our discussion here will parallel the presentation there.

We will consider a situation in which matter fields are localised to some spatial region $\Sigma$. In this region, there is a time-dependent source of energy and momentum $T_{\mu\nu}(\mathbf{x}', t)$, such as two orbiting black holes. Outside of this region, the energy-momentum tensor vanishes: $T_{\mu\nu}(\mathbf{x}', t) = 0$ for $\mathbf{x}' \notin \Sigma$. We want to know what the metric $h_{\mu\nu}$ looks like a long way from the region $\Sigma$. The solution to (5.26) outside of $\Sigma$ can be given using the (retarded) Green's function; it is

$$\bar{h}_{\mu\nu}(\mathbf{x}, t) = 4G \int_\Sigma d^3x' \; \frac{T_{\mu\nu}(\mathbf{x}', t_{\rm ret})}{|\mathbf{x} - \mathbf{x}'|} \tag{5.27}$$

here $t_{\rm ret}$ is the retarded time, given by

$$t_{\rm ret} = t - |\mathbf{x} - \mathbf{x}'|$$

It's not too hard to show that this solution satisfies the de Donder gauge condition $\partial^\mu \bar{h}_{\mu\nu} = 0$ provided that the energy momentum tensor is conserved, $\partial^\mu T_{\mu\nu} = 0$. The solution does not, however, automatically satisfy the temporal and traceless conditions (5.21). The solution (5.27) captures the causality of the wave equation: the gravitational field $\bar{h}_{\mu\nu}(\mathbf{x}, t)$ is influenced by the matter at position $\mathbf{x}'$ at the earlier time $t_{\rm ret}$, so that there is time for this influence to propagate from $\mathbf{x}'$ to $\mathbf{x}$.

We denote the size of the region $\Sigma$ as $d$. We're interested in what's happening at a point $\mathbf{x}$ which is a distance $r = |\mathbf{x}|$ away. If $|\mathbf{x} - \mathbf{x}'| \gg d$ for all $\mathbf{x}' \in \Sigma$ then we can approximate

$$|\mathbf{x} - \mathbf{x}'| = r - \frac{\mathbf{x}\cdot\mathbf{x}'}{r} + \ldots \Rightarrow \qquad \frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{r} + \frac{\mathbf{x}\cdot\mathbf{x}'}{r^3} + \ldots \tag{5.28}$$

We also have a factor of $|\mathbf{x} - \mathbf{x}'|$ that sits inside $t_{\rm ret} = t - |\mathbf{x} - \mathbf{x}'|$. This means that we should also Taylor expand the argument of the energy-momentum tensor

$$T_{\mu\nu}(\mathbf{x}', t_{\rm ret}) = T_{\mu\nu}(\mathbf{x}', t - r + \mathbf{x}\cdot\mathbf{x}'/r + \ldots)$$

Now we'd like to further expand out this argument. But, to do that, we need to know something about what the source is doing. We will assume that the motion of matter is *non-relativistic*, so that the energy momentum tensor doesn't change very much over the time $\tau \sim d$ that it takes light to cross the region $\Sigma$. For example, if we have a system comprised of two objects (say, neutron starts or black holes) orbiting each other with characteristic frequency $\omega$ then $T_{\mu\nu} \sim e^{-i\omega t}$ and the requirement that the motion is non-relativistic becomes $d \ll 1/\omega$. Then we can further Taylor expand the current to write

$$T_{\mu\nu}(\mathbf{x}', t_{\rm ret}) = T_{\mu\nu}(\mathbf{x}', t - r) + \dot{T}_{\mu\nu}(\mathbf{x}', t - r)\,\frac{\mathbf{x}\cdot\mathbf{x}'}{r} + \ldots \tag{5.29}$$

We have two Taylor expansions, (5.28) and (5.29). At leading order in $d/r$ we take the first term from both these expansions to find

$$\bar{h}_{\mu\nu}(\mathbf{x}, t) \approx \frac{4G}{r}\int_\Sigma d^3x'\; T_{\mu\nu}(\mathbf{x}', t - r)$$

We first look at the expressions for $\bar{h}_{00}$ and $\bar{h}_{0i}$. The first of these is

$$\bar{h}_{00}(\mathbf{x}, t) \approx \frac{4G}{r}E \quad \text{with} \quad E = \int_\Sigma d^3x'\; T_{00}(\mathbf{x}', t - r) \tag{5.30}$$

This is simply a recapitulation of the Newtonian limit (5.16), with the long distance gravitational potential given by $\Phi = -GE/r$ where $E$ is the total energy inside the region $\Sigma$. At the linear order to which we're working, current conservation $\partial^\mu T_{\mu\nu} = 0$ ensures that the energy $E$ inside $\Sigma$ is constant, so the time dependence drops out.

Similarly, we have

$$\bar{h}_{0i}(\mathbf{x}, t) \approx -\frac{4G}{r}P_i \quad \text{with} \quad P_i = -\int_\Sigma d^3x'\; T_{0i}(\mathbf{x}', t - r) \tag{5.31}$$

Here $P_i$ is the total momentum of the matter inside $\Sigma$ which, again, is conserved. We can always go to a rest frame where this matter is stationary in which case $P_i = 0$ and hence $\bar{h}_{0i} = 0$. This was the choice we implicitly made in describing the Newtonian limit (5.16).

Neither the expression for $\bar{h}_{00}$ nor $\bar{h}_{ij}$ captures the physics that we are interested in. The results only know about the conserved quantities inside the region $\Sigma$, not about how they're moving. However, things become more interesting when we look at the spatial components of the metric,

$$\bar{h}_{ij}(\mathbf{x},t) \approx \frac{4G}{r}\int_\Sigma d^3x'\; T_{ij}(\mathbf{x}',t-r)$$

with $i,j = 1,2,3$. Now the integral on the right-hand side is not a conserved quantity. However, it is possible to relate it to certain properties of the energy distribution inside $\Sigma$.

**Claim:**

$$\int_\Sigma d^3x'\; T_{ij}(\mathbf{x}',t) = \frac{1}{2}\ddot{I}_{ij}(t)$$

where $I_{ij}$ is the quadrupole moment of the energy,

$$I_{ij}(t) = \int_\Sigma d^3x\; T^{00}(\mathbf{x},t)\, x_i x_j \tag{5.32}$$

**Proof:** We start by writing

$$T^{ij} = \partial_k(T^{ik}x^j) - (\partial_k T^{ik})x^j = \partial_k(T^{ik}x^j) + \partial_0 T^{0i}x^j$$

where, in the second equality, we've used current conservation $\partial_\mu T^{\mu\nu} = 0$. (Note that current conservation in the full theory is $\nabla_\mu T^{\mu\nu} = 0$, but in our linearised analysis this reduces to $\partial_\mu T^{\mu\nu} = 0$.) For the $T^{0i}$ term, we play the same trick again. Symmetrising over $(ij)$, we have

$$T^{0(i}x^{j)} = \frac{1}{2}\partial_k(T^{0k}x^ix^j) - \frac{1}{2}(\partial_k T^{0k})x^ix^j = \frac{1}{2}\partial_k(T^{0k}x^ix^j) + \frac{1}{2}\partial_0 T^{00}x^ix^j$$

When we integrate this over $\Sigma$, we drop the terms that are total spatial derivatives. We're left with

$$\int_\Sigma d^3x'\; T^{ij}(\mathbf{x}',t) = \frac{1}{2}\partial_0^2\int_\Sigma d^3x'\; T^{00}(\mathbf{x}',t)x'^ix'^j$$

which is the claimed result. $\square$

We learn that, far from the source, the metric takes the form

$$\bar{h}_{ij}(\mathbf{x},t) \approx \frac{2G}{r}\ddot{I}_{ij}(t-r) \tag{5.33}$$

This is the physics that we want: if we shake the matter distribution in some way then, once the signal has had time to propagate, this will affect the metric. Because the equations are linear, if the matter shakes at some frequency $\omega$ the spacetime will respond by creating waves at parametrically same frequency. (In fact, we'll see a factor of 2 arises in the example of a binary system (5.36).)

In fact, we can now revisit the other components $\bar{h}_{00}$ and $\bar{h}_{0i}$. The gauge condition $\partial^\mu \bar{h}_{\mu\nu} = 0$ tells us that

$$\partial_0 \bar{h}_{0i} = \partial_j \bar{h}_{ji} \quad \text{and} \quad \partial_0 \bar{h}_{00} = \partial_i \bar{h}_{i0}$$

The first of these equations gives

$$\partial_0 \bar{h}_{0i} \approx \partial_j \left( \frac{2G}{r}\ddot{I}_{ij}(t-r) \right) = -\frac{2G\hat{x}_j}{r^2}\ddot{I}_{ij}(t-r) - \frac{2G\hat{x}_j}{r}\dddot{I}_{ij}(t-r) \tag{5.34}$$

where we've used the fact that $\partial_j r = x_j/r = \hat{x}_j$. Which of these two terms in (5.34) is bigger? As we get further from the source, we would expect the second, $1/r$, term to dominate over the first, $1/r^2$ term. But the second term has an extra time derivative, which means an extra factor of the characteristic frequency of the source, $\omega$. This means that the second term dominates provided that $r \gg 1/\omega$ or, in terms of the wavelength $\lambda$ of the emitted gravitational wave, $r \gg \lambda$. This is known as the *far-field zone* or, sometimes, the *radiation zone*. In this regime, we have

$$\bar{h}_{0i} \approx -\frac{2G\hat{x}_j}{r}\ddot{I}_{ij}(t-r)$$

where we've integrated (5.34). In general, the integration constant is given by the $P_i$ term that we previously saw in (5.31). In the answer above, we've set this integration constant to zero by choosing coordinates in which $P_i = 0$, meaning that the centre of mass of the source doesn't move. We can now repeat this to determine $\bar{h}_{00}$. The same argument means that we discard one term, and retain

$$\bar{h}_{00} = \frac{4G}{r}E + \frac{2G\hat{x}_i\hat{x}_j}{r}\ddot{I}_{ij}(t-r)$$

If we tried to compute these $\ddot{I}$ terms in $\bar{h}_{00}$ and $\bar{h}_{0i}$ directly from (5.27), we would have to go to higher order in the expansion. Implementing the gauge condition, as above, saves us this work.

### 5.3.2 An Example: Binary Systems

As an example, consider two stars (or neutron stars, or black holes) each with mass $M$, separated by distance $R$, orbiting in the $(x, y)$ plane. Using Newtonian gravity, the stars orbit with frequency

$$\omega^2 = \frac{2GM}{R^3} \tag{5.35}$$

If we treat these stars as point particles, then the energy density is simply a product of delta-functions

$$T^{00}(\mathbf{x}, t) = M\delta(z)\left[\delta\left(x - \frac{R}{2}\cos\omega t\right)\delta\left(y - \frac{R}{2}\sin\omega t\right) + \delta\left(x + \frac{R}{2}\cos\omega t\right)\delta\left(y + \frac{R}{2}\sin\omega t\right)\right]$$

The quadrupole (5.32) is then easily evaluated

$$\begin{aligned} I_{ij}(t) &= \frac{MR^2}{2}\begin{pmatrix} \cos^2\omega t & \cos\omega t\,\sin\omega t & 0 \\ \cos\omega t\,\sin\omega t & \sin^2\omega t & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \frac{MR^2}{4}\begin{pmatrix} 1 + \cos 2\omega t & \sin 2\omega t & 0 \\ \sin 2\omega t & 1 - \cos 2\omega t & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{5.36}$$

The resulting metric perturbation is then

$$\bar{h}_{ij} \approx -\frac{2GMR^2\omega^2}{r}\begin{pmatrix} \cos 2\omega t_{\rm ret} & \sin 2\omega t_{\rm ret} & 0 \\ \sin 2\omega t_{\rm ret} & -\cos 2\omega t_{\rm ret} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where $t_{\rm ret} = t - r$ is the retarded time.

This gravitational wave propagates out more or less radially. If we look along the $z$-axis, then the wave takes the same form as the plane wave (5.22) that we saw previously, now with combination of $H_+$ and $H_X$ polarisations, $\pi/2$ out of phase, also known as circular polarisation.

We can use this to give us a ballpark figure for the expected strength of gravitational waves. Using (5.35) to replace the frequency, we have

$$|h_{ij}| \sim \frac{G^2M^2}{Rr}$$

Clearly the signal is largest for large masses $M$, orbiting as close as possible so $R$ is small. The densest objects are black holes whose size is given by the Schwarzchild radius $R_s = 2GM$. As the black holes come close, we take $R \approx R_s$ to get

$$|h_{ij}| \sim \frac{GM}{r}$$

A black hole weighing a few solar masses has Schwarzchild radius $R_s \sim 10$ km. Now it's a question of how far away these black holes are. If two such black holes were orbiting in, say, the Andromeda galaxy which, at 2.5 million light years, has $r \approx 10^{18}$ km, we would get $h \sim 10^{-17}$. At a distance of a billion light-years, we're looking at $h \sim 10^{-20}$. These are small numbers. Nonetheless, as we mentioned previously, this is the sensitivity that has been achieved by gravitational wave detectors.

### 5.3.3 Comparison to Electromagnetism

For both electromagnetic and gravitational waves, there is a multipole expansion that determines the long distance wave behaviour in terms of the source. (Full details of the calculations in Maxwell theory can be found in the lectures on Electromagnetism.) In electromagnetism, the multipoles of the charge distribution $\rho(\mathbf{x})$ are the charge

$$Q = \int_\Sigma d^3x \; \rho(\mathbf{x})$$

the dipole

$$\mathbf{p} = \int_\Sigma d^3x \; \rho(\mathbf{x})\mathbf{x}$$

the quadrupole

$$\mathbb{Q}_{ij} = \int_\Sigma d^3x \; \rho(\mathbf{x}) \left(3x_i x_j - \delta_{ij} x^2\right)$$

and so on. Charge conservation tells us that $\dot{Q} = 0$: the total charge cannot change which means that there is no monopole contribution to electromagnetic waves. Instead the leading order contribution comes from the dipole. Indeed, repeating the calculation that we saw above in the context of Maxwell theory shows that the leading order contribution to electromagnetic waves

$$\mathbf{A}(\mathbf{x}, t) \approx \frac{\mu_0}{4\pi r}\, \dot{\mathbf{p}}(t - r) \tag{5.37}$$

We can compare this to the situation in gravity. The multipoles of the energy distribution $T_{00}(\mathbf{x})$ are the total energy

$$E = \int_\Sigma d^3x \; T_{00}(\mathbf{x})$$

the dipole which, for energy, is related to the centre of mass of the distribution

$$\mathbf{X} = \frac{1}{E}\int_\Sigma d^3x\; T_{00}(\mathbf{x})\mathbf{x}$$

the quadrupole

$$I_{ij}(t) = \int d^3x\; T_{00}(\mathbf{x},t)\, x_i x_j$$

The conservation of energy, $\dot{E} = 0$, is responsible for the lack of a monopole contribution to gravitational radiation. But, as we saw above, in contrast to electromagnetism, the dipole contribution also vanishes. This too can be traced to a conservation law: we have

$$E\dot{X}_i = \int_\Sigma d^3x\; (\partial_0 T_{00})x_i = \int_\Sigma d^3x\; (\partial_j T_{j0})x_i = -\int_\Sigma d^3x\; T_{i0} = P_i$$

where, in the penultimate equality, we have integrated by parts and, in the final equality, we have used the definition of the total momentum $P_i$ defined in (5.31). But conservation of momentum $\mathbf{P}$ means that the second time derivative of the dipole vanishes

$$E\ddot{\mathbf{X}} = \dot{\mathbf{P}} = 0$$

This is the physical reason that there's no gravitational dipole: it would violate the conservation of momentum.

In electromagnetism, there is another dipole contribution to the gauge potential: this is

$$\mathbf{A}^{\rm MD}(\mathbf{x},t) = -\frac{\mu_0}{4\pi r}\,\hat{\mathbf{x}}\times\dot{\mathbf{m}}(t-r)$$

where the magnetic dipole $\mathbf{m}$ is defined by

$$\mathbf{m} = \frac{1}{2}\int_\Sigma d^3x\; \mathbf{x}\times\mathbf{J}(\mathbf{x})$$

In our gravity, the analogous term comes from the $\dot{T}_{ij}$ in the expansion (5.29). The analog of the magnetic dipole in gravity is

$$J_i = \int_\Sigma d^3x\; \epsilon_{ijk}x_j T_{0k}$$

But this is again something familiar: it is the angular momentum of the system. This too is conserved, $\dot{\mathbf{J}} = 0$, which means that, again, the dipole contribution vanishes in gravity. The leading order effect is the quadrupole.

### 5.3.4 Power Radiated: The Quadrupole Formula

A source which emits gravitational waves will lose energy. We'd like to know how much energy is emitted. In other words, we'd like to understand how much energy is carried by the gravitational waves.

In the context of electromagnetism, it is fairly easy to calculate the analogous quantity. The energy current in electromagnetic waves is described by the $T^{0i}$ components of the energy-momentum tensor, better known as the Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}$$

To compute the power $\mathcal{P}$ emitted by an electromagnetic source, we simply integrate this energy flux over a sphere $\mathbf{S}^2$ that surrounds the source,

$$\mathcal{P} = \int_{\mathbf{S}^2} d^2\mathbf{r}\cdot\mathbf{S}$$

Evaluating this using the dipole approximation for electromagnetic waves (5.37), and doing a suitable average, we find the Larmor formula

$$\mathcal{P} = \frac{\mu_0}{6\pi c}\,|\ddot{\mathbf{p}}|^2$$

Our task in this section is to perform the same calculations for gravitational waves.

This is not as easy as it sounds. The problem is the one we addressed in Section 4.5.5: there is no local energy-momentum tensor for gravitational fields. This means that there is no analog of the Poynting vector for gravitational waves. It looks like we're scuppered.

There is, however, a way forward. The idea is that we will attempt to define an energy-momentum tensor $t_{\mu\nu}$ for gravitational waves which, in the linearised theory, obeys

$$\partial^\mu t_{\mu\nu} = 0$$

The problem is that, as we mentioned in Section 4.5.5, there is no way to achieve this in a diffeomorphism invariant way. In the full non-linear theory, this mean that $t_{\mu\nu}$ is not actually a tensor. In our linearised theory, it means that $t_{\mu\nu}$ will not be invariant under the gauge transformations (5.10). Nonetheless, we'll first define an appropriate $t_{\mu\nu}$, and then worry about the lack of gauge invariance later.

### A Quick and Dirty Approach: the Fierz-Pauli Action

When asked to construct an energy-momentum tensor for the metric perturbations, the first thing that springs to mind is to return to the Fierz-Pauli action (5.8). Viewed as an action describing a spin 2 field propagating in Minkowski space, we can then treat it as any other classical field theory and compute the energy-momentum tensor in the usual ways.

For example if we work in transverse traceless gauge, with $h = 0$ and $\partial^\mu h_{\mu\nu} = 0$ then, after an integration by parts, the Fierz-Pauli action becomes

$$S_{FP} = -\frac{1}{8\pi G}\int d^4x \; \frac{1}{4}\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu}$$

which looks like the action for a bunch of massless scalar fields. The energy density then takes the schematic form

$$t^{00} \sim \frac{1}{G}\dot{h}_{\mu\nu}\dot{h}^{\mu\nu}$$

There are also gradient terms but, for wave equations, these contribute in the same way as time derivatives. Strictly speaking, we should be working with the momentum $t^{0i}$, but this scales in the same way and calculation is somewhat easier if we work with $t^{00}$. Our previous expression (5.33) for the emitted gravitational wave wasn't in transverse-traceless gauge. If we were to massage it into this form, we have

$$h_{ij}(\mathbf{x}, t) \sim \frac{G}{r}\ddot{\mathcal{Q}}_{ij}(t - r)$$

where $\mathcal{Q}_{ij}$ is the traceless part of the quadrupole moment,

$$\mathcal{Q}_{ij} = I_{ij} - \frac{1}{3}I_{kk}\delta_{ij}$$

Putting this together, suggests that the energy density carried in gravitational waves is schematically of the form

$$t^{00} \sim \frac{G}{r^2}\dddot{\mathcal{Q}}_{ij}^2$$

Integrating over a sphere at a large distance, suggests that the energy lost in gravitational waves should depend on the square of the third derivative of the quadrupole,

$$\mathcal{P} \sim G\dddot{\mathcal{Q}}_{ij}^2$$

It turns out that this is indeed correct. A better treatment gives

$$\mathcal{P} = \frac{G}{5}\dddot{\mathcal{Q}}_{ij}\dddot{\mathcal{Q}}^{ij} \tag{5.38}$$

where, as in all previous formulae, $\dddot{\mathcal{Q}}_{ij}$ should be evaluated in retarded time $t_{\rm ret} = t - r$. This is the quadrupole formula, the gravitational equivalent of the Larmor formula.

Before the direct detection of gravitational waves, the quadrupole formula gave us the best observational evidence of their existence. The *Hulse-Taylor pulsar* is a binary neutron star system, discovered in 1974. One of these neutron stars is a pulsar, emitting a sharp beam every 59 ms. This can be used to very accurately track the orbit of the stars and show that the period – which is about 7.75 hours – is getting shorter by around 10 $\mu$s each year. This is in agreement with the quadrupole formula (5.38). Hulse and Taylor were awarded the 1993 Nobel prize for this discovery.

### Looking for a Better Approach

Any attempt to improve on the discussion above opens up a can of worms. The calculation needed to nail the factor of 1/5 is rather arduous. More importantly, however, there are also a number of conceptual issues that we need to overcome. Rather than explaining the detailed integrals that give the factor of 1/5, we'll instead focus on some of these conceptual ideas.

Our first task is to do a better job of defining $t_{\mu\nu}$. There are a number of ways to proceed.

- First, we could try to do a less shoddy job of computing the energy-momentum tensor $t_{\mu\nu}$ from the Fierz-Pauli action (5.8). This, it turns out, suffers a number of ambiguities. If, for example, we attempted to compute $t_{\mu\nu}$ as the Noether currents associated to spacetime translations, then we would find that the result is neither symmetric in $\mu$ and $\nu$, nor gauge invariant. That's not such a surprise as it's also true for Maxwell theory. We can then try to add an "improvement" term

  $$t_{\mu\nu} \rightarrow t_{\mu\nu} + \partial^\rho \Theta_{\rho\mu\nu}$$

  where $\Theta_{\rho\mu\nu} = -\Theta_{\mu\rho\nu}$ which ensures that $\partial^\mu \partial^\rho \Theta_{\rho\mu\nu} = 0$ and the extra term doesn't ruin conservation of the current. In Maxwell theory, such a term can be added to make the resulting energy-momentum tensor both symmetric and gauge invariant. For the Fierz-Pauli action, we can make it symmetric but not gauge invariant.

  A similar approach is to forget the origin of the Fierz-Pauli action and then attempt to write a generalisation of the action in "curved spacetime" by contracting indices with a metric $g^{\mu\nu}$ and replacing derivatives with $\nabla_\mu$. We could then evaluate the energy-momentum tensor using the usual formula (4.44), subsequently restricting to flat space. Here too there are ambiguities which now arise from the possibility of including terms like $R_{\mu\nu} h^{\mu\rho} h^\nu{}_\rho$ or $R_{\mu\nu\rho\sigma} h^{\mu\rho} h^{\nu\sigma}$ in the action. These vanish in Minkowski space, but give different energy-momentum tensors. For any choice, the result is again symmetric but not gauge invariant.

- Another approach is to take the lack of energy-conservation of the matter fields seriously, and try to interpret this as energy transferred into the gravitational field. To this end, let's look again at the covariant conservation $\nabla_\mu T^{\mu\nu} = 0$. As we stressed in Section 4.5.5, covariant conservation is not the same thing as actual conservation. In particular, we can rewrite the covariant conservation equation as

$$\begin{aligned}\nabla_\mu T^\mu{}_\nu &= \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}T^\mu{}_\nu\right) - \Gamma^\rho_{\mu\nu}T^\mu{}_\rho \\ &= \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}T^\mu{}_\nu\right) - \frac{1}{2}\partial_\nu g_{\mu\rho}T^{\mu\rho} = 0\end{aligned}$$

  where, to get the second line, we've invoked the symmetry of $T^{\mu\rho}$. Note that the simplification of the Christoffel symbol to $g_{\mu\rho,\nu}$ only happens when the $\nu$ index is down; this reflects the fact we're writing the equations in a non-covariant way. Next, we use the Einstein equation to replace $T^{\mu\rho}$ on the right-hand side by $\frac{1}{8\pi G}G^{\mu\rho}$. This gives

$$\partial_\mu\left(\sqrt{-g}T^\mu{}_\nu\right) = \frac{1}{16\pi G}\sqrt{-g}\,\partial_\nu g_{\mu\rho}\left(R^{\mu\rho} - \frac{1}{2}Rg^{\mu\rho}\right) = \frac{1}{16\pi G}\sqrt{-g}\,\partial_\nu g_{\mu\rho}R^{\mu\rho}$$

  The idea is to massage the right-hand side so that this expression becomes

$$\partial_\mu(\sqrt{-g}T^\mu{}_\nu) = -\partial_\mu(\sqrt{-g}t^\mu{}_\nu)$$

  for some $t^\mu{}_\nu$ which is referred to as the *Landau-Lifshitz pseudotensor*. This equation suggests that the sum of the matter energy $T^\mu{}_\nu$ and the gravitational energy $t^\mu{}_\nu$ is conserved. However, this statement should be treated with suspicion because it's coordinate dependent: the pseudotensor $t_{\mu\nu}$ is not a real tensor: its expression is long and horrible involving many terms, each of which is quadratic in $\Gamma$ and quadratic in $g$. (You can find it in (101.6) of Landau and Lifshitz, volume 2 but it's unlikely to give you a sense of enlightenment.) The expression for the pseudo-tensor is slightly nicer in the linearised theory, albeit only slightly.

- The final approach is perhaps the least intuitive, but has the advantage that it gives a straightforward and unambiguous path to find an appropriate non-tensor $t_{\mu\nu}$. Motivated by the expectation that any putative $t_{\mu\nu}$ will be quadratic in $h_{\mu\nu}$, we expand the Einstein equations to the next order. We keep $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Expanding to second order, the Einstein equations becomes

$$\left[R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right]^{(1)} + \left[R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right]^{(2)} = 8\pi GT_{\mu\nu}$$

where the subscript $(n)$ means restrict to terms of order $h^n$. We rewrite this as

$$\left[R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right]^{(1)} = 8\pi G\left(T_{\mu\nu} + t_{\mu\nu}\right) \tag{5.39}$$

with the second order expansion of the Einstein tensor now sitting suggestively on the right-hand side it is interpreted as the gravitational energy-momentum non-tensor

$$\begin{aligned} t_{\mu\nu} &= -\frac{1}{8\pi G}\left[R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right]^{(2)} \\ &= -\frac{1}{8\pi G}\left[R^{(2)}_{\mu\nu} - \frac{1}{2}R^{(2)}\eta_{\mu\nu} - \frac{1}{2}R^{(1)}h_{\mu\nu}\right] \end{aligned}$$

If we're far from the source then we can neglect the term $R^{(1)}$ since it vanishes by the equation of motion. (More precisely, it vanishes at linear order and so fails to contribute at the quadratic order that we care about.) We end up with seemingly simple expression

$$t_{\mu\nu} = -\frac{1}{8\pi G}\left[R^{(2)}_{\mu\nu} - \frac{1}{2}R^{(2)}\eta_{\mu\nu}\right] \tag{5.40}$$

The linearised Bianchi identity is $\partial^\mu\left[R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right]^{(1)} = 0$. But this means that if we are far from sources, so $T^{\mu\nu} = 0$, and the equation of motion (5.39) is satisfied, then we necessarily have $\partial^\mu t_{\mu\nu} = 0$ as befits a conserved current. All that's left is to evaluate the Ricci tensor to second order in the perturbation $h_{\mu\nu}$. This is painful. The answer turns out to be

$$\begin{aligned} R^{(2)}_{\mu\nu}[h] = \; & \frac{1}{2}h^{\rho\sigma}\partial_\mu\partial_\nu h_{\rho\sigma} - h^{\rho\sigma}\partial_\rho\partial_{(\mu}h_{\nu)\sigma} + \frac{1}{4}\partial_\mu h_{\rho\sigma}\partial_\nu h^{\rho\sigma} + \partial^\sigma h^\rho{}_\nu\partial_{[\sigma}h_{\rho]\mu} \\ & + \frac{1}{2}\partial_\sigma\left(h^{\sigma\rho}\partial_\rho h_{\mu\nu}\right) - \frac{1}{4}\partial^\rho h\partial_\rho h_{\mu\nu} - \left(\partial_\sigma h^{\rho\sigma} - \frac{1}{2}\partial^\rho h\right)\partial_{(\mu}h_{\nu)\rho} \end{aligned}$$

Pretty huh? Substituting this into the expression (5.40) gives an equally pretty expression for $t_{\mu\nu}$. Once again however, $t_{\mu\nu}$ is not gauge invariant.

We see that there are a number of different ways to construct an energy-momentum tensor $t_{\mu\nu}$ for gravitational waves. But none are gauge invariant. In order to relate this to something physical, we clearly have to construct something which is gauge invariant.

It is possible to extract something gauge invariant from $t_{\mu\nu}$ provided that our spacetime is asymptotically Minkowski. We could, for example, integrate $t^{00}$ over an infinite spatial hypersurface. This defines the *ADM energy* which can be shown to be constant in time. We'll revisit this in Section **??** in the full non-linear theory.

Alternatively, we could integrate $t^{0i}$ over a sphere at $\mathcal{I}^+$. This too gives a gauge invariant quantity, which is the time dependence of the so-called *Bondi energy*. This too can be defined in the full non-linear theory.

Here we give a less rigorous but slightly simpler construction. The gravitational wave, like any wave, varies over some typical length scale $\lambda$. We average over this oscillations by introducing a coarse-grained energy tensor

$$\langle t_{\mu\nu}\rangle = \int_V d^4x\ W(x-y)t_{\mu\nu}(y)$$

where the integral is over some region $V$ of typical size $a$. The weighting function $W(x)$ has the property that it varies smoothly over $V$ with $\int_V d^4x\ W(x) = 1$ and $W(x) = 0$ on $\partial V$. The coarse graining means that averages of total derivatives scale as $\langle \partial X\rangle \sim 1/a$. For large $a$, we can neglect such terms. Similarly, we can "integrate by parts" inside averages, so that $\langle X\partial Y\rangle = -\langle(\partial X)Y\rangle + \mathcal{O}(1/a)$. A fairly straightforward calculation shows that, in transverse-traceless gauge, the averaged energy-momentum tensor is simply

$$\langle t_{\mu\nu}\rangle = \frac{1}{32\pi G}\langle \partial_\mu h_{\rho\sigma}\partial_\nu h^{\rho\sigma}\rangle$$

where we neglect total derivatives. We can check that this is indeed conserved,

$$\partial^\mu\langle t_{\mu\nu}\rangle = \frac{1}{32\pi G}\langle(\Box h_{\rho\sigma})\partial_\nu h^{\rho\sigma} + \frac{1}{2}\partial_\nu\left(\partial_\mu h_{\rho\sigma}\partial^\mu h^{\rho\sigma}\right)\rangle = 0$$

The first term vanishes by the equation of motion, while the second is a total derivative and so can be neglected. More importantly, under a gauge transformation

$$\delta\langle t_{\mu\nu}\rangle = \frac{1}{16\pi G}\langle\partial_\mu h_{\rho\sigma}\partial_\nu(\partial^\rho\xi^\sigma + \partial^\sigma\xi^\rho)\rangle$$

But now we can integrate by parts and use the de Donder gauge condition $\partial^\rho h_{\rho\sigma} = 0$. We see that the averaged $\langle t_{\mu\nu}\rangle$ is gauge invariant, with $\delta\langle t_{\mu\nu}\rangle = 0$ up to total derivative term of order $\mathcal{O}(1/a)$. In other words, $\langle t_{\mu\nu}\rangle$ is almost gauge invariant. A better way of saying "almost gauge invariant" is "not gauge invariant". If we really want something gauge invariant, which we do, we must take $a\to\infty$, meaning that we average over all of spacetime.

Finally, we can compute the power emitted by gravitational wave at infinity by

$$\mathcal{P} = \int_{\mathbf{S}^2} d^2x\ \hat{n}_i\langle t^{0i}\rangle$$

with $\hat{n}_i$ a normal vector to $\mathbf{S}^2_\infty$. With some tedious integrals, we then find the answer (5.38).

### 5.3.5 Gravitational Wave Sources on the ✉

We can do some quick, back-of-the-envelope calculations to get a sense for how much energy is emitted by a gravitational wave source. Assuming Newtonian gravity is a good approximation, two masses $M$, separated by a distance $R$, will orbit with frequency

$$\omega^2 R \sim \frac{GM}{R^2}$$

The quadrupole is $\mathcal{Q} \sim MR^2$ and so $\dddot{\mathcal{Q}} \sim \omega^3 MR^2$. We learn that the power emitted scales as (5.38)

$$\mathcal{P} \sim G\dddot{\mathcal{Q}}^2 \sim \frac{G^4 M^5}{R^5} \tag{5.41}$$

To get numbers out of the this, we need to put the factors of $c$ back in. Recall that the Schwarzchild radius of an object is $R_s = 2GM/c^2$. and the dimensions of Newton's constant are $[G] = M^{-1}L^3T^{-2}$. So we can write this as

$$\mathcal{P} = \left(\frac{R_s}{R}\right)^5 L_{\text{Planck}} \tag{5.42}$$

where the *Planck luminosity* is

$$L_{\text{Planck}} = \frac{c^5}{G} \approx 3.6 \times 10^{52}\ \text{J}\,\text{s}^{-1}$$

This is a silly luminosity. The luminosity of the Sun is $L_\odot \approx 10^{-26} L_{\text{Planck}}$. With $10^{11}$ stars, the luminosity of the galaxy is $L_{\text{galaxy}} \approx 10^{-15} L_{\text{Planck}}$. There are roughly $10^{10}$ galaxies in the visible universe, which means that all the stars in the all the galaxies shine with a luminosity $\approx 10^{-5} L_{\text{Planck}}$.

Yet, when two black holes orbit and spiral towards each other, at the point where their separation is comparable to their Schwarzchild radius, the formula (5.42) tells us that the power they emit in gravitational waves is approximately $L_{\text{Planck}}$. For that brief moment before they collide, spiralling black holes emit more energy than all the stars in the visible universe.

Since the power emitted by colliding black holes is so ridiculously large, we might harbour some hope that we will still get a significant energy from more mundane systems. We could, for example, look at our solar system. The formula (5.42) assumes that the orbiting objects have the same mass. If two objects with masses $M_1 \gg M_2$ are in orbit, then (5.41) is replaced by

$$\mathcal{P} \sim \frac{G^4 M_1^3 M_2^2}{R^5}$$

(A derivation of this can be found on Examples Sheet 4.) Jupiter has a mass $10^{-3}M_\odot$ and orbits at a distance $\approx 10^9$ km from the Sun. Using the fact that the Schwarzchild radius of the Sun is $R_s \approx 3$ km, we find that the power emitted in gravitational waves by Jupiter is

$$\mathcal{P} \approx 10^{-50} L_{\rm Planck} \approx 10^{-24} L_\odot$$

This is completely negligible. We can trace this to the power of 5 in (5.42) which means the fall-off in power is quick: extreme events in the universe emit a ridiculous amount of energy in gravitational waves. Events involving objects that are merely heavy emit essentially zero.

Of course, the question that we all really want to ask is: how much gravitational radiation can we emit by shaking our arms around? Suppose that we go really crazy, doing jumping jacks and generally acting like a loon. For once, SI units are useful. The mass of our arms is few kg, moving a distance of around a metre, with a frequency around a second. So $\mathcal{Q} \approx 1\ \mathrm{kg\,m^2}$ and $\dddot{\mathcal{Q}} \approx 1\ \mathrm{kg\,m^2\,s^{-3}}$. The power is then

$$\mathcal{P} \sim \frac{G\dddot{Q}^2}{c^5} \approx 10^{-52}\ \mathrm{J\,s^{-1}}$$

To put this in perspective, let's remind ourselves that ultimately the world is quantum and although we have no hope of detecting individual gravitons it is surely the case that gravitational waves come in quanta with energy $E = \hbar\omega$. So we could ask: how long do we have to wave our arms before we emit a single graviton? The energy of a graviton with frequency $\omega \approx 1\,\mathrm{s}^{-1}$ is $E \approx 10^{-34}\, J$. So the calculation above tells us that we can expect to emit a single graviton if we wave our hands around for

$$T = 10^{18}\ \mathrm{s}$$

This is more or less the age of the universe. You may be many things, but you are not a factory for making gravitons.