Hilary term, 2025

C7.6: General Relativity 2

University of Oxford: Part C

Chris Couzens

Office N1.06

Mathematical Institute, University of Oxford

Andrew Wiles Building, Radcliffe Observatory Quarter

Woodstock Road

Oxford, OX2 6GG

Disclaimer: There are almost certainly typos in the notes, if something does not look correct or needs further explanation please let me know.

Please send any comments/corrections to Christopher.Couzens@maths.ox.ac.uk

Recommended books and resources

There are a large variety of good textbooks and lecture notes on general relativity. This course borrows from a number of them, in various different places, chiefly among them is the book by Sean Carroll [1], the book by Schutz [2] and the book by Wald [3]. Some useful lecture notes for the black hole part of the lectures are by Harvey Reall and by Fay Dowker. For the gravitational waves see grav waves notes.

For the background material one can read the GR1 lecture notes. This should cover all the necessary prerequisites that one would need to know about general relativity, for convenience a summary is provided in appendix A.

Contents

1 Introduction **8**

2 Linearised gravity and Gravitational waves **9**

2.1 Linearised Gravity 9

2.2 Newtonian Limit 13

2.3 Gravitational waves without sources 14

2.3.1 Transverse Traceless Gauge 15

2.3.2 Plane wave solutions 17

2.3.3 Geodesic deviation 19

2.3.4 The passing of a gravitational wave 21

2.4 The energy of Gravitational waves 24

2.4.1 How gravitational waves curve the background 27

2.4.2 The energy momentum tensor of gravitational waves 30

2.5 Gravitational waves from a source 34

2.5.1 Binary merger 41

2.6 Sources of gravitational waves 48

2.7 Laser Interferometers 50

3 Spherical cold stars and stellar collapse **52**

3.1 Tolman–Oppenheimer–Volkoff equations 54

3.2 Buchdahl's limit 58

3.3 Summary 62

4 What is a black hole: Causality and Penrose Diagrams **64**

4.1 Conformal compactification 64

4.1.1 Minkowski Space in two-dimensions 66

4.1.2 Minkowski Space in $d > 2$ 67

4.1.3 Rindler spacetime in 1+1 dimensions 70

4.1.4 Kruskal Space 72

4.2 Asymptotic flatness 72

4.3 Singularities 74

4.3.1 Different types of singularities 75

4.4 Null hypersurfaces 77

4.5 Definition of a black hole and the event horizon 81

5 Charged Black holes 84

5.1 Einstein gravity coupled to electromagnetism 84

5.2 Reissner–Nordström black hole 87

5.2.1 Super extremal RN: $e^2 > M^2$ 88

5.2.2 Sub-extremal RN: $M^2 > e^2$ 88

5.2.3 Extremal RN: $M^2 = e^2$ 96

5.3 Cauchy surfaces and horizons 98

5.4 Charges in curved spacetime 107

6 Rotating black holes 111

6.1 The Kerr–Newman solution 111

6.2 The Kerr solution 112

6.2.1 Singularities 115

6.3 Maximal extension of Kerr and its Penrose diagram 118

6.4 Komar Integrals and conserved quantities along geodesics 118

6.4.1 Lie derivative recap 120

6.4.2 Killing vectors 122

6.4.3 Conserved quantities along geodesics 124

6.4.4 Komar integrals 127

6.5 Ergosphere and Penrose process (or how to steal energy from a black hole) 130

7 Laws of black hole thermodynamics 136

7.1 The four laws of (black hole) themrodynamics 136

7.2 Zeroth law of black hole mechanics 137

7.3 First law 140

7.4 Second law 145

7.5 Third law 147

7.6 Why should black holes carry an entropy? 147

8 The black hole information paradox: the short version NOT EXAMINABLE 148

8.1 A short introduction to Hawking radiation 148

8.2 Information Paradox 151

9 Hawking temperature: the longer version(non-examinable) 153
9.1 Quantum mechanics 153
9.2 Quantum field theory 157
9.3 QFT in curved spacetime 159
9.4 Unruh effect 162
9.5 Hawking temperature 168
9.6 Black hole evaporation 172

A A short review of GR1 176
A.1 Manifolds 176
A.2 Riemannian geometry 180
A.3 Einstein's equation 185
A.4 Schwarzschild solution 185

Conventions

- We will use the god-given signature convention of mostly plus $(-,+,+,+)$. This may differ with the convention you have used in other courses, especially field theory courses. This convention is preferable when thinking about geometry as it gives positive spatial distances. For quantum field theory the other convention is preferable since it ensures that energies and frequencies are positive. You may map between the two conventions through *Wick rotation*, essentially allowing the coordinates to become complex.

- Spacetime indices will be taken to be greek letters from the middle of the alphabet: $\mu, \nu, \rho, ...$ and run over $0, 1, 2, 3$. Latin indices $i, j, k, ..$ run over the spatial directions and take values $1, 2, 3$.

- We employ Einstein summation convention, repeated indices are summed over, unless otherwise stated.

- We work in units where the speed of light c is set to 1. Occasionally it is instructive to reintroduce c which can be done by dimensional analysis.

- The Minkowski metric will be denoted by $\eta_{\mu\nu} = \text{diagonal}(-1, 1, 1, 1)_{\mu\nu}$.

- After introducing curvature we will take the metric to be $g_{\mu\nu}$ and the determinant will be $\det(g_{\mu\nu}) \equiv g$.

- The set of all vector fields on a manifold M is $\mathcal{X}(M)$.

Useful formulae

- The Lagrangian for the geodesic equation of a massive test particle is

$$\mathcal{L}\Big(\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}, x^\mu\Big) = \sqrt{-g_{\mu\nu}(x)\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}\,,$$

  with λ an arbitrary parameter along the worldline.

- The geodesic equation for an affinely parametrised massive particle is

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0\,, \qquad g_{\mu\nu}(x)\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = -1\,,$$

  where τ is the proper time. For light, the first equation takes the same form just replacing τ with an affine parameter. The second is modified by $-1 \to 0$.

- The Christoffel symbols (Levi–Civita connection) are

$$\Gamma^{\mu}{}_{\nu\rho} = \frac{1}{2} g^{\mu\sigma} \Big(\partial_\nu g_{\sigma\rho} + \partial_\rho g_{\sigma\nu} - \partial_\sigma g_{\nu\rho} \Big) \,.$$

- The Riemann tensor is

$$R^{\mu}{}_{\nu\rho\sigma} = \partial_\rho \Gamma^{\mu}{}_{\nu\sigma} - \partial_\sigma \Gamma^{\mu}{}_{\nu\rho} + \Gamma^{\mu}{}_{\rho\lambda} \Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\mu}{}_{\sigma\lambda} \Gamma^{\lambda}{}_{\nu\rho} \,. \tag{0.1}$$

  - Symmetries

$$R_{\mu\nu\rho\sigma} = -R_{\mu\nu\sigma\rho} \,,$$
$$R_{\mu\nu\rho\sigma} = R_{\sigma\rho\mu\nu} \,.$$

  - Bianchi identity 1

$$R^{\mu}{}_{\nu\rho\sigma} + R^{\mu}{}_{\rho\sigma\nu} + R^{\mu}{}_{\sigma\nu\rho} = 0 \,.$$

  - Bianchi Identity 2

$$\nabla_\mu R^{\sigma}{}_{\lambda\nu\rho} + \nabla_\nu R^{\sigma}{}_{\lambda\rho\mu} + \nabla_\rho R^{\sigma}{}_{\lambda\mu\nu} = 0 \,.$$

- Ricci tensor

$$R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu}$$

- Ricci scalar

$$R = R_{\mu\nu} g^{\mu\nu} \,.$$

- Einstein tensor

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu} \,.$$

- Einstein–Hilbert action plus cosmological constant,

$$S = \frac{1}{16\pi G} \int \mathrm{d}^4 x \sqrt{-g} \Big(R + \Lambda \Big) \,.$$

- Under a variation $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$ we have

$$\delta g^{\mu\nu} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\rho\sigma} \,,$$
$$\delta g = g g^{\mu\nu} \delta g_{\mu\nu} \,,$$
$$\delta R_{\mu\nu} = \nabla_\rho \delta \Gamma^{\rho}{}_{\mu\nu} - \nabla_\mu \delta \Gamma^{\rho}{}_{\rho\nu} \,.$$

- The *Levi–Civita symbol* is defined to be

$$\epsilon_{\mu_1....\mu_p} = \begin{cases} 1 & \mu_1....\mu_p \text{ even permutation of } S_p \\ -1 & \mu_1....\mu_p \text{ odd permutation of } S_p \\ 0 & \text{otherwise} \end{cases}$$

with S_p the permutation group of p elements. It satisfies:

$$\epsilon_{\mu_1...\mu_p}\epsilon^{\mu_1...\mu_k\nu_{k+1}...\nu_p} = k!\delta^{\nu_{k+1}....\nu_p}_{\mu_{k+1}...\mu_p}$$

where the indices are raised with the Kronecker delta not the metric! This arises because this is not a tensor but a tensor density. Recall that the volume form $\mathrm{vol}(M) = \sqrt{|g|}\mathrm{d}^n x$ is invariant under coordinate transformations. Introducing a set of coordinates we have that the components

$$\mathrm{vol}(M)_{\mu_1....\mu_n} = \sqrt{|g|}\epsilon_{\mu_1....\mu_n} \,,$$

must transform as a rank $(0, n)$ tensor. The determinant transforms as a scalar density and therefore we see that the Levi–Civita symbol must also transform as a tensor density of opposite weight.

To see this more clearly it is useful to note that the Levi–Civita symbol can be used to compute the determinant:

$$\epsilon_{\mu_1...\mu_p}A^{\mu_1}{}_{\nu_1}.....A^{\mu_p}{}_{\nu_p} = \det(A)\epsilon_{\nu_1....\nu_p} \,.$$

1 Introduction

These are lecture notes for the second course of General Relativity following on from the C7.5. A summary of some of the material covered in GR1 that will be useful is given in appendix A. We will begin this course by studying gravitational waves and linearised gravity before moving on to studying black holes in more detail. We will study the formation of black holes from stellar collapse, introduce the Kerr and Reissner–Nordstrom black holes and use these to study black hole thermodynamics.

2 Linearised gravity and Gravitational waves

In GR 1 we have studied a theory we claim describes gravity, as a first step it should be possible to take a limit in which we recover Newtonian gravity. We know that Newtonian gravity works well for slowly moving fields in weak gravitational backgrounds, if GR does not reduce to this then something must be wrong. We will first study the Newtonian limit of Einstein gravity before using our results to study gravitational waves. Both require us to linearise gravity around the Minkowski vacuum. What this means in practice is that we take the metric to be the Minkowski metric with a small perturbation, and then plug this ansatz into the Einstein equations keeping only the linear terms in our perturbation. We will see that we correctly reproduce Newtonian gravity after taking a suitable limit.

As an application of our linearised theory we will consider gravitational waves propagating in space. We begin by considering plane wave gravitational waves, these come in from infinity and go back out to infinity. They are vacuum solutions in that they do not require any energy momentum tensor, so are not particularly physical, however they exhibit all the interesting properties in a simple setting. We will then study more physical gravitational waves arising from a rotating binary pair before discussing briefly observations of gravitational waves.

There are many resources on gravitational waves. In writing the notes I have pulled material from a variety of sources. A non-exhaustive list is: [2, 4–6]. For some work on gravitational waves in astrophysics and cosmology, which are topics beyond this course, see [7].

2.1 Linearised Gravity

We want to consider weak gravitational fields, this means that the metric is roughly Minkowski, or nearly flat. The weakness of the gravitational field is then expressed by decomposing the metric as a perturbation around the Minkowski metric[1]

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,, \tag{2.1}$$

with $|h_{\mu\nu}| \ll 1$ small everywhere in spacetime. That is each of the components of the metric is small everywhere. This assumption allows us to ignore anything that is higher than first order in h. i.e. we will drop h^2 and higher order terms. In the solar system, for example, we

[1] One could choose to expand around another metric, for example Schwarzschild or include a cosmological constant. The linearisation around Minkowski is much simpler since the Riemann tensor of the Minkowski metric vanishes. For metrics which are not flat one obtains additional contributions. For those interested see [8] for example. We will in fact require some higher order terms later on, but we will come to this problem when it arises.

have $|h_{\mu\nu}| \sim |\phi| c^{-2} \sim 10^{-6}$. Here ϕ is the Newtonian potential of the sun as measured by the Earth.

The inverse metric (to first order) is then

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \,, \tag{2.2}$$

where $h^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}$. We can now raise and lower indices with η since the corrections would be of higher order in the perturbation. We can think of this linearised theory as describing a theory of a symmetric tensor field $h_{\mu\nu}$ propagating on a flat background spacetime. If we instead perform a perturbation around some other background metric, then the theory is that of the symmetric tensor field propagating on said curved background.

We want to consider the equations of motion for the perturbations, which come from examining Einstein's equations to linear order in h. We ultimately need to find the Einstein equations and so to begin, we should work out the Christoffel symbols:

$$\begin{aligned} \Gamma^{\rho}{}_{\mu\nu} &= \frac{1}{2} g^{\rho\sigma}\big(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\big) \\ &= \frac{1}{2}\eta^{\rho\sigma}\big(\partial_\mu h_{\sigma\nu} + \partial_\nu h_{\sigma\mu} - \partial_\sigma h_{\mu\nu}\big) + \mathcal{O}(h^2)\,. \end{aligned} \tag{2.3}$$

Since the Riemann tensor is of the form $R \sim \partial\Gamma + \Gamma\Gamma$ the first order contributions will come from the derivative terms and not the 'squared' terms. We have

$$\begin{aligned} R^{\sigma}{}_{\rho\mu\nu} &= \partial_\mu \Gamma^{\sigma}{}_{\nu\rho} - \partial_\nu \Gamma^{\sigma}{}_{\mu\rho} + \mathcal{O}(h^2) \\ &= \frac{1}{2}\eta^{\sigma\lambda}\big(\partial_\mu\partial_\rho h_{\nu\lambda} - \partial_\mu\partial_\lambda h_{\nu\rho} - \partial_\nu\partial_\rho h_{\mu\lambda} + \partial_\nu\partial_\lambda h_{\mu\rho}\big) + \mathcal{O}(h^2)\,. \end{aligned} \tag{2.4}$$

It follows that the Ricci tensor is

$$R_{\mu\nu} = \frac{1}{2}\big(\partial^\sigma\partial_\nu h_{\sigma\mu} + \partial^\sigma\partial_\nu h_{\mu\sigma} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h\big) + \mathcal{O}(h^2)\,, \tag{2.5}$$

where $h = h^{\mu}{}_{\mu}$ is the trace and $\Box = \partial^\mu\partial_\mu$. Moreover the Ricci scalar is

$$R = \partial^\mu\partial^\nu h_{\mu\nu} - \Box h + \mathcal{O}(h^2)\,. \tag{2.6}$$

Putting all of this together into the Einstein tensor we end up with

$$G_{\mu\nu} = \frac{1}{2}\Big[\partial^\sigma\partial_\nu h_{\mu\sigma} + \partial^\sigma\partial_\mu h_{\nu\sigma} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h - \eta_{\mu\nu}\big(\partial^\rho\partial^\sigma h_{\rho\sigma} - \Box h\big)\Big] + \mathcal{O}(h^2)\,. \tag{2.7}$$

Aside
As an aside the Einstein tensor can be obtained by varying the following Lagrangian with respect to $h_{\mu\nu}$,

$$\mathcal{L} = \frac{1}{2}\Big[(\partial_\mu h^{\mu\nu})\partial_\nu h + \frac{1}{2}\partial^\mu h^{\rho\sigma}\partial_\mu h_{\rho\sigma} - \partial^\mu h^{\rho\sigma}\partial_\rho h_{\mu\sigma} + \partial^\mu h \partial_\mu h\Big] \,. \tag{2.8}$$

The full linearised equations of motion are then

$$\frac{1}{2}\Big[\partial^\sigma\partial_\nu h_{\mu\sigma} + \partial^\sigma\partial_\mu h_{\nu\sigma} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h - \eta_{\mu\nu}\big(\partial^\rho\partial^\sigma h_{\rho\sigma} - \Box h\big)\Big] = 8\pi G_N T_{\mu\nu}\,, \tag{2.9}$$

where $T_{\mu\nu}$ is assumed to be small also. Note that the energy-momentum tensor must satisfy

$$\partial_\mu T^{\mu\nu} = 0\,, \tag{2.10}$$

which is the linearised version of the conservation equation.

Before we can proceed we must deal with gauge invariance. We have demanded that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ however this does not completely fix the coordinate system on spacetime. Let us consider an infinitesimal change of coordinates

$$x^\mu \to x^\mu - \xi^\mu \tag{2.11}$$

with ξ assumed to be small. The metric undergoes the infinitesimal change

$$\delta g_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) = \xi^\lambda\partial_\lambda g_{\mu\nu} + g_{\mu\rho}\partial_\nu\xi^\rho + g_{\nu\rho}\partial_\mu\xi^\rho\,. \tag{2.12}$$

This is precisely the Lie derivative of the metric. If we act with an infinitesimal diffeomorphism along the curve with tangent ξ then the metric changes as

$$\delta g_{\mu\nu} = (\mathcal{L}_\xi g)_{\mu\nu} = \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu\,. \tag{2.13}$$

When the metric takes the linearised form this should be understood as a transformation of $h_{\mu\nu}$. Since we assume that both h and ξ are small[2] it follows that we may replace covariant derivatives of g with covariant derivatives of η where the Christoffel symbols vanish. Therefore under an infinitesimal diffeomorphism we have that $h_{\mu\nu}$ changes as:

$$h_{\mu\nu} \to h_{\mu\nu} + (\mathcal{L}_\xi\eta)_{\mu\nu} = h_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu\,. \tag{2.14}$$

[2]If we did not restrict to small ξ then we could go to a region where $h_{\mu\nu}$ is not small by a coordinate transformation, clearly we do not want this as it would render our approximation incorrect.

Example 2.1: Gauge freedom in Electromagnetism
For those who have seen gauge theories this is precisely the form of a gauge transformation of Maxwell theory. There we shift the one-form A as $A \to A + \mathrm{d}\Lambda$ which leaves the field strength (or curvature of the gauge bundle) $F = \mathrm{d}A$ invariant. Similarly the above transformation leaves the linearised Riemann tensor invariant, this is the curvature of the bundle in this case.

When we do computations in gauge theories we typically pick a gauge to work in. The most common gauge to take is the Lorentz gauge

$$\partial^\mu A_\mu = 0 \,, \tag{2.15}$$

which reduces the Maxwell equation $\mathrm{d} \star F = \star J$ with source to the wave equation

$$\Box A_\nu = J_\nu \,. \tag{2.16}$$

There is a similar kind of gauge here called *de Donder gauge*. The gauge condition is

$$\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h = 0 \,. \tag{2.17}$$

Aside
To see that this is always possible, suppose that you are given a metric where

$$\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h = f_\nu \,, \tag{2.18}$$

then after a gauge transformation we have

$$\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h + \Box \xi_\nu = f_\nu \,, \tag{2.19}$$

and it amounts to finding ξ such that $\Box \xi_\nu = f_\nu$. With suitable conditions this is always possible.

The de Donder gauge greatly simplifies our linearised equations of motion leading to

$$\Box h_{\mu\nu} - \frac{1}{2}\Box h \eta_{\mu\nu} = -16\pi G_N T_{\mu\nu} \,. \tag{2.20}$$

To further simplify it is useful to define

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu} \,, \tag{2.21}$$

so that the linearised Einstein equation becomes

$$\Box \bar{h}_{\mu\nu} = -16\pi G_N T_{\mu\nu}\,, \tag{2.22}$$

and the de Donder gauge condition is

$$\partial^\mu \bar{h}_{\mu\nu} = 0\,. \tag{2.23}$$

To see that this is a sensible definition we see that from $\bar{h}_{\mu\nu}$ we can recover $h_{\mu\nu}$ since by taking the trace on both sides we have

$$\bar{h} = \eta^{\mu\nu}\bar{h}_{\mu\nu} = -h\,, \tag{2.24}$$

and so we can reconstruct $h_{\mu\nu}$ as

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\bar{h}\eta_{\mu\nu}\,. \tag{2.25}$$

Note that de Donder gauge does not fix all the gauge freedom. In terms of $\bar{h}$ a coordinate transformation acts as

$$\delta\bar{h}_{\mu\nu} = \partial_\mu\xi_\nu + \partial_\nu\xi_\mu - \partial^\rho\xi_\rho\eta_{\mu\nu}\,, \tag{2.26}$$

and any ξ_ν such that $\Box\xi_\nu = 0$ will preserve the de Donder gauge, as can be seen from (2.19). The de Donder gauge reduces the 10 independent components of the symmetric $h_{\mu\nu}$ matrix to only 6 independent components which must satisfy (2.22).

2.2 Newtonian Limit

We are now in a position to take the Newtonian limit of GR. Newtonian gravity is valid when the gravitational fields are too weak to produce velocities near the speed of light, and general relativity must give the same results as Newtonian gravity. The requirement of the velocities to be small is equivalent to placing bounds on the distribution of energy momentum. Typically this places bounds on the Energy momentum tensor so that the components of the energy momentum tensor obey $|T^{00}| \gg |T^{0i}|, |T^{ij}|$. When T^{0i} is non-zero we also require that it is bigger than the components of the purely spatial pieces.

We require a low-density slowly moving distribution of matter. For simplicity we will take a stationary matter configuration (independent of time) so that the Energy-momentum tensor is

$$T_{00} = \rho(\vec{x})\,, \tag{2.27}$$

with all other components vanishing. Via the stationary assumption we may replace the wave operator $\Box$ with the 3d Euclidean Laplacian $\Box = -\partial_t^2 + \partial_i^2 = -\partial_t^2 + \nabla^2 = \nabla^2$, since nothing depends on the time coordinate. Einstein's equations then become

$$\nabla^2 \bar{h}_{00} = -16\pi G_N \rho(\vec{x})\,, \qquad \nabla^2 \bar{h}_{0i} = 0\,, \quad \nabla^2 \bar{h}_{ij} = 0\,. \tag{2.28}$$

With suitable boundary conditions the solutions are

$$\bar{h}_{00} = -4\Phi(\vec{x}) \qquad \bar{h}_{0i} = \bar{h}_{ij} = 0\,, \tag{2.29}$$

where Φ is identified with the Newtonian potential obeying

$$\nabla^2 \Phi(\vec{x}) = 4\pi G_N \rho(\vec{x})\,. \tag{2.30}$$

Translating back to $h_{\mu\nu}$ we find

$$h_{00} = -2\Phi(\vec{x})\,, \quad h_{ij} = -2\Phi(\vec{x})\delta_{ij}\,, \quad h_{0i} = 0\,. \tag{2.31}$$

The final metric is then

$$\mathrm{d}s^2 = -\big(1 + 2\Phi(\vec{x})\big)\mathrm{d}t^2 + \big(1 - 2\Phi(\vec{x})\big)\mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x}\,. \tag{2.32}$$

One can now compute the geodesic equations with this metric to find the *same* equations of motion as in the Newtonian theory. You will do this in problem sheet 1, and conclude that general relativity reduces to Newtonian gravity. If one takes the Schwarzschild solution and expands around large r it takes the form (2.32) with $\Phi(\vec{x}) = -\frac{G_N M}{r}$, this is the metric one would then expect far away from a point mass.

2.3 Gravitational waves without sources

One can also study gravitational waves using the linearised equations of motion. Gravitational waves are modulations in the spacetime that propagate at the speed of light and induce variations in the length of objects they pass through. Moreover they carry energy away from their sources. As you must be away gravitational waves have had a lot of recent experimental interest due to the observations of gravitational waves by LIGO (and also Virgo in Italy and KAGRA in Japan).[3] Theorists have also taken an interest in these experimental results with

[3] LIGO (Laser Interferometer Graviational-wave Observatory) is a gravitational wave detector built in the USA. There are two sites, one in Washington and the other in Louisiana. Both are Michelson interferometers. These work by merging two sources of light to create an interference pattern. A light beam is split with the two beams each travelling down a 4km long arm which are perpendicular to each other and exactly the same

the hope that extra precision tests of GR and its quantum gravity extension can be performed using this data, though this latter hope is probably some way off with current technology. One can also combine gravitational waves and cosmology to also probe cosmological models!

The first detection consisted of the merger of a 30 solar mass black hole and 35 solar mass black hole to produce a 62 solar mass black hole. The astute reader should observe that $30 + 35 = 65 \neq 62$, some of the mass, 3 solar masses worth, must have been radiated away. This is an astonishing amount of energy and was emitted in a tiny fraction of a second. In those milliseconds this was more energy than emitted by all the stars in all the galaxies in the observable universe! This energy was radiated away through gravitational waves.

2.3.1 Transverse Traceless Gauge

Before we solve the equations of motion we will consider using the remaining gauge freedom to simplify the metric as much as possible. We are interested in studying the propagation of gravitational waves and their interaction with test masses. As such we are interested in the wave outside the source where the energy momentum tensor vanishes: $T_{\mu\nu} = 0$. The equation of motion outside the source is then

$$\Box \bar{h}_{\mu\nu} = 0 \, . \tag{2.33}$$

We can greatly simplify the form of the metric outside the source by using the remaining gauge degrees of freedom. We have already imposed the de Donder gauge however we can still perform a gauge transformation with parameters ξ_μ such that $\Box \xi_\mu = 0$ whilst preserving our gauge choice. Recall that the gauge transformation is

$$\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \partial^\rho \xi_\rho \eta_{\mu\nu} \equiv \bar{h}_{\mu\nu} + \xi_{\mu\nu} \, . \tag{2.34}$$

We can choose the four components of ξ_μ to impose four conditions on $\bar{h}_{\mu\nu}$. We can choose ξ^0 such that $\bar{h} = 0$, which implies that $\bar{h}_{\mu\nu} = h_{\mu\nu}$. We can then choose the three components

length. The light reaches the end of the arms and is reflected back by mirrors to the beam splitter where the light is combined again and one finds an interference pattern. LIGO is designed so that the two beams of light totally destructively interfere upon reaching the beam splitter and this is observed by a photodetector. The default observation is that no light is detected by the photodetector. When a gravitational wave passes it stretches space in one direction and simultaneously compresses it in the perpendicular direction and vice versa. This means that one arm becomes longer and one becomes shorter and when the beams of light reach the beam splitter they no longer totally destructively interfere, instead a signal is detected. As the wave passes each arm oscillates in length and the beams oscillate in and out of phase. The idea behind this is therefore very simple, however the difference can be as little as 1/1000th of the width of a proton and it is a wonder of engineering that we can do this. The first detection of gravitational waves was on the 14th September 2015, which was the merger of two $\sim$ 30 solar mass black holes merging about 1.3 billion light years from Earth. Since then there have been many other detections, including neutron star-neutron star mergers and a neutron star-black hole merger.

ξ^i such that $h^{0i} = 0$. The de Donder gauge condition for $\mu = 0$ implies

$$\partial^0 h_{00} + \partial^i h_{0i} = 0\,, \tag{2.35}$$

and therefore with $h_{0i} = 0$ we have:

$$\partial^0 h_{00} = 0\,, \tag{2.36}$$

and is therefore time independent. Since it is time-independent it corresponds to the static part of the gravitational interaction, that is the Newtonian potential term of the source which generated the gravitational wave. From the point of view of the gravitational wave $\partial^0 h_{00}$ implies that $h_{00} = 0$. Therefore we have used the remaining gauge degrees of freedom to set $h_{0\mu} = 0$ and we are left with only the spatial parts h_{ij} which satisfy the gauge choice $\partial^i h_{ij} = 0$ and the trace condition $h^i{}_i = 0$. Thus we have

$$\boxed{h^{0\mu} = 0\,, \quad h^i{}_i = 0\,, \quad \partial^i h_{ij} = 0\,.} \tag{2.37}$$

This is known as the *transverse-tracelss gauge* or TT gauge. By using the de Donder gauge and now our transverse-traceless gauge we have reduced the 10 independent components to only 2 physical degrees of freedom! We will denote the metric in TT gauge by h^{TT}_{ij}.

We emphasise that the TT gauge cannot be chosen when $T_{\mu\nu} \neq 0$. This would render the equation of motion unsolvable, one can only perform this gauge transformation if it is consistent with the energy momentum tensor!

Aside

There is a nice way of obtaining the TT gauge by using a projector. Given a plane wave solution $h_{\mu\nu}(x)$ propagating in the (unit) direction $\vec{n}$ outside sources and in de Donder gauge, one can put the wave in TT gauge as follows. First one introduces the symmetric transverse tensor

$$P_{ij}(\vec{n}) = \delta_{ij} - n_i n_j\,. \tag{2.38}$$

Since we take $\vec{n}$ to be unit norm we have that $n^i P_{ij}(\vec{n}) = 0$ and moreover that it is a projector and satisfies $P_{ij}P_{jk} = P_{ik}$. Using the projector P one can construct

$$\Lambda_{ij,kl}(\vec{n}) = P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl}\,. \tag{2.39}$$

This satisfies

$$\Lambda_{ij,kl}\Lambda_{kl,mn} = \Lambda_{ij,mn}\,, \tag{2.40}$$

and is transverse in all indices. Furthermore it is traceless with respect to the (i, j) indices and the (k, l) indices and symmetric under $(i, j) \leftrightarrow (k, l)$. Given a plane wave solution $h_{\mu\nu}$

in de Donder gauge, the gravitational wave in TT gauge is given by

$$h_{ij}^{TT} = \Lambda_{ij,kl} h_{kl} \,. \tag{2.41}$$

One can show, by using the de Donder gauge condition and the properties of the projector that this solves the TT gauge condition and the wave equation.

2.3.2 Plane wave solutions

Let us now consider a solution to the source free linearised equations of motion. We will consider plane wave solutions. They are somewhat unphysical, describing waves coming from infinity and heading back out, however they exhibit many of the features that we want to study and will allow us to build up to considering more physical processes which create gravitational waves.

We need to solve (2.22) with $T_{\mu\nu} = 0$,

$$\Box \bar{h}_{\mu\nu} = 0 \,. \tag{2.42}$$

This is nothing but the wave equation, which admits plane wave solutions. For the moment we will not impose the TT gauge and see how it arises. Let

$$\bar{h}_{\mu\nu}(x) = \mathrm{Re}\Big[H_{\mu\nu} \mathrm{e}^{\mathrm{i}k_\sigma x^\sigma}\Big] \,, \tag{2.43}$$

with $H_{\mu\nu}$ a complex constant symmetric matrix, which we call the polarisation matrix, and k^μ a real vector which is called the *wave vector*. The wave equation reduces to

$$k_\mu k^\mu = 0 \,, \tag{2.44}$$

and therefore the wave vector is null. This implies that the gravitational waves move at the speed of light relative to the background Minkowski metric. There is still more to do, we still need to impose the de Donder gauge condition, (2.17), this implies:

$$k^\nu H_{\nu\mu} = 0 \,, \tag{2.45}$$

which is the condition that the waves are transverse to the direction of propagation. We still have our residual gauge freedom when ξ satisfies $\Box \xi_\mu = 0$ that we used above to go to TT gauge. We can take $\xi_\mu = l_\mu \mathrm{e}^{\mathrm{i}k_\rho x^\rho}$ which satisfies the required harmonic condition. This shifts the polarisation matrix to:

$$H_{\mu\nu} \to H_{\mu\nu} + \mathrm{i}(k_\mu l_\nu + k_\nu l_\mu - k^\rho l_\rho \eta_{\mu\nu}) \,. \tag{2.46}$$

We can now choose l_μ such that we can take our TT gauge

$$H_{0\mu} = 0\,, \qquad \text{and} \qquad H^\mu{}_\mu = 0\,. \tag{2.47}$$

After all the gauge has been fixed we have two remaining independent polarisations of $H_{\mu\nu}$.

We can orientate our coordinate system so that the wave travels along the z direction, and therefore the wave vector is

$$k^\mu = (\omega, 0, 0, \omega)\,. \tag{2.48}$$

The transverse gauge condition and our gauge freedom implies that the polarisation matrix $H_{\mu\nu}$ takes the form

$$H_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.49}$$

with both h_+ and $h_\times$ complex. These are the two independent polarisation states. We could also consider the superposition of waves with different frequencies travelling along the z direction. The form of the polarisation matrix does not change, though the functions h_+ and $h_\times$ now depend on the different frequencies. The superposition of waves coming from different directions would ruin this nice splitting into a 2×2 block. The final metric for a plane wave gravitational wave propagating along the z direction is then

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}z^2 + (1 + \mathrm{Re}[h_+ \mathrm{e}^{\mathrm{i}\omega(z-t)}])\mathrm{d}x^2 + (1 - \mathrm{Re}[h_+ \mathrm{e}^{\mathrm{i}\omega(z-t)}])\mathrm{d}y^2 + 2\mathrm{Re}[h_\times \mathrm{e}^{\mathrm{i}\omega(z-t)}]\mathrm{d}x\mathrm{d}y \tag{2.50}$$

Aside

Light also has two polarization states. Recall that when considering light there is something called helicity, this is a projection of the spin onto the direction of momentum. We can also consider the helicity here, consider a rotation about the z-axis. Let R_θ be the matrix rotation by an angle of θ. The polarization states transform as

$$H'_{\mu\nu} = (R_\theta)_\mu{}^\sigma (R_\theta)_\nu{}^\rho H_{\sigma\rho}\,. \tag{2.51}$$

We have that

$$R_\theta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.52}$$

We then have that

$$\begin{aligned} H'_{11} &= \cos 2\theta H_{11} + \sin 2\theta H_{12}\,, \\ H'_{12} &= -\sin 2\theta H_{11} + \cos 2\theta H_{12}\,. \end{aligned} \tag{2.53}$$

If we define

$$H_\pm = H_{11} \mp \mathrm{i} H_{12}\,, \tag{2.54}$$

then

$$H'_\pm = \mathrm{e}^{\pm 2\mathrm{i}\theta} H_\pm\,, \tag{2.55}$$

and we see that the polarization states $H_\pm$ have *helicity* ± 2.

We have now constructed the metric for our gravitational wave, it remains to understand how we can detect it. The answer to this question is *tidal forces*, or more concretely *geodesic deviation*. We will deviate temporarily from our linearized gravity discussion to study/recall geodesic deviation.

2.3.3 Geodesic deviation

In Euclidean space or in Minkowski spacetime, geodesics which are initially parallel will remain parallel forever. On a general curved manifold this notion of parallel is not possible, instead we can study whether nearby geodesics move together or apart, and characterise their relative acceleration.

Consider a one-parameter family of geodesics with coordinates $x^\mu(\tau : \sigma)$. Here τ is the affine parameter along the geodesics, all of which are tangent to the vector field X. Thus, along the surface spanned by $x^\mu(\tau : \sigma)$ we have

$$\left.\frac{\partial x^\mu}{\partial \tau}\right|_\sigma = X^\mu\,. \tag{2.56}$$

The parameter σ labels the different geodesics, see figure 1. We can compute the tangent vector in the σ direction to be generated by a second vector field S so that

$$S^\mu = \left.\frac{\partial x^\mu}{\partial \sigma}\right|_\tau\,. \tag{2.57}$$

This tangent vector is known as the *deviation vector*, its job is to take us from one geodesic to a nearby geodesic with the same affine parameter τ. The family of geodesics sweep out a 2d surface embedded in the manifold. We have freedom to choose coordinates so that on the surface $S = \frac{\partial}{\partial \sigma}$ and $X = \frac{\partial}{\partial \tau}$ consequently we have $[X, S] = 0$.

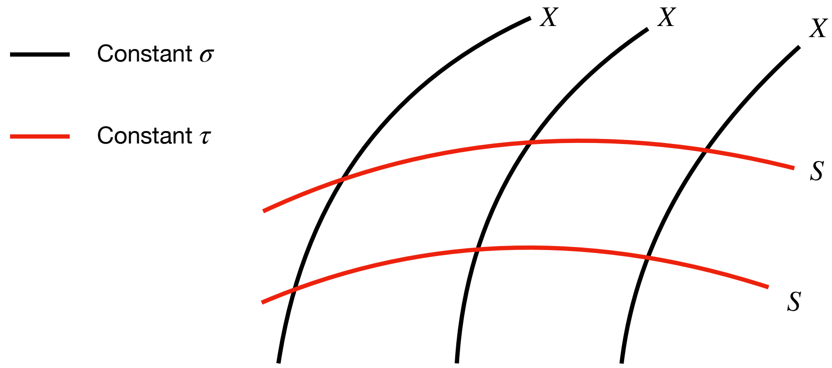

Figure 1: The black lines are geodesics generated by X while the red lines label constant τ and are generated by S with $[X, S] = 0$.

We can ask how neighbouring geodesics behave, do they converge, diverge, or remain the same distance apart? Consider a torsion free connection so that

$$\nabla_X S - \nabla_S X = [X, S] \,. \tag{2.58}$$

Since $[X, S] = 0$, we have

$$\nabla_X \nabla_X S = \nabla_X \nabla_S X = \nabla_S \nabla_X X + R(X, S)X \,, \tag{2.59}$$

where we have used the expression for the Riemann tensor in (0.1), or in coordinate free notation:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \,. \tag{2.60}$$

Since X is tangent to geodesics we have $\nabla_X X = 0$ and therefore we find

$$\nabla_X \nabla_X S = R(X, S)X \,. \tag{2.61}$$

In index notation we have

$$X^\nu \nabla_\nu \big(X^\rho \nabla_\rho S^\mu \big) = R^\mu{}_{\nu\rho\sigma} X^\nu X^\rho S^\sigma \,. \tag{2.62}$$

If we take an integral curve γ associated to X as before we have

$$\frac{D^2 S^\mu}{D\tau^2} = R^\mu{}_{\nu\rho\sigma} X^\nu X^\rho S^\sigma \,, \tag{2.63}$$

with $D/D\tau$ the covariant derivative along the curve γ, $D/D\tau \equiv \frac{\partial x^\mu}{\partial \tau} \nabla_\mu$. The left hand side tells us how the deviation vector S changes as we move along the geodesic and it measures the relative acceleration of neighbouring geodesics. From (2.63) we see that the relative acceleration of neighbouring geodesics is measured by the Riemann tensor. This is nothing other than *tidal forces*, recall that these showed up when we wanted to test the equivalence principle. Recall the thought experiment of people locked in a box, either accelerating in a rocket or on the Earth's surface. There is no local experiment we can do to tell these situations apart: however studying the non-local tidal forces does distinguish between these two setups.

Note that the relative acceleration vanishes for all families of geodesics if and only if the Riemann tensor vanishes, that is the manifold is flat. This is why geodesics which are initially parallel in Minkowski space remain parallel, because the manifold is flat.

2.3.4 The passing of a gravitational wave

The Physics behind the TT frame We now want to understand how geodesic deviation can be used to understand the passing of a gravitational wave. Effectively what we want to do is place a set of test particles and measure the separation between them as the gravitational wave passes. Mathematically this requires us to study the effect the gravitational waves have on neighbouring geodesics as it passes. Before we study this we will try to better understand physically what it means to be in the TT frame.

First let us look at the geodesic equation in the TT frame. Consider a test particle at rest at $\tau = 0$. The spatial components of the geodesic equation at $\tau = 0$ satisfy

$$\begin{aligned} \left.\frac{\mathrm{d}^2 x^i}{\mathrm{d}\tau^2}\right|_{\tau=0} &= -\Gamma^i{}_{\nu\rho}(x)\left.\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\right|_{\tau=0} \\ &= -\Gamma^i{}_{00}\left.\left(\frac{\mathrm{d}x^0}{\mathrm{d}\tau}\right)^2\right|_{\tau=0}, \end{aligned} \tag{2.64}$$

where in going to the second line we used that the mass is initially at rest. Expanding our Christoffel symbols as before we find that

$$\Gamma^i{}_{00} = \frac{1}{2}(2\partial_0 h_{0i} - \partial_i h_{00})\,. \tag{2.65}$$

Note that in TT gauge this vanishes! Therefore if at time $\tau = 0$ the particle is stationary then it remains stationary for all time! Particles which were at rest in the TT frame before the arrival of the wave remain at rest even after the wave has arrived. Strictly speaking this is only true to linear order in $h_{\mu\nu}$, if we include second order terms we do get a non-trivial contribution to $\Gamma^i{}_{00}$. However, for the gravitational waves we expect on Earth from distant sources we have $h = \mathcal{O}(10^{-21})$ and therefore going beyond linear order is of no interest whatsoever.

The TT frame coordinates are such that they stretch themselves in response to the arrival of the gravitational wave in such a way that the position of free test masses initially at rest do not change. This also means that the coordinate distance between two test masses initially at rest will not change as the wave passes either! You may be surprised about this fact at first, but this just illustrates in a nice way the fact that in GR the physical effects are not expressed by what happens to the coordinates since the theory is reparametrisation invariant. The gravitational wave does have a physical effect, we just chose coordinates such that they do not change as the wave passes. Physical effects can be detected by studying invariant quantities such as proper lengths, proper times and tidal forces. For example the proper distance does change with time, whereas the coordinate distance does not. This is the difference between using coordinate dependent observables and coordinate independent ones. The gravitational wave can be detected unambiguously with the coordinate independent observable but *not* the coordinate dependent one.

Measuring the passing of the wave To detect the wave we will using the geodesic deviation introduced above, using the family of geodesics $x^\mu(\tau : \sigma)$ introduced there. Recall that τ is the affine parameter and σ labels the geodesic. For simplicity consider the situation where, in the absence of the gravitational waves, the family of geodesics are in a rest frame, thus the tangent vector to the geodesic is $X^\mu = (1, 0, 0, 0)$. As the gravitational wave passes the geodesic will change as $X^\mu = (1, 0, 0, 0) + \mathcal{O}(h)$, but this effect will not concern us as it results only in sub-subleading effects. We can now input on the right-hand side of (2.63) our expression for the Riemann tensor (2.4). Recall that this is of order h already and it is for this reason the modification of X^μ is immaterial. Next observe that on the left-hand side of (2.63) we may replace the covariant derivative along the curve with derivative with respect to the affine parameter. Moreover since we are in the rest frame we can replace the proper time with the coordinate time.[4] We therefore have:

$$\frac{\mathrm{d}^2 S^\mu}{\mathrm{d}t^2} = R^\mu{}_{00\nu} S^\nu \,. \tag{2.66}$$

It remains to insert the Riemann tensor from equation (2.4) into the equation. We obtain

$$\frac{\mathrm{d}^2 S^\mu}{\mathrm{d}t^2} = \frac{1}{2} \frac{\mathrm{d}^2 h^\mu{}_\nu}{\mathrm{d}t^2} S^\nu \,, \tag{2.67}$$

where we impose the TT gauge condition. From the form of the polarisation vector, (2.49), we see that the gravitational wave affects neither S^0 nor S^3. The only effect is in the plane

[4] If one considers higher orders in the perturbation one cannot perform all these replacements with such impunity but at this linearised level this is all fine.

transverse to the direction of propagation. The result is remarkably simple and we see that the action on the separation vector is of the form of a Newtonian force acting on the particle.

To make our lives easier we can solve this in the $z = 0$-plane, though it is not difficult to track the z dependence. First consider the h_+ polarisation, setting $h_\times = 0$. We have:

$$\frac{\mathrm{d}^2 S^1}{\mathrm{d}t^2} = -\frac{\omega^2}{2}\mathrm{Re}\big[h_+ \mathrm{e}^{-\mathrm{i}\omega t}\big] S^1 \,, \quad \frac{\mathrm{d}^2 S^2}{\mathrm{d}t^2} = \frac{\omega^2}{2}\mathrm{Re}\big[h_+ \mathrm{e}^{-\mathrm{i}\omega t}\big] S^2 \,. \tag{2.68}$$

Since h_+ is small we can solve this perturbatively, we have

$$S^1(t) \approx S^1(0)\Big(1 + \frac{1}{2}\mathrm{Re}\big[h_+ \mathrm{e}^{-\mathrm{i}\omega t}\big] + \ldots\Big) \,, \quad S^2(t) \approx S^2(0)\Big(1 - \frac{1}{2}\mathrm{Re}\big[h_+ \mathrm{e}^{-\mathrm{i}\omega t}\big] + \ldots\Big) \,. \tag{2.69}$$

From this we can determine the way in which the geodesics are affected by the passing wave. We can think of the displacement vector as the distance from the origin to a neighbouring geodesic. Consider a collection of particles arranged around a circle of radius r in the x-y plane. The initial condition is such that $S^1(0)^2 + S^2(0)^2 = r^2$. We can now evaluate the position as the wave passes by. We see that due to the relative minus sign between the S^1 and S^2 term if the geodesic move outwards in one direction it necessarily moves inwards in the other orthogonal direction. We may now plot the circle as the wave passes by, the result is pictured in figure 2.

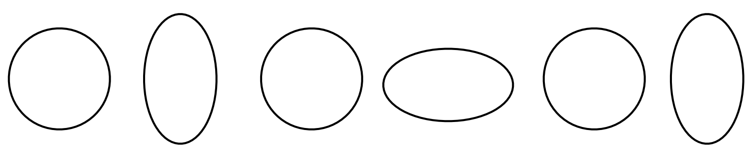

Figure 2: The displacement of a circle as a h_+ polarised gravitational waves passes.

Next consider the other choice of polarisation setting $h_+ = 0$ this time. Now the geodesic deviation equations are

$$\frac{\mathrm{d}^2 S^1}{\mathrm{d}t^2} = -\frac{\omega^2}{2}\mathrm{Re}[h_\times \mathrm{e}^{-\mathrm{i}\omega t} S^2] \,, \quad \frac{\mathrm{d}^2 S^2}{\mathrm{d}t^2} = -\frac{\omega^2}{2}\mathrm{Re}[h_\times \mathrm{e}^{-\mathrm{i}\omega t} S^1] \,, \tag{2.70}$$

which again admits a perturbative solution, in this case:

$$S^1(t) = S^1(0) + \frac{1}{2}S^2(0)\mathrm{Re}[h_\times \mathrm{e}^{-\mathrm{i}\omega t}] + \ldots \,, \quad S^2(t) = S^2(0) + \frac{1}{2}S^1(0)\mathrm{Re}[h_\times \mathrm{e}^{-\mathrm{i}\omega t}] + \ldots \,. \tag{2.71}$$

The difference between the previous case is a 45-degree rotation, one can easily see this by defining $S^1 \pm S^2$, i.e. taking the axes to be different 45-degree rotated, the resultant diagram is given in figure 3.

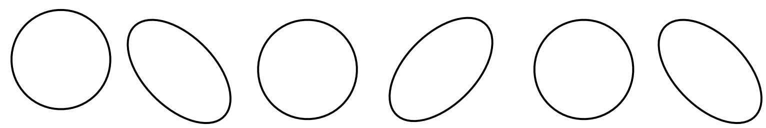

Figure 3: The displacement of a circle as a $h_\times$ polarised gravitational waves passes.

Before we conclude this section recall that the metric in (2.50) is an approximation since we performed a linearization around Minkowski space. One can in fact find an exact plane wave solution inspired by the above perturbative solution. Suppose that the wave is travelling in the z-direction as before. We see that the perturbative solution depends only on $u = t - z$, this suggests we should use an ansatz which depends only on u. It is also convenient to define the coordinate $v = t + z$, both of u and v define null coordinates, and in these coordinates the Minkowski metric becomes:

$$\mathrm{d}s^2 = -\mathrm{d}u\mathrm{d}v + \mathrm{d}x^2 + \mathrm{d}y^2 \,. \tag{2.72}$$

For simplicity let us consider only the + polarisation, in which case, in analogy to the perturbed metric we have the ansatz:

$$\mathrm{d}s^2 = -\mathrm{d}u\mathrm{d}v + f(u)^2\mathrm{d}x^2 + g(u)^2\mathrm{d}y^2 \,. \tag{2.73}$$

Exercise 1:

- Show that the metric (2.73) satisfies the vacuum Einstein equations if

$$\frac{\ddot{f}}{f} + \frac{\ddot{g}}{g} = 0 \,. \tag{2.74}$$

- Show that if g is nearly 1 then this agrees with the +-polarisation solution above.

- Show that even in the non-linear case that a particle initially at rest in these coordinates remains at rest.

This solution is one of a class called plane-fronted waves with parallel rays.

2.4 The energy of Gravitational waves

We now want to understand the energy and momentum carried by gravitational waves. Since we observed that the gravitational waves do some work on the test masses of the previous section they must carry some form of energy which they impart on the background. This is quite a subtle question to answer for a number of reasons.

- Firstly in GR there is no *local definition of energy density for the gravitational field.* Since the equivalence principle allows is to eliminate gravitational forces at any point (we can always introduce normal coordinates so that the metric at that point is the Minkowski metric), we could always go to such coordinates and we would find that there is no energy at that point.

- The second problematic point is that according to GR any source of energy generates curvature via a stress tensor. We have introduced our plane waves as a perturbation around flat space which is not curved. We therefore need to generalise the concept of a Gravitational wave as a perturbation of a generic background, not necessarily flat space.

There are two different roots that one can use to get an explicit expression for the energy-momentum tensor of gravitational waves. One is more geometrical while the other is more field theoretical.

1. Since General Relativity says that any form of energy induces curvature of spacetime we can work out whether gravitational waves source spacetime curvature. This will be through an energy-momentum tensor which we associated to the wave, and from this we can define the energy carried by the wave.

2. We can treat linearised gravity as a classical field theory of the field $h_{\mu\nu}$ and use Noether's theorem to compute the conserved quantity. This will again fix for us an energy momentum tensor.

We will look at the first option in these lectures, for those interested in the second method see [6]. To understand whether gravitational waves curve the background we need to generalise our starting point. Until now we have linearised the Einstein equations around the flat metric. This is insufficient to understand whether gravitational waves curve the background spacetime since by assumption we exclude that possibility from the beginning. We need to allow the background metric to be dynamical and then we can define the gravitational waves as perturbations over this curved dynamical metric. We write

$$g_{\mu\nu}(x) = \bar{g}_{\mu\nu}(x) + h_{\mu\nu}(x)\,, \qquad |h_{\mu\nu}| \ll 1\,. \tag{2.75}$$

However this is kind of dumb! How do we decide which part of $g_{\mu\nu}$ is the background and which are the fluctuations due to the gravitational wave? We could move pieces between the two with impunity. In the linearised theory where the background metric was uniquely fixed this was not a problem. A simple analogy one can make is trying to isolate what part of the vertical movement of a body of water is from a wave and which part belongs to the background.

There is a natural splitting between the spacetime background and the gravitational waves when there is a clear separation of scales. For example, if in some coordinate system we can write the metric as in (2.75), where $\bar{g}_{\mu\nu}$ has a typical scale of spatial variation L_B on top of which small amplitude perturbations are superimposed with a characteristic wavelength such that

$$\lambdabar \ll L_B\,, \tag{2.76}$$

where $\lambdabar = \lambda/(2\pi)$ is the reduced wavelength. In this case $h_{\mu\nu}$ has the interpretation of small ripples on a smooth background and is called the *short wave expansion*. Alternatively one could make a distinction in frequency space. In this case $\bar{g}_{\mu\nu}$ has frequencies up to a maximum value f_B while $h_{\mu\nu}$ is peaked around a frequency f such that

$$f \gg f_B\,. \tag{2.77}$$

In this case $h_{\mu\nu}$ is a high-frequency perturbation of a static or slowly varying background. It turns out in this case that in a suitable gauge that $h_{\mu\nu}$ obeys a wave equation and therefore its characteristic wavelength and frequency are related by $\lambda f = c$. On the other hand the scales L_B and f_B that characterise the background are a priori unrelated and therefore the two conditions above are independent. It suffices to impose one of them.

Having devised a method of splitting the two pieces we can ask the following questions:

- How does the high-frequency/short wavelength perturbation propagate through the background spacetime with metric $\bar{g}_{\mu\nu}$
- How does the perturbation affect the background metric? This will allow us to assign an energy momentum tensor to the gravitational wave and therefore give us a way of defining the energy of the gravitational wave.

Typically the separation of the metric uses the *short-wave expansion* however for detectors it is the second condition that is fulfilled. Ground based detectors have a size which is much smaller than the wavelength of the gravitational waves they are searching for. The reduced wave length of the typical wave that is detected is $\lambdabar \sim 50 - 500$ km. The detectors do not measure the spatial variations of the gravitational field but rather the *temporal* variations in their output as a gravitational wave passes.

2.4.1 How gravitational waves curve the background

We will now consider the situation in which in some reference frame we can make a splitting of the metric according to (2.75). This separation is then based on the fact that there is a clear distinction in scales either in space, in which case (2.76) applies, or in time in which case (2.77) applies. We begin by expanding the Einstein equations around the background $\bar{g}_{\mu\nu}$. In the expansion we have two small parameters: one is the amplitude $h \equiv \mathcal{O}(|h_{\mu\nu}|)$ and the second is either λ/L_B or f_B/f, depending on which of (2.76) or (2.77) applies.

We will expand the Einstein equations to quadratic order in $h_{\mu\nu}$. It is convenient to recast the Einstein equations in the form

$$R_{\mu\nu} = 8\pi G_N \left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) , \tag{2.78}$$

where $T_{\mu\nu}$ is the energy momentum tensor of matter and T its trace. We then expand the Ricci tensor to $\mathcal{O}(h^2)$ as

$$R_{\mu\nu} = \bar{R}_{\mu\nu} + R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \mathcal{O}(h^3) \, . \tag{2.79}$$

Here $\bar{R}_{\mu\nu}$ is constructed with $\bar{g}_{\mu\nu}$ only and the superscript denotes the power of h appearing in the expression. The crucial observation is to note the scales of the different terms. Since $\bar{R}_{\mu\nu}$ is constructed from $\bar{g}_{\mu\nu}$ it contains only low frequency modes. One the other hand $R^{(1)}_{\mu\nu}$ is linear in $h_{\mu\nu}$ and therefore contains only high-frequency modes. The second order term $R^{(2)}_{\mu\nu}$ contains both high and low frequency modes. For instance in the quadratic term $h_{\mu\nu}h_{\rho\sigma}$ a mode with high wave vector k_1 can combine with a mode with high wave vector $k_2 \sim= k_1$ so that the result is a low wave vector mode. One can therefore split the Einstein equations into two separate equations for the high and low frequency parts:

$$\bar{R}_{\mu\nu} = -[R^{(2)}_{\mu\nu}]^{\text{Low}} + 8\pi G_N \left[T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right]^{\text{Low}} , \tag{2.80}$$

$$R^{(1)}_{\mu\nu} = -[R^{(2)}_{\mu\nu}]^{\text{High}} + 8\pi G_N \left[T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right]^{\text{High}} . \tag{2.81}$$

The superscript "low" denotes projection onto long wavelengths or low frequencies depending on which of (2.76) or (2.77) applies. Similarly the "High" superscript denotes projection onto short wavelengths or high frequency depending on which of (2.76) or (2.77) applies.

One can compute the explicit expressions for $R^{(1)}_{\mu\nu}$ and $R^{(2)}_{\mu\nu}$. Let $\bar{\nabla}_\mu$ be the covariant derivative on $\bar{g}_{\mu\nu}$ then one finds that $R^{(1)}_{\mu\nu}$ is given by

$$R^{(1)}_{\mu\nu} = \frac{1}{2} (\bar{\nabla}^\rho \bar{\nabla}_\mu h_{\nu\rho} + \bar{\nabla}^\rho \bar{\nabla}_\nu h_{\mu\rho} - \bar{\nabla}^\rho \bar{\nabla}_\rho h_{\mu\nu} - \bar{\nabla}_\mu \bar{\nabla}_\nu h) \, . \tag{2.82}$$

At quadratic order one finds:

$$
\begin{aligned}
R^{(2)}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} g^{\alpha\beta} \bigg[& \frac{1}{2} \bar{\nabla}_\mu h_{\rho\alpha} \bar{\nabla}_\nu h_{\sigma\beta} + \bar{\nabla}_\rho h_{\nu\alpha} \bar{\nabla}_{[\sigma} h_{\beta]\mu} \\
& + h_{\rho\alpha} (\bar{\nabla}_\nu \bar{\nabla}_\mu h_{\sigma\beta} + \bar{\nabla}_\beta \bar{\nabla}_\sigma h_{\mu\nu} - \bar{\nabla}_\beta \bar{\nabla}_\nu h_{\mu\sigma} - \bar{\nabla}_\beta \bar{\nabla}_\mu h_{\nu\sigma}) \\
& + \Big(\frac{1}{2} \bar{\nabla}_\alpha h_{\rho\sigma} - \bar{\nabla}_\rho h_{\alpha\sigma} \Big) (\bar{\nabla}_\nu h_{\mu\beta} + \bar{\nabla}_\mu h_{\nu\beta} - \bar{\nabla}_\beta h_{\mu\nu}) \bigg] .
\end{aligned}
\tag{2.83}
$$

After some work it turns out that (2.81) is equivalent to $\bar{\nabla}^\rho \bar{\nabla}_\rho \bar{h}_{\mu\nu} = 0$ away from sources, which is nothing but the wave equation in curved space. In addition, the de Donder gauge condition should be replaced by $\bar{\nabla}^\nu \bar{h}_{\nu\mu} = 0$. We will not say much more about this since it is not the interesting part, that comes next, but you can read more about this in [6].

Let us now consider (2.80). When there is a clear cut separation between the length scales $\lambda\!\!\!^{-}$ of the gravitational waves and the length scale L_B of the background there is a simple method to perform the projection onto the long-wavelength modes. One begins by introducing a scale l such that $\lambda\!\!\!^{-} \ll l \ll L_B$ and we average over a spatial volume with sides of length l. In this way modes with wavelength L_B remain unaffected since they are basically constant over the volume used for averaging, while modes with reduced wavelength of order $\lambda\!\!\!^{-}$ are oscillating very fast and average to 0. Similarly if $h_{\mu\nu}$ is a high frequency perturbation of a quasi-static background we can introduce a timescale t which is much larger than the period $1/f$ and much smaller than the typical time scale $1/f_B$ of the background and the average over this time t. We can write (2.80) as

$$
\bar{R}_{\mu\nu} = - \left\langle R^{(2)}_{\mu\nu} \right\rangle + 8\pi G_N \left\langle T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right\rangle . \tag{2.84}
$$

Here $\langle .. \rangle$ denotes the spatial average over many wavelengths $\lambda\!\!\!^{-}$ if (2.76) applies and a temporal average over several periods $1/f$ if (2.77) applies.

Example 2.2: Averaging

As an example imagine we had the function $f(t) = \cos(\omega t)$. The frequency of this is $f = 2\pi/\omega$ and we could average this over a number of multiples of this period. The average of this would then be:

$$
\langle f(t) \rangle = \frac{1}{T} \int_0^T \mathrm{d}t f(t) . \tag{2.85}
$$

For the example function above we would have that the average vanishes! For half the period cos gives a positive contribution while for the other half it gives an equal negative contribution which cancel out over a period and so the average over many of these periods is 0. If instead we had $f(t) = \cos^2(\omega t)$ the average would be $\langle \cos^2(\omega t) \rangle = \frac{1}{2}$.

Aside

This averaging is nothing but a renormalization group transformation. One starts with the fundamental equations of the theory and integrates out the fluctuations that take place on a length scale smaller than l in order to obtain an effective theory that describes the physics at the length scale l. One can perform this in coordinate space, as above; in momentum space to integrate out the high momentum modes in order ot get a low energy effective action; or in frequency space in order to eliminate the fast temporal variations and to obtain the effective dynamics of the slowly varying degrees of freedom.

We can now define an effective energy momentum tensor $\overline{T}_{\mu\nu}$ from

$$\left\langle T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T \right\rangle = \overline{T}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}\overline{T}\,, \tag{2.86}$$

with $\overline{T} \equiv \overline{T}_{\mu\nu}\bar{g}^{\mu\nu}$. By construction $\overline{T}^{\mu\nu}$ is a purely low-frequency quantity and is the smoothed form of the energy momentum tensor $T_{\mu\nu}$. We also define the quantity $t_{\mu\nu}$ as

$$t_{\mu\nu} = -\frac{1}{8\pi G_N}\left\langle R^{(2)}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}R^{(2)} \right\rangle\,, \tag{2.87}$$

where $R^{(2)} = \bar{g}^{\mu\nu}R^{(2)}_{\mu\nu}$ and we define the trace as

$$\begin{aligned} t &= \bar{g}^{\mu\nu}t_{\mu\nu} \\ &= \frac{1}{8\pi G_N}\langle R^{(2)}\rangle\,. \end{aligned} \tag{2.88}$$

We may plug all this into (2.80) to obtain

$$\bar{R}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}\bar{R} = 8\pi G_N(\bar{T}_{\mu\nu} + t_{\mu\nu})\,. \tag{2.89}$$

This is the *coarse-grained* form of the Einstein equations and determine the dynamics of $\bar{g}_{\mu\nu}$ which is the long wave-length (low-frequency) part of the metric in terms of the long-wave-length (low frequency) part of the matter energy-momentum tensor $\overline{T}_{\mu\nu}$ and a tensor $t_{\mu\nu}$ which does not depend on external matter but only on the gravitational field itself and is quadratic in $h_{\mu\nu}$.

In summary:

- At a microscopic level there is no fundamental distinction between a background metric and fluctuations over it. The gravitational field is described by all its modes and its dynamics is fully determined by Einstein's equations.

- If some fluctuations $h_{\mu\nu}$ are clearly distinguishable from the background because their typical length scale $\lambdabar$ is much smaller than the typical length scale L_B that characterises the spatial variations of the background it is useful to introduce a macroscopic level of description which is valid at some length scale l such that $\lambdabar \ll l \ll L_B$. This effective theory is obtained by integrating out the short wavelength degrees of freedom which can be obtained by performing a spatial integral of the Einstein equations over a box of size l. Alternatively if the separation between fluctuations and background is based on the condition $f_B \ll f$ we can integrate out the fast varying degrees of freedom, performing a temporal average over several periods $1/f$ and we are left with the effective theory of the slowly varying degrees of freedom.

- The result of this procedure is (2.89). The left-hand side is the Einstein tensor for the slowly varying metric $\bar{g}_{\mu\nu}$. On the right-hand side we have the smoothed version of the matter energy momentum tensor. In addition we have a term $t_{\mu\nu}$ which is proportional to h^2. This shows that the effect of the gravitational waves on the background curvature is formally identical to that of matter with energy momentum tensor $t^{\mu\nu}$. We can therefore assign an energy momentum tensor to the gravitational waves!

- The gravitational waves energy momentum tensor $t_{\mu\nu}$ comes out automatically in an averaged form. This is because to the derive the effect of the gravitational waves on the background one is passing from the microscopic description to a macroscopic description.

2.4.2 The energy momentum tensor of gravitational waves

We now want to explicitly compute $t_{\mu\nu}$ and work out how to define the energy. We are going to be interested in the energy and momentum carried by the gravitational waves at large distances from the source. Since we are far away from the source we can approximate the background metric to be flat. In equation (2.83) we may therefore replace $\bar{\nabla}_\mu \to \partial_\mu$ and $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$. As we saw earlier the 4×4 symmetric matrix $h_{\mu\nu}$ has 10 degrees of freedom, of which only two were physical modes. Therefore $t_{\mu\nu}$ can in principle have contributions from both physical modes and gauge modes. This is not a contradiction since the Einstein tensor also depends on the coordinate system. The issue is how do we distinguish the contributions from the physical modes and those which are pure gauge? The former give the energy momentum tensor of the gravitational waves and describe physical effects while the latter will be associated to ripples in spacetime which are due to our coordinate choice and can be made to vanish by a better coordinate system.

The simplest method is to first impose the de Donder gauge. This immediately eliminates four of the spurious degrees of freedom. This leaves us with the two physical modes in h_{ij}^{TT} and the four gauge modes ξ_μ which satisfy $\nabla \xi_\mu = 0$. We can then fix one degree of freedom in ξ such that $h = 0$ and therefore $\bar{h}_{\mu\nu} = h_{\mu\nu}$ (see section 2.3.1). We may further simplify $R^{(2)}_{\mu\nu}$ by noticing that inside the average the spacetime derivative ∂_μ can be integrated by parts neglecting the boundary term.[5] Performing this and using the wave equation one finds the succinct result

$$\langle R^{(2)}_{\mu\nu} \rangle = -\frac{1}{4} \langle \partial_\mu h_{\alpha\beta} \partial_\nu h^{\alpha\beta} \rangle \,, \tag{2.90}$$

and that $\langle R^{(2)} \rangle$ vanishes. Therefore the final result is:

$$\boxed{t_{\mu\nu} = \frac{1}{32\pi G_N} \langle \partial_\mu h_{\alpha\beta} \partial_\nu h^{\alpha\beta} \rangle \,. \qquad (2.91)}$$

One can check that this is gauge invariant and therefore depends only on the physical modes h_{ij}^{TT} and therefore we can simply replace $h_{\mu\nu} \to h_{ij}^{TT}$. The gauge invariant energy density is

$$t^{00} = \frac{1}{32\pi G_N} \langle \dot{h}_{ij}^{TT} \dot{h}_{ij}^{TT} \rangle \,. \tag{2.92}$$

Example 2.3: Energy of a plane wave

Let us see how this works for our plane wave solutions in section 2.3.2. We have

$$h_{ij}^{TT} = \cos(\omega(t-z)) \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} \tag{2.93}$$

where we have taken h_+ and $h_\times$ to both be real. We have

$$t^{00} = \frac{1}{32\pi G_N} [2\omega^2 (h_+^2 + h_\times^2)] \langle \sin^2(\omega(t-z)) \rangle \tag{2.94}$$

We now need to perform the average. We will choose to perform an average over time. The frequency of the plane wave is $1/\omega$ and so our characteristic time scale will be $T = n 2\pi/\omega \gg 2\pi/\omega$, where $n \gg 1$. We have

$$\langle \sin^2(\omega(t-z)) \rangle \equiv \frac{1}{T} \int_0^T \sin^2(\omega(t-z)) = \frac{1}{2} \,, \tag{2.95}$$

[5] On generic functions an integration by parts of ∂_t is only possible if we have performed an integral over time, while integration by parts of ∂_i is only possible if we have integrated over space. Recall that in the de Donder gauge outside the source $h_{\mu\nu}$ satisfies the wave equation $\Box h_{\mu\nu} = 0$. For a solution propagating in the z direction all quantities are functions of $t - z$ and therefore one can replace $\partial_t f(t-z)$ with $-\partial_z f(t-z)$. If the integral is then over z we can use this to replace time derivatives with z derivatives, use integration by parts and then replace with time derivatives. Thus for solutions of the wave equation a spatial average allows us to integrate by parts over all directions not just spatial directions and similarly for a time average.

and therefore we find

$$t^{00} = \frac{1}{32\pi G_N}\omega^2(h_+^2 + h_\times^2)\,. \tag{2.96}$$

Now we should ask what does this compute!

Aside

Note from (2.89) that the left-had side is covariantly conserved with respect to $\bar{\nabla}$: $\bar{\nabla}^\mu(\bar{R}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}R) = 0$ and therefore we have:

$$\bar{\nabla}^\mu(\overline{T}_{\mu\nu} + t_{\mu\nu}) = 0\,. \tag{2.97}$$

The fact that it is the sum of the two terms that is covariantly conserved rather than the individual terms reflects the fact that there is an exchange of energy between the matter sources and the gravitational waves. Since at large distances the metric becomes flat space and $\overline{T}_{\mu\nu} = 0$ we have that far from the sources

$$\partial^\mu t_{\mu\nu} = 0\,. \tag{2.98}$$

The energy flux With the energy momentum tensor of the gravitational waves in hand we can compute the corresponding energy flux. This is the energy of gravitational waves flowing per unit time through a unit surface at a large distance from the source. The conservation of the energy momentum tensor $\partial^\mu t_{\mu\nu} = 0$ implies that

$$\int_V \mathrm{d}^3x(\partial^0 t_{00} + \partial^i t_{i0}) = 0\,, \tag{2.99}$$

with V a spatial volume in the far region bounded by some surface S. The gravitational energy inside the volume is

$$E_V = \int_V \mathrm{d}^3x\; t^{00}\,, \tag{2.100}$$

and therefore we have

$$\begin{aligned}\frac{\mathrm{d}E_V}{\mathrm{d}t} &= -\int_V \mathrm{d}^3x\;\partial_i t^{0i} \\ &= -\int_S \mathrm{d}A\; n_i t^{0i}\,,\end{aligned} \tag{2.101}$$

where n^i is an outward pointing normal to the surface and $\mathrm{d}A$ the surface element. Let S be a spherical surface at large distance r from the source. Its surface element is $\mathrm{d}A = r^2\mathrm{d}\Omega$ and its normal is $\vec{n} = \vec{r}$ in the radial direction. We have

$$\frac{\mathrm{d}E_V}{\mathrm{d}t} = -\int \mathrm{d}A\; t^{0r}\,, \tag{2.102}$$

with

$$t^{0r} = \frac{1}{32\pi G_N} \langle \partial^0 h_{ij}^{TT} \partial_r h_{ij}^{TT} \rangle \, . \tag{2.103}$$

At sufficiently large distances a gravitational wave propagating radially outward has perturbation of the form

$$h_{ij}^{TT}(t, r) = \frac{1}{r} f_{ij}(t - r) \, , \tag{2.104}$$

where f_{ij} is some function of the retarded time $t_{\rm ret} = t - r$. At large distances this implies that $t^{0r} = t^{00}$ and the energy satisfies

$$\frac{\mathrm{d}E_V}{\mathrm{d}t} = - \int \mathrm{d}A \; t^{00} \, . \tag{2.105}$$

This implies that the total energy inside the region decreases and therefore the outward propagating gravitational wave carries away and energy flux:

$$\begin{aligned} \frac{\mathrm{d}E}{\mathrm{d}A\mathrm{d}t} &= t^{00} \\ &= \frac{1}{32\pi G_N} \langle \dot{h}_{ij}^{TT} \dot{h}_{ij}^{TT} \rangle \, . \end{aligned} \tag{2.106}$$

or

$$\boxed{\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{r^2}{32\pi G_N} \int \mathrm{d}\Omega \langle \dot{h}_{ij}^{TT} \dot{h}_{ij}^{TT} \rangle \, .} \tag{2.107}$$

Momentum flux One can similarly compute the momentum of the gravitational waves inside a spherical shell V at large distances from the source to be

$$P_V^k = \int_V \mathrm{d}^3 x t^{0k} \, . \tag{2.108}$$

Considering, once again, a radially propagating gravitational wave one finds

$$\partial_t P_V^k = - \int_S \mathrm{d}A t^{0k} \, , \tag{2.109}$$

and therefore the momentum flux carried away by the outward propagating gravitational wave is

$$\frac{\mathrm{d}P^k}{\mathrm{d}A\mathrm{d}t} = t^{0k} \, , \tag{2.110}$$

and therefore

$$\boxed{\frac{\mathrm{d}P^k}{\mathrm{d}t} = - \frac{r^2}{32\pi G_N} \int \mathrm{d}\Omega \langle \dot{h}_{ij}^{TT} \partial^k h_{ij}^{TT} \rangle \, .} \tag{2.111}$$

The loss of linear momentum through GW emission for a Black hole binary merging to form another BH can lead to an astrophysically significant recoil velocity. This can lead to the merged BH being ejected from the host galaxy. This leads to important consequences on BH's mass growth through hierarchical mergers.

In conclusion we have found that in separating out Einstein's equations into low and high frequency parts that the low frequency part describes the effect of gravitational waves and the external matter on the background spacetime. The high frequency part gives a wave equation in curved space which describes the propagation of $h_{\mu\nu}$.

2.5 Gravitational waves from a source

The gravitational waves that we have constructed so far are plane wave solutions in the absence of a source. These come in from infinity and head back out to infinity. For something astrophysical this is not quite what we want, rather we would want gravitational waves that arise from a source (a merger for example) and then radiate out. With matter the equation to solve is (2.22), (we set $G_N = 1$ for now)

$$\Box \bar{h}_{\mu\nu} = -16\pi T_{\mu\nu}\,. \tag{2.112}$$

For those of you who have seen some quantum field theory and the Klein–Gordon equation or Electromagnetism this should not look too alien to you. We can solve this using a (retarded) Green's function. The Green's function satisfies

$$\Box_x G(x,y) = -4\pi\delta(x-y)\,, \tag{2.113}$$

and is given by:

$$G(x,y) = \frac{1}{|\vec{x}-\vec{y}|}\delta\left(y^0 - \left(x^0 - |\vec{x}-\vec{y}|\right)\right)\,. \tag{2.114}$$

Here the combination $x^0 - |\vec{x}-\vec{y}|$ is called the *retarded time*. Reinstating the factor of the speed of light one sees that this is the speed at which information spreads at the speed of light.

Aside

To compute the Green's function we first assume that there is some spherical symmetry (though this is not strictly necessary) so that the Green's function depends only on the quantities:

$$R = |\vec{x}-\vec{y}|\,, \qquad T = |x^0 - y^0|\,. \tag{2.115}$$

Recall that the delta function has the Fourier expression:

$$\delta(T) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \mathrm{e}^{-\mathrm{i}\omega T}\,, \tag{2.116}$$

and that we may expand the Green's function in Fourier modes as

$$G(T,R) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}\omega \tilde{G}(\omega,R)\mathrm{e}^{-\mathrm{i}\omega T}\,. \tag{2.117}$$

Substituting this into the defining equation we have

$$\begin{aligned}\Box\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}\omega \tilde{G}(\omega,R)\mathrm{e}^{-\mathrm{i}\omega T}\right] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}\omega (\nabla^2+\omega^2)\tilde{G}(\omega,R)\mathrm{e}^{-\mathrm{i}\omega T} \\ &= -4\pi\delta(R)\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \mathrm{e}^{-\mathrm{i}\omega T}\,,\end{aligned} \tag{2.118}$$

from which it follows that

$$\int_{-\infty}^{\infty} \mathrm{d}\omega \left[(\nabla^2+\omega^2)\tilde{G}(\omega,R) + 2\sqrt{2\pi}\delta(R)\right]\mathrm{e}^{-\mathrm{i}\omega T} = 0\,. \tag{2.119}$$

Inverting the Fourier transformation we have

$$(\nabla^2+\omega^2)\tilde{G}(\omega,R) + 2\sqrt{2\pi}\delta(R) = 0\,. \tag{2.120}$$

Solutions are given by

$$\tilde{G}(\omega,R) = \frac{1}{\sqrt{2\pi}}\frac{\mathrm{e}^{\pm\mathrm{i}\omega R}}{R}\,. \tag{2.121}$$

We now need to unwind the transformations to obtain an expression for G. We have

$$\begin{aligned}G_\pm(T,R) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}\omega \tilde{G}(\omega,R)\mathrm{e}^{-\mathrm{i}\omega T} \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \tilde{G}(\omega,R)\mathrm{e}^{-\mathrm{i}\omega(T\mp R)} \\ &= \frac{1}{R}\delta\left(T\mp R\right)\,.\end{aligned} \tag{2.122}$$

These are the *retarded* and *advanced* Green's functions respectively. The retarded includes a time delay by the amount of time it takes for the wave to propagate from $\vec{x}$ to $\vec{y}$. To see this assume that $y^0 > x^0$ (and reinstate factors of c) then we may write

$$G_-(x,y) = \frac{1}{|\vec{x}-\vec{y}|}\delta\left(y^0 - \left(x^0 - \frac{|\vec{x}-\vec{y}|}{c}\right)\right)\,. \tag{2.123}$$

The delta function is non-zero precisely when the wave has propagated from $\vec{x}$ to $\vec{y}$.

We can now use the Green's function to solve (2.22) in general, we have

$$\begin{aligned}\bar{h}_{\mu\nu}(x) &= 4\int \mathrm{d}^4y G(x,y)T_{\mu\nu}(y)\\ &= 4\int \mathrm{d}^3\vec{y}\frac{1}{|\vec{x}-\vec{y}|}T_{\mu\nu}(t-|\vec{x}-\vec{y}|,\vec{y})\,.\end{aligned} \tag{2.124}$$

The integral is over the past light-cone of the event x at which $\bar{h}_{\mu\nu}$ is evaluated. Outside of the source we can once again put this solution into the TT gauge similarly to section 2.3.1 and equation (2.39). We have that

$$\boxed{h_{ij}^{TT}(t,\vec{x}) = 4\Lambda_{ij,kl}(\vec{n})\int \mathrm{d}^3y\frac{1}{|\vec{x}-\vec{y}|}T_{kl}(t-|\vec{x}-\vec{y}|,\vec{y})\,.} \tag{2.125}$$

We have used the notation $\vec{n} = \vec{x}/|\vec{x}|$.

So far everything holds in general, however let us now assume that matter is confined to a compact region near the origin with a sphere of radius Δ centred on the origin containing all of the matter. Far away from the source, $r \equiv |\vec{x}| \gg |\vec{y}| \sim \Delta$ we may expand,

$$|\vec{x}-\vec{y}|^2 = r^2 - \vec{x}\cdot\vec{y} + \vec{y}^2 = r^2\left(1-\frac{2}{r}\vec{n}\cdot\vec{y} + \mathcal{O}(\Delta^2 r^{-2})\right)\,, \tag{2.126}$$

with $\vec{n}$ the unit norm vector such that $r\vec{n} = \vec{x}$. We therefore have:

$$\begin{aligned}|\vec{x}-\vec{y}| &= r - \vec{n}\cdot\vec{y} + \mathcal{O}(\Delta^2 r^{-1})\,,\\ T_{\mu\nu}(t-|\vec{x}-\vec{y}|,\vec{y}) &= T_{\mu\nu}(t-r,\vec{y}) + \vec{n}\cdot\vec{y}\;\partial_0 T_{\mu\nu}(t-r,\vec{y}) + \dots\,,\end{aligned} \tag{2.127}$$

From expanding the energy–momentum tensor we can in principle obtain two terms. The latter term is the time-scale on which the energy-momentum tensor is varying. We will assume a slowly varying source where the matter is moving non-relativistically, this allows us to drop the second term.[6]

There is a further simplification that we can make by using our gauge invariance. We saw that we can, without loss of generality, impose the gauge condition (2.23). Therefore the components involving a time index are not independent of the purely spatial components. We may thus focus on the purely spatial components and extract out the components with time legs by using the gauge fixing condition. Expanding our formula for h_{ij}^{TT} we find

$$h_{ij}^{TT}(x) \approx \frac{4}{r}\Lambda_{ij,kl}(\vec{n})\int \mathrm{d}^3y\; T_{kl}(t-r,\vec{y})\,. \tag{2.128}$$

[6]One can generalise to rapidly varying sources while still imposing weak sources. It is then useful to replace $T_{\mu\nu}$ with its Fourier transformation with respect to time and work with this. We will ignore this possibility here, the interested reader can look at section 5.5.2 of [5].

We now want to evaluate the above integral to obtain an expression for h^{TT}_{ij}. To do this we use the conservation of the energy-momentum tensor, which to this order, is simply $\partial_\mu T^{\mu\nu} = 0$.[7] We therefore have the two relations:

$$\partial_t T^{00} + \partial_i T^{i0} = 0\,, \qquad \partial_t T^{0i} + \partial_j T^{ji} = 0\,. \tag{2.129}$$

From these we obtain

$$\partial_t^2 T^{00} = \partial_m \partial_n T^{mn}\,. \tag{2.130}$$

Multiplying both sides by $x^i x^j$ we have

$$\begin{aligned}\partial_t^2(x^i x^j T^{00}) &= x^i x^j \partial_m \partial_n T^{mn} \\ &= \partial_m \partial_n (x^i x^j T^{mn}) - 2\partial_m (x^j T^{in} + x^i T^{jn}) + 2T^{ij}\,.\end{aligned} \tag{2.131}$$

We may use this identity to eliminate T^{ij} from our expression. We have

$$\begin{aligned}h^{TT}_{ij}(x) &\approx \Lambda_{ij,kl}(\vec{n})\frac{4}{r}\int \mathrm{d}^3y\; T_{kl}(t-r,\vec{y}) \\ &= \Lambda_{ij,kl}(\vec{n})\frac{2}{r}\int \mathrm{d}^3y \left[\partial_t^2(y^k y^l T^{tt}) + 2\partial_m(y^k T^{lm} + y^k T^{lm}) - \partial_m\partial_n(y^k y^l T^{mn})\right] \\ &= \frac{2}{r}\Lambda_{ij,kl}(\vec{n})\partial_t^2 \int \mathrm{d}^3y\; y^k y^l T^{tt}\,,\end{aligned} \tag{2.132}$$

where we have dropped total derivative terms in the last step. To this order we are considering what is called the *quadrupole radiation.*

Definition 1 *Second moment of the energy density*
The second moment of the energy density *is*

$$M^{ij}(t) = \int \mathrm{d}^3y y^i y^j T_{00}(t,\vec{y})\,. \tag{2.133}$$

This is a tensor in the usual Cartesian sense, transforming under rotations of the coordinates. We can also define the 0'th and 1'st moments, these are

$$M = \int \mathrm{d}^3y T_{00}\,, \qquad \text{and} \qquad M^i = \int \mathrm{d}^3y y^i T_{00}\,, \tag{2.134}$$

respectively. These are both conserved quantities which reflects the conservation of the mass and total momentum of the source.

[7] We will be raising and lowering the spatial indices with impunity in the following paragraphs. This is fine because at this order the metric is just the 3d identity matrix. We will still keep the usual raised and lowered index structure to make clear the Einstein summation convention.

In terms of the second moment of the energy density we have

$$h_{ij}^{TT}(t,\vec{x}) = \frac{2}{r}\Lambda_{ij,kl}(\vec{n})\ddot{M}^{kl}(t-r)\,, \tag{2.135}$$

to quadrupole order. To work out the final result we want to eliminate Λ. When the direction of propagation $\vec{n}$ is along the z-axis the projector P_{ij} is just the diagonal matrix $\mathrm{diag}(1,1,0)$. Therefore for an arbitrary 3×3 matrix one has:

$$\Lambda_{ij,kl}A_{kl} = \frac{1}{2}\begin{pmatrix} A_{11}-A_{22} & 2A_{12} & 0 \\ A_{21} & A_{22}-A_{11} & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} \tag{2.136}$$

One can then read off the two polarisations when $\vec{n}$ points in the z direction to be

$$\begin{aligned} h_+ &= \frac{G_N}{r}(\ddot{M}_{11}-\ddot{M}_{22})\,, \\ h_\times &= \frac{2G_N}{r}\ddot{M}_{12}\,, \end{aligned} \tag{2.137}$$

where it is understood that the right-hand side is evaluated at the retarded time $t-r$. To compute the amplitudes for a wave that in a frame with axes (x,y,z) propagates in a generic direction $\vec{n}$ we begin by introducing two additional unit vectors $\vec{u}$ and $\vec{v}$ which are all mutually orthogonal. In this (x',y',z') frame we can take $\vec{u}=\vec{x},\vec{v}=\vec{y}$ and $\vec{n}=\vec{z}$ and the wave propagates along the z' axis. Using the above we have that the physical polarisations are

$$\begin{aligned} h_+ &= \frac{G_N}{r}(\ddot{M}'_{11}-\ddot{M}'_{22})\,, \\ h_\times &= \frac{2G_N}{r}\ddot{M}'_{12}\,, \end{aligned} \tag{2.138}$$

where these are the components evaluated in the (x',y',z') frame. We can write this in the (x,y,z) frame by observing that in the (x',y',z') frame the vector $\vec{n}$ has coordinates $n'_i=(0,0,1)$ while in the (x,y,z) frame it has coordinates

$$n_i = (\sin\theta\sin\phi, \sin\theta\cos\phi, \cos\theta)\,. \tag{2.139}$$

The two vectors are related by a rotation $n_i = R_{ij}n'_j$ with

$$R = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}\,. \tag{2.140}$$

A tensor with two indices transforms as

$$A_{ij} = R_{ik}R_{jl}A'_{kl}\,, \tag{2.141}$$

and from this one can obtain the transformed results. In general the two polarisations are given by:

$$
\boxed{
\begin{aligned}
h_+(t;\theta,\phi) &= \frac{1}{r}\Big[\ddot{M}_{11}(\cos^2\phi - \sin^2\phi\cos^2\theta) + \ddot{M}_{22}(\sin^2\theta - \cos^2\phi\cos^2\theta) - \ddot{M}_{33}\sin^2\theta \\
&\quad - \ddot{M}_{12}\sin 2\phi(1+\cos^2\theta) + \ddot{M}_{13}\sin\phi\sin 2\theta + \ddot{M}_{23}\cos\phi\sin 2\theta\Big]\,, \qquad (2.142) \\
h_\times(t;\theta,\phi) &= \frac{1}{r}\Big[(\ddot{M}_{11} - \ddot{M}_{22})\sin 2\phi\cos\theta \\
&\quad + 2\ddot{M}_{12}\cos 2\phi\cos\theta - 2\ddot{M}_{13}\cos\phi\sin\theta + 2\ddot{M}_{23}\sin\phi\sin\theta\Big]\,. \qquad (2.143)
\end{aligned}
}
$$

This allows us to compute the angular distribution of the quadrupole radiation once the second moments of the energy density are given.

The second moment is a symmetric matrix and therefore we can decompose it in terms of its irreducible representations (the symmetric traceless part and the trace part):

$$
M^{kl} = \left(M^{kl} - \frac{1}{3}\delta^{kl}M_{ii}\right) + \frac{1}{3}\delta^{kl}M_{ii}\,. \tag{2.144}
$$

The first term is traceless by construction and since the Lambda tensor gives zero when contracted with δ_{kl} only the traceless part contributes. We let $T^{00} = \rho$ and to the order we are working this is just the mass density. We define the *quadrupole* moment to be

$$
\begin{aligned}
Q^{ij} &\equiv M^{ij} - \frac{1}{3}\delta^{ij}M_{kk} \\
&= \int \mathrm{d}^3x\ \rho(t,\vec{x})\Big(x^i x^j - \frac{1}{3}r^2\delta^{ij}\Big)\,.
\end{aligned} \tag{2.145}
$$

Our final result is that the transverse traceless components of the quadrupole metric perturbation are given by

$$
[h_{ij}^{TT}(t,\vec{x})]_{\text{quad}} = \frac{2G_N}{r}\Lambda_{ij,kl}(\vec{n})\ddot{Q}_{kl}(t-r)\,. \tag{2.146}
$$

Where recall that the Lambda tensor was constructed using the projector

$$
P_{ij} = \delta_{ij} - \vec{n}_i\vec{n}_j\,, \tag{2.147}
$$

which eliminates the vector components parallel to $\vec{n}$, leaving only the transverse components and

$$
\Lambda_{ij,kl}(\vec{n}) = P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl}\,. \tag{2.148}
$$

We can now compute the power radiated away by the gravitational waves per unit solid angle. From (2.107) we have that the radiated power per unit solid angle is:

$$\begin{aligned} \left(\frac{\mathrm{d}P}{\mathrm{d}\Omega}\right)_{\text{quad}} &= \frac{r^2}{32\pi G_N}\langle \dot{h}^{TT}_{ij}\dot{h}^{TT}_{ij}\rangle \\ &= \frac{G}{8\pi}\Lambda_{ij,kl}(\vec{n})\langle \dddot{Q}_{ij}\dddot{Q}_{kl}\rangle\,. \end{aligned} \tag{2.149}$$

We have used that h^{TT}_{ij} is given by (2.146) and that the Lambda tensor satisfies (2.40). As before the average is a temporal average over several characteristic periods of the gravitational waves and it is understood that $\dddot{Q}$ is evaluated at the retarded time $t-r$. We can perform the angular integral by using the following identities:

$$\begin{aligned} &\frac{1}{4\pi}\int \mathrm{dvol}(S^2) = 1\,, \quad \frac{1}{4\pi}\int \mathrm{dvol}(S^2)\vec{n}^i\vec{n}^j = \frac{1}{3}\delta^{ij}\,, \\ &\frac{1}{4\pi}\int \mathrm{dvol}(S^2)\vec{n}^i\vec{n}^j\vec{n}^k\vec{n}^l = \frac{1}{15}(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})\,, \end{aligned} \tag{2.150}$$

with $\vec{n}^i$ a unit norm vector pointing out of the sphere along the i'th Cartesian direction. One finds

$$\int \mathrm{d}\Omega\Lambda_{ij,kl} = \frac{2\pi}{15}(11\delta_{ik}\delta_{jl} - 4\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk})\,. \tag{2.151}$$

The total radiated power[8] in the quadrupole approximation is

$$P_{\text{quad}} = \frac{G_N}{5}\langle \dddot{Q}_i(t-r)\dddot{Q}_{ij}(t-r)\rangle\,, \tag{2.152}$$

where $\dddot{Q}_{ij}$ is again evaluated at the retarded time. This is the *quadrupole* formula for energy loss via gravitational wave emission, valid far from a non-relativistic source. For explicit computations it is more practical to use the second moment M_{ij} rather than the quadrupole moment it is given by

$$P_{\text{quad}} = \frac{G_N}{5}\left\langle \dddot{M}_{ij}\dddot{M}_{ij} - \frac{1}{3}(\dddot{M}_{kk})^2\right\rangle\,. \tag{2.153}$$

We can also compute the radiated angular momentum per unit time following a similar procedure as above and our energy momentum tensor for the gravitational wave. The final result is

$$\boxed{\left(\frac{\mathrm{d}J^i}{\mathrm{d}t}\right)_{\text{quad}} = \frac{2G_N}{5}\epsilon^{ikl}\langle \ddot{Q}_{km}\dddot{Q}_{lm}\rangle\,,} \tag{2.154}$$

[8] In astrophysics this is called the gravitational luminosity.

where the derivatives are once again evaluated at the retarded time.

These formulae have been verified experimentally. In 1974 Hulse and Taylor identified a binary Pulsar. This is a neutron star binary in which one of the stars is a pulsar, emitting a beam of radio waves in a certain direction like a light-house. The star rotates rapidly and the beam periodically points in our direction. We therefore receive pulses of radio waves with the period being very reliable. These pulses have been measured to very high accuracy and act like a clock. This can be used to determine the orbital period of the binary system. One can now check how the period is decreasing due to the energy loss by gravitational waves and finds that it is about 7μs per year. This small effect has been measured and agrees with the quadrupole formula to an accuracy of 0.3%. This gives very strong indirect evidence for the existence of gravitational waves and Hulse and Taylor received the Nobel prize in 1993 for their work.

Summary A lot has just happened, let us take a quick step back to review what we have done. We have shown that far away from a slowly moving source the perturbations of the metric to leading order are encoded in the second moment of the energy density M^{ij} and its derivatives. From here we can compute various things, one interesting observable is the total power radiated by the gravitational waves. We can also see that

2.5.1 Binary merger

We now want to apply everything we have learnt in the past few lectures to understand the gravitational waves emitted by a binary pair and their merger. Strictly speaking the estimates we have made about slowly moving objects and weak gravitational forces do not allow us to fully describe the binary pair in the later stages of a merger however it will give us a good upper bound on the time to coalesce.

Consider a binary system with masses m_1 and m_2 and positions $\vec{r}_1$ and $\vec{r}_2$ respectively. Define the reduced mass and total mass to be

$$\mu = \frac{m_1 m_2}{m_1 + m_2}\,, \quad m = m_1 + m_2\,, \tag{2.155}$$

respectively. In the Newtonian approximation and working in the centre of mass frame the dynamics reduces to a one-body problem with mass given by the reduced mass μ and equation of motion

$$\ddot{\vec{r}} = -\frac{Gm}{r^3}\vec{r}, \tag{2.156}$$

where $\vec{r} = \vec{r}_2 - \vec{r}_1$ is the relative coordinate. We will consider the simpler case of circular orbits though this can be generalised to elliptic orbits, see section 4 of [6]. The orbital frequency ω

is related to the orbital radius R by $Rv^2 = G_N m$ with $v = \omega R$ and therefore we have Kepler's law:

$$\omega^2 = \frac{G_N m}{R^3} \,. \tag{2.157}$$

We assume that the orbital motion is given by the usual Newtonian trajectory and neglect any back reaction due to the gravitational wave emission. By choosing coordinates appropriately we will take the motion to be in the x, y-plane. Then we have that the motion is given by

$$\begin{aligned} x_0(t) &= R\cos(\omega t + \tfrac{\pi}{2})\,, \\ y_0(t) &= R\sin(\omega t + \tfrac{\pi}{2})\,, \\ z_0(t) &= 0\,. \end{aligned} \tag{2.158}$$

The phase $\frac{\pi}{2}$ is a useful choice of the origin of time. The 00 component of the energy-momentum tensor is given by

$$T^{00} = \mu\delta(z)\delta(x - x_0(t))\delta(y - y_0(t))\,. \tag{2.159}$$

The second moment of the energy density is given by

$$M^{ij} = \int \mathrm{d}^3y \; y^i y^j T_{00}(t, \vec{y})\,, \tag{2.160}$$

and it is simple to find

$$\begin{aligned} M_{11} &= \mu R^2 \frac{1 - \cos(2\omega t)}{2}\,, \\ M_{22} &= \mu R^2 \frac{1 + \cos(2\omega t)}{2}\,, \\ M_{12} &= -\frac{\mu R^2}{2}\sin(2\omega t)\,, \end{aligned} \tag{2.161}$$

with all the other components vanishing. Therefore we have

$$\begin{aligned} \ddot{M}_{11} &= -\ddot{M}_{22} = 2\mu R^2\omega^2\cos(2\omega t)\,, \\ \ddot{M}_{12} &= 2\mu R^2\omega^2\sin(2\omega t)\,. \end{aligned} \tag{2.162}$$

We can plug these expressions into the equation for h^{TT}_{ij} in equation (2.135):

$$h^{TT}_{ij}(t, \vec{x}) = \frac{2}{r}\Lambda_{ij,kl}(\vec{n})\ddot{M}^{kl}(t - r)\,. \tag{2.163}$$

Before we proceed it is useful to recall what we want. We are interested in the polarisations as they are the ones that appear in the formulae for the energy and we have from equations

(2.142) and (2.143) their forms in terms of the simplified moments. We find:

$$\begin{aligned} h_+ &= \frac{4G_N\mu\omega^2R^2}{r}\frac{1+\cos^2\theta}{2}\cos(2\omega(t-r)+2\phi)\,, \\ h_\times &= \frac{4G_N\mu\omega^2R^2}{r}\cos\theta\sin(2\omega(t-r)+2\phi)\,. \end{aligned} \tag{2.164}$$

Notice that the quadrupole frequency radiation is at twice the frequency of the source. It is also interesting to observe that the dependence on ϕ appears only in the combination $2\omega(t-r)+2\phi$. This can be understood from the fact that the source is not invariant under rotations around the z-axis since at any given value of t the mass μ is at a specific position along the orbit which changes by a rotation around the z-axis. Thus the polarisations have a dependence on ϕ. Moreover, since a rotation of the source by an angle $\Delta\phi$ is the same as a time translation Δt with $\omega\Delta t = \Delta\phi$ it must appear in the combination $\omega(t-r)+\phi$.

From an observational point of view we only have access to the radiation which points from the system to our direction. The angle θ is therefore the angle of incident and measures the angle between the normal to the orbit and the line of sight. The distance r is to all practical purposes a constant for astrophysical sources. As long as we can neglect the proper motion of the source (this need not be the case), the angle ϕ is fixed. If we see the binary pair on edge $\theta = \frac{\pi}{2}$ then we only see the plus polarisation and the cross polarisation vanishes.

We can get a feel for the expected strength of these binary gravitational waves. Replacing the frequency by using (2.157) we have

$$|h_{ij}| \sim \frac{G_N^2M^2}{Rr}\,. \tag{2.165}$$

The largest signal will be obtained by increasing the mass M and making the binary pair orbit as closely as possible. The densest objects are black holes with Schwarzschild radius $R_s = 2G_NM$, for a solar mass black hole we have that $R_s \sim 10$km. We can then make them orbit as close as possible with $R > 2R_s$. We then have

$$|h_{ij}| \sim \frac{GM}{r}\,. \tag{2.166}$$

It remains to give a distance from these black holes. If they were orbiting in the nearest galaxy Andromeda, which is roughly 2.5 million light years away we would have $|h_{ij}| \sim 10^{-17}$. This is an incredibly small number, yet this sensitivity, and better, has been achieved!

Let us now compute the radiated power in the quadrupole approximation. We have

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{r^2}{16\pi G_N}\langle \dot{h}_+^2 + \dot{h}_\times^2\rangle\,. \tag{2.167}$$

Inserting our polarisations we find

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{4G_N\mu^2\omega^6}{\pi}\left\langle \left(\frac{1+\cos^2\theta}{2}\right)^2 \sin^2(2\omega(t-r)) + \cos^2\theta\cos^2(2\omega(t-r)) \right\rangle . \tag{2.168}$$

The average was a time average and we have

$$\langle\sin^2(2\omega t)\rangle = \langle\sin^2(2\omega t)\rangle = \frac{1}{2}\,, \tag{2.169}$$

and therefore we find that the radiation power is

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{2G_N\mu^2\omega^6}{\pi} g(\theta)\,, \tag{2.170}$$

where

$$g(\theta) = \left(\frac{1+\cos^2\theta}{2}\right)^2 + \cos^2\theta\,. \tag{2.171}$$

We can integrate over the solid angle to find that the total radiated power is

$$P_{\text{quad}} = \frac{32}{5} G_N\mu^2 R^4\omega^6\,. \tag{2.172}$$

It is useful to define the *chirp mass* via

$$M_c = \mu^{3/5} m^{2/5}\,, \tag{2.173}$$

and to define the gravitational wave frequency $\omega_{\text{gw}} = 2\omega$. The total radiated power then takes the form:

$$P = \frac{32}{5G_N}\left(\frac{G_N M_c \omega_{\text{gw}}}{2}\right)^{10/3}\,. \tag{2.174}$$

We have used the polarisations assuming that the motion of the sources is on a given fixed, circular Keplerian orbit. However we have seen that the emission of gravitational waves costs energy. The source for the radiated energy is the sum of the kinetic plus potential energy of the orbit, which is

$$\begin{aligned} E_{\text{orbit}} &= E_{\text{kin}} + E_{\text{pot}} \\ &= -\frac{G_N m_1 m_2}{2R}\,. \end{aligned} \tag{2.175}$$

To compensate the loss of energy due to the emission of gravitational waves the separation R must decrease, and therefore E_{orbit} becomes more and more negative.[9] According to Kepler's

[9] In our idealised setting of point-like particles the masses have no internal degrees of freedom from which energy can be extracted and the only possible source of energy is the orbital energy of the system. For a more realistic system of two stars, at least early in the coalescence the orbital frequency is much smaller than the frequencies of the normal modes of the the star and therefore the internal dynamics of the stars is decoupled from the orbital motion and all the energy supplied comes from the orbital energy of the system. For compact objects corrections that depend on the internal structure of the bodies enters at very high order in the expansion and only at very late times in the coalescence.

law if we decrease R then ω must increase. On the other hand if ω increases then the power radiated in gravitational waves increases as one can see from (2.174). We then have a runaway process which on a sufficiently long time-scale, leads to the coalescence of the binary system. When $\dot{\omega} \ll \omega^2$ we are in the so-called quasi-circular regime. Using Kepler's law we have

$$\dot{R} = -\frac{2}{3}(\omega R)\frac{\dot{\omega}}{\omega^2}\,, \tag{2.176}$$

and one sees that so long as we are in the quasi-circular regime then the change in the radius is much smaller than the tangential velocity ωR and the approximation of a circular orbit with a slowly varying radius is applicable. We will study the backreaction of the gravitational waves in this regime. Outside of this one must take into account further orders in the expansion. We may use Kepler's law to write the energy of the orbit as

$$E_{\text{orbit}} = -\left(\frac{G_N^2 M_c^5 \omega_{\text{gw}}}{32}\right)^{1/3}\,. \tag{2.177}$$

At fixed ω_{gw} the dependence on the masses is again only through the chirp mass. Equating the radiated power with minus the change in the orbital energy we find

$$\dot{\omega}_{\text{gw}} = \frac{12}{5}2^{1/3}(G_N M_c)^{5/3}\omega_{\text{gw}}^{11/3}\,. \tag{2.178}$$

It is convenient to write it in terms of $f_{\text{gw}} = \omega_{\text{gw}}/(2\pi)$ and then to solve finding

$$f_{\text{gw}}(\tau) = \frac{1}{\pi}\left(\frac{5}{256\tau}\right)^{3/8}(GM_c)^{-5/8}\,. \tag{2.179}$$

Here we have introduced the variable $t = t_{\text{coal}} - t$ which measures the time to coalescence. We find that the total time to coalescence is given by

$$\tau_0 = \frac{5}{256}\frac{R_0^4}{G_N^3 m^2 \mu}\,. \tag{2.180}$$

The divergence is cut off since when their separation becomes smaller than a critical distance the two star merge. This now tells you the frequency of the gravitational wave as a function of the time to merge! You can plug in various numbers and see what happens.

Example 2.4: Time to coalescene

For two objects with mass $1.4M_\odot$ the chirp mass is $1.21M_\odot$. At 10 Hz we get radiation emitted at about $\tau = 17$ min to coalescence. At 100 Hz we get radiation from the last 2 seconds and at 1 kHz we get radiation from the last few milliseconds. From Kepler's law we find that $f_{\text{gw}} = 1$ kHz then the separation is $R \simeq 33$ km. Such a small separation can be reached only for very compact bodies like neutron stars and black holes. Since the radius of a neutron star with $m = 1.4M_\odot$ is about 10 km our point-like approximation at this stage is inaccurate, though not meaningless.

We have already seen that the orbital radius shrinks. We have that

$$\frac{\dot{R}}{R} = -\frac{2}{3}\frac{\dot{\omega}_{\rm gw}}{\omega_{\rm gw}} = -\frac{1}{4\tau}\,. \tag{2.181}$$

Integrating we find

$$R(\tau) = R_0\left(\frac{t_{\rm coal} - t}{t_{\rm coal} - t_0}\right)\,, \tag{2.182}$$

where R_0 is the initial value of R at the initial time t_0. We plot this in figure 4.

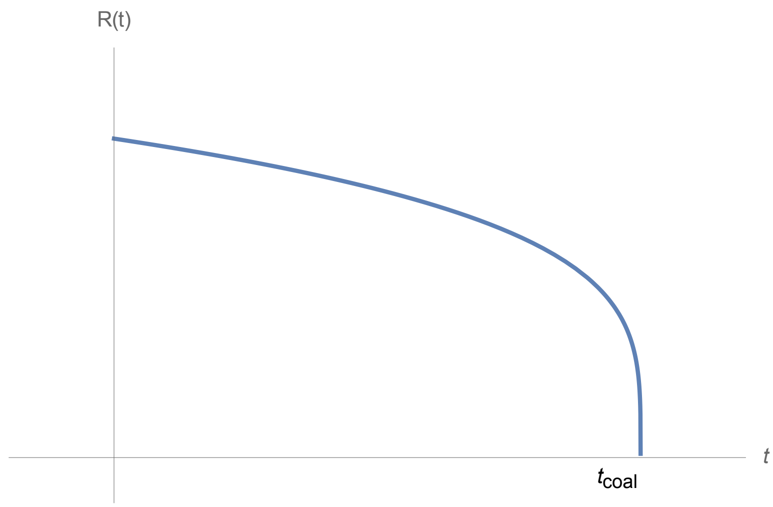

Figure 4: The change in the separation of the two masses.

We see that there is a long phase where R decreases smoothly before a plunge phase where our approximation breaks down. In this case our flat space approximation also breaks down.

So far we have only studied how the gravitational wave frequency evolves in time. We can now consider how the waveform changes. A particle that moves on a quasi-circular orbit in the (x, y) plane with a radius $R(t)$ and an angular velocity $\omega(t)$ has Cartesian coordinates

$$x(t) = R(t)\cos(\Phi(t)/2)\,, \quad y(t) = R(t)\sin(\Phi(t)/2)\,, \tag{2.183}$$

where

$$\Phi(t) = 2\int_{t_0}^{t} {\rm d}t'\,\omega(t') = \int_{t_0}^{t} {\rm d}t'\,\omega_{\rm gw}(t')\,. \tag{2.184}$$

We can now compute the gravitational wave production in the quadrupole approximation using the above. There are three difference

1. In the argument of the trigonometric functions $\omega_{\rm gw} t$ must be replaced by $\Phi(t)$.
2. In the factors in front of the trigonometric factors $\omega_{\rm grav}$ is replaced by $\omega_{\rm grav}(t)$.
3. We should include derivatives of $R(t)$ and $\omega_{\rm grav}$.

However as we have seen the radial velocity $\dot{R}$ is negligible as long as $\dot{\omega} \ll \omega^2$. The only changes are therefore the replacement of $\omega_{\rm gw} t$ by $\Phi(t)$ in the argument of the trigonometric functions and of $\omega_{\rm gw}$ by $\omega_{\rm gw}(t)$ in the prefactor evaluated at the retarded time. Then one finds that the polarisations are

$$\begin{aligned} h_+(t) &= \frac{4}{r}(G_N M_c)^{5/3}(\pi f_{\rm gw}(t_{\rm ret}))^{2/3}\frac{1+\cos^2\theta}{2}\cos\left(\Phi(t_{\rm ret})\right) , \\ h_\times(t) &= \frac{4}{r}(G_N M_c)^{5/3}(\pi f_{\rm gw}(t_{\rm ret}))^{2/3}\cos\theta \sin\left(\Phi(t_{\rm ret})\right) , \end{aligned} \tag{2.185}$$

where

$$\Phi(t) = -2(5G_N M_c)^{-5/8}\tau^{5/8} + \Phi_0 \, , \tag{2.186}$$

where Φ_0 is the integration constant equal to the value of Φ at coalescence. The final result, in terms of the time to coalescence time is

$$\begin{aligned} h_+(t) &= \frac{1}{r}(G_N M_c)^{5/4}\left(\frac{5}{\tau}\right)^{1/4}\frac{1+\cos^2\theta}{2}\cos\left(\Phi(\tau)\right) , \\ h_\times(t) &= \frac{1}{r}(G_N M_c)^{5/4}\left(\frac{5}{\tau}\right)^{1/4}\cos\theta\sin\left(\Phi(\tau)\right) . \end{aligned} \tag{2.187}$$

Plotting this we find figure 5.

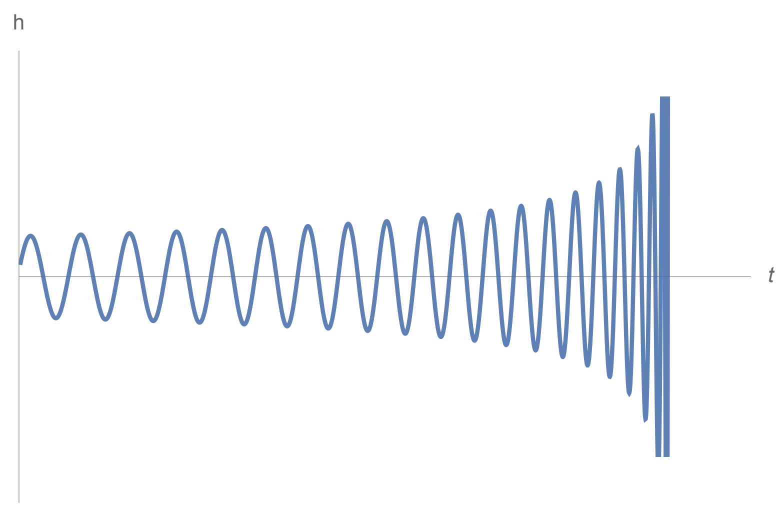

Figure 5: The time evolution of the gravitational wave amplitude in the inspiral phase of a binary system.

We see that both the frequency and amplitude increase as the coalescence is approached. This behaviour is known as chirping.

One can now do the same computations for a binary pair in elliptic orbits. The analysis is more difficult, see section 4.1.2 in [6] for example. One finds that the radiated power is

$$P = \frac{32G_N^4\mu^2m^3}{5a^5}\frac{1}{(1-e^2)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right), \tag{2.188}$$

where e is the eccentricity ($e = 0$ for the circular orbit). One finds that the change in the period, T is given by

$$\frac{\dot{T}}{T} = -\frac{96}{5}G_N^{5/3}\mu m^{2/3}(2\pi)^{8/3}T^{-8/3}f(e)\,. \tag{2.189}$$

This equation is the basis of the first experimental evidence for gravitational radiation. One can show that the radiation spectrum depends on the eccentricity, which allows for the determination of the eccentricity of the binary given the spectrum of gravitational waves. Furthermore the eccentricity of the binary pair decreases as the gravitational waves are emitted and the orbit becomes more circular. Unless some external interaction perturbs the binary system long before the two bodies approach the coalescence phase the ellipticity has become zero to high accuracy and the two bodies move on a circular orbit which shrinks adiabatically. For one of the original works on this understanding this circularisation see [9, 10].

2.6 Sources of gravitational waves

We have seen that for gravitational waves to be produced we require a time dependent quadrupole moment tensor. Astronomers and physicists have studied the various sources of gravitational waves in the universe. Four common groups of sources are: binary systems, spinning neutron stars, gravitational collapse and the Big Bang. Discussing each of these in detail is beyond the scope of the course however they each have important lessons for us.

Binary Systems This is the classic example of producing gravitational waves, and it is binary pairs that have been observed by LIGO. There is a lot of information that can be extracted from the gravitational waves emitted by a binary pair. The predicted waveforms from numerical simulations can be compared with the observed waveforms and gives a unique test of general relativity in the strongest possible gravitational fields. It also gives a way of studying black holes directly. After a merger of black holes or neutron stars has lead to a single black hole it will oscillate for a short time until it radiates away all its deformities and settles down to a smooth Kerr (see section 3) black hole. This *ringdown radiation* carries a distinctive signature that will distinguish the black holes from any neutron star or other possible origin.

A remarkable binary system was discovered by Hulse and Taylor in 1974. One of the stars is a pulsar with a period of 59 ms. The orbital period of the binary pair is 7.75 hours with the separation between the two stars about the radius of the Sun. The other star is not a pulsar and its presence has only been inferred rather than detected. A pulsar can be used as a very accurate clock and from observations the five parameters classifying the orbit can be obtained. These are the *inclination angle* of the orbit, the *longitude of the ascending node*, the *argument of the periastron*, the *orbital period* and the *eccentricity.* Two more parameters, defining the advance of the perihelion and the Doppler shift of the pulse period. The seven parameters allow a complete determination of the masses and orbital parameters. The orbital period changes slowly with time, shortening in duration as the two stars gradually approach each other. The inspiral is caused by the loss of orbital energy that has been carried off by the gravitational waves, you will calculate this in problem sheet 1. Therefore by monitoring the precise arrival times of the pulsar signals coming from the slowly decaying orbit the existence of gravitational radiation can be quantitatively confirmed and the quadrupole formula verified. This is despite the radiation itself not being visible.

Over 20 years of measurements the plots for the cumulative shift of the periastron are given in figure 6. The dots are the cumulative change in the time for the of periastron due to the progressively more rapid orbital period as the neutron stars inspiral due to the loss of energy due to the emission of gravitational waves. Plotting the observed values against the predictions from general relativity one finds excellent agreement. The line is not a fit to the data but the exact prediction from general relativity.

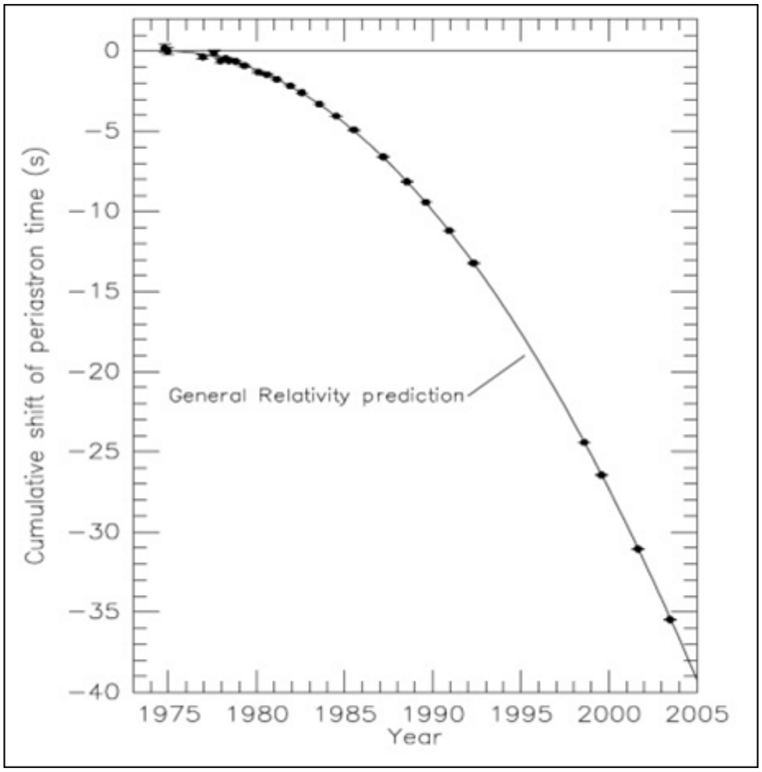

Figure 6: The change in the cumulative shift of periastron over a period of 30 years.

Cosmological gravitational waves In the early beginnings of the universe it is probable that there was a source of a random sea of gravitational radiation that forms a background that we observe today. The radiation originated in a host of individual events. The waves, now superimposed, have a very similar character to random noise. The expansion of the universe has cooled down the original radiation, the intensity of this radiation is still unknown. With more sensitive detectors it is expected that they will encounter a background noice which is isotropic in the sky. This is similar to the cosmic background radiation. For example gravitational waves from the Big Bang may lead to interesting tests of quantum gravity and the earliest moments of the universe if we can detect them. For the interested reader see [11] and references therein.

2.7 Laser Interferometers

Laser interferometry is based on the Michelson interferometer. Laser light is used to measure changes in the difference between two lengths of two perpendicular arms. Besides all manner of different sources of noise, such differences are induced by the strain of the a gravitational wave passing through the instrument. For LIGO the two arms have lengths of 4 km. Light sent from the laser light source to the beam splitter is divided evenly between the two arms. Having traversed the arms the light is reflected back to the splitter by mirrors at their far ends. On the return journey to the photodetector the roles of reflection and transmission in the splitter are interchanged for the two beams and the phase is inverted. The recombined beams heading towards the photodetector interfere destructively, while the beams heading back to the light source interfere constructively. The interference arms are adjusted so that the photodetector sees no light in general. A sufficiently strong gravitational wave passing through the interferometer can disturb this perfect balance and causes the two beams to no longer interfere destructively. One then measures a light signal which has the profile of the amplitude of the passing gravitational wave. This is an unbelievable feat of engineering. There are many sources of noise that must be dealt with, from vibrations from passing trucks to limitations in performing measurements due to quantum effects.

3 Spherical cold stars and stellar collapse

Birkhoff's theorem proves that the Schwarzschild solution is the unique asymptotically flat, spherically symmetric solution of Einstein's equations in the absence of matter and cosmological constant. As such, away from any spherically symmetric static object such as a star, planet or black hole the metric is the Schwarzschild metric. There are a few questions we may want to ask at this point. What is the metric inside a star where the Schwarzschild solution is no longer valid (since there is now a non-trivial contribution from the energy momentum tensor)? Does GR tell us anything about the different types of stars: hot stars, white dwarfs, neutron stars? In this section we answer these questions by studying the extension of the Schwarzschild solution to describe a *cold star*.

As opposed to a hot star, where there is a thermal source of pressure generated by nuclear reactions in its core, a cold star must be supported from collapse by a non-thermal pressure source. When a star forms by condensation of a dust cloud due to gravitational attraction the pressure increases which leads to an increase in temperature. When the dust cloud has collapsed far enough and has reached a critical temperature, nuclear fusion in the core begins. The dominant process is the conversion of four protons to form a helium-4 nucleus. The emission of photons and neutrinos at this stage provides a thermal radiation which balances against the collapse of the star due to gravity. As the Hydrogen fuel is depleted a helium core builds up and the pressure from thermal radiation decreases and the star begins to collapse again.

If the star is massive enough as the core contracts it once again heats up and if a critical temperature is reached, helium can be fused, giving a thermal pressure which halts the collapse. If the star is not big enough, the temperature which allows Helium to fuse is not reached and the star uses up its remaining fuel becoming a red dwarf. This process of a period of equilibrium followed by collapse can keep repeating with the formation of heavier nuclei in the core such as nickel and iron.

The crucial issue governing how far along this evolutionary sequence a star goes is whether electron degeneracy pressure becomes sufficient to support the star from further collapse. There is a critical mass M_C, (**??**), below which the collapse is halted by the electron degeneracy pressure. The Pauli exclusion principle states that two or more identical fermions[10] cannot occupy the same quantum state within a quantum system simultaneously.[11] Due to

[10] A fermion is a particle with half integer spin. Fermions obey Fermi–Dirac statistics. Quarks and leptons (electrons, muons and tau-ons and their neutrino versions) are examples of fermions.

[11] To get a feel of why this is true one needs to recall some facts about the wave-function in quantum

this a gas of cold fermions resists compression, producing a pressure known as *degeneracy pressure*. If the mass of the star is below the critical mass no further nuclear fusion will occur and the star will simply cool down forever in a stable *white dwarf* configuration. This is the fate of our sun. A white dwarf is much denser than a regular star: a matchbox sized piece of white dwarf material would weigh roughly the same as an elephant. Newtonian gravity is still applicable here and shows that a white dwarf cannot have a mass greater than the *Chandrasekhar limit*, 1.4 $\mathrm{M}_\odot$ with $M_\odot$ the mass of the Sun. A star more massive than this cannot end its life as a white dwarf unless it sheds some of its mass.

If M is greater than M_C then after a core of nickel and iron of mass $\sim M_C$ has formed it will be unable to support itself, electron degeneracy pressure is insufficient and no further nuclear fusion occurs. The core will undergo gravitational collapse once again. When the density of the core reaches nuclear density, the density of the nucleus of an atom, neutron degeneracy pressure and nuclear forces provide a significant cold matter pressure. At such high pressure one finds that beta decay is reversed, protons combine with electrons to produce neutrons. If the mass of the star is below the critical limit for cold matter M_{critical} $2M_\odot$ then the collapse will be halted leading to a *neutron star*. At this stage the Newtonian approximation is no longer applicable and one must use general relativity.

When the collapse of the core is halted or slowed at nuclear densities a shock wave is produced and this is expected to lead to the outer envelope of the star producing a supernova. The presence of pulsars (neutron stars with a hot spot rotating at high speed) at the sites of the Crab and Vela supernova remnants provides strong evidence that this supernovae are produced in conjunction with the collapse of the core of a star at the end-point of stellar evolution.

The final option is to have a star which has a mass larger than the critical mass M_{critical}. Equilibrium can never be achieved and complete gravitational collapse will occur. The end-point of such a collapse will be a Schwarzschild black hole. We find that for a massive enough star gravitational collapse into a black hole is inevitable.[12]

---

mechanics. We construct a state by acting on the ground state with operators. Operators which give bosons (integer spin field) satisfy commutation relations, while operators which give rise to fermions satisfy anti-commutation relations. If we want to insert the same (all quantum numbers the same) fermion at the same point we must act with the same operator but due to the anti-commutator relations this vanishes and therefore the wave-function vanishes.

[12] One can formulate this more concretely following Penrose and Hawing that collapse becomes inevitable once a *trapped surface* forms. A trapped surface is a two-dimensional surface for which both the out-going and in-going future directed geodesics orthogonal to the surface converge. For example consider spheres with r, t constant in the Schwarzschild metric, these are trapped surfaces for $r < R_{\text{Schwarzschild}}$.

In this section we will show that general relativity predicts a maximum mass for a cold star. To reach this conclusion we will assume that the star is spherically symmetric and static, recall that this is one of the assumptions that goes into Birkhoff's theorem. The interior of the star can be modelled by a perfect fluid and we then need to solve Einstein's solutions with a perfect fluid source and match onto the Schwarzschild solution outside the star.

3.1 Tolman–Oppenheimer–Volkoff equations

Since we have a static spacetime we have a timelike Killing vector field K with which we can foliate our spacetime with the surfaces Σ_t which are orthogonal to K. The orbits of SO(3) through a point $p \in \Sigma_t$ lie within Σ_t. This allows us to define coordinates (r, θ, ϕ) such that the most general metric with our given assumptions takes the form

$$\mathrm{d}s^2 = -\mathrm{e}^{2\Phi(r)}\mathrm{d}t^2 + \mathrm{e}^{2\Psi(r)}\mathrm{d}r^2 + r^2\mathrm{d}s^2(S^2)\,. \tag{3.1}$$

We now need to specify the energy-momentum tensor. Outside the star this vanishes and it remains to come up with a suitable ansatz within the star. We can describe this as a static perfect fluid. The energy momentum tensor for a perfect fluid takes the form

$$T_{\mu\nu} = (p+\rho)u_\mu u_\nu + p g_{\mu\nu}\,, \tag{3.2}$$

with u_μ the four-velocity of the fluid, normalised to $u_\mu u^\mu = -1$, ρ the energy density and p the pressure measured in the fluid's local rest frame. Since we are interested in time-independent and spherically symmetric stars the fluid should be at rest thus u points in the time-direction only and therefore the correctly normalised vector field u is

$$u = \mathrm{e}^{-\Phi(r)}\partial_t\,. \tag{3.3}$$

Moreover the time-independence and spherical symmetry imply that ρ and p only depend on r while the vanishing of the energy-momentum tensor outside of the star implies that ρ, p vanish when $r > R_c$ with R_c the radius of the star.

A fluid's equations of motion are determined by the conservation of the energy momentum tensor. This follows from the Einstein equations, ergo we need only consider the Einstein conditions in the following. Since the Einstein equations inherit the symmetries of the spacetime it follows that there are only three non-trivial independent conditions arising from the Einstein equations. We may take these to be the $tt, rr, \theta\theta$ components, (see the mathematica file in moodle which does this computation.)

The independent Einstein equations are

$$
\begin{aligned}
E_{tt} &= \frac{\mathrm{e}^{2\Phi}}{r^2}\left[\frac{\mathrm{d}}{\mathrm{d}r}\Big(r(1-\mathrm{e}^{-2\Psi})\Big) - 8\pi r^2\rho\right] = 0\,,\\
E_{rr} &= \frac{1}{r}\left[\mathrm{e}^{-2\Phi}\partial_r\mathrm{e}^{2\Phi} - \frac{\mathrm{e}^{2\Psi}-1}{r} - 8\pi r\mathrm{e}^{2\Psi}p\right] = 0\,,\\
E_{\theta\theta} &= \mathrm{e}^{-2\Psi}r\left[\mathrm{e}^{\Psi-\Phi}\partial_r\Big(r\mathrm{e}^{-\Psi}\partial_r\mathrm{e}^{\Phi}\Big) - \partial_r\Psi - 8\pi r\mathrm{e}^{2\Psi}p\right] = 0\,.
\end{aligned}
\tag{3.4}
$$

To proceed it is useful to introduce $m(r)$ via

$$
\mathrm{e}^{2\Psi(r)} = \left(1-\frac{2m(r)}{r}\right)^{-1}, \tag{3.5}
$$

with $2m(r) < r$. With this definition the tt component of the Einstein equation becomes

$$
\frac{\mathrm{d}m(r)}{\mathrm{d}r} = 4\pi r^2\rho(r)\,. \tag{3.6}
$$

Furthermore the rr component reduces to

$$
\frac{\mathrm{d}\Phi(r)}{\mathrm{d}r} = \frac{m(r)+4\pi r^3 p(r)}{r(r-2m(r))}\,. \tag{3.7}
$$

In the Newtonian limit we have $r^3p(r) \ll m(r)$ and $m(r) \ll r$ so (3.7) reduces to

$$
\frac{\mathrm{d}\Phi(r)}{\mathrm{d}r} \approx \frac{m(r)}{r^2}\,, \tag{3.8}
$$

this is just the spherically symmetric version of Poisson's equation for the Newtonian gravitational potential. We can see the other terms in (3.7) as relativistic corrections.

The final non-trivial component of the Einstein equations is the $\theta\theta$ component given above, however rather than using that equation, it is simpler to derive the final equation from the r-component of energy momentum conservation. This gives

$$
\frac{\mathrm{d}p(r)}{\mathrm{d}r} = -\big(p(r)+\rho(r)\big)\frac{m(r)+4\pi r^3p(r)}{r(r-2m(r))}\,. \tag{3.9}
$$

One can check that this is implied by $E_{\theta\theta} = 0$ above, see the mathematica file. In the Newtonian limit ($p \ll \rho, m(r) \ll r$) it reduces to the Newtonian hydrostatic equilibrium equation

$$
\frac{\mathrm{d}p(r)}{\mathrm{d}r} \approx -\frac{\rho(r)m(r)}{r^2}\,. \tag{3.10}
$$

Note that general relativity has little effect on the equilibrium configurations of stars with $p \ll \rho$ and $m(r) \ll r$. a Newtonian treatment is sufficient.

We have four unknown functions $\big(m(r), \Phi(r), \rho(r), p(r)\big)$ and only three equations so the system is currently underdetermined. The one remaining condition comes from the fact that we are interested in a cold star, one which has a vanishing temperature. Thermodynamics implies that T, ρ, p are not independent, and therefore we may write $p = p(\rho)$. Moreover, we should take $\rho > 0$ and $p > 0$ and that $p(\rho)$ is an increasing function of ρ.[13] The three equations (3.6), (3.7) and (3.9) are known as the *Tolman–Oppenheimer–Volkoff* equations.

Outside the star We know that in the absence of matter and with the imposed constraints, that the unique solution is the Schwarzschild solution:

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-2}\mathrm{d}r^2 + r^2\mathrm{d}s^2(S^2)\,, \tag{3.11}$$

with the constant M the total mass of the star. Recall that $R_s = 2M$ is the Schwarzschild radius where an event horizon is located. We must therefore take the star to have a radius larger than the Schwarzschild radius: $R_c > R_s$. Regular stars have $R_c \gg R_s$, for the Sun $R_s \approx 3$km while $R_c \approx 7 \times 10^5$km. We define the location of the surface of the star, R_c as the point where $p = 0$. Since the outside is described by the Schwarzschild solution and we require that the metrics patch together smoothly, we have that

$$M \equiv m(R_c)\,. \tag{3.12}$$

Inside the star We now want to consider the interior of the star, and patch it with the exterior solution above such that the full metric is smooth at the patching surface at $r = R_c$. We can integrate (3.6) to give

$$m(r) = 4\pi \int_0^r \rho(r')r'^2\mathrm{d}r' + m_*\,, \tag{3.13}$$

with m_* an integration constant.

At $r = 0$ the solution should be smooth and look like flat Minkowski space, the net gravitational attraction at the centre is zero and is therefore equivalent to Minkowski space. This implies that as $r \to 0$ we have $\mathrm{e}^{2\Psi(0)} = 1$. Comparing with (3.5) we see that this is equivalent to $m(0) = 0$. From our integrated solution, (3.13) we see that this implies that the integration constant vanishes, $m_* = 0$.

[13] If this were not the case then the star would be unstable since a fluctuation in some region that led to an increased energy density would lead to a decrease in pressure. This would cause the fluid to more into this region which would lead to a further increase in ρ and the fluctuation would continue to grow.

At $r = R_c$, for our interior solution to match with the Schwarzschild solution, we need to impose the boundary condition

$$M = 4\pi \int_0^{R_c} \rho(r) r^2 \mathrm{d}r \,. \tag{3.14}$$

Aside
There is a slight subtlety here in that the total energy of the matter should include the correct volume measure when integrating over a spacelike hypersurface, the energy for the spacelike hypersurface Σ_t is defined to be

$$E = \int_{\Sigma_t} \rho(r) \mathrm{dvol}(\Sigma_t) = \int_{\Sigma_t} \rho(r) \mathrm{e}^{\Psi(r)} r^2 \sin\theta \mathrm{d}r \wedge \mathrm{d}\theta \wedge \mathrm{d}\phi = 4\pi \int_0^{R_c} \rho(r) \mathrm{e}^{\Psi(r)} r^2 \mathrm{d}r \,. \tag{3.15}$$

Note that this differs with the total mass of the star due to the $\mathrm{e}^{\Psi(r)}$ factor. Since $\mathrm{e}^{\Psi(r)} > 1$ it follows that $E > M$ and one can associate the positive difference $E - M$ to be the gravitational binding energy of the star. This would be the amount of energy needed to disperse the matter to infinity, for spherical stars this is a well-defined concept but does not always make sense in GR.

Note that due to the constraint that $2m(r) < r$ for all r, which imposes that $\mathrm{e}^{\Psi(r)} > 0$ it follows that there is an upper bound on the possible mass of the star: $2M < R_c$. There is no Newtonian analogue of this condition. Reinstating the factors of c and G_N we have $2G_N M < c^2 R_c$ and in the $c \to \infty$ limit this is trivial, hence why this constraint is not seen in the Newtonian theory.

In order to solve the TOV equations we should use numerical integration. We view (3.6) and (3.9) as a coupled set of ODEs for $m(r)$ and $\rho(r)$ for some given equation of state. These can be solved, at least numerically on a computer once initial conditions for the mass and density are given. We have that $m(0) = 0$ and therefore we ned only specify a density $\rho_c = \rho(0)$ at the centre of the star.

Given these initial conditions we can numerically solve (3.6) and (3.9). Since the latter equation shows that p decreases with r there must be some point where the pressure vanishes, this is the surface of the star and the radius is determined by $p(R_c) = 0$. We can invert this to determine R_c as a function of ρ_c. From (3.14) we can determine M as a function of ρ_c. Finally we may determine $\Phi(r)$ inside the star by integrating (3.7) from the surface of the star with initial condition that $2\Phi(R_c) = \log(1 - 2M/R_c)$, i.e. it gives the Schwarzschild solution potential. Hence for a given equation of state, static, spherically symmetric cold stars are form a 1-parameter family of solutions labelled by the central density ρ_c.

3.2 Buchdahl's limit

We will now, using Einstein's equations, find bounds on the compactness of stars. We define the compactness of a spherically symmetric configuration to be the quotient of twice the total mass of the object with the stars radius:

$$C(R_c) = \frac{2M}{R_c} \,. \tag{3.16}$$

Typical Neutron stars have compactness around $C(R_c) \sim 0.4$ while at the extreme end $C(R_c) = 1$ is a black hole. We will derive the Buchdahl limit which states that $C(R_c) < 8/9$. Buchdahl proved this by making no strong hypothesis about the equation of state except that the matter is barotropic (density a function of pressure only).

Theorem 1 *Buchdahl's theorem*

Consider a solution to the TOV equations assuming a perfect fluid with an equation of state that fulfils the two following properties:

1. *The function* $\mathrm{e}^{2\Phi(r)}$ *is at least a* C^1 *function while* $\mathrm{e}^{2\Psi(r)}$ *is at least* C^0*. At* $r = R_c$ *the solution matches with the Schwarzschild solution.*
2. *The density is a monotonically decreasing function. Continuity and monotonicity assumptions along with the boundary conditions require that* $\rho(r) \geq 0$*.*

For any such solution the compactness satisfies the inequality $C(R_c) < 8/9$*.*

Demanding that the solution is well defined immediately enforces $2M/R_c \leq 1$ since $C(R_c) = 1$ would lead to a diverging pressure. This is of course just the black hole limit.

To proceed we will take suitable combinations of the Einstein equations and then integrate. Notice that because of the isotropic pressure both the rr and $\theta\theta$ Einstein equations depend only on p and therefore by taking a suitable linear combination we can eliminate the pressure. We take

$$\begin{aligned} 0 &= E_{\theta\theta} - r^2 \mathrm{e}^{-2\Psi} E_{rr} \\ &= r^3 \mathrm{e}^{-\Phi-\Psi} \left[\partial_r \left(\frac{\mathrm{e}^{-\Psi}}{r} \partial_r \mathrm{e}^{\Phi} \right) - \frac{\mathrm{e}^{\Psi+\Phi}}{2} \partial_r \left(\frac{1-\mathrm{e}^{-2\Psi}}{r^2} \right) \right] . \end{aligned} \tag{3.17}$$

Plugging in our redefinition of Ψ we have

$$\partial_r \left(\frac{\mathrm{e}^{-\Psi}}{r} \partial_r \mathrm{e}^{\Phi} \right) = \mathrm{e}^{\Phi+\Psi} \partial_r \left(\frac{m(r)}{r^3} \right) . \tag{3.18}$$

We now use our second assumption that ρ is monotonically decreasing. The right-hand side is the derivative of the average density and therefore it must be non-positive since the average density should also be a monotonically decreasing function. This implies that the left-hand side must be non-positive, hence

$$\partial_r \left(\frac{\mathrm{e}^{-\Psi}}{r} \partial_r \mathrm{e}^{\Phi} \right) \leq 0 \,. \tag{3.19}$$

We can integrate this inequality from the surface into the star to some smaller radius r which gives

$$\left(\frac{\mathrm{e}^{-\Psi}}{r} \partial_r \mathrm{e}^{\Phi} \right) \bigg|_{R_c} - \left(\frac{\mathrm{e}^{-\Psi}}{r} \partial_r \mathrm{e}^{\Phi} \right) \bigg|_{r} \leq 0 \,. \tag{3.20}$$

Rearranging and using the replacement for Ψ and Φ on the surface we find

$$\left(\frac{\mathrm{e}^{-\Psi}}{r} \partial_r \mathrm{e}^{\Phi} \right) \bigg|_{r} \geq \frac{M}{R_c^3} \,, \tag{3.21}$$

where we have used the continuity conditions on the surface to write the right-hand side. We can now multiply both sides by $r\mathrm{e}^{\Psi}$ and integrate again, this time from the surface to the centre at $r = 0$. One finds

$$\mathrm{e}^{\Phi} \Big|_{r=R_c} - \mathrm{e}^{\Phi} \Big|_{r=0} \geq \frac{M}{R_c^3} \int_0^{R_c} r \mathrm{e}^{\Psi} \mathrm{d}r \,. \tag{3.22}$$

We know the matching condition for Φ on the surface and also that e^{Φ} must be positive everywhere and therefore we have

$$\mathrm{e}^{\Phi(0)} \leq \sqrt{1 - C(R_c)} - \frac{M}{R_c^3} \int_0^{R_c} \mathrm{d}r \frac{r}{\sqrt{1 - \frac{2m(r)}{r}}} \,. \tag{3.23}$$

Since we assumed monotonicity on the density the smallest value of $m(r)$ is then the value it would have for a uniform density function thus we have the inequality

$$m(r) \geq \frac{M r^3}{R_c^3} \,. \tag{3.24}$$

The best upper bound is then achieved by inserting the uniform density value of $m(r)$ finding

$$\mathrm{e}^{\Phi(0)} \leq \frac{3}{2} \sqrt{1 - C(R_c)} - \frac{1}{2} \,. \tag{3.25}$$

We know that $\mathrm{e}^{\Phi(0)} > 0$ and therefore we find:

$$C(R_c) < \frac{8}{9} \,. \tag{3.26}$$

This is called the Buchdahl limit.

Remarks

- We began with a perfect fluid and an isotropic distribution of pressures. In the anisotropic case in which pressures in the angular directions are allowed to grow without bound there is no limit on the compactness of the object.

- The proof tells us which density profile saturates the bound, it is the uniform density profile.

- The proof required the monotonicity condition on the density. If this is lifted then the bound is not as sharp (or absent) depending on what other conditions one imposes. See [12] for examples where these assumptions are lifted and different bounds found.

Suppose we manage to construct a star in equilibrium with a radius $R_c = 9M/4$ and gave it a spherically symmetric push inward. It has no choice but to collapse inwards and can never reach a static state again. During the collapse the metric outside is just the Schwarzschild metric and therefore once it has fully collapsed the remaining metric is just the Schwarzschild solution in a vacuum, the metric of a black hole!

One can improve the bound further. Part of the issue is knowing the equation of state inside a high density star in thermal equilibrium. This is a strongly coupled theory and we are ignorant as to the exact details. Instead, we may know the equation of state in some region (envelope) $r_0 < r < R_c$ connected to the outside. The core is then the region $0 < r < r_0$ where the exact equation of state is unknown to us. At $r = r_0$ we know that the density is ρ_0 and inside the core we have $\rho > \rho_0$ and in the envelope $\rho < \rho_0$.

Exercise 2:
From equation (3.9) after some algebra, which you will do in sheet 2 and assuming $\rho \geq 0$ and $\rho'(r) \leq 0$, one finds that

$$\frac{m(r)}{r} \leq \frac{2}{9}\left[1 - 6\pi r^2 p(r) + \sqrt{1 + 6\pi r^2 p(r)}\right]. \tag{3.27}$$

Evaluating on the radius of the star where $p = 0$, one finds

$$R_c \geq \frac{9M}{4}. \tag{3.28}$$

Note that this is actually independent of the equation of state and so it applies equally to hot stars and cold stars which satisfy these assumptions. Stars of uniform constant density can get arbitrarily close to saturating the bound but as they get closer to the bound the pressure at the centre diverges.

We may then use (3.27) and evaluate it on the envelope boundary to the core. The mass of the core is $m_0 = m(r_0)$ with r_0 the boundary of the envelope. We thus find:

Since the density in the core is bigger than the density on the boundary with the envelope we must have that

$$m_0 \geq \frac{4\pi r_0^3 \rho_0}{3} \,. \tag{3.29}$$

Note that Newtonian gravity would also predict this inequality, however in GR we also have the additional constraint (3.27) which we should evaluate at $r = r_0$ where we know the equation of state and may therefore determine $p_0 = p\big(\rho(r_0)\big)$:

$$\frac{m_0}{r_0} < \frac{2}{9}\left[1 - 7\pi r_0^2 p_0 + \sqrt{1 + 6\pi r_0^2 p_0}\right] . \tag{3.30}$$

Since the RHS is a decreasing function of p_0 evaluating at $p_0 = 0$ we get the weaker bound

$$m_0 < \frac{4r_0}{9} \,. \tag{3.31}$$

These two inequalities define a finite region in the $m_0 - r_0$ plane. Hence, even though we are ignorant of the equation of state within the core, GR predicts that its mass cannot be arbitrarily large.

Using (3.29) to eliminate r_0 and plugging this into (3.31) we have

$$m_0 < \frac{4}{9\sqrt{3\pi\rho_0}} \,. \tag{3.32}$$

Hence, even though we do not know the equation of state inside the core, GR predicts that its mass cannot be indefinitely large. Experimentally we know the equation of state of cold matter at densities much higher than the density of atomic nucei so we take $\rho_0 = 5 \times 10^{14}\mathrm{g/cm}^3$. Plugging this into the above gives the bound $m_0 < 5M_\odot$.

If we are given a core with mass m_0 and radius r_0 we can solve (numerically) for the envelope region using the known equation of state and the equations for $m(r)$ and $p(r)$ with the initial conditions given by the core. If one plugs this into a computer programme one finds that the maximal mass M as a function of ρ_0, m_0. One can then vary this over the allowed region for (m_0, r_0) one finds that the largest mass is attained for the maximum of m_0. At this maximum the envelope contributes less than 1% of the total mass so the maximum mass of M is at almost the same as the maximum of m_0 and we have $M \leq 5M_\odot$.

This is an upper bound for *any* physically reasonable equation of state for $\rho > \rho_0$. Any equation of state will have a smaller upper bound than the one given here. One may put further constraints on what we call a physically reasonable equation of state. A natural demand is that the speed of sound through the mass should not exceed the speed of light, so that $\frac{\mathrm{d}p}{\mathrm{d}\rho} \leq 1$, then the upper bound is further reduced to about $3M_\odot$.

3.3 Summary

What have we learnt from this exercise? Firstly we see once again that GR predicts something that Newtonian gravity cannot, we find an upper bound on the maximal size of any cold star, independent of its composition. Secondly, this has an extremely important consequence for the ultimate fate of a star. Ordinary hot starts are supported against collapse under their own weight by ideal gas pressure resulting from their high temperature. This pressure is much higher than the pressure that can be produced by cold matter at comparable densities and so the above upper limits do not apply. However, since a hot star radiates energy, just look out the nearest window during the day, if this energy is not replenished hydrostatic equilibrium cannot be maintained. As the fuel source is used up the hydrostatic equilibrium is lost and it begins to contract until the cold matter pressure dominates the remaining thermal pressure. If the star was small enough a stable equilibrium may be reached using cold matter pressure and will remain like this forever. However if the mass is greater than the cold matter upper limit, equilibrium can never be achieved and the star would have to undergo complete gravitational collapse unless they shed some of their mass to bring their total mass below the upper bound.

This is a very active area of research. Trying to understand the equation of state of neutron stars remains an open problem. There are bounds on the possible equation of state, and with these bounds on the maximum mass a neutron star can have, see for example [13]. The core of the neutron star is described by a strongly coupled field theory: quantum chromodynamics (QCD) which is not amenable to perturbation techniques you are familiar with. Neutron stars occupy the low temperature, large chemical potential region of the QCD phase diagram. One technique that has had some success recently is to use holography: see [14] for a review.

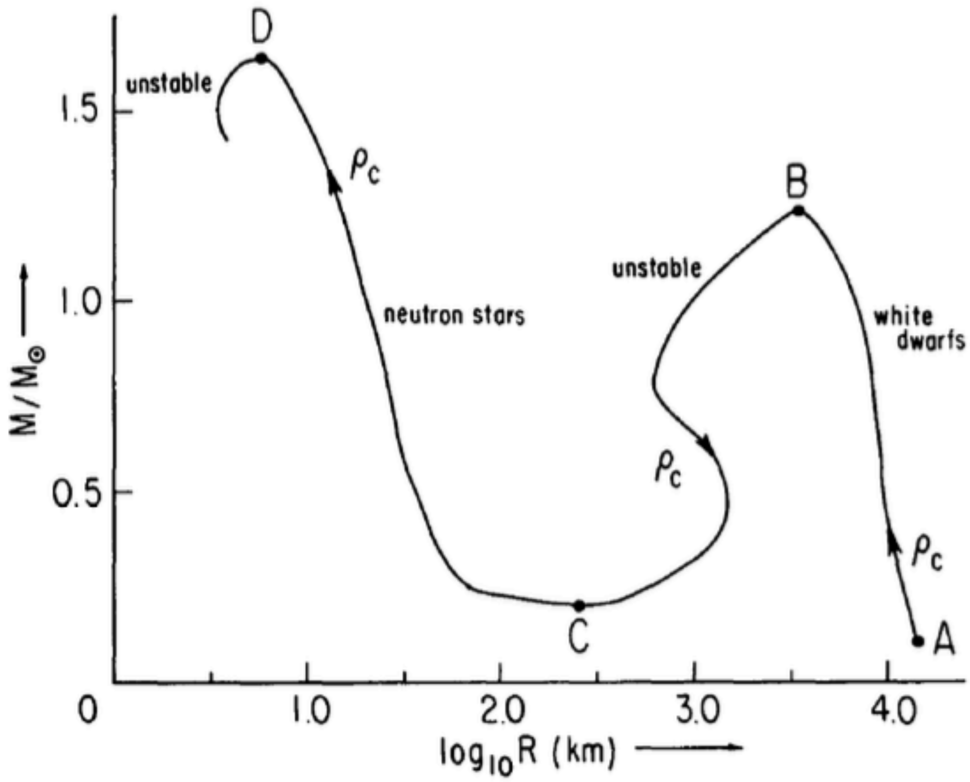

Figure 7: The equilibrium configurations of cold matter. Given an equation of state the equilibrium configuration is uniquely determined by the central density ρ_c. The radii and masses of these configurations are shown for values of ρ_c ranging from $\approx 10^5$g cm^{-3} at point A to $\approx 10^17$g cm^{-3} beyond point D. In the white dwarf regime the values of M and R_c depend somewhat on the assumed composition of the star. The neutron star regime is far more dependant on the assumptions that go into the equation of state, and interactions between the fundamental constituents of the matter. In the latter regime this is just a rough sketch of the qualitative features. The point B is the Chandrasekhar limit and beyond this the white dwarf must undergo further gravitational collapse to become a neutron star. It is at this point that the electron degeneracy pressure is insufficient to prevent gravitational collapse and therefore the equation of state changes past this point.
Figure taken from Wald based on a figure by Harrison, Thorne, Wakano and Wheeler.

4 What is a black hole: Causality and Penrose Diagrams

What is a black hole? From studying the Schwarzschild solution we can say some abstract words about the existence of a horizon, light not being able to escape and so forth. However these are just words and what we really want is a mathematical definition of what a black hole is. In this section we will answer this question.

We will begin by understanding how we can draw finite diagrams of our four-dimensional spaces in order to probe the causal structure of the spacetime. To do this we will conformally compact our spacetime and from this draw a so-called *Penrose diagram*. We will then turn our attention to investigating what singularities we are allowed and how we can detect them. We know that there are singularities in the Schwarzschild metric in Schwarzschild coordinates, the singularity at $r = 0$ is physical while the one at $r = 2M$ is not physical. We understand that the second singularity is a coordinate singularity and by performing a change of coordinates to Eddington–Finkelstein coordinates it may be removed. We want to understand is how we can detect such singular points and understand what type of singularity they are.

4.1 Conformal compactification

Let us consider a spacetime M. One of the postulates that we demand General Relativity satisfies is that it is causal. A signal can be sent between two distinct points if and only if the points can be joined by a non-spacelike curve. Our goal in this section is to investigate the properties of causality on spacetime. Given that our spacetimes are generically infinite in extent this can be difficult to understand on a piece of paper. There is a useful way of resolving this issue called *conformal compactification*.

Definition 2 *Conformal transformation A* conformal transformation *is a map from a spacetime (M, g) to a spacetime $(M, \tilde{g})$ such that*

$$\tilde{g}_{\mu\nu}(x) = \Omega(x)^2 g_{\mu\nu}(x)\,, \tag{4.1}$$

where $\Omega(x)$ is a smooth function of the spacetime coordinates and $\Omega(x) \neq 0$ for all $x \in M$.

One reason why conformal transformations are useful is because they preserve the causal structure of spacetime. Consider a vector V^μ on M, not necessarily a geodesic. Then since $\Omega(x)^2 > 0$ it follows that

$$\begin{aligned} g_{\mu\nu}V^\mu V^\nu > 0 \quad &\Leftrightarrow \quad \tilde{g}_{\mu\nu}V^\mu V^\nu > 0\,, \\ g_{\mu\nu}V^\mu V^\nu = 0 \quad &\Leftrightarrow \quad \tilde{g}_{\mu\nu}V^\mu V^\nu = 0\,, \\ g_{\mu\nu}V^\mu V^\nu < 0 \quad &\Leftrightarrow \quad \tilde{g}_{\mu\nu}V^\mu V^\nu < 0\,. \end{aligned} \tag{4.2}$$

Hence curves which are timelike, null or spacelike with respect to one metric remain timelike, null or spacelike respectively in the conformally rescaled metric. Moreover one can show that two spacetimes whose metrics are related by a conformal transformation have the same null geodesics. However, timelike and spacelike geodesics in one metric will not necessarily be geodesics in the other.

Exercise 3:
Problem sheet 1.

- Show that under a conformal transformation null geodesics remain null geodesics.
- Show that timelike geodesics need not be geodesics in the conformally transformed metric.

We may use this to our advantage when studying the causal structure of spacetime. By using a suitably chosen conformal factor we may bring "infinity" to a finite coordinate distance. This allows us to draw the causal structure on a finite piece of paper. This object is known as a *Penrose diagram* and encodes the causal structure of the spacetime.

The general procedure for drawing a Penrose diagram is to perform the following steps.

- First change coordinates on (M, g) such that "infinity" is brought to finite coordinate distance. This then allows us to draw the spacetime on a finite piece of paper. The points at "infinity" will become the edges of the finite diagram. Typically the metric will diverge at these points.
- To remedy the divergences we perform a conformal transformation on g to obtain $\tilde{g}$ which is regular on the edges. The new pair $(M, \tilde{g})$ is a good representation of the original spacetime (M, g) for understanding the causal structure: they have the exact same causal structure.
- It is customary to add the points at infinity to the spacetime to form a new manifold $\tilde{M}$ (with boundary now). The resulting spacetime $(\tilde{M}, \tilde{g})$ is often called the *conformal compactification* of (M, g).

Note that this has some limitations. Conformal transformations generically change the curvature tensors so that $\tilde{R}_{\mu\nu\rho\sigma} \neq R_{\mu\nu\rho\sigma}$, $\tilde{R}_{\mu\nu} \neq R_{\mu\nu}$, $\tilde{R} \neq R$ ... and so forth, therefore, the conformally compactified spacetime is unphysical, it does not satisfy the Einstein field equations anymore. Moreover, as you saw from problem sheet 1, timelike and spacelike geodesics of (M, g) are not geodesics in $(M, \tilde{g})$. The utility of the conformal compactification is for understanding the causal structure.

To understand this better let us consider some examples.

4.1.1 Minkowski Space in two-dimensions

Our first example is Minkowski space in two-dimensions. The metric in rectangular coordinates is given by

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2\,, \tag{4.3}$$

where $-\infty < t, x < \infty$. The null geodesics are given by $t \pm x =$constant. We may introduce light-cone coordinates $u = t - x$ and $v = t + x$ which makes the null geodesics pretty simple. In these coordinates the metric becomes

$$\mathrm{d}s^2 = -\mathrm{d}u\mathrm{d}v\,. \tag{4.4}$$

The coordinates are still infinite and so we have not really done much yet. To proceed we want to shrink infinity down to a finite distance away. Define

$$u = \tan\tilde{u}\,, \qquad v = \tan\tilde{v}\,, \tag{4.5}$$

where $-\frac{\pi}{2} < \tilde{u}, \tilde{v} < \frac{\pi}{2}$. Note that the range is open because strictly $u, v \to \pm\infty$ are not in the spacetime. The line-element with these coordinates is now

$$\mathrm{d}s^2 = -\frac{1}{\cos^2\tilde{u}\cos^2\tilde{v}}\mathrm{d}\tilde{u}\mathrm{d}\tilde{v}\,. \tag{4.6}$$

It diverges as $\tilde{u}, \tilde{v} \to \pm\frac{\pi}{2}$. We can now define a new metric conformally related to the one above. The obvious conformal factor to use is chosen to remove the prefactor. We take

$$\tilde{g} = \cos^2\tilde{u}\cos^2\tilde{v}g = -\mathrm{d}\tilde{u}\mathrm{d}\tilde{v}\,. \tag{4.7}$$

This metric is now regular at the points at infinity where either $\tilde{u}$ $\tilde{v}$ are equal to $\pm\frac{\pi}{2}$. Since it is regular there we may now add these points to the spacetime. The resulting spacetime $(\tilde{M}, \tilde{g})$ is the *conformal compactification* of (M, g). We may now draw this, see figure 8

The two points $(\tilde{u}, \tilde{v}) = (-\frac{\pi}{2}, -\frac{\pi}{2})$ and $(\tilde{u}, \tilde{v}) = (\frac{\pi}{2}, \frac{\pi}{2})$ are denoted by $i^{\mp}$ respectively. All past and future directed timelike curves end up at $i^{\mp}$ so we refer to i^-/i^+ as past/future timelike infinity. Future directed null geodesics either end up at $\tilde{v} = \frac{\pi}{2}$ with constant $|\tilde{u}| < \frac{\pi}{2}$ or at $\tilde{u} = \frac{\pi}{2}$ with constant $|\tilde{v}| < \frac{\pi}{2}$. This set of points is denoted by $\mathscr{I}^+$ (called scri-plus) and referred to as *future null infinity.* An analogous definition holds for past null infinity $\mathscr{I}^-$ (scri-minus). Spacelike infinity, i^0 denotes the set of end-points of spacelike geodesics, which correspond to $(\tilde{u}, \tilde{v}) = (\frac{\pi}{2}, -\frac{\pi}{2})$ and $(\tilde{u}, \tilde{v}) = (-\frac{\pi}{2}, \frac{\pi}{2})$.

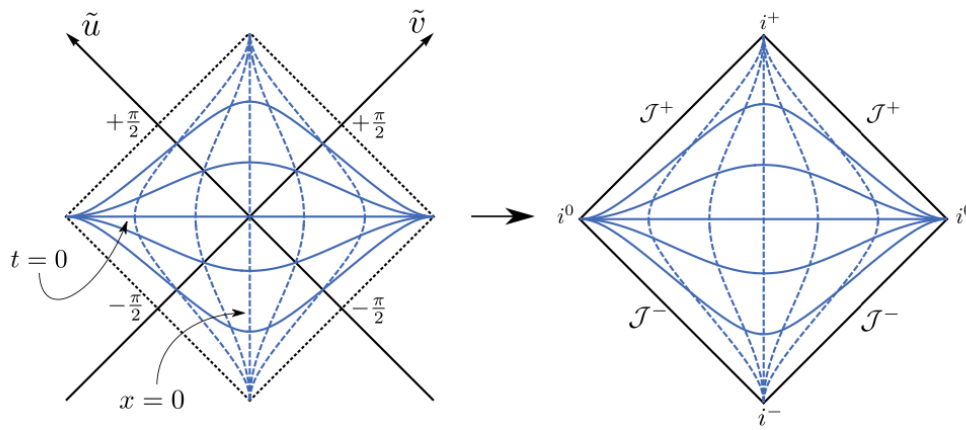

Figure 8: **Left:** On the left we have Minkowski space, (M, g) in $(\tilde{u}, \tilde{v})$ coordinates. The boundaries $\tilde{u}, \tilde{v} = \pm\frac{\pi}{2}$ are not part of M and g diverges there. Lines with $r =$ const are given by dashed lines, while the solid lines are those with $t =$ const. **Right:** On the right is the Penrose diagram of the conformally compactified spacetime. Future past timelike infinity $i^{\pm}$, future/past null infinity is denoted $\mathcal{J}^{\pm}$ while spacelike infinity is denoted i^0.

4.1.2 Minkowski Space in $d > 2$

We have just seen the Penrose diagram for $d = 2$, it turns out that this is some-what special in dimension, Minkowski space in $d > 2$ is somewhat different. Consider Minkowski space in $d > 2$ dimensions. We may use the "rectangular" metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \sum_{i=1}^{d-1} (\mathrm{d}x^i)^2 \,, \tag{4.8}$$

where the coordinates have ranges $t \in (-\infty, \infty)$, $x^i \in (-\infty, \infty)$. To proceed we may a change of coordinates going to spherical polar coordinates so that the spacelike part of the metric is equivalent to

$$\sum_{i=1}^{d-1} (\mathrm{d}x^i)^2 = \mathrm{d}r^2 + r^2 \mathrm{d}s^2(S^{d-2}) \,, \tag{4.9}$$

with S^{d-2} the unit $(d-2)$-dimensional sphere and $\mathrm{d}s^2(S^{d-2})$ the round metric on it. This exhibits the spacetime as a cone centred at $x^i = 0$. We take $r \geq 0$. In these coordinates the Minkowski metric is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2 \mathrm{d}s^2(S^{d-2}) \,. \tag{4.10}$$

We can define light-cone coordinates

$$u = t - r \,, \quad v = t + r \,, \tag{4.11}$$

which puts the metric into the form

$$\mathrm{d}s^2 = -\mathrm{d}u\mathrm{d}v + \frac{(v-u)^2}{4}\mathrm{d}s^2(S^{d-2})\,. \tag{4.12}$$

Note that since $r \geq 0$ we have $u \leq v$. We now want to bring infinity to finite coordinate length, to do this we change coordinates to

$$u = \tan\tilde{u}\,, \quad v = \tan\tilde{v}\,, \tag{4.13}$$

where

$$\tilde{u} \in \Big(-\frac{\pi}{2}, \frac{\pi}{2}\Big)\,, \qquad \tilde{v} \in \Big(-\frac{\pi}{2}, \frac{\pi}{2}\Big)\,. \tag{4.14}$$

Note that the range is open since the points at $\pm\infty$ in the original coordinates are not part of the spacetime. We still need to impose that $\tilde{u} \leq \tilde{v}$. In these coordinates the metric reads

$$\mathrm{d}s^2 = -\frac{1}{4\cos^2\tilde{u}\cos^2\tilde{v}}\Big[-4\mathrm{d}\tilde{u}\mathrm{d}\tilde{v} + \sin^2(\tilde{v}-\tilde{u})\mathrm{d}s^2\big(S^{d-2}\big)\Big]\,. \tag{4.15}$$

We may now use a conformal transformation to remove the overall pre-factor and we are left with

$$\tilde{g} = 4\cos^2\tilde{u}\cos^2\tilde{v}g = -4\mathrm{d}\tilde{u}\mathrm{d}\tilde{v} + \sin^2(\tilde{v}-\tilde{u})\mathrm{d}s^2\big(S^{d-2}\big)\,. \tag{4.16}$$

As before, after the conformal transformation $\tilde{u}, \tilde{v} = \pm\frac{\pi}{2}$ is no longer a problem and we may compactify the space by including these points. We therefore have the coordinate ranges $-\frac{\pi}{2} \leq \tilde{u} \leq \tilde{v} \leq \frac{\pi}{2}$. At fixed point on the sphere the metric is the same as that of 2d Minkowski space, the difference is in the ranges of $\tilde{u}, \tilde{v}$. We only include the half which is right of the vertical line. Every point on the sphere represents a $d-2$ dimensional sphere of radius $\sin(\tilde{v}-\tilde{u})$. The Penrose diagram is drawn in figure 9

In 4d, we can picture this differently. Define the coordinates $T = \tilde{v}+\tilde{u}$ and $\chi = \tilde{v}-\tilde{u}$. The coordinate ranges are then $-\pi < T < \pi$ and $0 < \chi < \pi$, with the added constraint $|T| + \chi \leq \pi$. The metric reads

$$\hat{g} = -\mathrm{d}T^2 + \mathrm{d}\chi^2 + \sin^2\chi\mathrm{d}s^2(S^2)\,. \tag{4.17}$$

The spatial part is just the round metric of a three-sphere. This therefore represents a static universe with spherical spatial slices corresponding to a finite portion of the Einstein static universe. See the right-hand side of figure 9 there this is plotted. Note that the vertical direction of the cylinder is T while the angular direction is χ. At each point there is a

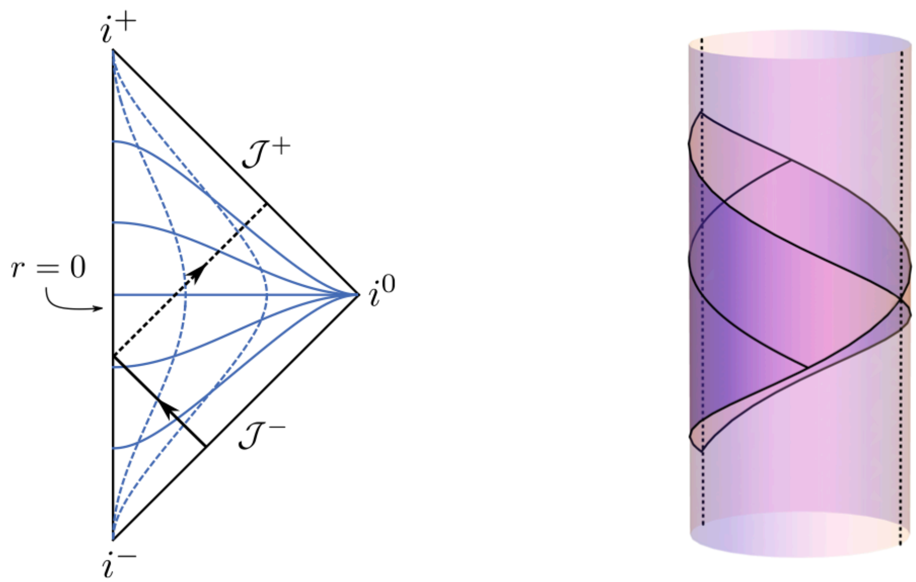

Figure 9: **Left:** On the left we have Minkowski space in general dimension > 2. Each point represents a $d-2$-dimensional sphere. As the null geodesic passes through $r = 0$ it emerges on another copy of the Penrose diagram whose points represent the anti-podes (diametrically opposite point) on the spheres. **Right:** The right digram shows the conformal compactification for $d = 4$ as a portion of the Einstein static universe. The curved line represents that same null geodesic as on the left-hand-side.

two-sphere with radius $\sin^2\chi$. We have

$$\begin{aligned}
i^+ =&\ \text{future timelike infinity}\ (T=\pi, \chi=0)\,,\\
i^0 =&\ \text{spatial infinity}\ (T=0, \chi=\pi)\,,\\
i^- =&\ \text{past timelike infinity}\ (T=-\pi, \chi=0)\,,\\
\mathscr{I}^+ =&\ \text{future null infinity}(T=\pi-\chi, 0<\chi<\pi)\,,\\
\mathscr{I}^- =&\ \text{past null infinity}(T=-\pi+\chi, 0<\chi<\pi)\,.
\end{aligned} \tag{4.18}$$

Note that $i^\pm, i^0$ are actually points since $\chi = 0$ and $\chi = \pi$ are the north and south poles of S^3. Meanwhile $\mathscr{I}^\pm$ are null surfaces with the topology of $\mathbb{R}\times S^2$.

There are a number of features to observe. Radial null geodesics are at $\pm 45°$ in the diagram. All timelike geodesics begin at i^- and end at i^+. All null geodesics begin at $\mathscr{I}^-$ and end at $\mathscr{I}^+$.

4.1.3 Rindler spacetime in 1+1 dimensions

Rindler space is a subregion of Minkowski space associated with observers who are eternally accelerated at a constant rate. It appears often when looking at the near-horizon region of black holes. Consider the two-dimensional Minkowski metric and an observer moving at a uniform acceleration of magnitude α^{-1} in the x-direction. Their trajectory is

$$t(\tau) = \alpha \sinh\left(\frac{\tau}{\alpha}\right), \qquad x(\tau) = \alpha \cosh\left(\frac{\tau}{\alpha}\right), \tag{4.19}$$

which has constant acceleration α,

$$a^\mu a_\mu = \alpha^2, \qquad a^\mu = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2}. \tag{4.20}$$

Note that the trajectory of the observer satisfies

$$x^2(\tau) - t^2(\tau) = \xi^2, \tag{4.21}$$

which describes a hyperboloid asymptoting to null paths $x = -t$ in the past and $x = t$ in the future. The accelerated observer travels from past null infinity to future null infinity, rather than timelike infinity as would be reached by geodesic observers. The region $x \leq t$ is forever hidden to them which makes the line $x = t$ a horizon to these observers. This horizon is of a different flavour to the Schwarzschild horizon since that is an observer independent object while this horizon is associated with a special family of observers, see figure 10.

Rindler space corresponds to the right wedge $x > |t|$ foliated by the worldlines of the accelerated observers.

We can choose new coordinates (η, ξ) on 2d Minkowski space that is adapted to uniformly accelerated motion. Let

$$t = \xi \sinh(\eta), \quad x = \xi \cosh(\eta), \tag{4.22}$$

with coordinate range $0 < \xi < \infty$ and $-\infty < \eta < \infty$. In these coordinates the Minkowski metric in (η, ξ) coordinates is

$$\mathrm{d}s^2 = -\xi^2 \mathrm{d}\eta^2 + \mathrm{d}\xi^2. \tag{4.23}$$

The proper time measured by an accelerated observer, i.e. a stationary ($\xi =$constant) observer in Rindler coordinates is $\tau = \xi\eta$. Since Rindler space is just a subregion of Minkowski space the Penrose diagram is just a piece of figure 8.

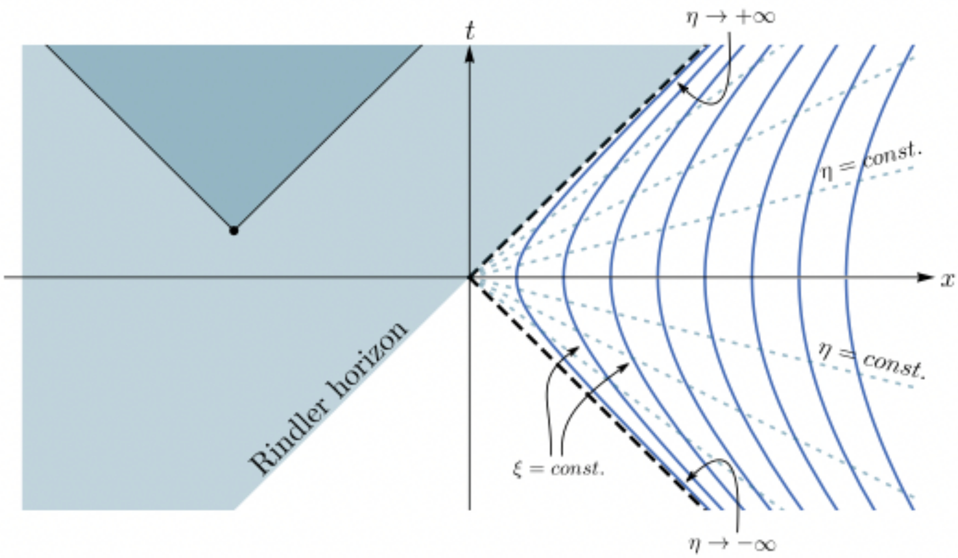

Figure 10: Eternally accelerating observers in Minkowski space. Their worldlines are in blue and labelled by ξ. Events in the shaded region such as the black dot are hidden to them. The Rindler horizon is the boundary between the shaded and unshaded regions. Rindler space is the right wedge bounded by the dashed black lines which are null. The straight lines are lines of constant Rindler time.

Exercise 4:

Consider the Schwarzschild metric in Schwarzschild coordinates. Let the horizon be at $r = r_h$ and make the change of coordinates:

$$r = r_h + \frac{1}{4r_h}\rho^2 \,. \tag{4.24}$$

Expand the metric around small ρ (keeping only the terms which are leading for each $\mathrm{d}x^2$ term) to show that one obtains:

$$\mathrm{d}s^2 = \left[-\kappa^2\rho^2\mathrm{d}t^2 + \mathrm{d}\rho^2 + (4M^2)\mathrm{d}s^2(S^2)\right](1 + \mathcal{O}(\rho^2)) \,. \tag{4.25}$$

The near-horizon geometry is therefore the direct product of 2d Minkowski with a two-sphere with the metric on Minkowski space being in Rindler coordinates. Here κ is the surface gravity which is given by

$$\kappa = \frac{1}{4M} \,. \tag{4.26}$$

4.1.4 Kruskal Space

Recall that we could extend the Schwarzschild solution beyond the horizon by using Kruskal coordinates. The metric in these coordinates reads

$$\mathrm{d}s^2 = -\frac{32M^3}{r}\exp\Big(-\frac{r}{2M}\Big)\mathrm{d}U\mathrm{d}V + r^2\mathrm{d}s^2(S^2)\,. \tag{4.27}$$

Recall that the range of the coordinates is $-\infty < U, V < \infty$. We need to define a new set of null coordinates to bring infinity to a finite coordinate distance. We transform as

$$U = \tan\tilde{U}\,, \quad V = \tan\tilde{V}\,, \tag{4.28}$$

such that $-\frac{\pi}{2} < \tilde{U}, \tilde{V} < \frac{\pi}{2}$. The line element becomes

$$\mathrm{d}s^2 = \frac{1}{4\cos^2\tilde{U}\cos^2\tilde{V}}\Bigg[-\frac{128M^3}{r}\exp\Big(-\frac{r}{2M}\Big)\mathrm{d}\tilde{U}\mathrm{d}\tilde{V} + r^2\cos^2\tilde{U}\cos^2\tilde{V}\mathrm{d}s^2(S^2)\Bigg]\,. \tag{4.29}$$

We perform the usual conformal transformation

$$\tilde{g} = 4\cos^2\tilde{U}\cos^2\tilde{V}g = -\frac{128M^3}{r}\exp\Big(-\frac{r}{2M}\Big)\mathrm{d}\tilde{U}\mathrm{d}\tilde{V} + r^2\cos^2\tilde{U}\cos^2\tilde{V}\mathrm{d}s^2(S^2)\,. \tag{4.30}$$

The curvature singularity at $r = 0$ is at $UV = 1$ in U, V coordinates and now corresponds to

$$1 = UV = \tan\tilde{U}\tan\tilde{V} \quad \Leftrightarrow \quad \sin\tilde{U}\sin\tilde{V} - \cos\tilde{U}\cos\tilde{V} = 0 \quad \Leftrightarrow \quad \cos(\tilde{U}+\tilde{V}) = 0\,. \tag{4.31}$$

This implies that it is located at $\tilde{U} + \tilde{V} = \pm\frac{\pi}{2}$. To make this simpler it is useful to define $\tilde{U} = T - X$ and $\tilde{V} = T + X$. The Penrose diagram includes the points at infinity and the singularity, we draw it in figure 11.

In contrast to the conformal compactification of Minkowski space the conformally related metric is singular at $i^{\pm}$. Lines of constant r meet at $i^{\pm}$ and this includes the curvature singularity at $r = 0$. Less obviously, it turns out that one cannot choose Ω to make the conformally rescaled metric smooth at i^0.

We can also plot the Penrose diagram of a spherically symmetric collapsing star. The interior of the star is excluded since the stress energy tensor does not vanish there. We end up with only the two regions 1 and 3 of Kruskal spacetime, there is no white hole region anymore.

4.2 Asymptotic flatness

The Schwarzschild solution is an asymptotically flat spacetime. Somewhat colloquially we define an asymptotically flat spacetime to be one that "looks like Minkowski spacetime at infinity". The remainder of this section is to give more rigour to this statement.

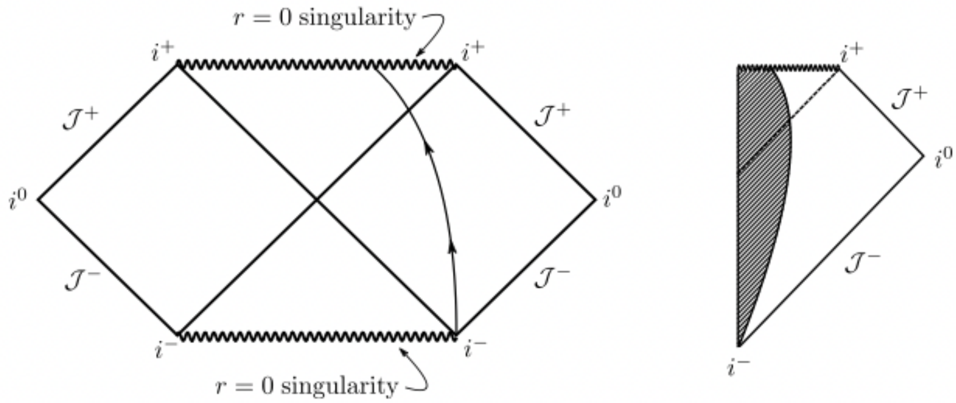

Figure 11: **Left:** The Penrose diagram for Kruskal spacetime. The possible trajectory of the surface of a collapsing star is plotted, the parts to the left correspond to the interior of the star and is described by a metric (at fixed time slice) to the metric we constructed in section 3. **Right:** The Penrose diagram for a collapsing star. The curved surface represents the surface of the star with the shaded area corresponding to the interior of the star. The horizon corresponds to the dashed line and appears in spacetime once the star has collapsed sufficiently.

Infinity in Minkowski spacetime consists of the regions $\mathscr{I}^\pm$, $i^\pm$ and i^0. The points $i^\pm$ are singular points in the conformal compactification of Kruskal spacetime. Since we want to include the latter as asymptotically flat we cannot use $i^\pm$ in our definition of asymptotic flatness. Moreover i^0 is not smooth in Kruskal spacetime so we will not include i^0 either.[14] We are then left with $\mathscr{I}^\pm$ to define an asymptotically flat spacetime.

Recall that a manifold with a boundary is defined in a similar way to a manifold without boundary, the only difference is that the charts are now maps $\varphi : M \to \mathbb{R}^n/\mathbb{Z}_2 = \{(x^1, ...x^n) : x^1 \geq 0\}$. The boundary ∂M is defined to be the set of points which have $x^1 = 0$ in some chart.

Definition 3 *Asymptotically flat at null infinity*

A time-orientable spacetime (M, g) *is* asymptotically flat at null infinity *if there exists a spacetime* $(\tilde{M}, \tilde{g})$ *such that*

1. *There exists a positive function* Ω *on* M *such that* $(\tilde{M}, \tilde{g})$ *is an extension of* $(M, \Omega^2 g)$*.*

[14] One can refine the definition of asymptotic flatness to include i^0, however for simplicity we will just remove it.

2. *Within $\tilde{M}$, M can be extended to obtain a manifold with a boundary $M \cup \partial M$.*

3. *Ω can be extended to a function on $\tilde{M}$ such that $\Omega = 0$ and $\mathrm{d}\Omega \neq 0$ on ∂M.*

4. *∂M is the disjoint union of two components $\mathscr{I}^+$ and $\mathscr{I}^-$, each diffeomorphic to $\mathbb{R} \times S^2$.*

5. *No past (future) directed causal curve starting in M $\mathscr{I}^+$ ($\mathscr{I}^-$).*

6. *$\mathscr{I}^\pm$ are complete (see below).*

Let us unpack this definition slightly. Conditions 1-3 are just the requirement that we can perform a conformal compactification of the spacetime. The requirement that $\mathrm{d}\Omega \neq 0$ on ∂M just means that Ω has a first order singularity on ∂M, note that this is the case of the examples considered above. This is necessary to make sure the metric approaches Minkowski space at the appropriate rate near $\mathscr{I}^\pm$. The remaining conditions 4-6 ensure that infinity has the same structure at null infinity as Minkowski spacetime. Condition 4 says that i^0 should exist, condition 5 says we can define a future and a past.

Then

$$g = -2\mathrm{d}u\mathrm{d}r + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) + ..., \tag{4.32}$$

for large r. The ellipses denote subleading terms at large r. The leading terms are simply the Minkowski metric. If one converts to the inertial frame coordinates (t, r, θ, ϕ) so that the leading order metric is $\mathrm{diag}(-1, 1, 1, 1)$ then the correction terms are of the form r^{-1} and suppressed in the large r limit. We see that the metric of an asymptotically metric approaches the Minkowski metric at $\mathscr{I}^+$.

We can finally understand condition 6 of the definition of an asymptotically flat spacetime.

Definition 4 *Complete*

$\mathscr{I}^+$ is complete *if the generators of $\mathscr{I}^+$ are complete. The generators are complete if the affine parameter extends to $\pm\infty$. A similar definition holds for $\mathscr{I}^-$.*

4.3 Singularities

We have seen that a spherically symmetric gravitational collapse results in the formation of a singularity. One can ask whether this is an artefact of the spherical symmetry or if it is something more generic? In Newtonian gravity the collapse of a spherically symmetric ball of matter produces a singularity with infinite density at the origin, however a tiny perturbation of the spherical symmetry does not lead to a singularity, rather a bouncing solution which removes the singularity periodically. One could ask whether this is the same in GR. Work by

Roger Penrose answered this question and showed that singularities are a generic prediction of general relativity.[15]

4.3.1 Different types of singularities

We have seen two different types of singularity when we considered the Schwarzschild solution. We have defined a metric singularity to arise in some basis if its components are not smooth or the determinant vanishes (so that it is not invertible at that point). We learnt that a *coordinate singularity* can be eliminated by a change of coordinates, for example $r = 2M$ in the Schwarzschild spacetime in Schwarzschild coordinates. These singularities are unphysical since they can be removed by a better choice of coordinates. If it is not possible to remove the singularity by a change of coordinates then we have a physical singularity. A *scalar curvature singularity* is a singularity where some scalar constructed from the Riemann tensor blows up. Since it is a scalar it is diffeomorphism invariant and thus if it diverges in one coordinate system it diverges in all.

Conical singularities Not all physical singularities are curvature singularities. Consider the manifold $M = \mathbb{R}^2$ and introduce plane polar coordinates (r, ϕ) with $\phi \sim \phi + 2\pi$ and define the Riemannian metric

$$g = \mathrm{d}r^2 + \lambda^2 r^2 \mathrm{d}\phi^2\,, \tag{4.33}$$

with $\lambda > 0$. The metric determinant vanishes at $r = 0$, however for $\lambda = 1$ this is just Euclidean space in polar coordinates. We can convert to Cartesian coordinates to see that $r = 0$ is just a coordinate singularity. However for $\lambda \neq 1$ this is no longer the case. Define $\phi' = \lambda\phi$, then the metric is

$$g = \mathrm{d}r^2 + r^2 \mathrm{d}\phi'^2\,, \tag{4.34}$$

which is locally isometric to Euclidean space and therefore curvature singularity free. However, it is not globally isometric to Euclidean space because the period of ϕ' is $2\pi\lambda$ rather than 2π. Consider a circle with radius $r = \epsilon$, this has

$$\frac{\text{circumference}}{\text{radius}} = \frac{2\phi\lambda\epsilon}{\epsilon} = 2\pi\lambda\,, \tag{4.35}$$

which does not tend to 2π as $\epsilon \to 0$. Recall that any smooth Riemannian manifold is locally flat, i.e. one recovers results of Euclidean geometry on sufficiently small length scales. The above shows that this is not true for the above metric for small circles around $r = 0$ and

[15] It is sometimes said that GR predicts its own downfall. This is because GR predicts singularities but is ill equipped to deal with them. To fully understand them we need a theory of quantum gravity, which GR is not.

therefore the metric cannot be extended smoothly to $r = 0$. This is an example of a *conical singularity*.

A problem in defining singularities is that they are not places, they do not belong to the spacetime manifold because we define spacetime as the pair (M, g) where g is a smooth Lorentzian metric. This is the reason we remove $r = 0$ from the Kruskal spacetime, the metric is no longer smooth there. Similarly in the above example if we want to have a smooth manifold we should take $M = \mathbb{R}^2/(0,0)$ so that $r = 0$ is not part of the spacetime M. In both of these examples the existence of the singularity implies that some geodesics cannot be extended to arbitrarily large affine parameter because they end at the singularity (in these case both at $r = 0$). We will use this property to define what we mean by a *singular* spacetime.

First eliminate the trivial case where a geodesic ends because we haven't taken the range of its affine parameter to be large enough. A curve is a smooth map $\gamma : (a, b) \to M$. Sometimes a curve can be *extended*, that is it is part of a bigger curve. If this happens then the first curve will have an endpoint, which is defined as follows.

Definition 5 *Future endpoint*

The point $p \in M$ is a future end-point *of a future-directed causal curve $\gamma : (a, b) \to M$ if, for any neighbourhood O of p there exists t_0 such that $\gamma(t) \in O$ for all $t > t_0$. We say that γ is* future inextendible *if it has no future endpoint. Similarly for past endpoints and past in-extendibility. The curve γ is* in-extendible *if it is both future and past in-extendible.*

Example 4.1: An in-extendible spacetime

Let (M, g) be Minkowski spacetime and let $\gamma : (-\infty, 0) \to M$ be $\gamma(t) = (t, 0, 0, 0)$. Then the origin is a future end-point of γ. However if we instead let (M, g) be Minkowski spacetime with the origin removed then γ is future in-extendible.

Definition 6 *Complete*

A geodesic is complete *if an affine parameter for the geodesic extends to $\pm\infty$. A* spacetime is geodesically complete *if all in-extendible causal geodesics are complete.*

Example 4.2: Geodesically complete and incomplete spacetimes

Minkowski spacetime is geodesically complete as is the spacetime describing a spherically symmetric star. Kruskal spacetime on the other hand is geodesically *incomplete* because some geodesics reach $r = 0$ in finite affine parameter and hence cannot be extended to infinite affine parameter.

A spacetime which is extendible, like the Schwarzschild solution in Schwarzschild coordinates, is also incomplete. The incompleteness arises because we are not considering the full spacetime. One should therefore consider the maximal extension of spacetime.

Example 4.3: Rindler is geodesically incomplete

Consider the Rindler metric

$$\mathrm{d}s^2 = -\xi^2 \mathrm{d}\eta^2 + \mathrm{d}\eta^2\,, \tag{4.36}$$

which has a singularity at $\xi = 0$. Null geodesics for this metric satisfy:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\lambda}(\dot{t}\xi^2) = 0\,, \\ \dot{\xi}^2 - \xi^2 \dot{t}^2 = 0\,. \end{aligned} \tag{4.37}$$

We may solve the first using that it defines a conserved quantity: thus

$$E = \dot{t}\xi^2\,. \tag{4.38}$$

We therefore find that

$$\dot{\xi} = \pm\frac{E}{\xi}\,, \tag{4.39}$$

which has solution:

$$\xi = \sqrt{c \pm 2E\lambda}\,. \tag{4.40}$$

We see that $\xi = 0$ is reached within finite affine parameter and therefore Rindler is geodesically incomplete. We may extend Rindler space to obtain Minkowski space which is of course geodesically complete.

We define a spacetime to be singular if it is:

Definition 7 *Singular*

A spacetime is singular *if it is both geodesically incomplete and in-extendible.*

This is true for the Kruskal spacetime, and the Kruskal extension of the RN and Kerr–Newman black holes that we will study later in the course.

4.4 Null hypersurfaces

Definition 8 *Hypersurface*

A hypersurface *is a $d-1$ dimensional space living within a d-dimensional spacetime. It can be defined in terms of a single real equation.*

Example 4.4: Sphere as a hypersurface

We may define a sphere as a hypersurface within $\mathbb{R}^d$ as:

$$S^{d-1} = \{x_i \in \mathbb{R}^d : \sum_{i=1}^{d} x_i^2 = R^2\}\,. \tag{4.41}$$

Definition 9 *Hypersurface-orthogonal*

Let Σ be a hypersurface in M specified by the equation $f(x) = 0$, with $f : M \to \mathbb{R}$ a smooth function. We require $\mathrm{d}f \neq 0$ on Σ, then $\mathrm{d}f$ is normal to Σ.[16] *The dual vector to $\mathrm{d}f$, let us call it ξ, is said to be* hypersurface orthogonal*.*

Definition 10 *Timelike, spacelike and null hypersurfaces*

Let ξ be the dual vector to the normal of the hypersurface. Then we have:

- *If ξ is timelike then the hypersurface is a* spacelike hypersurface*.*
- *If ξ is spacelike then the hypersurface is a* timelike hypersurface*.*
- *If ξ is null then the hypersurface is a* null hypersurface*.*

Aside

Given a normal to a hypersurface if follows that any other normal to the hypersurface can be written as $n = g\mathrm{d}f + fn'$ with g a smooth function which does not vanish anywhere on Σ, and n' a smooth one-form.Therefore we necessarily have

$$\mathrm{d}n = \mathrm{d}g \wedge \mathrm{d}f + \mathrm{d}f \wedge n' + f\mathrm{d}n' \quad \Rightarrow \quad \mathrm{d}n\big|_{\Sigma} = (\mathrm{d}g - n') \wedge \mathrm{d}f \quad \Rightarrow \quad n \wedge \mathrm{d}n\big|_{\Sigma} = 0\,. \tag{4.42}$$

Conversely Frobenius' theorem implies.

Theorem 2 Frobenius*:*

If n is a non-zero one-form such that $n \wedge \mathrm{d}n = 0$ everywhere, then there exist functions f, g such that $n = g\mathrm{d}f$ and therefore n is normal to the surfaces defined by $f(x) =$constant, and therefore hypersurface-orthogonal.

Example 4.5: Null hypersurface in Schwarzschild

Consider surfaces of constant r in Schwarzschild spacetime. The one-form $n = \mathrm{d}r$ is normal

[16] To see this consider a curve $\gamma(\lambda) \subset \Sigma$. By definition $f(\gamma(\lambda)) = 0$ for all λ, thus, $0 = \partial_\lambda f(\gamma(\lambda)) = \dot\gamma^\mu(\lambda)\partial_\mu f(\gamma(\lambda)) = \mathrm{d}f(\dot\gamma(\lambda))$. The latter is equivalent to $\mathrm{d}f$ being normal to the hypersurface.

to such surfaces. The norm is

$$n^2 = 1 - \frac{2M}{r} \,. \tag{4.43}$$

We see that the hypersurface $r = 2M$ is a null hypersurface.

Definition 11 *Killing horizon*

A null hypersurface Σ is a Killing horizon *of a Killing vector ξ if ξ is normal to Σ on Σ.*

For every Killing horizon we can associate a quantity called the *surface gravity*. Given the Killing horizon we have an associated Killing vector, ξ which is null on the horizon. In an asymptotically flat spacetime ξ is normalised so that asymptotically it goes to $\xi^2 \to -1$. Since ξ is a normal vector to the Killing horizon it obeys the geodesic equation

$$\xi^\mu \nabla_\mu \xi^\nu = \kappa \xi^\nu \,. \tag{4.44}$$

It turns out that κ is constant over the horizon. It parametrises how far away the Killing vector ξ is from being an affinely parametrised geodesic.

Example 4.6: Schwarzschild Killing horizon

Since the horizon corresponds to a coordinate singularity in Schwarzschild coordinates we need to work in coordinates which extend beyond the horizon. The easiest to use are the ingoing Eddington–Finkelstein coordinates. We have that the metric takes the form

$$\mathrm{d}s^2 = -f(r)\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r + r^2\mathrm{d}s^2(S^2) \,, \qquad f(r) = 1 - \frac{2M}{r} \,. \tag{4.45}$$

The inverse metric, in matrix form is

$$g^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & f(r) & 0 & 0 \\ 0 & 0 & r^{-2} & 0 \\ 0 & 0 & 0 & \frac{1}{r^2 \sin^2\theta} \end{pmatrix} \tag{4.46}$$

The hypersurface is defined to be $r = 2M$ and therefore the most general normal is given by

$$n = h\mathrm{d}r + (r - 2M)n' \,, \tag{4.47}$$

with $h(r = 2M) \neq 0$ and n' an arbitrary one-form smooth on $r = 2M$. We will continue to work with the simpler $\mathrm{d}r$ for this computation however when considering more general black holes it can be useful to consider the function h and one-form n'. Then we have that the

$$\xi^\mu = g^{\mu\nu} n_\nu = (1, f(r), 0, 0) \,. \tag{4.48}$$

On $r = 2M$ this reads

$$\xi = \partial_v \,, \tag{4.49}$$

which is indeed a Killing vector. It has norm

$$|\xi|^2 = -f(r) \,, \tag{4.50}$$

and therefore it is time-like outside the horizon and null on it.

We can compute the surface gravity on the horizon. We have

$$\xi^\mu \nabla_\mu \xi^\nu = \xi^\mu \partial_\mu \xi^\nu + \xi^\mu \Gamma^\nu{}_{\mu\sigma} \xi^\sigma = \Gamma^v{}_{vv} \,. \tag{4.51}$$

We can compute this component of the Levi–Civita connection simply as

$$\Gamma^v{}_{vv} = \frac{1}{2} g^{v\sigma} (2\partial_v g_{\sigma v} - \partial_\sigma g_{vv}) - \frac{M}{r^2} \,. \tag{4.52}$$

Therefore we have

$$\nabla_\xi \xi^\nu = \frac{1}{4M} \xi^\nu \,, \tag{4.53}$$

and thus

$$\kappa = \frac{1}{4M} \,. \tag{4.54}$$

This will later be understood as the temperature of the black hole.

Let n_μ be normal to a null hypersurface $\mathcal{N}$. Then any (non-zero) vector X^μ tangent to the hypersurface obeys $n_\mu X^\mu = 0$. Therefore, either X^μ is spacelike or X^μ is parallel to n^μ. In particular note that n^μ is tangent to the hypersurface, since it is null, hence on $\mathcal{N}$ the integral curves of n^μ lie within $\mathcal{N}$.

Proposition: The integral curves of n are null geodesics. We call them the generators of $\mathcal{N}$.

Proof: Let $\mathcal{N}$ be given by the equation f =constant for some function f with $\mathrm{d}f \neq 0$ on $\mathcal{N}$. Then we have $n = h\mathrm{d}f$ for h some function which does not vanish on f =constant. Let

$N = \mathrm{d}f$, the integral curves of n and N are the same up to a choice of reparametrisation.[17] Since $\mathcal{N}$ is null we have that $N^\mu N_\mu = 0$ on $\mathcal{N}$ which implies that the gradient of this function is normal to $\mathcal{N}$:

$$\nabla_\mu (N^\nu N_\nu)\Big|_{\mathcal{N}} = 2\alpha N_\mu \,, \tag{4.55}$$

with α some function on $\mathcal{N}$. Now since $\nabla_\mu N_\nu = \nabla_\mu \nabla_\nu f = \nabla_\nu \nabla_\mu f = \nabla_\nu N_\mu$ we have

$$N^\nu \nabla_\mu N_\nu = N^\nu \nabla_\nu N_\mu \quad \Rightarrow \quad N^\nu \nabla_\nu N_\mu\Big|_{\mathcal{N}} = \alpha N_\mu \,. \tag{4.56}$$

This is nothing but the geodesic equation for a non-affinely parametrised geodesic. Hence on $\mathcal{N}$ the integral curves of N, and therefore also n are null geodesics.

Example 4.7: Kruskal

Consider Kruskal spacetime, with metric (A.69). Let $N = \mathrm{d}U$, this is null everywhere (since $g^{UU} = 0$) and is normal to a family of null hypersurfaces defined by U =constant. Since $N^2 = 0$ everywhere it follows that N is tangent to affinely parametrised null geodesics. Raising an index gives

$$N^\mu = -\frac{r}{16M^3} \mathrm{e}^{\frac{r}{2M}} \left(\frac{\partial}{\partial V}\right)^\mu \,. \tag{4.57}$$

Let $\mathcal{N}$ be the surface $U = 0$. Since $U = 0$ corresponds to $r = 2M$ on $\mathcal{N}$ we have that N is simply a constant multiple of $\frac{\partial}{\partial V}$. Thus V is an affine parameter for the generators of $\mathcal{N}$. Similarly U is an affine parameter of for the generators of the null hypersurface $V = 0$.

Black holes are characterised by the fact that you can enter them but never exit. The most important feature is therefore not the singularity but rather the event horizon. An event horizon is a hypersurface separating those spacetime points that are connected to infinity by a timelike path from those that are not. In the following section we want to make this statement more mathematically rigorous.

4.5 Definition of a black hole and the event horizon

It remains to understand mathematically what a black hole is.

[17] To see this consider the integral curves defined by n^μ:

$$n^\mu = \frac{\mathrm{d}\tilde{x}^\mu(\tilde{\lambda})}{\mathrm{d}\tilde{\lambda}} \,.$$

We have that the integral curves for N are then

$$N^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = h^{-1} n^\mu = h^{-1} \frac{\mathrm{d}\tilde{x}^\mu}{\mathrm{d}\tilde{\lambda}} = \left[h^{-1} \frac{\mathrm{d}\lambda}{\mathrm{d}\tilde{\lambda}}\right] \frac{\mathrm{d}\tilde{x}^\mu}{\mathrm{d}\lambda} \,.$$

By choosing the parameter $\lambda(\tilde{\lambda})$ so that $\frac{\mathrm{d}\lambda}{\mathrm{d}\tilde{\lambda}} = h$ we may make the bracket in the last term become unity and therefore we have shown that the integral curves are the same up to a choice of reparametrisation.

Definition 12 *Causal curve*

A causal curve *is any path which is timelike or null everywhere.*

Definition 13 *Causal future, chronological future*

Given any subset S of a manifold M, we can define the causal future *of S denoted $J^+(S)$ to be the set of points that can be reached from S following a future directed causal curve.*

The chronological future *$I^+(S)$ is the set of points that can be reached by following a future directed timelike curve. A point p will always be in its causal future $J^+(S)$ but not necessarily its own chronological future $I^+(p)$, though it could be.*

The causal past J^- and chronological past I^- are defined analogously.

We can now define a black hole and its event horizon. Consider a manifold with metric (M, g) and its conformal compactification $(\bar{M}, \bar{g})$. Recall that the causal past J^- of a region is the set of all points we can reach from that region by moving along a past-directed timelike paths. We can define the causal past of scri-plus $J^-(\mathscr{I}^+) \subset \bar{M}$. The set of points of M that can send a signal to $\mathscr{I}^+$ is $M \cap J^-(\mathscr{I}^+)$. We define the black hole region to be the complement of this region, and the future event horizon to be the boundary of the black hole region:

Definition 14 *Black hole region, future event horizon*
Let (M, g) be a spacetime that is asymptotically flat at null infinity. The Black hole region *is*

$$\mathcal{B} = M \backslash [M \cap J^-(\mathscr{I}^+)] \,, \tag{4.58}$$

where $J^-(\mathscr{I}^+)$ is defined using the unphysical spacetime $(\bar{M}, \bar{g})$. The future event horizon *is $\mathcal{H}^+ = \partial\mathcal{B}$.*

Definition 15 *White hole region, past event horizon Similarly the* white hole region *is*

$$\mathcal{W} = M \backslash [M \cap J^+(\mathscr{I}^-)] \,, \tag{4.59}$$

and the past event horizon *is $\mathcal{H}^- = \partial\mathcal{W}$.*

Theorem 3 *Hawking 1972*
In a stationary, analytic, asymptotically flat vacuum black hole spacetime, $\mathcal{H}^+$ is a Killing horizon.

Definition 16 *Naked Singularity*

A naked singularity *is a singularity from which signals can reach $\mathscr{I}^+$, i.e. one that is not hidden behind an event horizon.*

Conjecture *Strong cosmic conjecture*

Naked singularities cannot form in gravitational collapse from generic initially non-singular states in an asymptotically flat spacetime obeying the dominant energy conditions.

5 Charged Black holes

At this point we have almost beaten to death the Schwarzschild solution, we need some new solutions to play with. There is a generalisation to the Schwarzschild solution that we can study: we can give it some electric and magnetic charges. This will retain the static and spherically symmetric properties of the Schwarzschild solution but introduce a gauge field which couples gravity with electromagnetism. This charged black hole is known as the *Reissner–Nordström* (RN) black hole.

In nature large imbalances of charge do not occur, it is favourable for the charged object to attract particles of opposite charge and gradually lose its charge. We would therefore expect matter undergoing gravitational collapse to be neutral and so the presence of charged black holes in nature does not seem particularly relevant. Nevertheless the solution exhibits some interesting features. Moreover, for those doing string theory, RN black holes occasionally appear, though probably not in your course.

5.1 Einstein gravity coupled to electromagnetism

We want to couple Einstein gravity to Electromagnetism. Recall that the general prescription for coupling matter to gravity is through minimal coupling.[18] Minimal coupling says we replace the Minkowski metric with the curved metric of spacetime, we replace regular derivatives with covariant derivatives and add in the correct volume measure.

Electromagnetism in terms of forms

Recall that Electromagnetism is governed by Maxwell's equations:

$$\begin{aligned} \nabla \times \vec{B} - \partial_t \vec{E} &= \vec{J}\,, \\ \nabla \cdot \vec{E} &= \rho\,, \\ \nabla \times \vec{E} + \partial_t \vec{B} &= 0\,, \\ \nabla \cdot \vec{B} &= 0\,. \end{aligned} \tag{5.1}$$

Here $\vec{B}$ and $\vec{E}$ are the electric and magnetic field 3-vectors, $\vec{J}$ is a current, ρ is the charge density. These equations are invariant under Lorentz transformations, even though they do not look invariant. We can write these in a manifestly invariant way by introducing the two-form field strength F and its one-form potential A.

[18] One can also add non-minimal terms but we will not consider these here.

Writing the Maxwell's equations in component notation we have

$$\begin{aligned} \epsilon^{ijk}\partial_j B_k - \partial_0 E^i &= J^i\,, \\ \partial_i E^i &= J^0\,, \\ \epsilon^{ijk}\partial_j E_k + \partial_0 B^i &= 0\,, \\ \partial_i B^i &= 0\,. \end{aligned} \tag{5.2}$$

We have introduced the current 4-vector $J = (\rho, \vec{J})$ to rewrite the first two conditions. Let us define the field strength tensor $F_{\mu\nu}$ to be

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}_{\mu\nu} \tag{5.3}$$

We have

$$F^{0i} = E^i\,, \qquad F^{ij} = \epsilon^{ijk}B_k\,. \tag{5.4}$$

Therefore the first two equations in (5.3) can be rewritten as

$$\begin{aligned} \partial_j F^{ij} - \partial_0 F^{0i} &= J^i\,, \\ \partial_i F^{0i} &= J^0\,, \end{aligned} \tag{5.5}$$

which may be rewritten as

$$\partial_\mu F^{\mu\nu} = -J^\nu\,. \tag{5.6}$$

Similarly the bottom two equations in (5.3) may be rewritten as

$$\partial_{[\mu}F_{\nu\lambda]} = 0\,. \tag{5.7}$$

Writing F as a two-form we have the two equations

$$\mathrm{d}\star F = -J\,, \qquad \mathrm{d}F = 0\,. \tag{5.8}$$

The first equation is known as the Maxwell equation, while the second is the Bianchi identity. Since $\mathrm{d}F = 0$ this means that locally F can be written as a closed form,

$$F = \mathrm{d}A\,, \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{5.9}$$

The one-form A is known as the gauge field. Note that it is not unique, $A + \mathrm{d}\Lambda$ gives the same field strength F when Λ is a smooth function. Adding the term $\mathrm{d}\Lambda$ to the potential is known as a gauge transformation, it is a redundancy/symmetry in our description of the the theory. Physical quantities will generally be expressed in terms of the field strength F. On the other hand we view the gauge field as the dynamical field of the theory, i.e. the field we vary an action with respect to.

We can write an action for electromagnetism by using the gauge field A and defining the field strength F to be $F = \mathrm{d}A$. Then the action giving rise to Maxwell's equations with sources is

$$S_{\text{Maxwell}} = \int \mathrm{d}^4x \; \mathcal{L}_{\text{EM}} = \int \mathrm{d}^4x \Big[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu \Big] . \tag{5.10}$$

We have

$$\frac{\partial \mathcal{L}_{\text{EM}}}{\partial A_\nu} = J^\nu , \tag{5.11}$$

and

$$\frac{\partial \mathcal{L}_{\text{EM}}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu} . \tag{5.12}$$

Putting everything together, the Euler Lagrange equations give

$$\partial_\mu F^{\mu\nu} = -J^\nu , \tag{5.13}$$

as we found above from Maxwell's equations. The Bianchi identity arises because we define $F = \mathrm{d}A$ and by using that $\mathrm{d}^2 = 0$.

The Lagrangian for electromagnetism in the absence of sources in flat space is

$$\mathcal{L}_{\text{EM}} = -F_{\mu\nu} F_{\rho\sigma} \eta^{\mu\rho} \eta^{\nu\sigma} . \tag{5.14}$$

To couple this to gravity we will replace the Minkowski metric with the curved metric, add the volume measure and replace derivatives with covariant derivatives. Derivatives appear in the field strength as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \longrightarrow \nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu , \tag{5.15}$$

where the latter follows when using the Levi–Civita connection. The action for Einstein–Maxwell theory is then

$$S = \frac{1}{16\pi} \int \mathrm{d}^4x \sqrt{-g} \Big(R - F_{\mu\nu} F_{\rho\sigma} g^{\mu\rho} g^{\nu\sigma} \Big) \equiv \frac{1}{16\pi} \int \mathrm{d}^4x \sqrt{-g} \Big(R - F_{\mu\nu} F^{\mu\nu} \Big) . \tag{5.16}$$

The equations of motion derived from the variation of the Einstein–Maxwell action are

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} &= 2 \Big(F_{\mu\rho} F_\nu{}^\rho - \frac{1}{4} g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \Big) , \\ \nabla_\mu F^{\mu\nu} &= 0 , \end{aligned} \tag{5.17}$$

and should be accompanied by the Bianchi identity $\mathrm{d}F = 0$. Note that the second equation of motion is called the Maxwell equation and is equivalent to

$$\mathrm{d} \star F = 0 . \tag{5.18}$$

Exercise 5:
Check the equations of motion are indeed those derived from the action.

5.2 Reissner–Nordström black hole

There is a generalisation of Birkhoff's theorem for four-dimensional Einstein–Maxwell theory.

Theorem 4 *The Unique spherically symmetric solution of the Einstein–Maxwell equations with* non-constant area radius function r *is the* Reissner–Nordström *solution:*

$$\begin{aligned} \mathrm{d}s^2 &= -f(r)\mathrm{d}t^2 + f(r)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}s^2(S^2)\,, \\ A &= -\frac{Q}{r}\mathrm{d}t - P\cos\theta\mathrm{d}\phi\,, \\ f(r) &= 1 - \frac{2M}{r} + \frac{e^2}{r^2}\,, \qquad e^2 = Q^2 + P^2\,. \end{aligned} \tag{5.19}$$

The solution has three parameters: M, P, Q. We will show later that these are the mass, magnetic charge and electric charge of the solution. Note that there is no evidence for the existence of magnetic monopoles (which the P describes) in nature, however it is a valid solution of the equations of motion.

Aside
Note that there is a statement that the area function should be non-constant. If this is relaxed there is an additional solution which corresponds to $\mathrm{AdS}_2 \times S^2$. This is the near-horizon limit of the extremal Reissner–Nordstrom solution.

There are several properties which are similar to the Schwarzschild solution. The solution is static with the timelike Killing vector ∂_t. It is also asymptotically flat, like the Schwarzschild solution and has a curvature singularity at $r = 0$. Note that we may smoothly recover the Schwarzschild solution by sending $Q, P \to 0$ which turns the gauge field off.

To discuss the properties of the solution we need to study the possible degenerations of the metric. To do this it is convenient to define

$$\Delta(r) = r^2 f(r) = r^2 - 2Mr + e^2 = (r - r_+)(r - r_-)\,, \qquad r_\pm = M \pm \sqrt{M^2 - e^2}\,. \tag{5.20}$$

The metric then takes the form

$$\mathrm{d}s^2 = -\frac{\Delta(r)}{r^2}\mathrm{d}t^2 + \frac{r^2}{\Delta(r)}\mathrm{d}r^2 + r^2\mathrm{d}s^2(S^2)\,. \tag{5.21}$$

One can check that the Kretschmann scalar is

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = \frac{8(7e^3 - 12e^2Mr + 6M^2r^2)}{r^8}\,, \tag{5.22}$$

and clearly $r = 0$ is a genuine curvature singularity. Following the string cosmic censorship conjecture we would like to hide the singularity behind a horizon, much in the same way that the curvature singularity is hidden in the Schwarzschild solution behind a horizon. From our discussion on black hole horizons we must look for a Killing horizon. This will be determined by having a root of $\Delta(r)$ so that the hypersurface normal to $\mathrm{d}r$ at the root is a Killing horizon. From the form of the metric it is clear that metric has (at least) a coordinate singularity at this point. There are then three distinct behaviours for the metric depending on the possible roots $r_\pm$, which in turn are determined by the sign of $M^2 - e^2$.

5.2.1 Super extremal RN: $e^2 > M^2$

If $M^2 - e^2 < 0$ the roots $r_\pm$ are complex and therefore $\Delta(r)$ does not have any real zeros. Thus the curvature singularity is not hidden behind a horizon and we have a naked singularity. There is no obstruction to an observer travelling to the singularity, studying it and then returning to us to tell us all about it. If one studies the geodesics one finds that the naked singularity is repulsive, timelike geodesics never intersect $r = 0$, rather they approach $r = 0$ but reverse course and move away. Null geodesics can reach the singularity as can non-geodesic timelike curves.

As $r \to \infty$ the solution approaches flat spacetime and the causal structure looks normal everywhere. The conformal diagram will therefore be just like that of Minkowski space, except now $r = 0$ is a singularity. The nakedness of the singularity should be offensive to you. We should never expect to find a black hole with $M^2 < e^2$ as a result of gravitational collapse. Roughly, the condition states that the total energy of the hole is less than the contribution to the energy of the electromagnetic fields alone, and therefore we must have something with negative mass. We consider this unphysical. The Penrose diagram is given in figure 12. Constructing the Penrose diagram proceeds in much the same way as for the Schwarzschild black hole with negative mass in problem sheet 2.

5.2.2 Sub-extremal RN: $M^2 > e^2$

In this case Δ has two real simple roots and there are consequently two coordinate singularities. The surfaces defined by $r = r_\pm$ are both null hypersurfaces and are both Killing horizons. The singularity at $r = 0$ is a timelike line (contrast this with Schwarzschild where it was spacelike).

To see that they are coordinate singularities we can proceed in a similar manner as we did for the Schwarzschild solution and define tortoise like coordinates. Let us begin with $r > r_+$

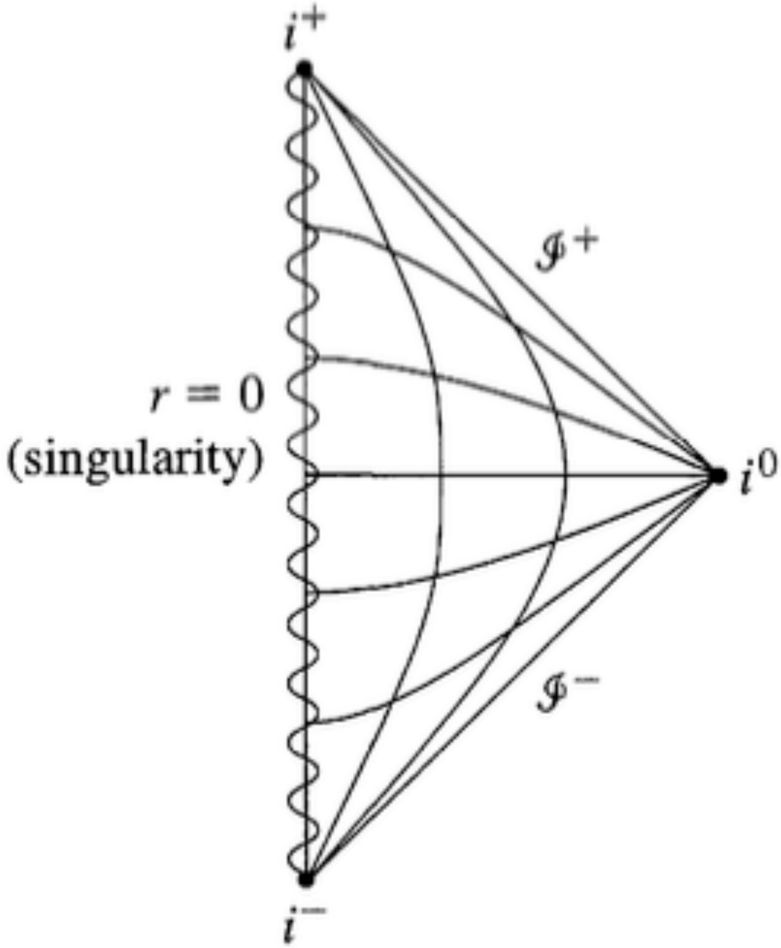

Figure 12: The Penrose diagram for the super-extremal Reissner–Nordström solution.

and define

$$\mathrm{d}r_* = \frac{r^2}{\Delta(r)}\mathrm{d}r\,. \tag{5.23}$$

This is the usual trick of defining a new coordinate in which the null radial geodesics are straight lines and is the generalisation of the Tortoise coordinate for the RN solution. Integrating gives

$$r_* = r + \frac{1}{2\kappa_+}\log\frac{r-r_+}{r_+} + \frac{1}{2\kappa_-}\log\frac{r-r_-}{r_-} + \text{const}\,, \tag{5.24}$$

where

$$\kappa_\pm = \frac{r_\pm - r_\mp}{2r_\pm^2}\,, \tag{5.25}$$

and we have added constants (the $r_\pm$ denominators in the log terms) for later simplicity. Now define the Eddington–Finkelstein coordinates

$$u = t - r_* \,, \qquad v = t + r_* \,. \tag{5.26}$$

In ingoing Eddington–Finkelstein coordinates the RN metric is

$$\mathrm{d}s^2 = -\frac{\Delta(r)}{r^2}\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r + r^2\mathrm{d}s^2(S^2) \,. \tag{5.27}$$

This is now smooth for any $r > 0$ hence we can analytically continue the metric into the new region $0 < r < r_+$. There remains a curvature singularity at $r = 0$ and, as we will see momentarily, there are null hypersurfaces at $r = r_\pm$.

Null hypersurfaces In the ingoing-Eddington–Finkelstein coordinates it is not obvious that the surfaces $r = r\pm$ are null hypersurfaces, in fact they are Killing horizons. First consider the normal to these hypersurfaces defined by $r = r_\pm$, we may take

$$n = \mathrm{d}r \,. \tag{5.28}$$

Recall that the normals are not uniquely defined. In fact we may take a more general normal such that $n = h\mathrm{d}r + (r - r\pm)n'$ where $h(r_\pm) \neq 0$ and n' is a smooth one-form. Working with the simple normal $n = \mathrm{d}r$ we can compute the dual vector field to be

$$\xi = g^{\mu\nu}n_\nu\partial_\mu = \partial_v + \frac{\Delta(r)}{r^2}\partial_r \,. \tag{5.29}$$

It follows that the norm is given by

$$|\xi|^2 = \frac{\Delta(r)}{r^2} \,, \tag{5.30}$$

which vanishes for $r = r_\pm$ and thus the hypersurface is null. Moreover, on the null hypersurface we have $\xi|_\Sigma = \partial_v$ which is a Killing vector. Therefore both $r = r_\pm$ define Killing horizons. No point in the region $r < r_+$ can send a signal to $\mathscr{I}^+$, hence it describes a black hole. This behaves in the same way as the Schwarzschild horizon. The black hole region is $r \leq r_+$ and the future event horizon is the null hypersurface $r = r_+$.

Maximal extension of sub-extremal RN Though the outer horizon behaves in the same way to Schwarzschild, the inner horizon and singularity at $r = 0$ behave differently. Between $r_- < r < r_+$ we have that ∂_t is space-like, as in Schwarzschild, but becomes time-like again between $0 < r < r_-$. This means that once you have gone past $r = r_+$ you are forced towards

$r = r_-$ since it lies in your future. However, once you have passed through $r = r_-$ there is no longer a requirement to go towards $r = 0$, this is a time-like line and therefore not necessarily in your future: it can be avoided. This is different to Schwarzschild where the singularity was space-like and you must hit it.

To understand the global structure better we need to define Kruskal-like coordinates for our spacetime. We define

$$U^{\pm} = -\mathrm{e}^{-\kappa_{\pm} u}\,, \qquad V^{\pm} = \pm \mathrm{e}^{\kappa_{\pm} v}\,, \tag{5.31}$$

Starting in the region $r > r_+$ we use coordinates (U^+, V^+, θ, ϕ) to obtain the metric

$$\mathrm{d}s^2 = -\frac{r_+ r_-}{\kappa_+^2 r^2} \mathrm{e}^{-2\kappa_+ r} \left(\frac{r - r_-}{r_-}\right)^{1+\kappa_+/|\kappa_-|} \mathrm{d}U^+ \mathrm{d}V^+ + r^2 \mathrm{d}s^2(S^2)\,, \tag{5.32}$$

where $r(U^+, V^+)$ is defined implicitly by

$$-U^+ V^+ = \mathrm{e}^{2\kappa_+ r} \left(\frac{r - r_+}{r_+}\right) \left(\frac{r_-}{r - r_-}\right)^{\kappa_+/|\kappa_-|}. \tag{5.33}$$

Notice that the metric is ill-defined at $r = r_-$ and therefore we should keep U^+ and V^+ so that $r > r_-$. The RHS of (5.33) is a monotonically increasing function of r from $r > r_-$, to see this use that $\kappa_+ > |\kappa_-|$. Initially we have $U^+ < 0$ and $V^+ > 0$ which gives $r > r_+$, but now we can analytically continue to $U^+ \geq 0$ or $V^+ \leq 0$. In particular the metric is smooth and non-degenerate when $U^+ = 0$ or $V^+ = 0$. We obtain a diagram very similar to the Kruskal diagram we had for Schwarzschild, see figure 13.

Just as for Kruskal we have a pair of null hypersurfaces ($r = r_+$) which intersect in the bifurcation 2-sphere located at $U^+ = V^+ = 0$. Surfaces of constant t are Einstein–Rosen bridges which connect regions I and IV. The major difference to the Schwarzschild solution is that we no longer have a curvature singularity in regions II and III because $r(U^+, V^+) \geq r_-$. However we know from our ED coordinates that it is possible to extend our metric into the $r < r_-$ region, hence the above spacetime must be extendable. Indeed, radial null geodesics reach $r = r_-$ in finite proper affine time and therefore the spacetime is extendible. To extend our metric further we need another change of coordinates.

To do this we should start in region II and use ingoing EF coordinates (v, r, θ, ϕ), since we know that these cover regions I and II. We can now define a retarded time coordinate u in region II. Define the time coordinate $t = v - r_*$ in region II with r_* as defined in (5.24). The metric in coordinates (t, r, θ, ϕ) takes the static RN form given above with $r_- < r < r_+$. Now define $u = t - r_* = v - 2r_*$. Having defined u in region II we can now define Kruskal

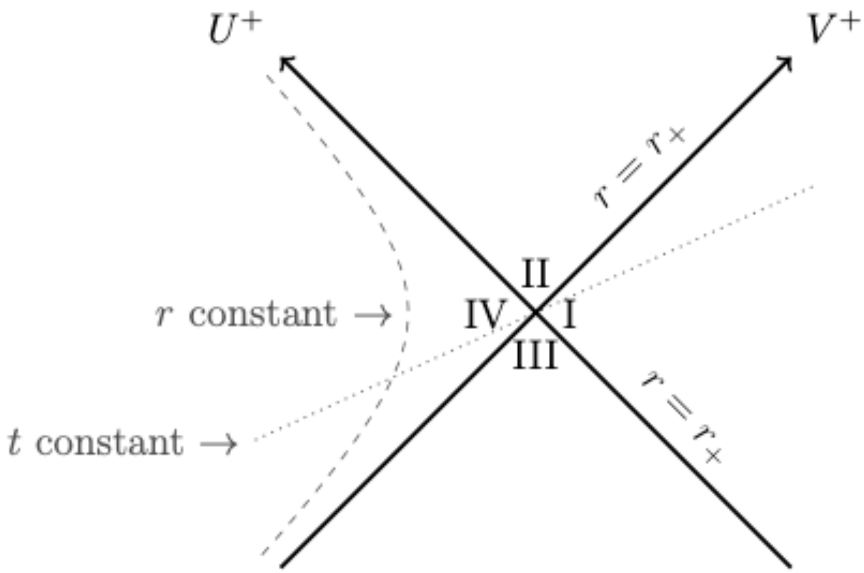

Figure 13: The Reissner–Nordström solution in (U^+, V^+) coordinates.

coordinates $U^- < 0$ and $V^- < 0$ in region II using the formula above in equation 5.31. In these coordinates the metric is

$$\mathrm{d}s^2 = -\frac{r_+ r_-}{\kappa_-^2 r^2} \mathrm{e}^{2|\kappa_-|r} \left(\frac{r_+ - r}{r_+}\right)^{1+|\kappa_-|/\kappa_+} \mathrm{d}U^- \mathrm{d}V^- + r^2 \mathrm{d}s^2(S^2)\,, \tag{5.34}$$

where $r(U^-, V^-) < r_+$ is given implicitly by

$$U^- V^- = \mathrm{e}^{-2|\kappa_-|r} \left(\frac{r - r_-}{r_-}\right) \left(\frac{r_+}{r_+ - r}\right)^{|\kappa_-|/\kappa_+} . \tag{5.35}$$

Notice that the metric is ill defined at $r = 0$ and $r = r_+$. The former is our singularity and the latter is just the outer horizon which is a coordinate singularity in these coordinates. We may as before analytically continue U^- and V^- so that $U^- \geq 0$ and $V^- \geq 0$ which gives the diagram 14.

We now have the regions V and VI in which $0 < r < r_-$. These regions contain the curvature singularity at $r = 0$ $(U^- V^- = -1)$[19] which is timelike. Region III' is isometric to region III and so by introducing new coordinates $(U^{+\prime}, V^{+\prime})$ this can be analytically continued to the future to give further new regions I', II', and IV' as shown in figure 15. In this diagram

[19]In Schwarzschild the singularity was at $UV = 1$.

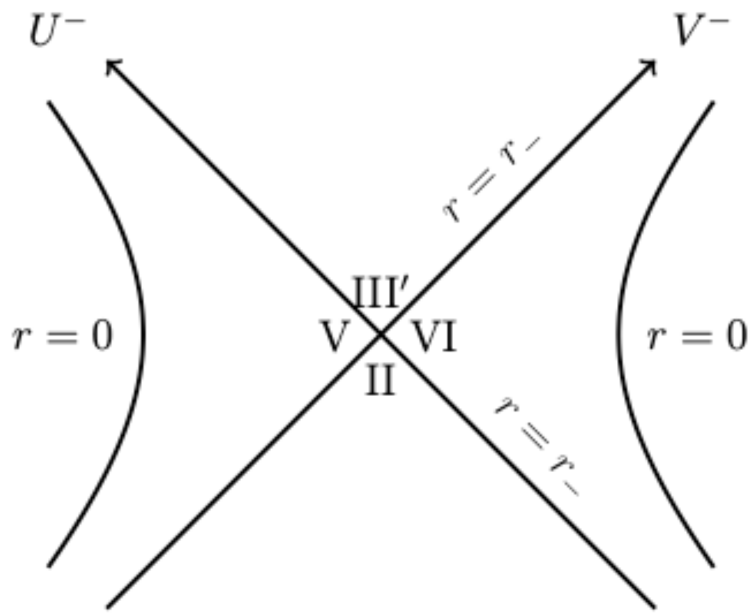

Figure 14: The Reissner–Nordström solution in (U^-, V^-) coordinates.

the regions I' and IV' are new asymptotically flat regions isometric to I and IV. We may repeat this procedure indefinitely to the future and past, so that the maximal analytic extension of the RN solution contains infinitely many regions. The resulting Penrose diagram is given in figure 16. It extends to infinity in both the past and future.

This seems a bit crazy, infinite universes, what is happening here? Notice that if you are an observer falling into the black hole from far away, r_+ is just like the Schwarzschild horizon. At this radius r switches from being a spacelike coordinate to a timelike one and therefore you necessarily move in the direction of decreasing r. Witnesses outside the black hole see the same phenomena that they would for the Schwarzschild solution, the infalling observer is seen to move more and more slowly and is increasingly redshifted.

The inevitable fall from r_+ to ever-decreasing radii only lasts until you reach the null surface at $r = r_-$ where r switches from being a timelike coordinate back to being spacelike. You need not continue travelling on a trajectory of decreasing r and therefore your inevitable doom of hitting the singularity can be stopped. Indeed $r = 0$ is a timelike line and you are therefore and therefore not necessarily in your future. At this point you can continue on to $r = 0$ or begin to move in the direction of increasing r back through the null surface at $r = r_-$.

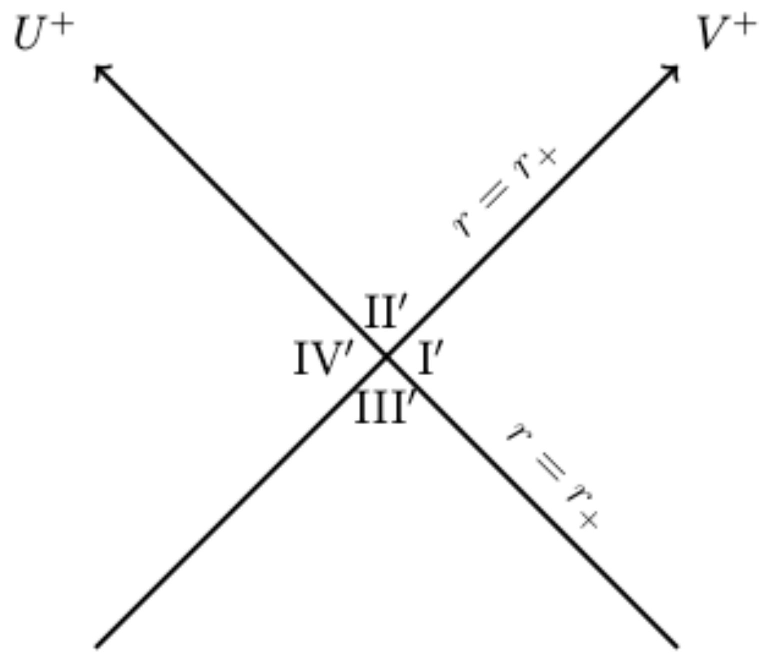

Figure 15: The regions I', II', IV' of the Reissner–Norström solution.

Then r will once again be a timelike coordinate, however now the orientation is reversed and you must travel in the direction of increasing r until you are spat out of the event horizon at $r = r_+$. This is like emerging from a white hole into the rest of the universe. From here you can choose to go back into the black hole, this time a different one to the one you initially entered. You may then repeat this over and over again to your hearts content.

How much of this story is actually science over science fiction? Well, not much. Viewing the universe from the point of an observer inside the black hole who is about to cross the event-horizon at $r = r_-$ you notice that the observer can look back in time to see the entire history of the external universe, at least as seen from the black hole. They see this infinitely long history in a finite proper time thus any signal that gets to them as they approach $r = r_-$ is infinitely blue-shifted. Therefore it is likely that any non-spherically symmetric perturbation that comes into an RN black hole will violently disturb the geometry. For this reason it is difficult to say exactly what the actual geometry inside the horizon looks like, but there is no good reason why it must contain an infinite number of asymptotically flat regions connecting

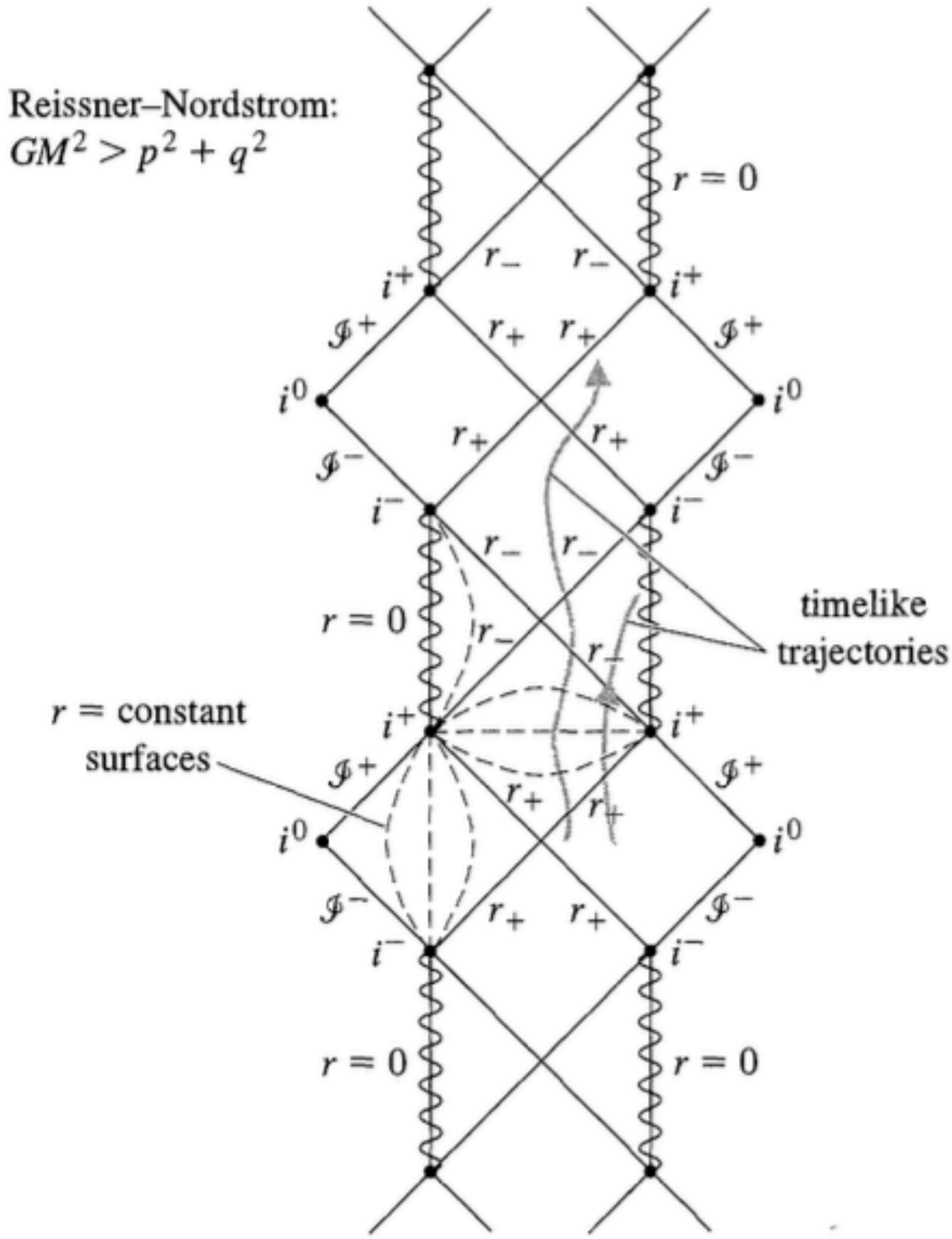

Figure 16: The Penrose diagram of the RN black hole.

to each other via various wormholes.

5.2.3 Extremal RN: $M^2 = e^2$

Finally let us consider the extremal RH when the two roots become equal and we obtain a double root. The metric of the RN extremal solution is

$$\mathrm{d}s^2 = -\left(1 - \frac{M}{r}\right)^2 \mathrm{d}t^2 + \left(1 - \frac{M}{r}\right)^{-2} \mathrm{d}r^2 + r^2 \mathrm{d}s^2(S^2)\,, \tag{5.36}$$

which has a coordinate singularity at $r = r_+ = r_- = M$.

The coordinate r is never time-like, it becomes null on the horizon at $r = r_+ = r_-$ but is spacelike either side of the horizon. The singularity at $r = 0$ is once again a timelike line and as in the previous cases may be avoided. You may avoid the singularity and continue to move to the future to extra copies of the asymptotically flat region, the singularity is always to the "left". The Penrose diagram is given in figure 17.

Aside

Multiple Extremal RN black holes

Extremal black holes appear frequently when considering supergravity theories (gravity plus supersymmetry). A rule of thumb, with some exceptions, is that supersymmetry implies extremal. As we will see extremal implies that the temperature of the black hole vanishes. The solution seems unstable since adding a little matter will take us to the sub extremal solution. In the extremal case the mass is balanced by the charge, this can be reformulated when considering a supersymmetric theory as saturating the Bogomol'nyi–Prasad–Sommerfield bound. Two extremal black holes with the same sign charges will attract each other gravitationally but repel each other electromagnetically and the two forces precisely cancel. We can find exact solutions to the coupled Einstein–Maxwell equations representing any number of such black holes in a stationary configuration.

To see this it is useful to first rewrite the RN solution and to focus on just electric charges for simplicity. Define the radial coordinate

$$\rho = r - M \tag{5.37}$$

then the metric takes the isotropic form

$$\mathrm{d}s^2 = -H(\rho)^{-2}\mathrm{d}t^2 + H(\rho)^2\Big[\mathrm{d}\rho^2 + \rho^2\mathrm{d}s^2(S^2)\Big]\,, \tag{5.38}$$

where

$$H(\rho) = 1 + \frac{M}{\rho}\,. \tag{5.39}$$

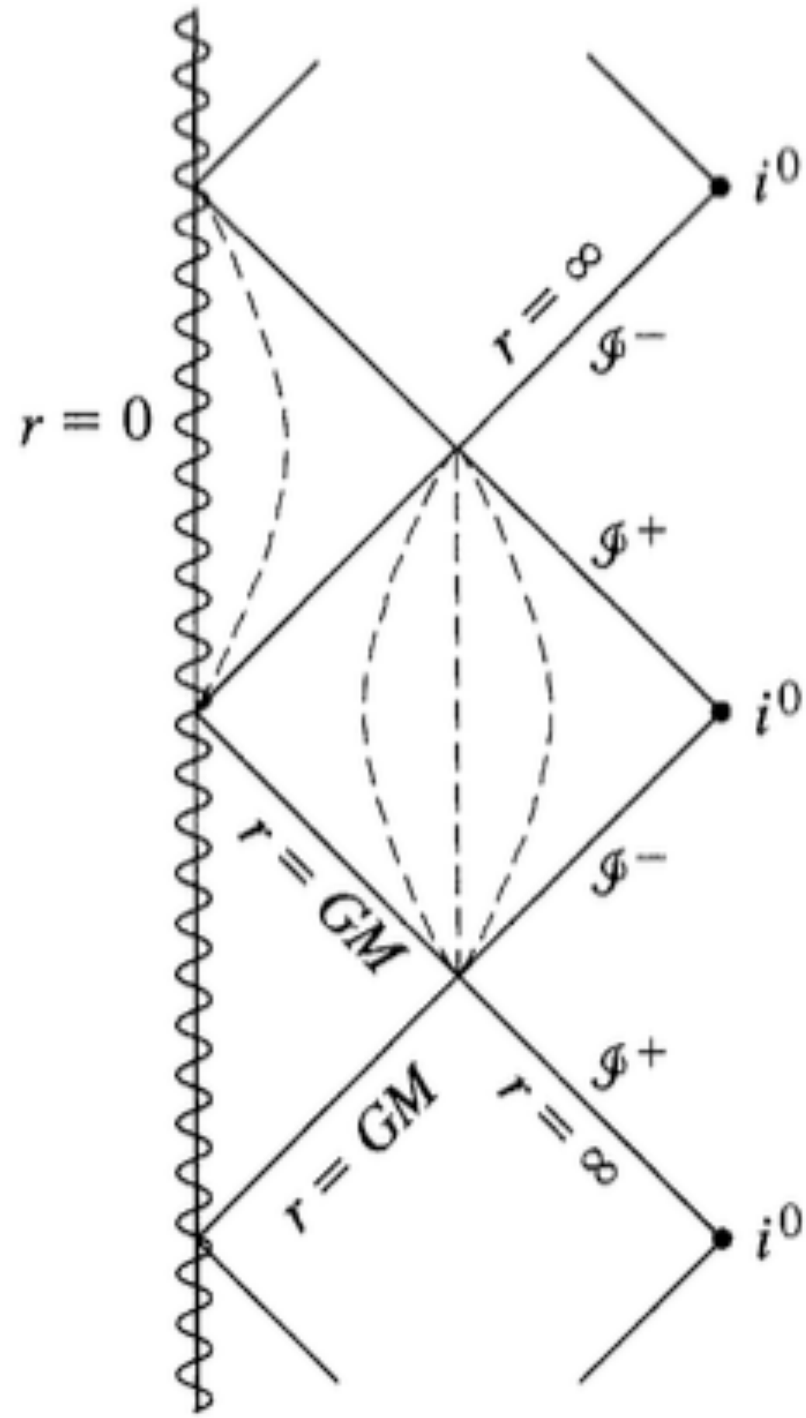

Figure 17: The Penrose diagram of the extremal RN black hole.

Since the bracketed part of the metric is just the metric on $\mathbb{R}^3$ we may rewrite the metric as

$$\mathrm{d}s^2 = -H(\vec{x})^{-2}\mathrm{d}t^2 + H(\vec{x})^2\Big[\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\Big]\,, \tag{5.40}$$

with

$$H(\vec{x}) = 1 + \frac{M}{|\vec{x}|}\,. \tag{5.41}$$

In the original components the electric field of the extremal solution can be expressed in terms of a vector potential A as

$$F_{rt} = \frac{Q}{r^2} = \partial_r A_t\,, \qquad A_t = -\frac{Q}{r}\,. \tag{5.42}$$

We may rewrite this as

$$A_t = H^{-1} - 1\,. \tag{5.43}$$

We can now forget that H takes the form above and just plug the metric into the field equations and we find that we have a solution provided

$$\nabla^2 H = 0\,, \tag{5.44}$$

with ∇^2 the Laplacian on $\mathbb{R}^3$. It is straightforward to write down all solutions that are well behaved at infinity, they take the form

$$H = 1 + \sum_{a=1}^{N} \frac{M_a}{|\vec{x} - \vec{x}_a|}\,, \tag{5.45}$$

for some set of N spatial points $\vec{x}_a$. These are the locations of the N extremal RN black holes with masses M_a and electric charges $Q_a = M_a$.

5.3 Cauchy surfaces and horizons

We have now seen the Penrose diagrams of the RN black hole and observed that the maximal extension has some very distinct features compared to the maximal extension of the Schwarzschild black hole: most obviously from the Penrose diagrams one sees that there is an infinite tower of these diagrams glued together. One should wonder how physical the maximal extension really is. A second worrying feature is that one can follow an ingoing radial geodesic towards the singularity and end up in the region between $r = 0$ and $r = r_-$ but not be destined to hit $r = 0$. There, one can either carry on travelling towards the singularity or exit via the white-hole region. There is no way of ensuring, in advance of entering the black hole, what the final outcome is. We are used to classical theories being fully deterministic.

If I throw a pen in your direction you can in principle work out exactly where it will end up if you know the initial conditions. GR seems to give rise to something which is not fully deterministic.

Many physical questions can be rephrased as an initial value problem. Given the state of a system at some moment in time what will the state of the system be at some later time. The fact that this has a definitive answer is due to causality: future events can be understood as consequences of initial conditions plus the laws of physics. Initial value problems are as common in GR as in Newtonian physics or special relativity, however the dynamical nature of the spacetime background introduces new ways in which an initial value formulation could break down. The reason for this lack of determinism in the RN black hole is because the horizon at $r = r_-$ is a *Cauchy horizon.*

Definition 17 *Achronal, Future domain of dependence*

A subset $S \subset M$ is called achronal *if no two points in S are connected by a time-like curve. For example any edgeless spacelike hypersurface in Minkowski space is achronal. Given a closed (the complement is open) achronal set we define the* future domain of dependence *of S, $D^+(S)$ to be the set of all points p such that* every *past moving inextendible causal curve through p must intersect S.*

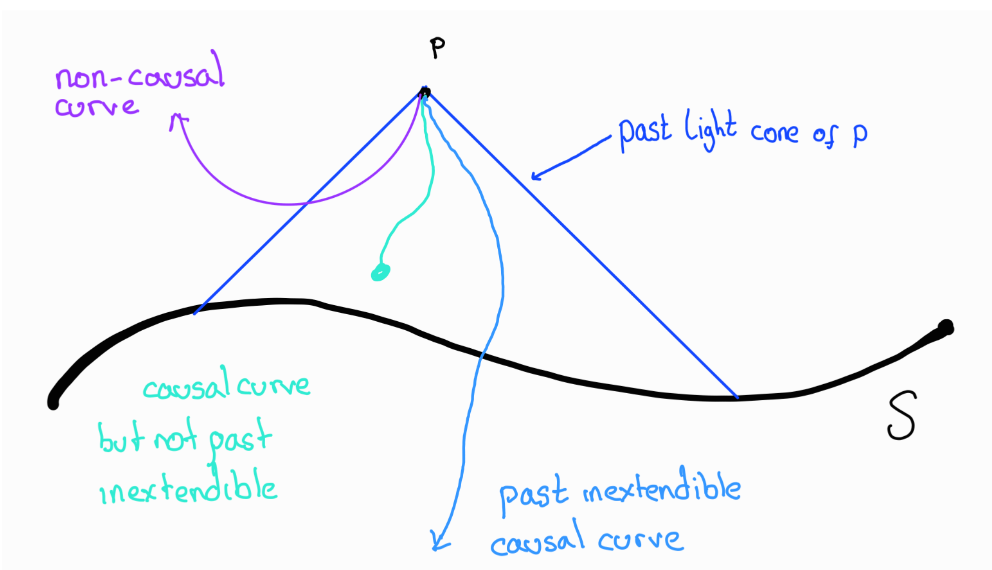

Figure 18: Past directed curves passing through p. The dark blue lines denote the past light cone. The light blue curve is an inextendible causal curve, the aqua curve is causal but can be extended. The purple curve is not causal. When considering the domain of dependence we only consider curves of the light blue type. We see that all past directed causal curves passing through p intersect S and therefore p is in the domain of dependence of S.

> **Remark**
> Recall that by inextendible we mean that the curve goes on forever (in affine parameter) and does not end at some finite point. It follows that elements of S are elements of $D^+(S)$.

A similar definition of the past domain of dependence, $D^-(S)$ holds by replacing future with past.

Definition 18 *Future/past Cauchy horizon*

We define the boundary of $D^+(S)$ to be the future Cauchy horizon $H^+(S)$ *and likewise the boundary of $D^-(S)$ to be the* past Cauchy horizon $H^-(S)$.

> **Remark**
> It follows that these boundaries are necessarily null surfaces.

We have sketched these different objects in figure 19.

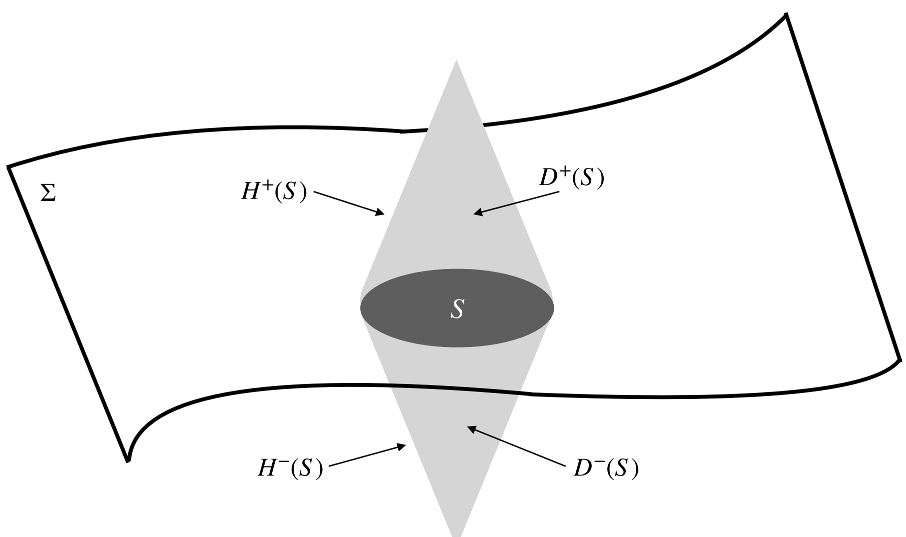

Figure 19: A depiction of the domains of dependence of the set S on the achronal surface Σ.

If nothing moves faster than light, signals cannot propagate outside the lightcone of any point p. Therefore if every curve that remains inside the lightcone must intersect S then information specified on S should be sufficient to predict what the situation is at p. That is, initial data for matter fields on S can be used to solve for the matter fields at p.

Definition 19 *Domain of dependence*

The set of all points for which we can predict what happens by knowing what happens on S is the union $D(S) = D^+(S) \cup D^-(S)$ is called the domain of dependence*.*

Definition 20 *Cauchy surface*

A closed achronal surface Σ is said to be a Cauchy surface *if the domain of dependence $D(\Sigma)$ is the entire manifold.*

Remark

Information given on the Cauchy surface can be used to predict what happens throughout all of spacetime. If a spacetime has a Cauchy surface (it need not) it is said to be *globally hyperbolic*.

Having defined a lot of things let us see some examples.

Example 5.1: A line in 2d Minkowski

Consider a bounded achronal line in 2d Minkowski space. The resultant domain of influence is compact.

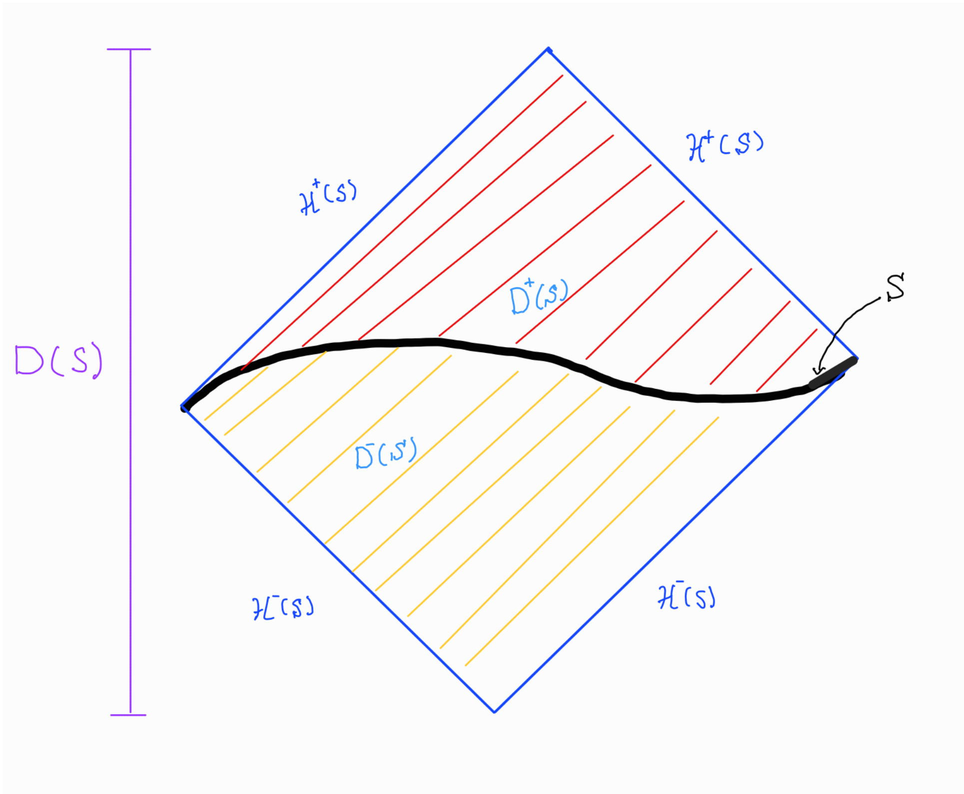

Figure 20: The domain of dependence of a surface S in 2d Minkowski space. The line S is spacelike everywhere. The dark blue lines are at 45 degrees giving rise to the lightcones at the end-points of S. Note that there are Cauchy horizons.

Example 5.2: Unbounded line

In this example consider the unbounded hyperboloid:

$$S := \left\{ (t,x) \in \mathbb{R}^2 \middle| t = -\sqrt{1+x^2} \right\} . \tag{5.46}$$

Observe that as $|x| \to \infty$ the line S asymptotes to the past lightcone $t = -|x|$. The domains are therefore:

$$\begin{aligned} D^+(S) &= \left\{ (t,x) \in \mathbb{R}^2 \middle| -\sqrt{1+x^2} \leq t < -|x| \right\} , \\ D^-(S) &= \left\{ (t,x) \in \mathbb{R}^2 \middle| t \leq -\sqrt{1+x^2} \right\} . \end{aligned} \tag{5.47}$$

Note that $D^+(S)$ is not closed since lightlike curves on $t = -|x|$ do not meet S. The future domain of dependence is sketched in figure 21.

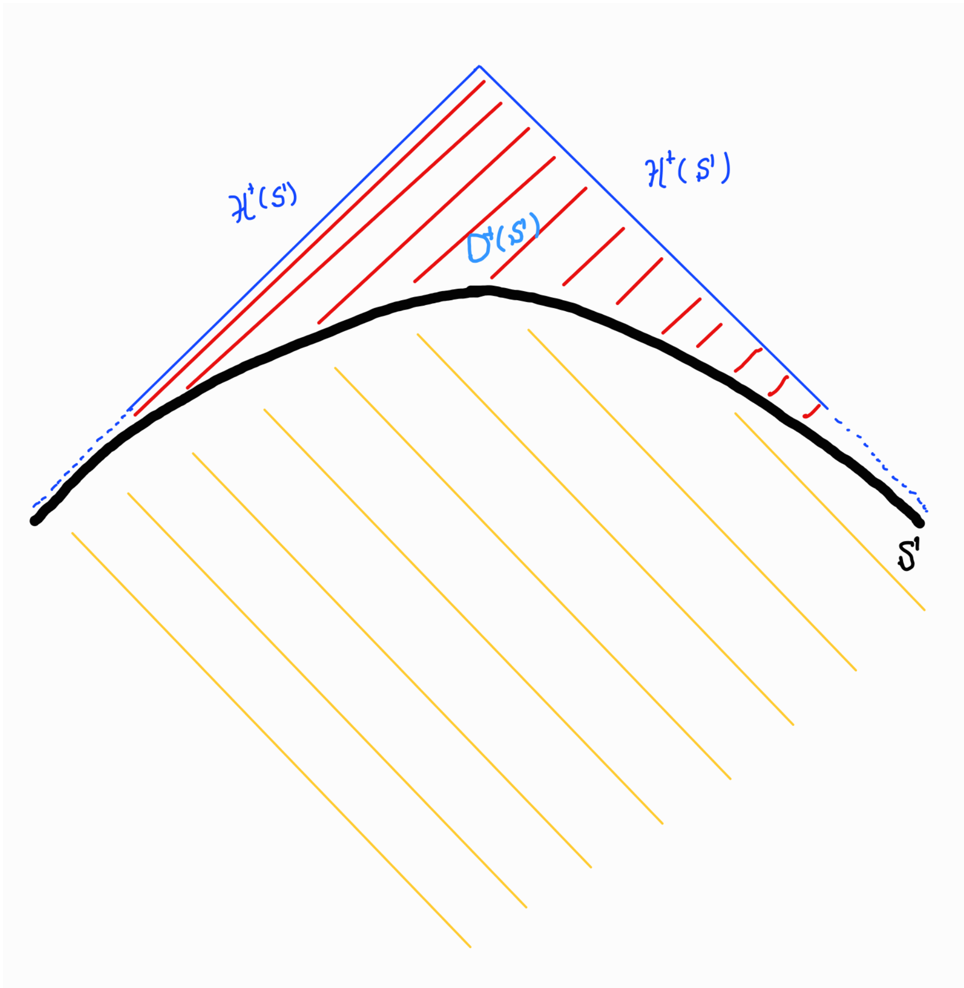

Figure 21: The domain of dependence of a surface S in 2d Minkowski space.

Example 5.3: Removing a point on the surface

So far we have taken a nice continuous line for S. Consider removing a point from the line S. This leads to the future domain of dependence being further restricted. To see why consider the lightcone emanating from the removed point. Any points within this lightcone have to be removed from the domain of dependence. We therefore end up with a figure such as figure 22.

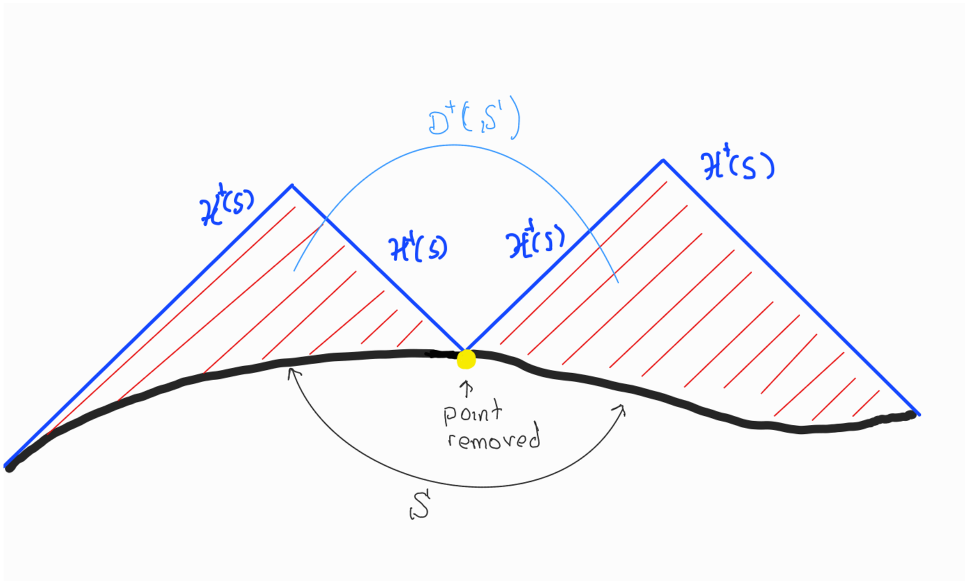

Figure 22: The domain of dependence of a surface S in 2d Minkowski space with a point removed from S. This should be compared with figure 20. By removing a point from S we must remove a large region from the future domain of dependence.

Example 5.4: Removing a point from spacetime

Above we have considered moving a point from S. We have a similar restriction when removing a point from spacetime. We see that there will be causal curves where the removed point is in their past or future. These will never reach S and therefore they cannot be included in the domain of dependence, see figure 22 for an example.

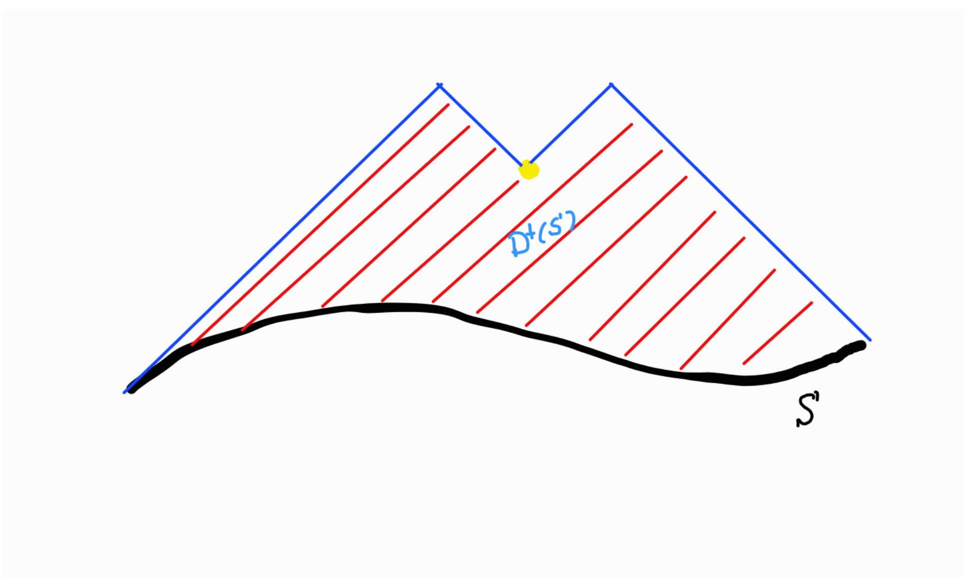

Figure 23: The domain of dependence of a surface S in 2d Minkowski space. The yellow dot denotes a point that has been removed from spacetime. This drastically changes the domain of dependence. One should compare to figure 20.

None of the examples above were Cauchy surfaces. For Minkowski we can simply take a spatial hypersurface, for example $t = 0$ which is a Cauchy surface, clearly this is not unique.

Example 5.5: Cauchy surfaces in Kruskal

Slightly more non-trivially Kruskal spacetime (the maximally extended Schwarzschild solution) also admits and infinite family of Cauchy surfaces. See figure 24.

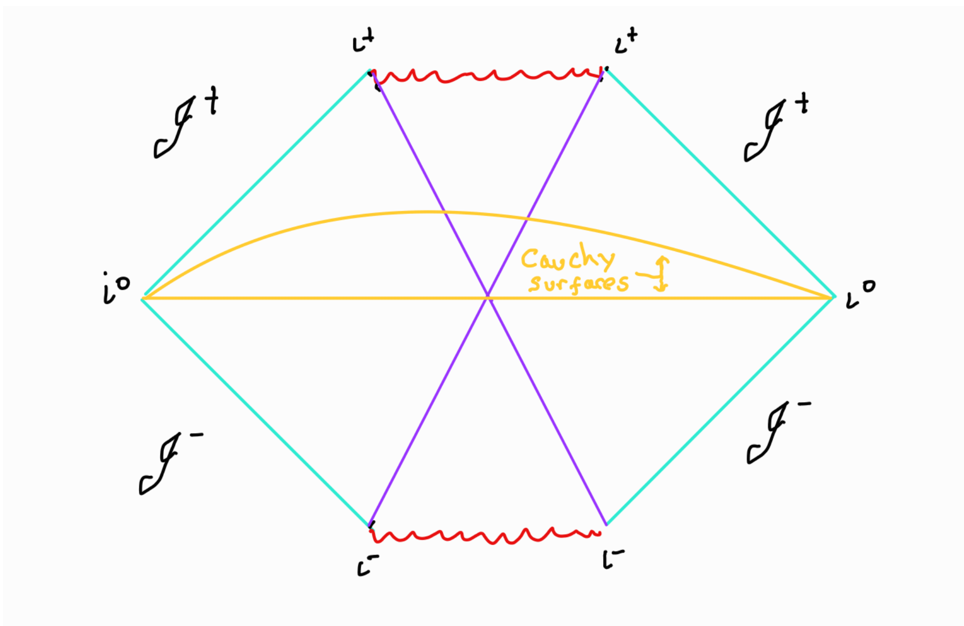

Figure 24: Two distinct choices of Cauchy surface for the Kruskal spacetime.

Example 5.6: Cauchy horizons in Reissner–Nordstrom

The Reissner Nordstrom solution does not admit a Cauchy surface and has Cauchy horizons at the inner horizon.

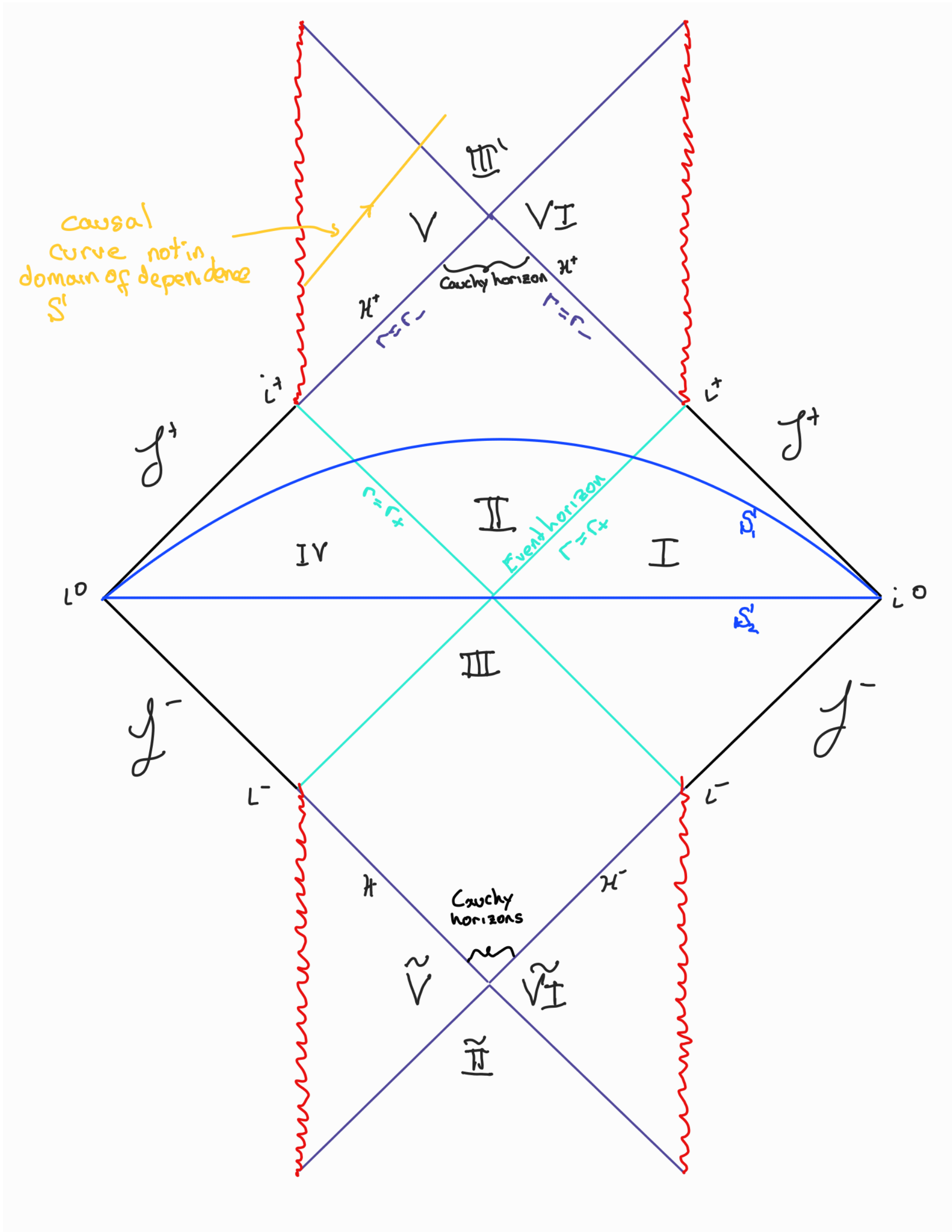

Figure 25: Some wannabe Cauchy surfaces are drawn in dark blue, however all of them admit the same Cauchy horizons drawn in purple. We can see that there are causal curves which are not in the domain of dependence in region V and VI, drawn in yellow.

What are we to make of this lack of determinism and the Cauchy horizon? The maximal extension corresponds to an *eternal black hole* as opposed to an astrophysical black hole since it would not arise in the collapse of matter to form a black hole. The Cauchy horizon is believed to turn into a curvature singularity under small perturbations, including even an observer trying to cross it. This is a major point in favour of Penrose's strong cosmic censorship hypothesis. Penrose's argument for instability of the Cauchy horizon uses a blue-shift effect. There is an infinite blue-shift. Imagine two observers, Alice and Bob. Bob has the misfortune of falling past the black hole horizon, and arrives at the Cauchy horizon in finite proper time. A signal sent by Alice at constant frequency into the horizon will be infinitely blue-shifted when received by Bob as Bob approaches his Cauchy horizon crossing time. In terms of geometric optics this is an instability in the behaviour of linear wave equations like the scalar wave equation on a fixed background. One can view the scalar Laplace equation as a naive linearisation of the vacuum equations (recall our linearisation of GR). The Cauchy horizon leads to a solution which is unstable, and one may presume once the non-linear effects are taken into account the instability forces spacetime to break down before the Cauchy horizon. This has lead to the expectation that upon small perturbations not only is the Cauchy horizon unstable but rather it a spacelike singularity is generated across which the metric is inextendible, i.e. new physics beyond GR is needed. What happens with these horizons is still a very active topic of research.

5.4 Charges in curved spacetime

We have constructed a black hole which is electromagnetically charged, it should therefore carry some form of charge: we want to understand how to compute electric and magnetic charges in gravity. We expect that in the RN black hole the parameters Q and P should be interpreted as the electric and magnetic charges respectively. Consider Maxwell's equation in the presence of a current density J. The equations of motion are:

$$\mathrm{d} \star F = -4\pi \star J \,, \quad \mathrm{d}F = 0 \,. \tag{5.48}$$

The first implies that $\mathrm{d} \star J = 0$, which in components is equivalent to $\nabla_\mu J^\mu = 0$. This is the definition of a conserved current.

Consider a spacelike hypersurface Σ, for example $t = 0$ in RN. We define the total electric charge on Σ to be

$$Q = -\int_\Sigma \star J \,. \tag{5.49}$$

Using Maxwell's equations we can write

$$Q = \frac{1}{4\pi} \int_\Sigma \mathrm{d} \star F \,, \tag{5.50}$$

and if we assume that Σ has a boundary $\partial\Sigma$, then Stoke's theorem gives

$$Q = \frac{1}{4\pi} \int_{\partial\Sigma} \star F \,. \tag{5.51}$$

This is the analogue of Gauss' law $Q \sim \int \vec{E} \cdot \mathrm{d}\vec{S}$. It is telling us about the amount of flux going out of Σ.

Example 5.7: Minkowski space

Consider the Minkowski spacetime in spherical polar coordinates and choose the orientation so that the volume form is

$$\mathrm{dvol} = r^2 \sin\theta \mathrm{d}t \wedge \mathrm{d}r \wedge \mathrm{d}\theta \wedge \mathrm{d}\phi \,. \tag{5.52}$$

Take Σ to be the surface at fixed $t = 0$.[a] We may view Σ as the boundary of the region $t \leq 0$ then Stoke's theorem fixes the orientation of Σ as $r^2 \sin\theta \mathrm{d}r \wedge \mathrm{d}\theta \wedge \mathrm{d}\phi$. Let Σ_R be the region of Σ with $r \leq R$, the boundary is then the two-sphere with radius R: S^2_R. Stokes' theorem fixes the orientation of the two-sphere to be $\mathrm{d}\theta \wedge \mathrm{d}\phi$. Consider the Coulomb potential

$$A = -\frac{q}{r}\mathrm{d}t \,, \qquad F = -\frac{q}{r^2}\mathrm{d}t \wedge \mathrm{d}r \,. \tag{5.53}$$

Taking the Hodge dual gives

$$\star F = q \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi \,, \tag{5.54}$$

and hence the charge on Σ_R is

$$S[\Sigma_R] = \frac{1}{4\pi} \int_{S^2_R} \star F = \frac{1}{4\pi} q \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi = q \,. \tag{5.55}$$

Our definition of Q gives the expected result.

[a] We could of course choose any value of t, there is no reason to take this but for it being notationally simpler.

For an asymptotically flat hypersurface in Minkowski spacetime we can take the limit $R \to \infty$ to express the total charge on Σ as an integral at infinity. Motivated by this we define the charges at asymptotic infinity to be

$$Q = \frac{1}{4\pi} \lim_{r\to\infty} \int_{S^2_r} \star F \,, \qquad P = \frac{1}{4\pi} \lim_{r\to\infty} \int_{S^2_r} F \,, \tag{5.56}$$

where S^2_r is the sphere with radius r.

Note that even when there is no charged matter, $J = 0$ we can still obtain a non-trivial charge, for example the RN solution above. The total charge on a spacelike hypersurface vanishes, since $J = 0$, however when we convert the integral to a surface integral at infinity we obtain two terms because the surface has two asymptotically flat ends. The charges on each of these boundary pieces can be non-zero, so long as they cancel each other when summed.

It remains to be seen why we call this a conserved charge. Consider two spacelike surfaces Σ_1 and Σ_2. Consider the cylindrical surface, V which is bounded by Σ_1 and Σ_2, and large enough to contain all of the sources, see figure 26. From this latter condition it follows that $J = 0$ on the boundaries and outside V. We then have

$$\begin{aligned} 0 &= \int_V \mathrm{d} \star J \\ &= \int_{\partial V} \star J \\ &= \int_{\Sigma_1} \star J - \int_{\Sigma_2} \star J \\ &= \frac{1}{4\pi} \int_{\partial \Sigma_1} \star F - \frac{1}{4\pi} \int_{\partial \Sigma_2} \star F \\ &= Q[\Sigma_1] - Q[\Sigma_2] \,. \end{aligned} \tag{5.57}$$

We can also define magnetic charges similarly. Since they are already defined on the spacelike hypersurface we just need to integrate F, as opposed to $\star F$. A similar argument for showing that it is conserved holds too.

Charges for RN black hole Let us compute the charges for the RN black hole. We have

$$F = \frac{Q}{r^2} \mathrm{d}r \wedge \mathrm{d}t + P \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi \,. \tag{5.58}$$

The magnetic charge is defined to be

$$P[S^2] = \frac{1}{4\pi} \int_{S^2} F = \frac{P}{4\pi} \int_{S^2} \mathrm{dvol}(S^2) = P \,. \tag{5.59}$$

For the electric charge we need the Hodge dual. We have

$$\star F = \frac{1}{r^2} \mathrm{d}t \wedge \mathrm{d}r + Q \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi \,. \tag{5.60}$$

Therefore the electric charge is

$$Q[S^2] = \frac{1}{4\pi} \int_{S^2} Q \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi = Q \,. \tag{5.61}$$

We find that indeed the parameters Q and P are the electric and magnetic charges.

This is one type of conserved charge that we can compute. We will see another example in the following section which allows us to compute the mass and angular momentum of a black hole.

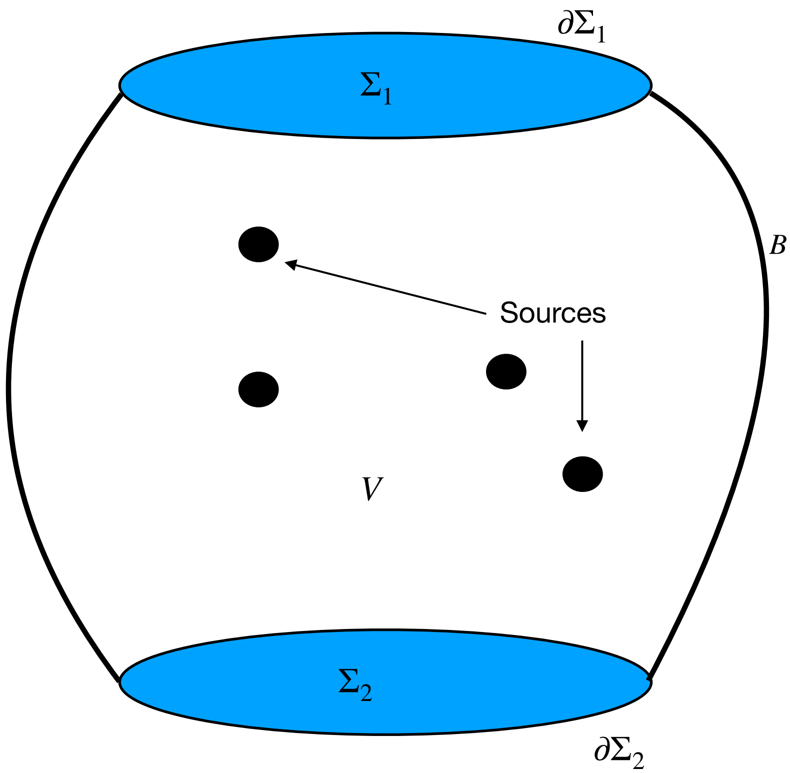

Figure 26: A schematic picture of the region V containing all the soources bounded by the two spacelike hypersurfaces Σ_i and also B. We require that the current vanishes on the black vertical boundary B or alternatively send this to infinity and impose suitable boundary conditions there.

6 Rotating black holes

All the solutions we have seen so far have been static and spherically symmetric, though these are nice testing grounds for us to learn things from they are not likely to be objects that we will see in our universe. Observational evidence seems to suggest that black holes should rotate. In this section we will study rotating black holes, in particular the Kerr black hole. This is arguably the most important solution of general relativity, largely because of the no-hair theorem, according to which all stationary black holes in the universe are Kerr black holes.

Since the black holes are rotating we must give up our spherical symmetry, they can however be axisymmetric: symmetric under rotations about an axis. Moreover, we must give up our metric being static, and reduce to the weaker stationary class of metric.

Definition 21 *Stationary*

A spacetime is stationary *if it admits an everywhere timelike Killing vector K.*

Definition 22 *Static*

A spacetime is static *if it admits a hypersurface-orthogonal timelike Killing vector field.*

> **Remark**
> Note: Static implies stationary, but the converse is not true.

This follows since if we were to run time in the opposite direction we must see rotation in the opposite direction, clearly this cannot be static, we should then impose the weaker stationary condition. These generalisations make the metric a lot more complicated. Although the Schwarzschild solution and Reissner–Nordström solutions were discovered shortly after general relativity was invented, the metric we will study, known as the Kerr(–Newman) metric was first found in 1963. Kerr originally found the rotating metric without any charges but was later extended by Newman to include charges.

6.1 The Kerr–Newman solution

The Kerr–Newman solution in *Boyer-Lindquist coordinates* is

$$\begin{aligned}
\mathrm{d}s^2 &= -\frac{\Delta(r) - a^2\sin^2\theta}{\rho(r,\theta)^2}\mathrm{d}t^2 - \frac{2a\sin^2\theta(r^2+a^2-\Delta(r))}{\rho(r,\theta)^2}\mathrm{d}t\mathrm{d}\phi \\
&\quad + \frac{(r^2+a^2)^2 - a^2\sin^2\theta\Delta(r)}{\rho(r,\theta)^2}\sin^2\theta\mathrm{d}\phi^2 + \frac{\rho(r,\theta)^2}{\Delta(r)}\mathrm{d}r^2 + \rho^2(r,\theta)\mathrm{d}\theta^2\,, \qquad (6.1) \\
A &= -\frac{1}{\rho(r,\theta)^2}\Big(Qr(\mathrm{d}t - a\sin^2\theta\mathrm{d}\phi) + P\cos\theta(a\mathrm{d}t - (r^2+a^2)\mathrm{d}\phi)\Big)\,.
\end{aligned}$$

The functions are

$$\rho(r,\theta)^2 = r^2 + a^2\cos^2\theta\,, \quad \Delta(r) = r^2 - 2Mr + a^2 + e^2\,, \quad e^2 = Q^2 + P^2\,. \tag{6.2}$$

At large r the above coordinates reduce to the spherical polar coordinates in Minkowski spacetime, θ, ϕ have the usual interpretation as angles on S^2, so we have $0 < \theta < \pi$ and $\phi \in [0, 2\pi]$. Notice that there is no problem with taking $r \in (-\infty, \infty)$, in contrast with the other black holes we have considered, the surface $r = 0$ is no longer a singularity.

The Kerr–Newman solution depends on 4 parameters, a, M, Q and P. You may guess that M is the mass, Q the electric charge, P the magnetic charge and a related to the angular momentum. We will show how to compute the angular momentum and mass soon, but for the moment let us just give the result. The parameter a is the angular momentum per unit mass,

$$a = \frac{J}{M}\,, \tag{6.3}$$

with J the Komar angular momentum. This is the unique stationary black hole solution of the Einstein–Maxwell theory. An equilibrium black hole in the presence of an electromagnetic field is therefore fully characterised by the four numbers M, J, Q and P. This goes by the name of the *no hair theorem.*

Note that the metric can be rearranged into the form

$$\mathrm{d}s^2 = -\frac{\Delta(r)}{\rho^2(r,\theta)}\left(\mathrm{d}t - a\sin^2\theta\mathrm{d}\phi\right)^2 + \frac{\rho^2(r,\theta)}{\Delta(r)}\mathrm{d}r^2 + \frac{\sin^2\theta}{\rho^2(r,\theta)}\left(a\mathrm{d}t - (r^2+a^2)\mathrm{d}\phi\right)^2 + \rho^2(r,\theta)\mathrm{d}\theta^2\,, \tag{6.4}$$

which makes clear that the $a = 0$ limit recovers the Reissner–Nordstrom solution of the previous section.

6.2 The Kerr solution

Since all of the essential phenomena persist in the absence of charge we will set $Q = P = 0$ in the remainder of this section. If we set $a \to 0$ the metric reduces to the Schwarzschild solution as one would expect.

Remark

If instead we keep a fixed but set $M \to 0$ then we recover flat space, but in funky coordinates:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \frac{r^2 + a^2\cos^2\theta}{r^2+a^2}\mathrm{d}r^2 + (r^2 + a^2\cos^2\theta)\mathrm{d}\theta^2 + (r^2+a^2)\sin^2\theta\mathrm{d}\phi^2\,. \tag{6.5}$$

The spatial part of the metric is flat three-dimensional space written in ellipsoidal coordinates, see figure 27. They are related to Cartesian coordinates in three-dimensional space by

$$
\begin{aligned}
x &= \sqrt{r^2 + a^2} \sin\theta \cos\phi \,, \\
y &= \sqrt{r^2 + a^2} \sin\theta \sin\phi \,, \\
z &= r \cos\theta \,.
\end{aligned} \tag{6.6}
$$

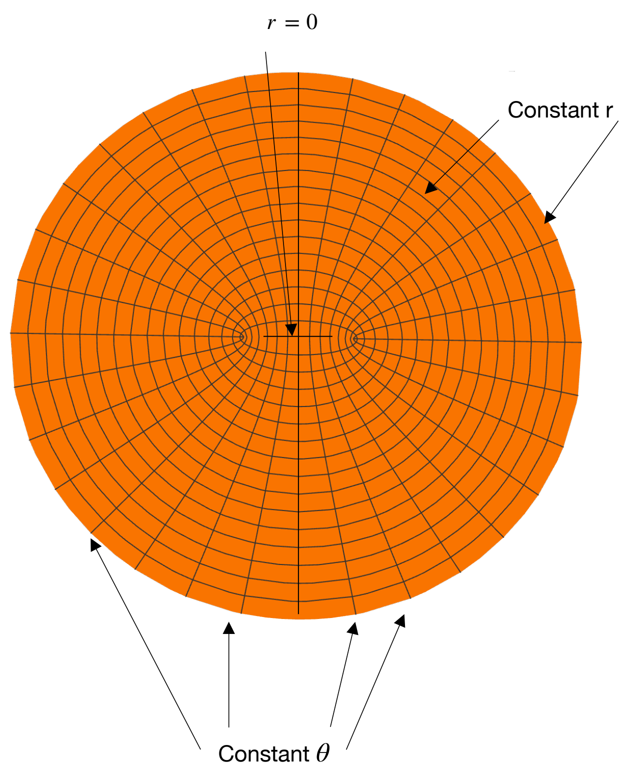

Figure 27: The structure of the ellipsoidal coordinates of the Kerr metric. The region $r = 0$ is a two-dimensional disc with length $2a$.

The Kerr spacetime is asymptotically flat. For $r \to \infty$ we have that the asymptotics are the same as Schwarzschild with mass M. Conversely for $r \to -\infty$ we have that it approaches Schwarzschild with a negative mass.

The metric admits two Killing vectors: both of which are manifest since the metric is independent of both t and ϕ. The Killing vectors are $K = \partial_t$ and $R = \partial_\phi$, which are necessary for the spacetime to be stationary and axis-symmetric. In contrast to Schwarzschild K is not orthogonal to t =constant hypersurfaces, this is why the metric is stationary and not static. Intuitively this makes sense since the black hole is rotating, so not static, but it is spinning in exactly the same place at all times so it is stationary. R expresses the axial symmetry of the solution, this is the symmetry around the axis of rotation.

Remark

Besides the Killing vectors the Kerr metric also has a Killing tensor. A Killing tensor is any symmetric $(0, n)$ tensor $\sigma_{\mu_1 \dots \mu_n}$ satisfying

$$\nabla_{(\nu}\sigma_{\mu_1 \dots \mu_n)} = 0 \,. \tag{6.7}$$

For the Kerr geometry we can define the $(0, 2)$ tensor

$$\sigma_{\mu\nu} = 2\rho^2 l_{(\mu}n_{\nu)} + r^2 g_{\mu\nu} \,, \tag{6.8}$$

where

$$l^\mu = \frac{1}{\Delta}(r^2 + a^2, \Delta, 0, a) \,, \quad n^\mu = \frac{1}{2\rho^2}(r^2 + a^2, -\Delta, 0, a) \,. \tag{6.9}$$

Both vectors are null and satisfy

$$l^\mu l_\mu = 0 \,, \quad n^\mu n_\mu = 0 \,, \quad l^\mu n_\mu = -1 \,. \tag{6.10}$$

One can use this to show that the geodesic equation is integrable, that is we can solve exactly the geodesic equations using integrals of motion. For Schwarzschild we have the three conserved quantities associated to the energy, angular momentum and the Lagrangian after taking an affine parameter. Each of these three conserved quantities exist in the Kerr background too. However, when considering Schwarzschild we can use the spherical symmetry to further constrain the motion in a plane, this makes geodesics in a Schwarzschild background integrable. Such an argument is no longer possible since we do not have spherical symmetry, intuitively we expect there to be some difference between geodesics in the equatorial plane and those following a more general path. The additional conserved quantity associated to the Killing tensor replaces the role of restricting to planar motion.

The coordinates have been chosen so that the event horizons occur at those fixed values of r for which $g^{rr} = 0$. Since $g^{rr} = \Delta/\rho^2$ we have zeroes when

$$\Delta(r) = r^2 - 2Mr + a^2 = 0 \,. \tag{6.11}$$

We may then write

$$\Delta = (r - r_+)(r - r_-) \,, \quad r_\pm = M \pm \sqrt{M^2 - a^2} \,. \tag{6.12}$$

Similar to the Reissner–Nordstrom black hole the solution has three branches depending on the roots of $\Delta(r)$.

- $M^2 - a^2 < 0$ or $|J| < 1$ is the slowly rotating/ underspinning case.
- $M^2 - a^2 = 0$ or $|J| = 1$ is the extremal case.
- $M^2 - a^2 > 0$ or $|J| > 1$ is the rapidly rotating/ over spinning case.

The over-spinning solution has a naked singularity while the extremal solution is unstable. We will just study the slowly rotating case where $M^2 > a^2$ from now on.

6.2.1 Singularities

Let us analyse the various singularities appearing in the metric. As an indicator of singularities we compute the Kretschmann tensor:

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = \frac{48M^2}{\rho^{12}}(r^6 - 15a^2r^4\cos^2\theta + 15a^4r^2\cos^4\theta - a^6\cos^6\theta)\,. \tag{6.13}$$

Clearly this only diverges when $\rho = 0$ which implies $r = 0$ and $\theta = \frac{\pi}{2}$. There are also the usual singularities associated to $\theta = 0, \pi$ but these are just the usual coordinate singularities arising from using spherical polar coordinates and so we will ignore these. The final possible singularities are at the two roots $r_\pm$.

First let us show that $r = r_\pm$ are just coordinate singularities. To do this we define *Kerr coordinates* (v, r, θ, χ) for $r > r_+$ by

$$\mathrm{d}v = \mathrm{d}t + \frac{r^2 + a^2}{\Delta(r)}\mathrm{d}r\,, \quad \mathrm{d}\chi = \mathrm{d}\phi + \frac{a}{\Delta(r)}\mathrm{d}r\,. \tag{6.14}$$

In the new coordinates we have $\chi \sim \chi + 2\pi$ and the Killing vectors are

$$K = \frac{\partial}{\partial v}\,, \qquad R = \frac{\partial}{\partial \chi}\,. \tag{6.15}$$

The new metric in these coordinates is

$$\begin{aligned} \mathrm{d}s^2 = &-\frac{\Delta(r) - a^2\sin^2\theta}{\rho^2}\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r - \frac{2a\sin^2\theta(r^2 + a^2 - \Delta(r))}{\rho^2}\mathrm{d}v\mathrm{d}\chi \\ &- 2a\sin^2\theta\mathrm{d}\chi\mathrm{d}r + \frac{(r^2+a^2)^2 - \Delta(r)a^2\sin^2\theta}{\rho^2}\sin^2\theta\mathrm{d}\chi^2 + \rho^2\mathrm{d}\theta^2\,. \end{aligned} \tag{6.16}$$

This change of coordinates shows that the metric is non-degenerate at $r = r_\pm$. We can analytically continue through the surface $r = r_\pm$ into a new region with $-\infty < r < r_\pm$.

The surfaces $r = r_\pm$ are null hypersurfaces with normal

$$\xi^\mu_\pm = K^\mu + \Omega_\pm R^\mu \,, \tag{6.17}$$

where

$$\Omega_\pm = \frac{a}{r_\pm^2 + a^2} \,, \tag{6.18}$$

is the angular velocity.

Note that one-form

$$\xi = \frac{\rho(r,\theta)}{r^2 + a^2} \mathrm{d}r \,, \tag{6.19}$$

vanishes on the $r = r_\pm$ surfaces and is therefore normal to these hypersurfaces. The dual vector field is

$$\xi = \partial_v + \frac{\Delta(r)}{r^2 + a^2} \partial_r + \frac{a}{r^2 + a^2} \partial_\chi \,, \tag{6.20}$$

which agrees with $\xi_\pm$ above on the horizon where $\Delta(r_\pm) = 0$. The norm of ξ is

$$\xi^\mu \xi_\mu = \frac{\rho^2(r,\theta)\Delta(r)}{(r^2 + a^2)^2} \,, \tag{6.21}$$

which clearly vanishes at $r = r_\pm$ and therefore the vector $\xi_\pm$ is a null Killing vector on $r = r_\pm$. The region $r \leq r_+$ part of the black hole region of this spacetime with $r = r_+$ the future event horizon $\mathcal{H}^+$. In Boyer–Lindquist coordinates the Killing vector is

$$\xi = \frac{\partial}{\partial t} + \Omega_+ \frac{\partial}{\partial \phi} \,. \tag{6.22}$$

Observe that $\xi^\mu \partial_\mu(\phi - \Omega_+ t) = 0$ and therefore $\phi = \Omega_+ t + \text{const}$ on integral curves of ξ^μ. Conversely, integral curves of K have $\phi = \text{const}$. We see that particles moving on orbits of ξ rotate with angular velocity Ω_+ with respect to a stationary observer (someone on an orbit of K). In particular they rotate with this angular velocity with respect to a stationary observer at infinity. Since ξ is tangent to the generators of $\mathcal{H}^+$ (the null geodesics defined as normal to the surface), these generators also rotate with angular velocity Ω_+ with respect to a stationary observer at infinity. Therefore we interpret Ω_+ as the angular velocity of the black hole.

The remaining singular point is now where $\rho = 0$, i.e. at the point

$$r = 0 \,, \qquad \theta = \frac{\pi}{2} \,. \tag{6.23}$$

To understand the singularity at this point we should fix ourselves to a constant time slice and go to the point where $\rho = 0$. Working in Kerr coordinates we have:

$$\mathrm{d}s^2\Big|_{v,r,\theta \text{ Fixed}} = \frac{(r^2 + a^2)^2 - \Delta(r) a^2 \sin^2\theta}{\rho^2} \sin^2\theta \mathrm{d}\chi^2 \tag{6.24}$$

We see that as we take $r \to 0$ we are left with

$$\mathrm{d}s^2\Big|_{v,r,\theta \text{ Fixed}} = a^2 \sin^2\theta \mathrm{d}\chi^2. \tag{6.25}$$

This then defines a disc parametrised by θ and χ. When we also take $\theta = \frac{\pi}{2}$ we end up with the metric $\mathrm{d}s^2 = a^2\mathrm{d}\chi^2$ which is the metric on a ring of radius a. Therefore, in the Kerr metric, the curvature singularity has the structure of a ring. The rotation has softened the Schwarzschild singularity, spreading it out over a ring. If you travel towards $r = 0$ from any other angle other than $\theta = \frac{\pi}{2}$ you will not encounter the singularity and will instead pass through and enter a new asymptotically flat region, i.e. $r \to -\infty$. This is not an identical copy of the spacetime you came from though, instead it is described by the Kerr metric with $r < 0$ (effectively $M \to -M$). As a result Δ never vanishes and there are no horizons in this space.

This spacetime with $r < 0$ has an unusual feature. One finds that $R = \partial_\phi$ becomes time-like near the singularity, the metric at fixed $t, r < 0$ and $\theta = \frac{\pi}{2}$ is

$$\mathrm{d}s^2 = \left(r^2 + a^2 + \frac{2Ma}{r}\right)\mathrm{d}\chi^2\,, \tag{6.26}$$

which close enough to the singularity is negative. Since χ is 2π-periodic we end up with *closed timelike curves* CTCs. You may sometimes hear these referred to as time-machines. It is a curve that is everywhere timelike and that eventually returns to where it started in spacetime. You can then travel on this CTC and meet yourself in the past!

This region is unphysical. Much like in the case of sub-extremal RN the inner horizon at $r = r_-$ is a Cauchy horizon and becomes a curvature singularity in the presence of the smallest perturbations to the Kerr metric: at the inner horizon perturbations are infinitely blueshifted, which leads to divergences in the curvature scalars.

When we considered Schwarzschild we saw that it describes the metric outside a spherical star. This was a consequence of Birkhoff's theorem. In contrast the Kerr solution does *not* describe the spacetime outside a rotating star. This solution is expected to describe only the final state of gravitational collapse. One can't obtain a solution describing gravitational collapse to form a Kerr black hole by simply gluing in a ball of collapsing matter as was possible for Schwarzschild. Additionally, the spacetime during collapse would not even be stationary as gravitational waves must be emitted.

Theorem 5 *Carter 1971, Robinson 1975*

If (M, g) is a stationary, axisymmetric, asymptotically flat vacuum spacetime suitably regular on, and outside a connected event horizon then (M, g) is a member of the 2-parameter Kerr family of solutions. The parameters are the angular momentum and mass.

This result says that the final state of gravitational collapse is generically a Kerr black hole and is fully characterised by just 2 numbers. In contrast the initial state can be arbitrarily complicated. Nearly all information about the initial state is lost during gravitational collapse: either by radiation to infinity, or by falling into the black hole, and just 2 numbers are required to describe the final state on and outside the event horizon. There is an extension of this theorem for the 4-parameter Kerr–Newman solution.

6.3 Maximal extension of Kerr and its Penrose diagram

We now want to understand the Penrose diagram for the Kerr solution. It is now more difficult to draw the Penrose diagram because the metric is no longer spherically symmetric. Since the curvature singularity will appear only for $\theta = \frac{\pi}{2}$ the Penrose diagram will look different for $\theta \neq \frac{\pi}{2}$ and $\theta = \frac{\pi}{2}$. To represent both cases it is customary to draw a Penrose diagram that is an amalgamation of the Penrose diagram for an observer falling in from the north pole and along the equatorial plane at fixed χ in Kerr coordinates. Notice that $\chi = \text{const}$ means that ϕ is not constant so the observer falling in at $\theta = \frac{\pi}{2}$ rotates about the polar axis. See figure 28 for the Penrose diagram.

6.4 Komar Integrals and conserved quantities along geodesics

In the above we have claimed that the Kerr black hole is rotating and has angular momentum $J = aM$, we would like to back up this claim. This relies on us defining a Komar integral, which is essentially a charge associated to a Killing vector.

We have seen that we can define conserved electric and magnetic charges given a gauge field, one can understand the need for a charge associated to a Killing vector by playing a little game with Kaluza–Klein reduction.

Consider Einstein gravity in five-dimensions without a cosmological constant. Let us take an ansatz for the metric of the form

$$\mathrm{d}s^2 = \phi^2(x)\big(\mathrm{d}\psi + A_\mu \mathrm{d}x^\mu\big)^2 + g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \,, \tag{6.27}$$

where ∂_ψ is a Killing vector and the one-form A is defined only on the base with coordinates x. Note that gauge transformations are just coordinate transformations in this formalism.

We can now plug this into the five-dimensional vacuum Einstein equations. One

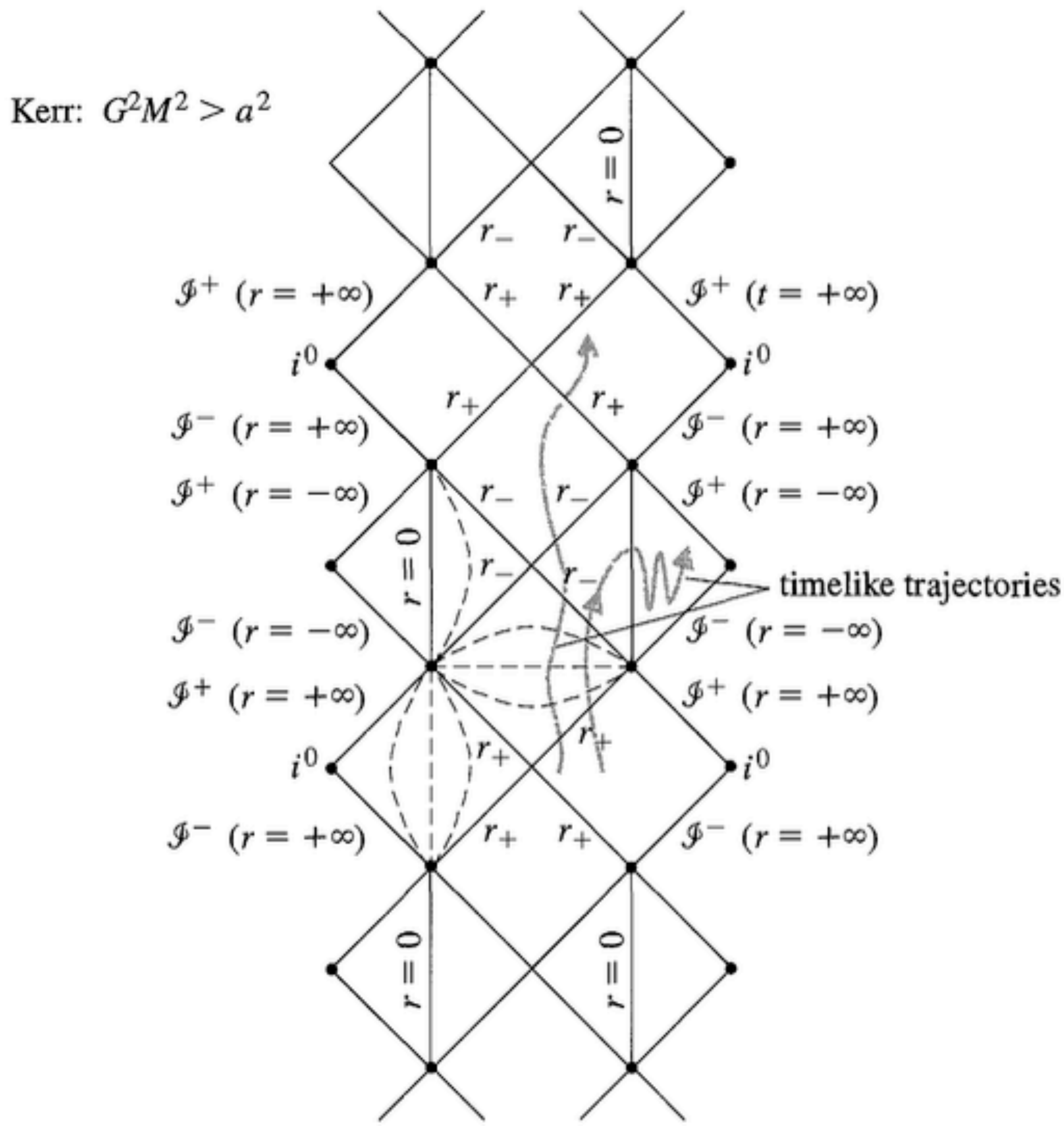

Figure 28: The Penrose diagram for sub-extremal Kerr. There are and infinite number of copies of the region outside the black hole. The singularity at $r = 0$ only appears for $\theta = \frac{\pi}{2}$ and is absent for other values of θ. The regions beyond the singularity are where we have CTCs.

finds that there are three conditions one must impose in order for the metric to satisfy the five-dimensional Einstein vacuum equations:

$$
\begin{aligned}
\Box\phi &= \frac{1}{4}\phi^3 F^{\mu\nu}F_{\mu\nu}\,, \\
\nabla_\mu\Big(\phi^3 F^{\mu\nu}\Big) &= 0\,, \\
R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R &= \frac{1}{2}\phi^2\Big(F_{\mu\rho}F_\nu{}^\rho - \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\Big) + \frac{1}{\phi}(\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\Box\phi)\,,
\end{aligned}
\tag{6.28}
$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and everything is a four-dimensional object defined by the metric $g_{\mu\nu}$. For a constant ϕ we can see the Maxwell equation and Einstein equation of the Einstein–Maxwell theory, of course setting ϕ =constant imposes a non-trivial relation on the F but let us forget about this for the moment.

We see that if 5-dimensional spacetime has a circle which is small, then we see a four-dimensional spacetime which is Einstein gravity plus electromagnetism. Now we know that in the four-dimensional theory we can define electric (and magnetic) charges, but there should be some remnant of these electric charges in the five-dimensional theory. In the five-dimensional theory it must enter through the gauge field A and therefore it is connected to the Killing vector ∂_ψ: there must be a way of defining a conserved charge to a Killing vector which is the analogue of the electric charge in the dimensionally reduced theory.

6.4.1 Lie derivative recap

Let (M, g) be a Lorentzian manifold with metric g. Given a smooth vector-field X on M we define an integral curve $\gamma(\lambda) : \mathbb{R} \to M$ to be a curve whose tangent vector is equal to X at every point $p \in \gamma$. That is we demand

$$
X^\mu\Big|_p = \frac{\mathrm{d}}{\mathrm{d}\lambda}x^\mu(\lambda)\Big|_p\,. \tag{6.29}
$$

This is equivalent to solving a set of first order ODEs with fixed initial conditions, and therefore there is a unique solution at least locally.

Let $\gamma(\lambda, p)$ be the integral curve of X which passes through the point p when $\lambda = 0$. The map $\gamma : \mathbb{R} \times M \to M$ defines the flow generated by X. The flow defines an abelian group since one can show that $\sigma(\lambda_1, \sigma(\lambda_2, p)) = \sigma(\lambda_1 + \lambda_2, p)$. Let $\sigma_\lambda(p) = \sigma(\lambda, p)$ then

$$
\begin{aligned}
&\sigma_\lambda\big(\sigma_\tau(p)\big) = \sigma_{\lambda+\tau}(p)\,, \\
&\sigma_0 = \text{Unit element}\,, \\
&\sigma_{-\lambda} = (\sigma_\lambda)^{-1}\,.
\end{aligned}
\tag{6.30}
$$

This allows us to move points along the curve, in particular by using the flow we can move tensors from one point on the flow to another, recall that this goes by the name of *push-forward* or *pull back* depending on what object we are acting on.[20] This allows us to define the Lie derivative along the vector field X. For a vector Y we have

$$\mathcal{L}_X Y|_p = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\big(\sigma_{-\epsilon}(p)\big)_* Y\big|_{\sigma_\epsilon(p)} - Y\big|_p\Big] \,. \tag{6.31}$$

One can show that

$$\mathcal{L}_X Y = [X,Y] \,, \tag{6.32}$$

with $[\ ,\ ] : \mathcal{X}(M)\times\mathcal{X}(M)\to\mathcal{X}(M)$ the Lie bracket

$$[X,Y] = \Big(X^\nu\partial_\nu Y^\mu - Y^\nu\partial_\nu X^\mu\Big)\partial_\mu \,. \tag{6.33}$$

The Lie derivative can be extended to any tensor with appropriate generalisation. For tensors one must use a combination of the push-forward and pull-back. Of primary interest to us here is the Lie derivative of the metric. We have

$$\mathcal{L}_X g = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[(\sigma_\epsilon(p))^* g|_{\sigma_\epsilon(p)} - g|_{\sigma_\epsilon(p)}\Big] \,. \tag{6.34}$$

Note that the pull back uses σ_ϵ rather than $\sigma_{-\epsilon}$, this is not a typo. In coordinates we have

$$(\mathcal{L}_X g)_{\mu\nu} = X^\sigma\partial_\sigma g_{\mu\nu} + g_{\sigma\nu}\partial_\mu X^\sigma + g_{\mu\sigma}\partial_\nu X^\sigma \,, \tag{6.35}$$

which by using the Levi–Civita connection can be rewritten as

$$(\mathcal{L}_X g)_{\mu\nu} = \nabla_\mu X_\nu + \nabla_\nu X_\mu \,. \tag{6.36}$$

More generally, let T be a tensor of rank (q,r), then the Lie derivative along the vector field X in local coordinates is

$$\begin{aligned}\mathcal{L}_X T^{\mu_1\ldots\mu_q}{}_{\nu_1\ldots\nu_r} &= X^\sigma\partial_\sigma T^{\mu_1\ldots\mu_q}{}_{\nu_1\ldots\nu_r} - \big(\partial_\sigma X^{\mu_1}\big)T^{\lambda\ldots\mu_q}{}_{\nu_1\ldots\nu_r} - \ldots - \big(\partial_\sigma X^{\mu_q}\big)T^{\mu_1\ldots\sigma}{}_{\nu_1\ldots\nu_r}\\ &\quad + \big(\partial_{\nu_1}X^\sigma\big)T^{\mu_1\ldots\mu_q}{}_{\sigma\ldots\nu_r} + \ldots + \big(\partial_{\nu_r}X^\sigma\big)T^{\mu_1\ldots\mu_q}{}_{\nu_1\ldots\sigma} \,.\end{aligned} \tag{6.37}$$

To make this more manifestly tensorial one can replace the partial derivatives with *any* torsion free connection[21], not necessarily the Levi–Civita connection. One can show that the Lie derivative satisfies

$$\begin{aligned}\mathcal{L}_X(T+S) &= \mathcal{L}_X T + \mathcal{L}_X S \,,\\ \mathcal{L}_X(T\otimes S) &= \Big(\mathcal{L}_X T\Big)\otimes S + T\otimes\Big(\mathcal{L}_X S\Big) \,,\\ \mathcal{L}_{[X,Y]} &= \mathcal{L}_X\mathcal{L}_Y - \mathcal{L}_Y\mathcal{L}_X \,,\\ \mathcal{L}_X f &= X[f] \,,\end{aligned} \tag{6.38}$$

where X, Y are vector fields, T and S are arbitrary tensors, and f is a function.

[20] Given a smooth function $f : M \to N$ the push forward $f_* : T_p(M) \to T_{f(p)}(N)$ acts on a vector field V as $(f_*V)[g] = V[g\circ f]$. The pullback $f^* : T^*_{f(p)}(N) \to T^*_p(M)$, acts as $\langle f^*\omega, V\rangle = \langle \omega, f_*V\rangle$.

[21] Recall that a connection is torsion free if the connection coefficients satisfy $\Gamma^\mu{}_{\nu\rho} = \Gamma^\mu{}_{\rho\nu}$.

6.4.2 Killing vectors

Recall that under a diffeomorphism generated by the vector field X,

$$x^\mu \to \tilde{x}^\mu = x^\mu - X^\mu \,. \tag{6.39}$$

the change in the metric is given by the Lie derivative of the metric under the vector field:

$$\delta g_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) = \Big(\mathcal{L}_X g\Big)_{\mu\nu} \,. \tag{6.40}$$

Aside

To see this let us transform the metric under the above coordinate transformation and expand working to linear order.

$$g_{\mu\nu}(x) \to \tilde{g}_{\mu\nu}(\tilde{x}) \equiv \frac{\partial x^\rho}{\partial \tilde{x}^\mu}\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma}(x) \,. \tag{6.41}$$

Then we have

$$\frac{\partial x^\rho}{\partial \tilde{x}^\mu} = \delta^\rho_\mu + \frac{\partial}{\partial \tilde{x}^\rho} X^\mu = \delta^\rho_\mu + \frac{\partial}{\partial x^\rho} X^\mu + \mathcal{O}(X^2) \,, \tag{6.42}$$

and therefore to leading order

$$\tilde{g}_{\mu\nu}(\tilde{x}) = g_{\mu\nu}(x) + \partial_\mu X^\rho g_{\rho\nu} + \partial_\nu X^\rho g_{\mu\rho} \tag{6.43}$$

Now since we are interested in $\tilde{g}_{\mu\nu}(x)$ and not $\tilde{g}_{\mu\nu}(\tilde{x})$ we have

$$\tilde{g}_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(\tilde{x} + X) = \tilde{g}_{\mu\nu}(\tilde{x}) + X^\rho \partial_\rho g_{\mu\nu} + \mathcal{O}(X^2) \,. \tag{6.44}$$

Plugging this expansion in we find

$$\begin{aligned} \delta g_{\mu\nu}(x) &= \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) \\ &= X^\sigma \partial_\sigma g_{\mu\nu} + g_{\sigma\nu}\partial_\mu X^\sigma + g_{\mu\sigma}\partial_\nu X^\sigma \,. \end{aligned} \tag{6.45}$$

One can now lower the index on X^σ

$$\begin{aligned} \delta g_{\mu\nu}(x) &= g^{\sigma\rho} X_\rho \partial_\sigma g_{\mu\nu} + \partial_\mu X_\nu - \partial_\mu g_{\sigma\nu} X^\sigma + \partial_\nu X_\mu - \partial_\nu g_{\sigma\mu} X^\sigma \\ &= \partial_\mu X_\nu + \partial_\nu X_\mu - X_\rho g^{\sigma\rho}(\partial_\nu g_{\sigma\mu} + \partial_\mu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu}) \\ &= \nabla_\mu X_\nu + \nabla_\nu X_\mu \,, \end{aligned} \tag{6.46}$$

with ∇ the Levi–Civita connection.

The Lie derivative on a rank $(2,0)$-tensor acts as

$$(\mathcal{L}_X g)_{\mu\nu} = X^\sigma \partial_\sigma g_{\mu\nu} + g_{\sigma\nu}\partial_\mu X^\sigma + g_{\mu\sigma}\partial_\nu X^\sigma \,. \tag{6.47}$$

Using the aside above we see that the Lie derivative of the metric along X can be rewritten as

$$(\mathcal{L}_X g)_{\mu\nu} = \nabla_\mu X_\nu + \nabla_\nu X_\mu \,. \tag{6.48}$$

There are special vector fields, known as *Killing vectors*, which preserve the form of the metric after the coordinate transformation, i.e. $\delta g_{\mu\nu} = 0$. We can define Killing vectors to be vector fields that obey:

$$\nabla_\mu X_\nu + \nabla_\nu X_\mu = 0 \,. \tag{6.49}$$

They are vectors which define flows along which the metric does not change. We say that it generates an *isometry* of the spacetime and that the metric has a *symmetry*.

There is an upper limit on the number of Killing vectors a manifold can have. In n dimensions the maximum number of linearly independent (by constant coefficients) Killing vectors is

$$\frac{n(n+1)}{2} \,. \tag{6.50}$$

Spaces which admit the maximum number of Killing vectors are known as *maximally symmetric spaces*.

Definition 23 *Maximally symmetric space*

An n-dimensional space with the maximum number of Killing vectors, $\frac{n(n+1)}{2}$, is called a maximally symmetric space*.*

Not examinable To prove this one needs to use that for a Killing vector K we have

$$\nabla_\mu \nabla_\nu K^\sigma = R^\sigma{}_{\nu\mu\rho} K^\rho \,. \tag{6.51}$$

Then we view the Killing equation (6.49) as a set of first order PDEs for the n functions K^μ. We can now find a solution as a series expansion around some arbitrary point p in the manifold. We would have

$$K^\mu(x) = K^\mu(p) + (x^\nu - p^\nu)\partial_\nu K^\mu\Big|_{x=p} + \frac{1}{2}(x^\nu - p^\nu)(x^\sigma - p^\sigma)\nabla_\nu\nabla_\sigma K^\mu\Big|_{x=p} + \ldots . \tag{6.52}$$

However since (6.51) allows us to express the second derivative of K at p in terms of $K(p)$ and $\partial_\mu K(p)$ it follows that we may eliminate second derivative terms from the expansion. In fact we may go further, whacking (6.51) with another derivative allows us to express the third derivative of K in terms of $K_\nu(p)$ and $\nabla_\mu K_\nu(p)$ too. We can do this infinitely many times to obtain expressions for all higher derivative terms. Therefore the solution is determined uniquely by the initial conditions $K^\mu(p)$ and $\nabla_\mu K^\nu|_{x=p}$. The general solution is then of the form

$$K_\mu(x) = A_\mu{}^\nu(x,p) K_\nu(p) + B_\mu{}^{\nu\rho}(x,p) \nabla_\nu K_\rho \Big|_{x=p} \,, \tag{6.53}$$

where A and B are complicated functions depending on the initial point p and the metric and its derivatives but independent of the initial data of the Killing vector. Therefore we have shown that every Killing vector can be determined in terms of the initial conditions $K_\mu(p)$ and $\nabla_\mu K_\nu\big|_{x=p}$. There are n-independent components of $K_\mu(p)$ and $\frac{n(n-1)}{2}$ independent components of $\nabla_\mu K_\nu\big|_{x=p}$. The latter comes about because the initial conditions must satisfy the Killing equation, which fixes $\nabla_\mu K_\nu\big|_{x=p}$ to be a $n \times n$ anti-symmetric matrix which has $\frac{n(n-1)}{2}$ independent components. This gives the claimed total of $\frac{n(n+1)}{2}$ Killing vectors.

If a manifold is maximally symmetric it means that the curvature is the same in all directions. The Riemann tensor can in fact be fixed in terms of the constant Ricci scalar and takes the form

$$R_{\mu\nu\rho\sigma} = \frac{R}{n(n-1)} \left(g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho} \right) . \tag{6.54}$$

This means that locally the space is determined by the Ricci scalar.[22]

Example 6.1: Maximally symmetric spaces

Examples of maximally symmetric Euclidean spaces are flat Euclidean plane, spheres and hyperbolic space. Lorentizan maximally symmetric spaces include Minkowski space and (anti-) de-Sitter space.

6.4.3 Conserved quantities along geodesics

We have seen when studying the geodesics of Schwarzschild that there are some conserved quantities. The underlying mathematics behind why these quantities are conserved is lacking, here we remedy that.

Consider the action

$$S = \int \mathrm{d}\lambda \sqrt{\left| g_{\mu\nu}\big(x(\lambda)\big) \frac{\mathrm{d}x^\mu(\lambda)}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu(\lambda)}{\mathrm{d}\lambda} \right|} \,. \tag{6.55}$$

For simplicity let us assume that λ is an affine parameter which allows us to consider the action with the square root removed. From GR1 we know that geodesics are the curves which extremise the action, that is geodesics are curves, $x^\mu(\lambda)$, which when deformed by a small amount $\delta x^\mu(\lambda)$, the change in the action vanishes.

[22]For example both a torus and the Euclidean plane are flat, and hence the Riemann tensor vanishes, however they are very different spaces, one is compact while the other is non-compact. The Ricci scalar therefore does not capture the global difference of the two spaces.

Consider deforming the curve via

$$x^\mu \to x^\mu + \epsilon X^\mu \,. \tag{6.56}$$

The change in the action is

$$\begin{aligned}\delta S &= S(x^\mu + \epsilon X^\mu) - S(x^\mu) \\ &= \epsilon \int \mathrm{d}\lambda \Big[x^\rho \partial_\rho g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu + g_{\mu\nu} \big(\dot{X}^\mu \dot{x}^\nu + \dot{x}^\mu \dot{X}^\nu \big) \Big] + \mathcal{O}(\epsilon^2) \\ &= \epsilon \int \mathrm{d}\lambda \; \dot{x}^\rho \dot{x}^\sigma \Big[X^\mu \partial_\mu g_{\rho\sigma} + g_{\nu\sigma} \partial_\rho X^\nu + g_{\nu\rho} \partial_\sigma X^\nu \Big] + \mathcal{O}(\epsilon^2) \\ &= \epsilon \int \mathrm{d}\lambda \; \dot{x}^\rho \dot{x}^\sigma \Big[\nabla_\rho X_\sigma + \nabla_\sigma X_\rho \Big] + \mathcal{O}(\epsilon^2) \,.\end{aligned} \tag{6.57}$$

We have used that $\dot{X}^\mu = \dot{x}^\rho \partial_\rho X^\mu$. We see that if X is a Killing vector field then we have a symmetry of the action. We know from Noether's theorem that there must be a conserved charge.

Given an action S with Lagrangian density $\mathcal{L}$,

$$S = \int \mathrm{d}\lambda \; \mathcal{L}\big(x(\lambda)\big) \,, \tag{6.58}$$

define the conjugate momentum p_μ to be

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \,. \tag{6.59}$$

Then for any Killing vector X,

$$Q = X^\mu p_\mu \,, \tag{6.60}$$

is a conserved quantity along the geodesic.

Proof: Consider a small variation $\delta x^\mu = \epsilon X^\mu$ generated by the Killing vector field X as above. As shown above such variations leave the action invariant: $\delta S = 0$, which is equivalent to

$$\frac{\partial \mathcal{L}}{\partial x^\mu} X^\mu + \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \dot{X}^\mu = 0 \,. \tag{6.61}$$

Along a geodesic the Euler–Lagrange equations are satisfied:

$$\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = 0 \,. \tag{6.62}$$

Therefore along the geodesic (6.61) implies

$$0 = \Big(\frac{\mathrm{d}}{\mathrm{d}\lambda} p_\mu \Big) X^\mu + p_\mu \frac{\mathrm{d}}{\mathrm{d}\lambda} X^\mu = \frac{\mathrm{d}}{\mathrm{d}\lambda} \Big(p_\mu X^\mu \Big) = \frac{\mathrm{d}}{\mathrm{d}\lambda} Q \,. \tag{6.63}$$

Note that Q is conserved only along the geodesic, for a path which is not a geodesic this is *not* conserved. We can see immediately why this must be the case in the derivation above since we used the Euler–Lagrange equations.

Example 6.2: Kerr conserved quantities

Kerr has two Killing vectors and so each will give rise to a conserved quantity. The two Killing vectors are $K = \partial_t$ and $R = \partial_\phi$. We therefore have that there are two conserved quantities: $E = -p_t$ and $l = p_\phi$. Explicit computation gives

$$\begin{aligned} E = -K_\mu p^\mu &= m\left[\left(1 - \frac{2Mr}{\rho^2}\right)\frac{\mathrm{d}t}{\mathrm{d}\tau} + \frac{2Mar}{\rho^2}\sin^2\theta\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right], \\ l = R_\mu p^\mu &= m\left[-\frac{2Mar}{\rho^2}\sin^2\theta\frac{\mathrm{d}t}{\mathrm{d}\tau} + \frac{(r^2+a^2)^2 - \Delta(r)a^2\sin^2\theta}{\rho^2}\sin^2\theta\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right]. \end{aligned} \tag{6.64}$$

It is interesting to note that one can set the angular momentum to vanish but still have a rotation. Observe that $l = 0$ fixes:

$$\begin{aligned} \rho^2\frac{\mathrm{d}t}{\mathrm{d}\lambda} &= \frac{E}{\Delta}(\rho^2(r^2+a^2) + 2a^2mr\sin^2\theta), \\ \rho^2\frac{\mathrm{d}\phi}{\mathrm{d}\lambda} &= \frac{2amrE}{\Delta}. \end{aligned} \tag{6.65}$$

So we find that on the geodesic:

$$\frac{\mathrm{d}\phi}{\mathrm{d}\lambda} = \frac{2amr}{\rho^2(r^2+a^2) + 2a^2mr\sin^2\theta}. \tag{6.66}$$

Hence for a geodesic with zero angular momentum about the symmetry axis has a rotational motion around that axis. Moreover it is independent of the geodesics characteristics, the mass, energy and so forth. This is a manifestation of the Lense–Thirring effect, also known as the dragging of inertial frames or frame dragging.

One can also define Killing tensors. These are symmetric tensors which satisfy

$$\nabla_{(\mu}X_{\nu_1\ldots\nu_q)} = 0. \tag{6.67}$$

If X is such a tensor then

$$Q = p^{\nu_1}\ldots p^{\nu_q}X_{\nu_1\ldots\nu_q}, \tag{6.68}$$

is conserved along the geodesic. The existence of such Killing tensors is non-trivial and very special.[23]

[23]What we really mean is Killing tensors which are distinct with the metric and products of Killing vectors. Each of these satisfies (6.67) but confers no additional information.

6.4.4 Komar integrals

Let k be a Killing vector, and recall that this implies $\nabla_{(\mu}k_{\nu)} = 0$. Therefore $\nabla_\mu k_\nu$ is anti-symmetric. We can define the two-form

$$K_{\mu\nu} = \nabla_\mu k_\nu \,, \qquad K = \mathrm{d}k \,, \tag{6.69}$$

were we have abused notation to write k for the form and also the vector field. For any vector X recall that we have

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) X^\sigma = R^\sigma{}_{\rho\mu\nu} X^\rho \,. \tag{6.70}$$

Let us use this with the Killing vector and contract the σ and μ indices, we have

$$\begin{aligned} \nabla_\mu \nabla_\nu k^\mu - \nabla_\nu \nabla_\mu k^\mu &= R_{\rho\nu} k^\rho \\ &= \nabla_\mu \nabla_\nu k^\mu \\ &= \nabla^\mu K_{\nu\mu} \,, \end{aligned} \tag{6.71}$$

and therefore we have

Exercise 6:

$$\nabla^\mu K_{\mu\nu} = -R_{\nu\mu} k^\mu \,. \tag{6.72}$$

In form notation we have

$$\mathrm{d} \star \mathrm{d}k = 8\pi G_N \star J = 2 \star R_{\mu\nu} k^\mu \mathrm{d}x^\nu \,. \tag{6.73}$$

This should look reminiscent of how we defined electric charges in the previous section. We may rewrite the above using Einstein's equations:

$$R_{\mu\nu} = 8\pi G_N \Big(T_{\mu\nu} - \frac{1}{2} T^\rho{}_\rho g_{\mu\nu} \Big) \,, \tag{6.74}$$

to find the current

$$J_\mu = 2 \Big(T_{\mu\nu} - \frac{1}{2} T^\rho{}_\rho g_{\mu\nu} \Big) k^\nu \,. \tag{6.75}$$

Thus $\mathrm{d} \star J = 0$. In analogy to how we defined a charge in electromagnetism, on a spatial hypersurface Σ, we may define the conserved charge

$$Q_k(B) = -\int_\Sigma \star J = \frac{1}{8\pi} \int_\Sigma \mathrm{d} \star \mathrm{d}k = \frac{1}{8\pi} \int_{\partial\Sigma} \star \mathrm{d}k \tag{6.76}$$

We define the charge to be taken at asymptotic infinity.

Definition 24 *Komar mass*

Let Σ be a spacelike hypersurface with boundary S^2_r in an asymptotically flat stationary spacetime, with time-like Killing vector k. The Komar mass *(or Komar energy) is*

$$M_{Komar} = -\frac{1}{8\pi} \lim_{r\to\infty} \int_{S^2_r} \star \mathrm{d}k \,. \tag{6.77}$$

This is a measure of the total energy of the spacetime. This energy comes from both matter and the gravitational field. You have seen in GR1 exercises that even when computing the Komar mass for the Schwarzschild solution, which is a vacuum solution with no matter, we find a non-zero Komar mass which is equal to M.

Since the only property of k we used was that it is a Killing vector we can also define the angular momentum in a similar way.

Definition 25 *Komar angular momentum*

Let Σ be a spacelike hypersurface with boundary S^2_r in an asymptotically flat stationary spacetime with killing axisymmetric vector k. Then the Komar angular momentum *is*

$$J_{Komar} = \frac{1}{16\pi} \lim_{r\to\infty} \int_{S^2_r} \star \mathrm{d}k \,. \tag{6.78}$$

Example 6.3: Komar mass and angular momentum of the Kerr solution

Let us now put our new conserved quantities to the test for the Kerr solution. For the Komar mass the relevant Killing vector is $k = \partial_t$. We first need to lower an index, which gives:

$$k = -\frac{\Delta(r)}{\rho^2}(\mathrm{d}t - a\sin^2\theta \mathrm{d}\phi) + \frac{a\sin^2\theta}{\rho^2}(a\mathrm{d}t - (r^2+a^2)\mathrm{d}\phi) \,. \tag{6.79}$$

We now need to take the exterior derivative, giving

$$\mathrm{d}k = \frac{2M}{\rho^4}\left((r^2 - a^2\cos^2\theta)(\mathrm{d}t - a\sin^2\theta\mathrm{d}\phi)\wedge \mathrm{d}r - 2ar\cos\theta\sin\theta(a\mathrm{d}t - (r^2+a^2)\mathrm{d}\phi)\wedge \mathrm{d}\theta\right) \,. \tag{6.80}$$

Computing the Hodge star we have:

$$\star\mathrm{d}k = -\frac{2M}{\rho^4}\Big(2ar\cos\theta(\mathrm{d}t - a\sin^2\theta\mathrm{d}\phi)\wedge\mathrm{d}r + (r^2-a^2\cos^2\theta)\sin\theta(a\mathrm{d}t-(r^2+a^2)\mathrm{d}\phi)\wedge\mathrm{d}\theta\Big) \,. \tag{6.81}$$

We now want to pull this form back to the two-sphere at infinity, this just means extracting out the piece with legs $\mathrm{d}\theta\wedge\mathrm{d}\phi$. We have

$$\star\mathrm{d}k|_{S^2_r} = -\frac{2M(a^2+r^2)(r^2-a^2\cos^2\theta)\sin\theta}{\rho^4}\mathrm{d}\theta\wedge\mathrm{d}\phi \,. \tag{6.82}$$

We now want to integrate this over the sphere at infinity. We can either try to integrate this brute force and then take the limit or we can take the limit here and then perform the integration. For either method we find:

$$M_{\text{Komar}} = M\,, \tag{6.83}$$

which is of course what we expected.

For the angular momentum we have that the Killing vector is $R = \partial_\phi$. Lowering the index we find:

$$R = -\frac{1}{\rho^2}\Big(a\sin^2\theta\Delta(\mathrm{d}t - a\sin^2\theta\mathrm{d}\phi) - (r^2+a^2)\sin^2\theta(a\mathrm{d}t - (r^2+a^2)\mathrm{d}\phi)\Big)\,. \tag{6.84}$$

Computing the exterior derivative we find:

$$\begin{aligned}\mathrm{d}R = \frac{2}{\rho^4}\Big[&aM\sin^2\theta\left(r^2 - a^2\cos^2\theta\right)\mathrm{d}r\wedge\mathrm{d}t + 2aMr\sin\theta\cos\theta\left(a^2+r^2\right)\mathrm{d}t\wedge\mathrm{d}\theta\\ &+\sin^2\theta\left(a^4\cos^4\theta(r-M) + a^2\cos^2(\theta)\left(a^2M + r^2(M+2r)\right) - a^2Mr^2 + r^5\right)\mathrm{d}r\wedge\mathrm{d}\phi\\ &+\sin\theta\cos\theta\Big((a^2+r^2)^3 + \Delta\big(\rho^4 - (a^2+r^2)^2\big)\Big)\mathrm{d}\theta\wedge\mathrm{d}\phi\Big]\,,\end{aligned} \tag{6.85}$$

and taking the Hodge dual we find:

$$\begin{aligned}\star\mathrm{d}R = \frac{2}{\rho^4}\Big[&\left(2Mr(a^2+r^2) - 2Mr\rho^2 + \rho^4\right)\cos\theta\mathrm{d}t\wedge\mathrm{d}r + 2aMr\sin^4\theta\cos\theta\mathrm{d}r\wedge\mathrm{d}\phi\\ &+(2Mr^2(a^2+r^2) - M(a^2+r^2)\rho^2 + (M-r)\rho^4)\sin\theta^3\mathrm{d}t\wedge\mathrm{d}\theta\\ &+aM(2r^2(a^2+r^2) + (r^2-a^2)\rho^2)\sin\theta^3\mathrm{d}\theta\wedge\mathrm{d}\phi\Big]\,.\end{aligned} \tag{6.86}$$

As before we extract out the term which has legs along the two-sphere and integrate this over the two-sphere before sending $r\to\infty$. The relevant term is

$$\star\mathrm{d}R\Big|_{S_r^2} = \frac{2aM(2r^2(a^2+r^2) + (r^2-a^2)\rho^2)}{\rho^4}\sin\theta^3\mathrm{d}\theta\wedge\mathrm{d}\phi\,. \tag{6.87}$$

Integrating we find:

$$J = \frac{1}{16\pi}\lim\int_{S_r^2}\mathrm{d}R = aM\,. \tag{6.88}$$

This justifies our earlier claim, (6.3), that a is the reduced angular momentum. It also makes more clear the three different regimes for the roots. The slowly rotating case corresponds to the angular momentum satisfying $0 < |J| < 1$, the extremal case is $|J| = 1$ and the rapidly rotating case is $|J| > 1$. Here we should reinstate factors of the speed of light where there is a 1, showing that the extremal case is when the black hole rotates with at the speed of light.

6.5 Ergosphere and Penrose process (or how to steal energy from a black hole)

By definition a black hole is a region of space where no matter nor light can escape from. It may come as a surprise that we can extract energy from a black hole if it has something called an ergosphere. To better understand what an ergosphere is it is instructive to first consider stationary observers. Recall that a stationary observe is someone who remains stationary with respect to infinity. This means that they do not allow the spatial coordinates to vary at all. Since we are considering a black hole, in order for these stationary observers to remain stationary a force must be applied to stop them from falling into the black hole. Given a sufficiently powerful force, a stationary observer can hover arbitrarily close to the Schwarzschild horizon. For a rotating black hole there is a limit to how close a stationary observer can get. The four-velocity of a stationary observer is, by definition, $X^\mu = (\dot{t}, 0, 0, 0)$. The condition for the four-velocity to be timelike and parametrised by proper time is

$$-1 = X^\mu X_\mu = g_{tt}\dot{t}^2 = -\left(1 - \frac{2Mr}{\rho^2}\right)\dot{t}^2\,.$$

Where this condition is satisfied this determines $\dot{t}$. For Schwarzschild this is valid arbitrarily close to the horizon, though if you pass the horizon you can no longer be stationary. In contrast, for the Kerr black hole, on the surface

$$r = r_e(\theta) = M + \sqrt{M^2 - a^2\cos^2\theta}\,, \tag{6.89}$$

$g_{tt} = 0$ and therefore we cannot solve (6.5). Note that this is **not** the horizon, we can still cross this surface and emerge back out to infinity if we want. Worse still, for $r_+ < r < r_e(\theta)$, it is not hard to see that $g_{tt} > 0$ and therefore (6.5) cannot be solved there either! This implies that no stationary observer can exist in this region: however much force you apply.

An alternative, but complimentary way of seeing this, is by studying the norm of the time-like Killing vector at infinity. The norm of the Killing vector $K = \partial_t$ is

$$K^\mu K_\mu = -\frac{1}{\rho^2}(\Delta - a^2\sin^2\theta)\,, \tag{6.90}$$

which we see does not vanish on the horizon, instead on the horizon it is spacelike. This Killing vector is already spacelike at the outer horizon, except at the north and south poles at $\theta = 0, \pi$ where it is null. The locus of points where $K^\mu K_\mu = 0$ is called the *stationary limit surface*[24] and is given by

$$(r - M)^2 = M^2 - a^2\cos^2\theta\,, \tag{6.91}$$

[24]From our earlier discussion the name stationary limit surface makes sense since it is the surface beyond which a stationary observer $r = \theta = \phi =$ const can no longer exist. Passing through the ergosphere the observer must necessarily rotate, and therefore ϕ is no longer constant.

while the outer event horizon is given by

$$(r_+ - M)^2 = M^2 - a^2\,. \tag{6.92}$$

Thus there is a region between these two surfaces, which is called the *ergosphere*, where K is spacelike, see figure 29. Therefore since in the ergosphere ∂_t is not time-like one cannot travel along its integral curves and remain stationary with respect to observers at infinity. A stationary observer is someone whose 4-velocity is parallel to K, since this is spacelike in the ergosphere they cannot be stationary. Recall that in order to be timelike we need to satisfy $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = -1$ inside the ergosphere. However each of the terms of the metric are positive definite inside the ergosphere except the term $g_{t\phi}$, and therefore $\dot{\phi} \neq 0$ and so must rotate. Since $\dot{t} > 0$ for a future directed worldline, we must have $\dot{\phi} > 0$ and therefore the timelike worldline is dragged around in the direction of the rotation of the black hole. This effect is an example of *frame dragging*.

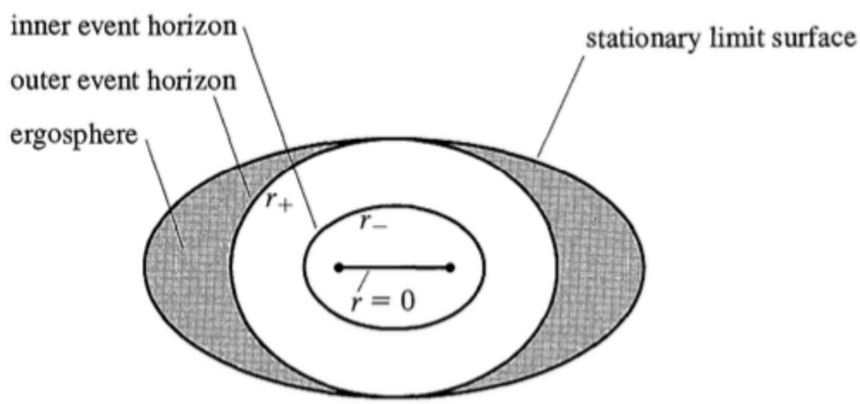

Figure 29: The horizon structure around the Kerr solution. The event horizons are null surfaces that separate points past which it is impossible to return to a certain region of space. The stationary limit surface, is timelike everywhere except where it is tangent to the event horizon at the poles. It represents the place past which it is impossible to be a stationary observer. The ergosphere between the stationary limit surface and the outer event horizon is a region in which it is possible to enter and leave again but not to remain stationary.

We may exploit this to obtain energy from the black hole. Consider a particle with 4-momentum $p^\mu = m\dot{x}^\mu$ with m the rest of the particle. Recall that the existence of Killing vectors implies the existence of conserved quantities along geodesics. We have the two conserved quantities:

$$
\begin{aligned}
E = -K_\mu p^\mu &= m\left[\left(1 - \frac{2Mr}{\rho^2}\right)\frac{\mathrm{d}t}{\mathrm{d}\tau} + \frac{2Mar}{\rho^2}\sin^2\theta\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right], \\
l = R_\mu p^\mu &= m\left[-\frac{2Mar}{\rho^2}\sin^2\theta\frac{\mathrm{d}t}{\mathrm{d}\tau} + \frac{(r^2+a^2)^2 - \Delta(r)a^2\sin^2\theta}{\rho^2}\sin^2\theta\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right].
\end{aligned}
\tag{6.93}
$$

These differ slightly with the definitions before where we had energy and angular momentum per unit mass, here we have multiplied by the mass of the particle. They are of course still conserved. The minus sign in E is because at infinity both K and p are timelike and so their inner product is negative and we want energy to be positive.

Let the particle approach a Kerr black hole along a geodesic. The energy of the particle according to a stationary observer at infinity is conserved along the geodesic. Inside the ergosphere, since K becomes spacelike we can imagine particles for which

$$E = -K_\mu p^\mu < 0\,. \tag{6.94}$$

This may bother you slightly that there is a particle with negative energy however, one can find that all particles have positive energy outside the ergosphere, those with negative energy must remain in the ergosphere or be accelerated until its energy is positive if it is to escape.

This allows for a way of extracting energy from a rotating black hole. Let us start away far from the black hole and throw something into the black hole along a geodesic. Let us denote the 4-momentum to be p_0, then its total energy that we measure is

$$E_0 = -p_0^\mu K_\mu\,, \tag{6.95}$$

which is conserved. Let the object enter the ergosphere. We arrange for the object to eject a mass, in a smart way, whilst in the ergosphere. Conservation of momentum gives

$$p_0 = p_1 + p_2\,, \tag{6.96}$$

with p_1 the momentum of the object and p_2 the momentum of the ejected mass. Contracting with the Killing vector K we have the expected relation

$$E_0 = E_1 + E_2\,. \tag{6.97}$$

If we arrange for $E_2 < 0$ by a clever choice of way of ejecting the mass, then we must have $E_1 > E_0$. Penrose showed that the ejected mass with negative energy must fall into the black hole, while the object can now escape with more energy than it initially began with. This is the *Penrose process* and is a method for extracting energy from a rotating black hole. The energy imparted to the particle that escapes to infinity is drawn from the rotational energy of the black hole.

So can a rotating black hole be used as an infinite source of energy? There is no such thing as a free lunch (though cafe pi occasionally has free lunch samples), so the energy must comes from somewhere, and the only candidate is that it comes from the black hole. The Penrose process extracts energy from the black hole by decreasing the black holes angular momentum. When the mass is ejected we need to it to be ejected with opposite direction rotation compared to the black hole's rotation. Recall that we saw that the event horizon was a Killing horizon with Killing vector

$$\xi^\mu = K^\mu + \Omega_+ R^\mu \,. \tag{6.98}$$

On the outer event horizon this indeed becomes null and is future directed and tangent to the generators of the horizon. The statement that the object with momentum p_2 crosses the event horizon by moving forward in time, is simply that

$$p_2^\mu \xi_\mu < 0 \,. \tag{6.99}$$

Plugging in the definitions of E and l, we see that this is equivalent to

$$l_2 < \frac{E_2}{\Omega_+} \,. \tag{6.100}$$

Since E_2 is negative and Ω_+ positive it follows that $l_2 < 0$ and therefore the particle has negative angular momentum, it is moving against the rotation of the black hole.

Once our object has escaped the ergosphere and the mass has fallen inside the event horizon the mass and the angular momentum of the black hole are changed. They are now the initial values plus the negative contributions from the in-falling mass:

$$\delta M = E_2 \,, \qquad \delta J = l_2 \,, \tag{6.101}$$

with $J = Ma$ the angular momentum of the black hole. The inequality (6.100) then translates into a limit on the amount the angular momentum can decrease

$$\delta J < \delta M \Omega_+^{-1} \,. \tag{6.102}$$

The ideal process would be when we have equality, in this case the mass thrown into the black hole becomes more and more null (since in this limit we have $p_2^\mu \xi_\mu \to 0$).

There is now a slight curiosity that appears, we can use the Penrose process to reduce the mass of the black hole, however there is a non-decreasing quantity: the area of the horizon. Let us compute the area of the event horizon at $r = r_+$. To do this we look at the induced metric on the horizon by setting t =const $r = r_+$. The induced metric is

$$\begin{aligned} ds^2(\text{horizon}) = \gamma_{ij} dx^i dx^j = ds^2(dt = 0\,, dr = 0\,, r = r_+) \\ \frac{(r_+^2 + a^2)^2}{r_+^2 + a^2\cos^2\theta}\sin^2\theta d\phi^2 + (r_+^2 + a^2\cos^2\theta)d\theta^2\,, \end{aligned} \tag{6.103}$$

The area of the horizon is then simply

$$A = \int \text{dvol(horizon)}\,. \tag{6.104}$$

For the metric at hand the determinant is

$$\begin{aligned} \det(\gamma) &= \frac{(r_+^2 + a^2)^2}{r_+^2 + a^2\cos^2\theta}\sin^2\theta \times (r_+^2 + a^2\cos^2\theta) = (r_+^2 + a^2)^2\sin^2\theta\,, \\ \text{dvol(horizon)} &= (r_+^2 + a^2)\sin\theta d\theta \wedge d\phi\,. \end{aligned} \tag{6.105}$$

The area is then

$$A_{\text{horizon}} = (r_+^2 + a^2)\int \sin\theta d\phi d\theta = 4\pi(r_+^2 + a^2)\,. \tag{6.106}$$

To show that this does not decrease we work with the so called *irreducible mass* defined by

$$M^2_{\text{irreducible}} = \frac{A}{16\pi}\,. \tag{6.107}$$

Then we have

$$\begin{aligned} M^2_{\text{irreducible}} &= \frac{r_+^2 + a^2}{4} \\ &= \frac{1}{2}\left(M^2 + \sqrt{M^4 - M^2a^2}\right) \\ &= \frac{1}{2}\left(M^2 + \sqrt{M^4 - J^2}\right)\,. \end{aligned} \tag{6.108}$$

We can differentiate to obtain how $M_{\text{irreducible}}$ is affected by changes in the mass or angular momentum:

$$\delta M_{\text{irreducible}} = \frac{a}{4M_{\text{irreducible}}\sqrt{M^2 - a^2}}(\Omega_H^{-1}\delta M - \delta J)\,. \tag{6.109}$$

We see that the inequality (6.102) becomes

$$\delta M_{\text{irreducible}} > 0\,. \tag{6.110}$$

The irreducible mass can never be reduced, hence the name. It follows that the maximum amount of energy that can be extracted from the black hole is

$$\max(E) = M - M_{\text{irreducible}} = M - \frac{1}{\sqrt{2}}\sqrt{M^2 + \sqrt{M^4 - J^2}}\,. \tag{6.111}$$

The result after a complete extraction of this amount of energy is a Schwarzschild solution with mass $M_{\text{irreducible}}$. The most efficient process is to start with an extremal Kerr black hole and then one can extract out approximately 29% of its total energy.

The irreducibility of $M_{\text{irreducible}}$ immediately shows that the surface area is non-decreasing. We have

$$\delta A = \frac{8\pi a}{\Omega_H \sqrt{M^2 - a^2}}\big(\delta M - \Omega_H \delta J\big)\,. \tag{6.112}$$

This may be recast as

$$\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J\,, \tag{6.113}$$

where κ is

$$\kappa = \frac{\sqrt{M^2 - a^2}}{2M(M + \sqrt{M^2 - a^2})}\,. \tag{6.114}$$

The quantity κ is the surface gravity of the Kerr solution. This is the force that an observer at infinity would have to exert in order to keep a unit mass at the horizon.

Recall that for every Killing horizon we can associate a quantity called the surface gravity. Given the Killing horizon we have an associated Killing vector, ξ which is null on the horizon. Since ξ is a normal vector to the Killing horizon it obeys the geodesic equation

$$\xi^\mu \nabla_\mu \xi^\nu = \kappa \xi^\nu\,. \tag{6.115}$$

It turns out that κ is constant over a Killing horizon.

Equation (6.113) first started people thinking about a correspondence between the laws of thermodynamics and black holes. The first law of thermodynamics is

$$\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V\,, \tag{6.116}$$

where T is the temperature, S the entropy, p the pressure and V the volume, thus $p\mathrm{d}V$ is the work done on the system. It is then natural to think of the term $\Omega_H \delta J$ as the work we do on the black hole by throwing our mass into the black hole. It is then natural to construct the dictionary

$$E \leftrightarrow M\,, \quad S \leftrightarrow \frac{A}{4G_N}\,, \quad T \leftrightarrow \frac{\kappa}{2\pi}\,. \tag{6.117}$$

This observation leads nicely on towards studying black hole thermodynamics.

7 Laws of black hole thermodynamics

In 1973 Bardeen, Carter and Hawking (BCH) wrote a paper, [15], in which they considered

The Four Laws of Black Hole Mechanics

J. M. Bardeen*

Department of Physics, Yale University, New Haven, Connecticut, USA

B. Carter and S. W. Hawking

Institute of Astronomy, University of Cambridge, England

Received January 24, 1973

Abstract. Expressions are derived for the mass of a stationary axisymmetric solution of the Einstein equations containing a black hole surrounded by matter and for the difference in mass between two neighboring such solutions. Two of the quantities which appear in these expressions, namely the area A of the event horizon and the "surface gravity" κ of the black hole, have a close analogy with entropy and temperature respectively. This analogy suggests the formulation of four laws of black hole mechanics which correspond to and in some ways transcend the four laws of thermodynamics.

Figure 30: The BCH paper.

stationary, axisymmetric black holes. They found that black holes obeyed laws reminiscent of the laws of thermodynamics. At the time they thought it was just an analogy. There seem to be some glaring flaws in this analogy: since nothing can escape from a black hole the temperature must vanish, secondly, the entropy is dimensionless whereas the horizon area is a length squared, a final perceived flaw is that the area of every black hole is separately non-decreasing, whereas only the total entropy is non-decreasing in thermodynamics. The resolution to all these flaws lies in a theory of quantum gravity which GR is not. Recall that going to a quantum theory was also the resolution for apparent paradoxes in thermodynamics, for example black body radiation. We will not study quantum gravity, this is an active area of research and even after decades of research we do not know what the correct theory is, though that is not to say that progress has not been made. Instead, we will present the classical laws of black hole thermodynamics and a little semi-classical analysis.

7.1 The four laws of (black hole) themrodynamics

There are four laws of black hole thermodynamics which should be contrasted with the laws of thermodynamics, see table 1.

Law	Thermodynamics	Black holes
0^{th}	The temperature T is constant throughout a system in thermal equilibrium.	The surface gravity κ is constant over the event horizon of a stationary black hole.
1^{st}	$\mathrm{d}E = T\mathrm{d}S + \sum_i \mu_i \mathrm{d}N_i$	$\mathrm{d}M = \frac{1}{8\pi}\kappa \mathrm{d}A + \Omega_H \mathrm{d}J + \Phi_H \mathrm{d}Q$
2^{nd}	$\mathrm{d}S \geq 0$	$\mathrm{d}A \geq 0$
3^{rd}	T cannot be reduced to zero by a finite number of operations.	κ cannot be reduced to zero by a finite number of operations.

Table 1: The four laws of black hole thermodynamics and the four laws of thermodynamics.

It may seem strange to say that a black hole has a temperature since nothing can escape from a black hole and therefore they cannot radiate. This would also mean that they cannot have a physical entropy. Once quantum effects are taken into account it turns out that a black hole can have a temperature. Moreover, as pointed out by Jacob Bekenstein the second law of thermodynamics would be violated if black holes did not have an entropy. One could throw in arbitrary objects into the black hole which have a large entropy and thus lower the entropy of the exterior universe. In order to save the 2nd law of thermodynamics it is essential for a black hole to have an entropy and moreover it must be proportional to the surface area of the horizon. Bekenstein's generalised second law states that

$$\mathrm{d}S_{\text{total}} = \mathrm{d}\big(S_{\text{external}} + S_{\text{BH}}\big) \geq 0. \tag{7.1}$$

In 1974 Hawking announced that black holes are hot (and people studying them even more) and radiate just like any hot body with a temperature

$$T_H = \frac{\hbar\kappa}{2\pi c k_B}\,, \tag{7.2}$$

from which it follows that a black holes has an entropy, which was later shown to be given by

$$S_{\text{BH}} = \frac{Ac^3}{4G_N\hbar}\,. \tag{7.3}$$

which is known as the *Bekenstein–Hawking* entropy.

In the remainder of the course our goal is to understand the laws of black hole thermodynamics as presented above. We will not present proofs for each of the laws, indeed not all have proofs, just overwhelming evidence.

7.2 Zeroth law of black hole mechanics

Proposition

Consider a null geodesic congruence that contains the generators of a Killing horizon $\mathcal{N}$. Then $\theta = \sigma = \omega = 0$ on $\mathcal{N}$.

Proof: We have already seen that $\omega = 0$ since the generators are hypersurface orthogonal. Let ξ be a Killing vector field normal to $\mathcal{N}$. On $\mathcal{N}$ we can write $\xi^\mu = hU^\mu$ where U^μ is tangent to the affinely parametrised generators of $\mathcal{N}$ and h is a function on $\mathcal{N}$. Le $\mathcal{N}$ be specified by an equation $f = 0$. Then we can write $U^\mu = h^- \xi^\mu + fV^\mu$ where V^μ is a smooth vector field. We can then calculate

$$B_{\mu\nu} = \nabla_\nu U_\mu = (\partial_\nu h^{-1})\xi_\mu + h^{-1}\nabla_\nu \xi_\mu + \partial_\nu f V_\mu + f\nabla_\nu V_\mu \,, \tag{7.4}$$

evaluating on $\mathcal{N}$ and using Killing's equation gives

$$B_{(\mu\nu)}\Big|_{\mathcal{N}} = \Big[\xi_{(\mu}\partial_{\nu)}h^{-1} + V_{(\mu}\partial_{\nu)}f\Big]\Big|_{\mathcal{N}} . \tag{7.5}$$

Since both ξ_μ and $\partial_\mu f$ are parallel to U_μ on $\mathcal{N}$ when we project onto $T_\perp$ both terms are eliminated and we have

$$\hat{B}_{\mu\nu}\Big|_{\mathcal{N}} = 0 \tag{7.6}$$

and thus $\theta = \sigma = 0$ on $\mathcal{N}$.

Theorem 6 *Zeroth law of black hole mechanics*
The surface gravity κ is constant on the future event horizon of a stationary black hole spacetime obeying the dominant energy condition.

Proof: Using Hawking's theorem we have that $\mathcal{H}^+$ is a Killing horizon with respect to some Killing vector ξ. We know that $\theta = 0$ along the generators of $\mathcal{H}^+$, and therefore $\frac{\mathrm{d}\theta}{\mathrm{d}\lambda} = 0$ along these generators. Moreover, we have just seen that on $\mathcal{H}^+$ we have $\sigma = \omega = 0$. Therefore Raychaudhuri's equation gives

$$0 = R_{\mu\nu}\xi^\mu\xi^\nu\Big|_{\mathcal{H}^+} = 8\pi\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^\rho{}_\rho\right)\xi^\mu\xi^\nu\Big|_{\mathcal{H}^+} = 8\pi T_{\mu\nu}\xi^\mu\xi^\nu\Big|_{\mathcal{H}^+} \tag{7.7}$$

where we have used Einstein's equation and that ξ is null on $\mathcal{H}^+$. This implies

$$J_\mu\xi^\mu\Big|_{\mathcal{H}^+} = 0\,, \qquad \text{where} \quad J_\mu \equiv -T_{\mu\nu}\xi^\nu \,. \tag{7.8}$$

Since ξ is a future-directed causal vector field, then by the dominant energy condition, so is J_μ (unless it is zero). Thus, J^μ is parallel to ξ^μ on $\mathcal{H}^+$ and consequently

$$0 = \xi_{[\mu}J_{\nu]}\Big|_{\mathcal{H}^+} = -\xi_{[\mu}T_{\nu]\rho}\xi^\rho\Big|_{\mathcal{H}^+} = -\frac{1}{8\pi}\xi_{[\mu}R_{\nu]\rho}\xi^\rho\Big|_{\mathcal{H}^+}\,, \tag{7.9}$$

where we have used Einstein's equation in the final step. One problem sheet 4 you are asked to show that this is equivalent to

$$0 = \frac{1}{8\pi}\xi_{[\mu}\partial_{\nu]}\kappa \,. \tag{7.10}$$

Therefore $\partial_\nu \kappa$ is proportional to ξ_ν and therefore for any vector field $\tilde{T}$ tangent to $\mathcal{H}^+$ it follows that $\tilde{T}^\mu \partial_\mu \kappa = 0$. Therefore κ is constant on $\mathcal{H}^+$ provided $\mathcal{H}^+$ is connected.

Let us identify what we need for proving that the surface gravity is constant on the horizon. We must be very careful with the formulae we use for the surface gravity and acting on them with derivatives since some only hold on the horizon. Since $\xi^2\big|_{\mathcal{H}^+} = 0$ we have that $\nabla_\mu\big(\xi^2\big)$ is normal to the horizon and therefore there is a function κ on the horizon such that

$$\nabla_\mu(\xi^2) = -2\kappa\xi_\mu\,. \tag{7.11}$$

We may rewrite this as

$$\xi^\mu \nabla_\nu \xi_\mu = -\xi^\mu \nabla_\mu \xi_\nu = -\kappa \xi_\nu\,, \tag{7.12}$$

which is just the geodesic equation in a non-affine parametrisation. The above derivation of the expression for the surface gravity makes clear that it holds on the Killing surface. This means that applying derivatives to the above expression is somewhat subtle, we can only differentiate on the Killing surface and not normal to it. Instead observe that if $\epsilon_{\mu\nu\rho\sigma}$ is the 4d volume element then $\epsilon^{\mu\nu\rho\sigma}\xi_\sigma$ is tangent to the horizon since $\epsilon^{\mu\nu\rho\sigma}\xi_\sigma\xi_\rho = 0$. Therefore we may use this to project the differential operator onto the horizon by acting with $\epsilon^{\mu\nu\rho\sigma}\xi_\rho\nabla_\sigma$ and then this may be applied to any object defined on the horizon. Equivalently we may act with $\xi_{[\mu}\nabla_{\nu]}$ on any object. Now applying this to (7.12) we obtain

$$\begin{aligned}
\xi_{[\rho}\nabla_{\sigma]}(\kappa\xi_\nu) &= \xi_\nu \xi_{[\rho}\nabla_{\sigma]}\kappa + \kappa\xi_{[\rho}\nabla_{\sigma]}\xi_\nu \\
&= \xi_{[\rho}\nabla_{\sigma]}\big(\xi^\mu\nabla_\mu\xi_\nu\big) \\
&= \big(\xi_{[\rho}\nabla_{\sigma]}\xi^\mu\big)\big(\nabla_\mu\xi_\nu\big) + \xi^\mu\xi_{[\rho}\nabla_{\sigma]}\nabla_\mu\xi_\nu \\
&= \big(\xi_{[\rho}\nabla_{\sigma]}\xi^\mu\big)\big(\nabla_\mu\xi_\nu\big) + \xi^\mu R_{\mu\nu[\rho}{}^\tau\xi_{\sigma]}\xi_\tau
\end{aligned} \tag{7.13}$$

We may simplify the first term by using the condition that ξ is hypersurface orthogonal and hence satisfies $\xi_{[\mu}\nabla_\nu\xi_{\rho]} = 0$. We find

$$\begin{aligned}
\big(\xi_{[\rho}\nabla_{\sigma]}\xi^\mu\big)\big(\nabla_\mu\xi_\nu\big) &= -\frac{1}{2}\big(\xi^\mu\nabla_\rho\xi_\sigma\big)\nabla_\mu\xi_\nu \\
&= -\frac{1}{2}\kappa\xi_\nu\nabla_\rho\xi_\sigma \\
&= \kappa\xi_{[\rho}\nabla_{\sigma]}\xi_\nu
\end{aligned} \tag{7.14}$$

This cancels the second term of the first row of (7.14). We therefore have

$$\xi_\nu\xi_{[\rho}\nabla_{\sigma]}\kappa = \xi^\mu R_{\nu\mu[\sigma}{}^\tau\xi_{\rho]}\xi_\tau \tag{7.15}$$

Since ξ is hypersurface orthogonal we have

$$\xi_\rho \nabla_\mu \xi_\nu = -2\xi_{[\mu}\nabla_{\nu]}\xi_\rho \,, \tag{7.16}$$

and acting on this with with $\xi_{[\sigma}\nabla_{\tau]}$ we obtain

$$\big(\xi_{[\sigma}\nabla_{\tau]}\xi_\rho\big)\nabla_\mu\xi_\nu + \xi_\rho\xi_{[\sigma}\nabla_{\tau]}\nabla_\mu\xi_\nu = -2\big(\xi_{[\sigma}\nabla_{\tau]}\xi_{[\mu}\big)\nabla_{\nu]}\xi_\rho - 2\big(\xi_{[\sigma}\nabla_{\tau]}\nabla_{[\nu}\xi_{|\rho|}\big)\xi_{\mu]} \,. \tag{7.17}$$

Application of (7.16) results in

$$-\xi_\rho R_{\mu\nu[\tau}{}^{\lambda}\xi_{\sigma]}\xi_\lambda = 2\xi_{[\mu}R_{\nu]\rho\sigma}{}^{\lambda}\xi^\rho\xi_\lambda \,. \tag{7.18}$$

Contracting over the ρ and τ indices gives

$$-\xi_{[\mu}R_{\nu]}{}^{\lambda}\xi_\lambda\xi_\sigma = \xi_{[\mu}R_{\nu]\rho\sigma}{}^{\lambda}\xi^\rho\xi_\lambda \,, \tag{7.19}$$

with the right-hand-side being the expression we required above. We therefore find

$$\xi_{[\mu}\nabla_{\nu]}\kappa = -\xi_{[\mu}R_{\nu]}{}^{\rho}\xi_\rho \,. \tag{7.20}$$

Plugging this into the formulae above gives the required result. You are asked to perform these steps in problem sheet 4.

7.3 First law

We have already seen a form of the first law when we considered the irreducible mass of the Kerr solution. We will give a somewhat heuristic argument here of the first law and then check it in more detail for the black holes we have studied previously. Consider the Killing vector associated to the Killing horizon, it takes the form $\xi = K + \Omega_H R$ where K generates time translations and R generates the axisymmetry. The corresponding charge is a combination of the mass and the angular momentum:

$$Q_\xi = -\frac{1}{8\pi}\int_{S^2_\infty}\star\mathrm{d}\xi = -\frac{1}{8\pi}\int_{S^2_\infty}\star\mathrm{d}K - \frac{\Omega_H}{8\pi}\int_{S^2_\infty}\star\mathrm{d}R = M - 2\Omega_H J \,. \tag{7.21}$$

We can also evaluate Q_ξ in another way. Let Σ be a spacelike hypersurface intersecting the horizon $\mathcal{H}^+$ on a two-sphere S^2_H which together with the two-sphere S^2_∞ at spatial infinity forms the boundary of Σ. Using Stoke's theorem we have:

$$\begin{aligned} Q_\xi &= -\frac{1}{8\pi}\int_{S^2_H}\star\mathrm{d}\xi - \frac{1}{8\pi}\int_\Sigma \mathrm{d}\star\mathrm{d}\xi \\ &= -\frac{1}{8\pi}\int_{S^2_H}\star\mathrm{d}\xi + 2\int_\Sigma\Big(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^\rho{}_\rho\Big)\xi^\nu\star\mathrm{d}x^\mu \,, \end{aligned} \tag{7.22}$$

where in the last step we used

$$\star \mathrm{d} \star \mathrm{d}\xi = 8\pi J\,, \qquad J = 2\Big(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^{\rho}{}_{\rho}\Big)\xi^{\nu}\mathrm{d}x^{\nu} \tag{7.23}$$

The integral over S^2_H may be regarded as the contribution from the black hole while the one over Σ is a combination of the mass and angular momentum of the matter and radiation outside the horizon. In order to treat the integral over S^2_H we observe that the volume form on S^2_H, can be written as

$$\mathrm{dvol}(S^2_H) = \star(n \wedge \xi)\,, \tag{7.24}$$

evaluated at the horizon. Here n^{μ} is another null vector normal to S^2_H, normalised so that $n^{\mu}\xi_{\mu} = -1$. Therefore

$$\begin{aligned}
\int_{S^2_H} \star\mathrm{d}\xi &= \int_{S^2_H} \mathrm{dvol}(S^2_H)\Big(\star\big(n\wedge\xi\big)\Big)^{\mu\nu}(\star\mathrm{d}\xi)_{\mu\nu} \\
&= 2\int_{S^2_H} \mathrm{dvol}(S^2_H) n^{\nu}\xi^{\mu}\nabla_{\mu}\xi_{\nu} \\
&= -2\kappa\int_{S^2_H}\mathrm{dvol}(S^2_H) \\
&= -2\kappa A_H\,.
\end{aligned} \tag{7.25}$$

Plugging this into (7.22) we arrive at

$$M = \frac{\kappa A_H}{4\pi} + 2\Omega_H J + 2\int_{\Sigma}\Big(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^{\rho}{}_{\rho}\Big)\xi^{\nu}\star\mathrm{d}x^{\mu} \tag{7.26}$$

If we are in pure GR, then $T_{\mu\nu} = 0$ and our spacetime is the Kerr black hole and the formula reads

$$M = \frac{\kappa A}{4\pi} + 2\Omega_H J\,. \tag{7.27}$$

This is Smarr's formula for the mass of a Kerr black hole. A formula for δM in the vacuum case can be obtained by varying (7.27)

$$\delta M = \frac{1}{4\pi}\big(A_H\delta\kappa + \kappa\delta A_H\big) + 2\big(J\delta\Omega_H + \Omega_H\delta J\big)\,. \tag{7.28}$$

An alternative computation gives

$$\delta M = -\frac{1}{4\pi}A_H\delta\kappa - 2J\delta\Omega_H\,. \tag{7.29}$$

Adding the two equations gives

$$\delta M = \frac{1}{8\pi}\kappa\delta A_H + \Omega_H\delta J\,. \tag{7.30}$$

In the case where there is an electric charge, we need to define the electric potential

$$\Phi_H = \xi^\mu A_\mu\Big|_{\mathcal{H}^+} - \xi^\mu A_\mu\Big|_{\infty} . \tag{7.31}$$

For asymptotically flat spacetimes we have that $A_\mu \to 0$ as we tend to ∞ and so the second term drops out. The 1st law with electric charge is then

$$\delta M = \frac{1}{8\pi}\kappa\delta A_H + \Omega_H \delta J + \Phi_H \delta Q \tag{7.32}$$

Aside

A cute way of seeing that this must be true in GR in a vacuum is to use the uniqueness theorems for the Kerr solution which say that $M = M(A, J)$ (in the absence of charge). Now in the units we are using both A and J have dimensions of M^2 so that the function $M(A, J)$ must be homogeneous of degree $1/2$. The Euler theorem of homogeneous functions then implies that

$$\begin{aligned} A\frac{\partial M}{\partial A} + J\frac{\partial M}{\partial J} &= \frac{1}{2}M \\ &= \frac{\kappa}{8\pi}A + \Omega_H J \,, \end{aligned} \tag{7.33}$$

where in the second line Smarr's formula has been used. Rearranging one finds

$$A\left(\frac{\partial M}{\partial A} - \frac{\kappa}{8\pi}\right) + J\left(\frac{\partial M}{\partial J} - \Omega_+\right) = 0 \,. \tag{7.34}$$

Since A and J are free parameters, we can tune then however we want we must have:

$$\frac{\partial M}{\partial A} = \frac{\kappa}{8\pi} \,, \qquad \frac{\partial M}{\partial J} = \Omega_+ \,, \tag{7.35}$$

which proves the statement.

Example 7.1: Kerr–Newman

Let us check this for the electrically charged Kerr–Newman solution:

$$\begin{aligned} \mathrm{d}s^2 &= -\frac{\Delta(r) - a^2\sin^2\theta}{\rho(r,\theta)^2}\mathrm{d}t^2 - \frac{2a\sin^2\theta(r^2+a^2-\Delta(r))}{\rho(r,\theta)^2}\mathrm{d}t\mathrm{d}\phi \\ &\quad + \frac{(r^2+a^2)^2 - a^2\sin^2\theta\Delta(r)}{\rho(r,\theta)^2}\sin^2\theta\mathrm{d}\phi^2 + \frac{\rho(r,\theta)^2}{\Delta(r)}\mathrm{d}r^2 + \rho^2(r,\theta)\mathrm{d}\theta^2 \,, \\ A &= -\frac{1}{\rho(r,\theta)^2}\Big(Qr(\mathrm{d}t - a\sin^2\theta\mathrm{d}\phi)\Big) \,. \end{aligned} \tag{7.36}$$

The functions are

$$\rho(r,\theta)^2 = r^2 + a^2\cos^2\theta \,, \quad \Delta(r) = r^2 - 2Mr + a^2 + e^2 \,, \quad e^2 = Q^2 + P^2 \,. \tag{7.37}$$

Recall that the Kerr–Newman solution is the unique stationary black hole solution of the Einstein–Maxwell theory. Let us compute the quantities that we will need to check the relation. The outer Killing horizon is at $r = r_+$ with

$$r_\pm = M \pm \sqrt{M^2 - a^2 - Q^2} \,. \tag{7.38}$$

First let us consider the surface area of the horizon. We fix an arbitrary time $t = t_0$ and look at the induced metric on the intersection $t = t_0$ and $r = r_+$, we find

$$\mathrm{d}s^2(H) = \gamma_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \rho(r_+,\theta)^2 \mathrm{d}\theta^2 + \frac{r_+^2 + a^2}{\rho(r_+,\theta)^2}\sin^2\theta \mathrm{d}\phi^2 \,. \tag{7.39}$$

The volume form is

$$\mathrm{dvol}(\gamma) = (r_+^2 + a^2)\sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi \,, \tag{7.40}$$

and so the surface area is

$$A_H = \int_{\mathcal{H}^+} \mathrm{dvol}(\gamma) = \int_0^{2\pi} \mathrm{d}\phi \int_0^\pi \mathrm{d}\theta (r_+^2 + a^2)\sin\theta = 4\pi(r_+^2 + a^2) \,. \tag{7.41}$$

Next let us consider the surface gravity. We first need to find the Killing vector which is null on the horizon and then use this to compute the surface gravity. Since the horizon is a Killing horizon we know that it must be of the form

$$\xi = K + \Omega_+ R \,, \tag{7.42}$$

where K and R are the generators of time translations and the axis symmetry respectively. Note that since a Killing vector remains a Killing vector under a constant rescaling there is an arbitrariness in how we pick such a Killing vector. We normalise such that K has coefficient 1, this is so that at asymptotic infinity the norm of the Killing vector is -1 in accordance with K being a timelike Killing vector. Now this needs to have zero norm on the horizon. The norm is

$$\begin{aligned} \xi^2\Big|_{\mathcal{N}^+} &= \frac{a^2\sin^2\theta}{r_+^2 + a^2\cos^2\theta} - \frac{2a\sin^2\theta(r_+^2+a^2)}{r_+^2 + a^2\cos^2\theta}\Omega_H + \frac{(r_+^2+a^2)^2}{r_+^2 + a^2\cos^2\theta}\sin^2\theta\Omega_H^2 \\ &= \frac{\sin^2\theta}{r_+^2 + a^2\cos^2\theta}\Big(a^2 - 2a(r_+^2+a^2)\Omega_H + (r_+^2+a^2)^2\Omega_H^2\Big) \,, \end{aligned} \tag{7.43}$$

and for this to vanish we need

$$\Omega_+ = \frac{a}{r_+^2 + a^2} \,. \tag{7.44}$$

This is the angular velocity of the black hole.

We can now try to compute the surface gravity. In order to use the formula

$$\nabla_\mu(\xi^2) = -2\kappa\xi_\mu \,. \tag{7.45}$$

we need to use coordinates in which the horizon is not a coordinate singularity. Rather than changing coordinates we will instead use an alternative formula for the surface gravity

$$\kappa^2 = \lim_{r\to r_+} \frac{g^{\mu\nu}\partial_\nu(\xi^2)\partial_\mu(\xi^2)}{4\xi^2} \,. \tag{7.46}$$

After a slightly painful computation we find

$$\kappa = \frac{r_+ - r_-}{2(r_+^2 + a^2)} \,. \tag{7.47}$$

Consider the electric potential. We have

$$\begin{aligned} \Phi_H = \xi^\mu A_\mu\Big|_{\mathcal{H}_+} &= \frac{Qr_+}{r_+^2 + a^2\cos^2\theta}(1 - \Omega_+ a\sin^2\theta) \\ &= \frac{Qr_+}{r_+^2 + a^2} \,. \end{aligned} \tag{7.48}$$

Finally let us remember that the electric charge is Q and the angular momentum is $J = aM$. Putting everything together we have

$$\begin{aligned} A_H &= 4\pi\Big(\big(M + \sqrt{M^2 - a^2 - Q^2}\big)^2 + a^2\Big) \\ &= 4\pi\Big(2M^2 - Q^2 + 2M\sqrt{M^2 - Q^2 - a^2}\Big) \,. \end{aligned} \tag{7.49}$$

Since M, Q and J are independent parameters this implies

$$\delta A = \frac{\partial A}{\partial M}\delta M + \frac{\partial A}{\partial Q}\delta Q + \frac{\partial A}{\partial J}\delta J \,. \tag{7.50}$$

After some explicit computation (which you will do in problem sheet 4) and a little rearranging we find

$$\delta M = \frac{1}{8\pi}\kappa\delta A + \Omega_H\delta J + \Phi_H\delta Q \,. \tag{7.51}$$

We see that the proof is deceptively simple, all the hard work goes into proving the uniqueness theorems. You need to know that the black hole settles down to another Kerr–Newman black hole and not some other spacetime. It is worth noting that there exist proofs of the first law known as physical process proofs that do not assume this.

7.4 Second law

The second law states that in any physical process the *area of the event horizon can never decrease*. This is a very surprising feature of these complicated nonlinear PDEs which Hawking proved using just the Einstein equation, the weak energy condition and cosmic censorship.

Let us give a sketch of the proof. Consider the congruence of the horizon and take a cross sectional area A_H at some value of the affine parameter λ along the geodesics. Then the expansion θ satisfies

$$\frac{\mathrm{d}A_H}{\mathrm{d}\lambda} = \theta A_H \,. \tag{7.52}$$

If we imagine the theorem is violated so that the area decreases then we must have $\theta < 0$ somewhere on the event horizon. Since the generators are geodesics the evolution of the expansion is governed by Raychaudhuri's equation. Recall that if $\theta < 0$ and the null energy condition is satisfied then $\theta \to -\infty$ in finite λ. This causes a caustic, see figure 31. Since the points p and q are timelike separated, this contradicts the assumption that the null curves are the generators of an event horizon, as no two points on the event horizon can be timelike separated. Thus by contradiction the cross sectional area of an event horizon cannot decrease. Note that the proof assumes Einstein's equations, they are not used in an essential way.

Example 7.2: Splitting black holes

The second law has some profound implications for what physical processes can occur. Consider a Schwarzschild black hole of mass M. We could ask whether a black hole can split into two black holes of smaller mass? The second law forbids this.

Let the masses of the new black holes be m_1 and m_2. Conservation of energy implies $M = m_1 + m_2$. The surface area of a Schwarzschild black hole is $A = 4\pi M^2$. We have that the entropy of the final state is $A_f = A_1 + A_2 = 4\pi(m_1^2 + m_2^2)$ and the entropy of the initial state is $A_i = 4\pi M^2 = 4\pi(m_1 + m_2)^2 = 4\pi(m_1^2 + m_2^2 + 2m_1m_2)$. It is clear that $A_i > A_f$ and therefore this process violates the second law. Black holes cannot split in two!

Example 7.3: Merging black holes

Consider the opposite scenario: two Schwarzschild black holes merging. After the merger they will form another Schwarzschild black hole but may also emit gravitational waves. Is there an upper-bound on the amount of gravitational waves that can be emitted? The 2nd law provides such a bound.

Let the mass of the two initial black holes be M_1 and M_2 and the final black hole have mass M_3. Let the energy of the radiated gravitational waves be m. Then conservation of

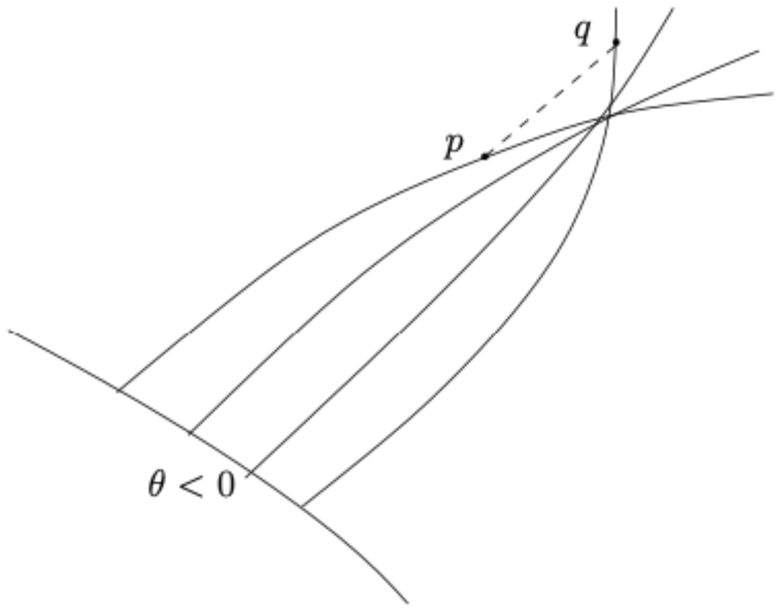

Figure 31: A family of null geodesics with $\theta < 0$ initially will form a caustic; the dotted curve connecting p and q lies within the local light cone, so these points are timelike separated.

energy gives:

$$M_1 + M_2 = M_3 + m. \tag{7.53}$$

The second law implies that

$$M_3^2 \geq M_1^2 + M_2^2\,. \tag{7.54}$$

Therefore the ratio between the energy of the emitted radiation and the initial energy is

$$\frac{m}{M_1 + M_2} = \frac{M_1 + M_2 - M_3}{M_1 + M_2} = 1 - \frac{M_3}{M_1 + M_2}\,. \tag{7.55}$$

The second law allows us to eliminate M_3 for an inequality, thus

$$\frac{m}{M_1 + M_2} = 1 - \frac{M_3}{M_1 + M_2} \leq 1 - \frac{\sqrt{M_1^2 + M_2^2}}{M_1 + M_2}\,. \tag{7.56}$$

This is maximised when the entropy does not change, that is

$$M_3^2 = M_1^2 + M_2^2\,, \tag{7.57}$$

and therefore the maximum possible amount radiated is

$$\frac{m}{M_1+M_2} = 1 - \frac{\sqrt{M_1^2+M_2^2}}{M_1+M_2}\,. \tag{7.58}$$

We can now extremise this with respect to the two masses to see what the optimal mass distribution is to produce the most efficient conversion of energy to gravitational waves. Given the symmetry it is not hard to see that this is precisely $M_1 = M_2$. The most efficient process occurs for colliding black holes of the same mass and the ratio becomes

$$\left.\frac{m}{M_1+M_2}\right|_{\max} = 1 - \frac{1}{\sqrt{2}} \sim 0.29. \tag{7.59}$$

7.5 Third law

Of all the laws this is on the least firm ground. When the surface gravity of a black hole vanishes it is called extremal. For the Kerr–Newman the extremal condition is equivalent to $M^2 = a^2 + Q^2 + P^2$. For Kerr and electrically charged Kerr black holes one can try to throw matter into the black hole and make it extremal. One finds that it gets harder and harder for the matter to make the black hole become closer to being an extremal black hole.

7.6 Why should black holes carry an entropy?

We see that we need to accept that black holes have an entropy for the mathematics to hold, but what are physical grounds for the existence of black hole entropy?

Black holes are formed from the collapse of matter which carries entropy. However the matter that has contributed to form a black hole is not visible from an observer watching from outside the event horizon. So the observer must conclude either that the entropy disappears in the formation and growth of black holes and thus that the second principle of thermodynamics is violated or that the black holes themselves carry entropy.[25]

In general relativity, black hole solutions are fully characterised by few conserved quantities such as the mass, the angular momentum and the electric charge. Black holes do not have hair. However there are many ways of forming a black hole with assigned values of these charges. From this perspective black holes are macroscopic thermodynamic objects with many microstates, corresponding to the different possible ways of forming the same macroscopic solution. Enumerating these microstates leads to an entropy.

[25] Famously Bekenstein's advisor Wheeler, asked "what happens if we throw a cup of tea into a black hole?"

8 The black hole information paradox: short version NOT EXAMINABLE

This has been a challenge since Hawking's original paper detailing the Hawking temperature. This remains an active area of research with different approaches used to try to resolve the paradox, despite this there still remains much for us to understand. The big question is *what is quantum gravity*? String theory gives a good solution to the renormalizability problem, along with many other good things, however it is only a perturbative expansion in powers of the coupling. We know that quantum theories exhibit many fascinating and important phenomena that are not visible in perturbation theory: for example quark confinement, chiral symmetry breaking, electroweak baryon and lepton number violation all occur in the Standard model. Quantum gravity should also have interesting aspects not accessible from perturbation theory. Similar questions in QFT 'were resolved' by considering the path integral defined via the renomalization group. A partial answer for string theory can be obtained by using the AdS/CFT duality. This relates string theory in certain backgrounds (AdS) to quantum field theories. The latter we know how to define and so this duality gives a window into studying quantum gravity. However this has limitations, for example it cannot be used to describe cosmological spacetimes. We are still missing some things.

8.1 A short introduction to Hawking radiation

We want to better understand how a black hole can emit radiation. Consider an inertial observer falling through the future horizon of a Schwarzschild black hole. This takes a finite amount of proper time, but from the viewpoint of an asymptotic observer they never see them on the horizon. Consider the Schwarzschild metric

$$\mathrm{d}s^2 = -\left(1-\frac{r_s}{r}\right)\mathrm{d}t^2 + \left(1-\frac{r_s}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}s^2(S^2)\,, \tag{8.1}$$

and consider the dynamics close to the horizon: $r = r_s + \delta$. The small δ behaviour is

$$\mathrm{d}s^2 \simeq -\frac{\delta}{r_s}\mathrm{d}t^2 + \frac{r_s}{\delta}\mathrm{d}\delta^2 + r_s^2\mathrm{d}s^2(S^2)\,. \tag{8.2}$$

Defining $4r_s\delta = \rho^2$ this becomes:

$$\mathrm{d}s^2 \simeq -\frac{\rho^2}{4r_s^2}\mathrm{d}t^2 + \mathrm{d}\rho^2 + r_s^2\mathrm{d}s^2(S^2)\,. \tag{8.3}$$

The first two terms are just Minkowski space. Defining

$$X = \rho\cosh\frac{t}{2r_s}\,, \qquad T = \rho\sinh\frac{t}{2r_s}\,, \tag{8.4}$$

the metric is

$$\begin{aligned}\mathrm{d}s^2 &\simeq -\,\mathrm{d}T^2 + \mathrm{d}X^2 + r_s^2\mathrm{d}s^2(S^2)\\ &= -\,\mathrm{d}U\mathrm{d}V + r_s^2\mathrm{d}s^2(S^2)\,,\end{aligned} \tag{8.5}$$

with

$$U = T - X = -\rho\mathrm{e}^{-t/(2r_s)}\,, \quad V = T + X = \rho\mathrm{e}^{t/(2r_s)}\,. \tag{8.6}$$

The Schwarzschild coordinates t, r only cover the region $X > |T|$, that is $U < 0 < V$. This can then be extended into the black hole with the analogue of Kruskal coordinates.

In terms of the coordinates U, V introduced above, V is approximately constant while U goes through the zero linearly in the proper time τ. Thus

$$\mathrm{d}\tau \approx \mathrm{e}^{-t/r_s}\mathrm{d}t\,. \tag{8.7}$$

Therefore, while the coordinate τ smoothly crosses the ingoing horizon at $U = 0, V = \text{const}$, the coordinate t stops at the horizon. An observer using the τ coordinate will cross the horizon freely while the observer using the coordinate t will interpret space as ending at the horizon. Hawking radiation essentially arises because of the relation between these two coordinates.

An infalling observer will expand a quantum field in modes of frequency ν with respect to τ, while an outside observer will use t to expand their fields with t-frequency ω. This is not the same expansion and it leads to positive and negative frequency modes getting mixed. Consider foliating the near-horizon geometry with smooth slices by taking $U + V$ as the time coordinate. In this foliation the geometry is changing adiabatically on a time scale r_s^{-1} but the modes we are discussing have a much higher frequency and therefore the geometry is changing slowly compared to ν. The adiabatic principle implies that this mode must be in the ground state to high accuracy in the modes of the infalling observer. (If the Hamiltonian for a quantum system is changing at a rate slow compared to the spacing between energy levels then the probability for the system in become excited is exponentially small.)

To simplify our lives let us drop the S^2, though it can be included quite simply. We will treat it as a $1 + 1$ dimensional system with a massless scalar propagating in the background with metric:

$$\begin{aligned}\mathrm{d}s^2 &= -\left(1 - \frac{r_s}{r}\right)\mathrm{d}u\mathrm{d}v\\ &= -\frac{4r_s^2}{r}\mathrm{e}^{-r/r_s}\mathrm{d}U\mathrm{d}V\,.\end{aligned} \tag{8.8}$$

The coordinates u, v are only valid in quadrant 1 and are the null coordinates for the asymptotic observer. The Klein–Gordon equation becomes:

$$\partial_u\partial_v\phi = 0 = \partial_U\partial_V\phi\,, \tag{8.9}$$

which have solutions $\phi = f(u)+g(v)$, and similar for U, V, which are the left and right moving part. We only consider the right-moving part and expand in modes:

$$
\begin{aligned}
\phi_R &= \int_0^\infty \frac{\mathrm{d}\nu}{2\pi\sqrt{2\nu}} \left(a_\nu \mathrm{e}^{-\mathrm{i}\nu U} + a_\nu^\dagger \mathrm{e}^{\mathrm{i}\nu U}\right) \\
&= \int_0^\infty \frac{\mathrm{d}\omega}{2\pi\sqrt{2\omega}} \left(b_\omega \mathrm{e}^{-\mathrm{i}\omega u} + b_\omega^\dagger \mathrm{e}^{\mathrm{i}\omega u}\right) .
\end{aligned} \tag{8.10}
$$

The non-zero commutators are

$$
[a_\nu, a_{\nu'}^\dagger] = 2\pi\delta(\nu - \nu') , \qquad [b_\omega, b_{\omega'}^\dagger] = 2\pi\delta(\omega - \omega') . \tag{8.11}
$$

The coordinate U is smooth across the horizon and therefore the a_ν modes are good for infalling observers. For the asymptotic observer they want to use the b_ω modes of definite frequency with respect to the time translation symmetry. The relation between these is:

$$
b_\omega = \int_0^\infty \frac{\mathrm{d}\nu}{2\pi} (\alpha_{\omega\ \nu} a_\nu + \beta_{\omega\ \nu} a_\nu^\dagger) , \tag{8.12}
$$

where

$$
\begin{aligned}
\alpha_{\omega\ \nu} &= 2r_s \sqrt{\frac{\omega}{\nu}} (2r_s\nu)^{2\mathrm{i}r_s\omega} \mathrm{e}^{\pi r_s \omega} \Gamma(-2\mathrm{i}r_s\omega) , \\
\beta_{\omega\ \nu} &= 2r_s \sqrt{\frac{\omega}{\nu}} (2r_s\nu)^{2\mathrm{i}r_s\omega} \mathrm{e}^{-\pi r_s \omega} \Gamma(-2\mathrm{i}r_s\omega) ,
\end{aligned} \tag{8.13}
$$

By the adiabatic principle the horizon crossing modes approach the black hole vacuum state and satisfy $a_\nu|0\rangle = 0$. The eternal modes (b's) satisfy

$$
\begin{aligned}
\langle 0|b_\omega^\dagger b_{\omega'}|0\rangle &= 2\sqrt{\omega\omega'} \int_0^\infty \frac{\mathrm{d}\nu}{2\pi\sqrt{2\nu}} \frac{\mathrm{d}\nu'}{2\pi\sqrt{2\nu'}} \beta^*_{\omega\ \nu} \beta_{\omega'\ \nu'} \langle 0|a_\nu a_{\nu'}^\dagger|0\rangle \\
&= 2\sqrt{\omega\omega'} \int_0^\infty \frac{\mathrm{d}\nu}{4\pi\nu} \beta^*_{\omega\ \nu} \beta_{\omega'\ \nu'} \\
&= \frac{2\pi}{\mathrm{e}^{4\pi r_s \omega} - 1} \delta(\omega - \omega') \\
&= \frac{2\pi}{\mathrm{e}^{\omega/T_H} - 1} \delta(\omega - \omega') .
\end{aligned} \tag{8.14}
$$

This is a black body spectrum of the expected temperature.

Pair production One has neglected to use that the b's are only defined in the first quadrant, so the ϕ_R that we construct is only valid outside the horizon in region 1 of the Penrose diagram. While the expression for b's in terms of a's is compete the inverse relation will not be, it also involves other operators $\tilde{b}_\omega$ which have support only inside the black hole (region 2). One has

$$
a_\nu = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \left(\alpha^*_{\omega\ \nu} b_\omega - \beta^*_{\omega\ \nu} b_\omega^\dagger + \tilde{\alpha}^*_{\omega\ \nu} - \tilde{\beta}^*_{\omega\ \nu} \tilde{b}_\omega^\dagger\right) . \tag{8.15}
$$

One can use this to write the a vacuum as

$$|0\rangle_a \propto \exp\left[\int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\omega/(2T_H)} b_\omega^\dagger \tilde{b}_\omega^\dagger\right] |0\rangle_{b,\tilde{b}} \,. \tag{8.16}$$

The role of the $b_\omega^\dagger$ operator is to raise the energy while that of the $\tilde{b}_\omega^\dagger$ is to lower the energy:

$$[H, b_\omega^\dagger] = \omega b_\omega^\dagger \,, \qquad [H, \tilde{b}_\omega^\dagger] = -\omega \tilde{b}_\omega^\dagger \,. \tag{8.17}$$

This negative energy arises because what we are calling energy is really the conserved charge associated to the Killing vector that looks like time translation outside the horizon. On the horizon this changes signature, and once inside it labels a momentum for the interior modes, therefore either signs are permitted.

In the 1+1 dimensional model the massless scalar field separates into right-moving and left-moving modes. These can scatter into one another and this can be an important effect. One then has

$$b_\omega = R_\omega + T_\omega \int_0^\infty \frac{\mathrm{d}\nu}{2\pi} (\alpha_{\omega\ \nu} a_\nu + \beta_{\omega\ \nu} a_\nu^\dagger) \,, \tag{8.18}$$

where c_ω are the left-moving modes coming in from spatial infinity $\mathscr{I}^-$, R_ω is the amplitude for them to reflect before reaching the horizon and T_ω is the transmission amplitude, with $|R_\omega|^2 + |T_\omega|^2 = 1$. The Hawking flux is then reduced by a greybody factor,

$$\langle 0| b_\omega^\dagger b_{\omega'} |0\rangle = |T_\omega|^2 \frac{2\pi}{\mathrm{e}^{\omega/T_H} - 1} \delta(\omega - \omega') \,. \tag{8.19}$$

Above 1+1 dimensions the transmission amplitude falls exponentially with the angular momentum.

This description gives rise to the interpretation of the Hawking emission process as particle pair creation close to the horizon with a negative energy particle falling into the black hole and a positive energy particle escaping to infinity.

8.2 Information Paradox

Black hole evaporation leads to a serious problem with unitarity. Imagine throwing quantum bits into the black hole at a rate such that their energy just equals that of the outgoing radiation. The black hole's mass and its horizon area stay constant. The number of possible states of the black hole grows without bound and we lose connection with the area being the entropy. To recover it we somehow need that the bits deep inside can escape with the Hawking radiation, or at least imprint their state on it. This is forbidden by causality though, once it passes through the horizon it cannot affect anything outside.

This is not quite a crisis yet, maybe our statistical interpretation of the black hole entropy needs to be given up. Consider a black hole that was formed by a shell of matter in some pure quantum state $\Psi\rangle$. After the initial collapse settles down the state of the outgoing modes is given by the exterior density matrix

$$\rho_{\text{exterior}} = |\Omega\rangle\langle\Omega|_{\text{ingoing}} \otimes \left(\otimes_{\omega l m} \left((1 - \mathrm{e}^{-\beta\omega}) \sum_n \mathrm{e}^{-\beta\omega n} |n\rangle\langle n|_{\omega l m} \right) \right)_{\text{outgoing}} \tag{8.20}$$

As time goes on the quantum state of the radiation outside becomes more and more mixed, its entanglement entropy is increasing. This is not too bad since this is just parametrising our ignorance of what is happening behind the horizon. However as the black hole evaporates it decreases in size until at some point it becomes Planckian. Until this point the entanglement entropy of the radiation field outside continues to increase. At this point one of three things must happen:

- The evaporation stops and the Planck sized object just sits around. This is called a *remnant*. For the total state to remain pure, as required by unitarity, the remnant must have an extraordinary amount of entanglement entropy. This would exceed the Bekenstein–Hawking value which would violate the state counting statistical mechanics interpretation of the BH entropy.

- The black hole finishes evaporating into ordinary quanta such as photons and gravitions. Energy conservation prevents the final burst from containing nearly enough entanglemnt entropy to purify the earlier radiation and therefore the end result of the evaporation process is a mixed state of the radiation whose entropy is of the order of the initial black hole entropy.

- Information is conserved via the Hawking radiation. This seems to require superluminal transport of information from the black hole interior. We have studied QG using an effective field theory, for the information to get out we require that the effective theory breaks down even where curvatures are small.

Hawking argued for (2), and the process of black hole formation and evaporation is not unitary. This also requires the violation of the principle of energy conservation. We don't usually like to throw our physical principles away, so can it be avoided?

This is an active research direction, and has a long and technical history to go into in anymore detail. The moral of the story though is that GR is not enough and we need a theory of quantum gravity.

9 Hawking temperature: the longer version(non-examinable)

General relativity is not a complete theory. For one, the singularity theorem provides evidence that the theory is incomplete. More convincingly, GR is a classical theory while the world is fundamentally quantum mechanical. Trying to understand quantum gravity is one of the leading avenues of research in high energy theory. Though there has been much progress, a full understanding of quantum gravity remains elusive.

There are two parts to GR: spacetime curvature and its influence on matter and the dynamics. of the metric in response to a varying energy momentum tensor. Lacking a true theory of quantum gravity we may still use the first part, saying that the quantum mechanical matter propagates in a curved background which we will hold fixed. Rather than obeying some dynamical equations, we take the metric to be fixed.

To begin let us review some quantum mechanics and quantum field theory before defining quantum field theory in curved space.

9.1 Quantum mechanics

Quantum mechanics is profoundly different from classical mechanics, despite this both try to answer the same three fundamental questions.

- The state of the system is represented as an element of a *Hilbert space*. Mathematically a Hilbert space is just a complex vector space equipped with a complex-valued inner product with the property that taking the inner product of two states in the opposite order is equivalent to complex conjugation. We denote elements of the Hilbert space as $|\psi\rangle$ and elements of the dual space as $\langle\psi|$ so that the inner product of $|\psi_1\rangle$ and $|\psi_2\rangle$ is $\langle\psi_2|\psi_1\rangle$ and obeys
$$\langle\psi_2|\psi_1\rangle^* = \langle\psi_1|\psi_2\rangle\,. \tag{9.1}$$
In quantum mechanics the Hilbert space of interest are very often infinite-dimensional. For example, if a classical system is represented by coordinate x and momentum p, the Hilbert space could be taken to consist of all square-integrable complex-valued functions of x, or equivalently all square-integrable complex valued functions of p but not both at once.

- Observables are represented by *self-adjoint operators* on the HIlbert space. An operator is Hermitian if
$$A^\dagger = A\,, \tag{9.2}$$
where
$$\langle\psi_2|A\psi_1\rangle = \langle A^\dagger\psi_2|\psi_1\rangle\,, \tag{9.3}$$

for all states $|\psi_1\rangle$, $|\psi_2\rangle$. Many operators will not be Hermitian, but observables should be real and this requires the operator to be Hermitian. In general such operators do not commute. This means that we cannot simultaneously specify the precise values of everything we might want to measure. There will be a maximally set of commuting observables which would represent all we can say about a system at once.

- Evolution of hte system may be represented in one of two ways: as unitary evolution of the state vector in Hilbert space in the *Schrodinger picture*, or by keeping the state fixed and allowing observables to evolve according to equations of motion called the *Heisenberg picture*.

Consider a harmonic oscillator. This has Lagrangian

$$L = \frac{1}{2}\dot{x}^2 - \frac{1}{2}\omega^2 x^2\,, \tag{9.4}$$

which has equation of motion

$$\ddot{x} + \omega^2 x = 0\,. \tag{9.5}$$

In the Schrodinger picture, where states are represented by complex-valued wave functions that evolve with time, such as $\psi(x,t)$. The wave function is really the set of components of the state vector $|\psi\rangle$ expressed in the delta function position basis $|x\rangle$ so that $|\psi(t)\rangle = \int \mathrm{d}x\psi(x,t)|x\rangle$. Canonical quantisation consists of imposing the canonical commutation relation

$$[\hat{x}, \hat{p}] = \mathrm{i}\,, \tag{9.6}$$

on the coordinate operator $\hat{x}$ and its conjugate momentum $\hat{p}$. For states represented as wave functions depending on t and x, the operator $\hat{x}$ is simply multiplication by x, so the commutation relation can be implemented by fixing

$$\hat{p} = -\mathrm{i}\partial_x\,. \tag{9.7}$$

The Hamiltonian operator is

$$H = -\frac{1}{2}\partial_x^2 + \frac{1}{2}\omega^2 x^2\,, \tag{9.8}$$

and the equation of motion is the Schrodinger equation

$$\mathrm{i}\partial_t\psi = H\psi\,. \tag{9.9}$$

Since the Hamiltonian is time independent the solutions separate into functions of space and functions of time, $\psi(x,t) = f(t)g(x)$. The solutions then come in a discrete set labelled by an integer $n \geq 0$ and we find

$$\psi_n(x,t) = \mathrm{e}^{-\frac{\omega x^2}{2}} H_n(\sqrt{\omega}x)\mathrm{e}^{-\mathrm{i}E_n t}\,, \tag{9.10}$$

where H_n is a Hermite polynomial of degree n and

$$E_n = \Big(n + \frac{1}{2}\Big)\omega\,. \tag{9.11}$$

These states are all eigenfunctions of H and E_n is an energy eigenvalue. An arbitrary state of the oscillator will consist of a superposition of the energy eigenstates,

$$\psi(x,t) = \sum_n c_n \psi_n(x,t)\,, \tag{9.12}$$

for some set of appropriately normalised coefficients c_n.

Note that there is a discrete spectrum of energy eigenstates, this is a quantum property. There is a ground state of lowest energy plus a set of excited states labelled by their energy eigenvalue. The ground state has a nonvanishing energy

$$E_0 = \frac{1}{2}\omega\,, \tag{9.13}$$

which is sometimes called the zero-point energy. The classical system would have had zero energy representing a particle with $x = p = 0$. The quantum zero-point energy can be traced to the Heisenberg uncertainty principle, which forbids us from localizing a state simultaneously in both position and momentum. There is a consequently a minimum amount of jiggle in the oscillator leading to a non-zero ground state energy.

An alternative way to solve the simple harmonic oscillator is to introduce creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$ defined by

$$\hat{a} = \frac{1}{\sqrt{2\omega}}(\omega\hat{x} + \mathrm{i}\hat{p})\,, \quad \hat{a}^\dagger = \frac{1}{\sqrt{2\omega}}(\omega\hat{x} - \mathrm{i}\hat{p})\,. \tag{9.14}$$

From the commutation relations for $\hat{x}$ and $\hat{p}$ we find

$$[\hat{a}, \hat{a}^\dagger] = 1\,, \tag{9.15}$$

and the Hamiltonian becomes

$$H = \omega\Big(\hat{a}\hat{a}^\dagger + \frac{1}{2}\Big)\,. \tag{9.16}$$

The creation and annihilation operators satisfy

$$[H, \hat{a}] = -\omega\hat{a}\,, \qquad [H, \hat{a}^\dagger] = \omega\hat{a}^\dagger\,. \tag{9.17}$$

We define the number operator

$$\hat{n} = \hat{a}^\dagger \hat{a} \,. \tag{9.18}$$

Consider an eigenstate $|n\rangle$ of the number operator,

$$\hat{n}|n\rangle = n|n\rangle \,. \tag{9.19}$$

By playing with the commutation relations we have

$$\begin{aligned} \hat{n}\hat{a}^\dagger|n\rangle &= (n+1)\hat{a}^\dagger|n\rangle \\ \hat{n}\hat{a}|n\rangle &= (n-1)\hat{a}|n\rangle \,, \end{aligned} \tag{9.20}$$

thus when acting with $\hat{a}^\dagger$ on $|n\rangle$ we obtain another eigenstate of $\hat{n}$ with eigenvalue raised by one and $\hat{a}$ gives an eigenstate with eigenvalue lowered by 1. n takes integral values from 0 to ∞ and therefore there must be a vacuum state with

$$\hat{a}|0\rangle = |0\rangle \,. \tag{9.21}$$

By acting with $\hat{a}^\dagger$ we can construct all of the eigenstates

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle \,. \tag{9.22}$$

The basis states are taken to be tome independent so a physical system observing Schrödinger's equation will be described by a state

$$|\psi(t)\rangle = \sum_n c_n \mathrm{e}^{-\mathrm{i}E_n t}|n\rangle \,, \tag{9.23}$$

with c_n constant coefficients.

In order to transition more smoothly to quantum field theory it is useful to also have the Heisenberg picture in which the states are fixed and the operators evolve with time. Any state can be written formally as some fixed initial state acted on by a unitary time evolution operator

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \,, \tag{9.24}$$

where

$$U(t) = \mathrm{e}^{-\mathrm{i}\int H \mathrm{d}t} \,. \tag{9.25}$$

The Schrödinger picture expression for the matrix element of a time-independent operator, A between two time-dependent states can be written in Heisenberg picture in terms of a time dependent operator $A(t)$ and time independent states as

$$\begin{aligned} \langle\psi_2(t)|A|\psi_1(t)\rangle &= \langle\psi_2(0)|U^\dagger(t)AU(t)|\psi_1(0)\rangle \\ &= \langle\psi_2|A(t)|\psi_1\rangle \,, \end{aligned} \tag{9.26}$$

with

$$A(t) = U^\dagger(t) A U(t)\,. \tag{9.27}$$

Such an operator satisfies the Heisenberg equation of motion:

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = \mathrm{i}[H, A(t)]\,, \tag{9.28}$$

which replaces the role of Schrödinger's equation in this picture.

9.2 Quantum field theory

Quantum field theory is a particular example of a quantum mechanical system in which we quantise a field (a function or tensor field defined on spacetime). Let us first consider a free scalar field in flat space. This has action

$$S = \int \mathrm{d}^n x \Big[-\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2 \Big] \equiv \int \mathrm{d}^n x \mathcal{L}\,. \tag{9.29}$$

The equation of motion is the Klein–Gordon equation,

$$\Box\phi - m^2\phi = 0\,. \tag{9.30}$$

To translate into a Hamiltonian picture one defines the conjugate momentum to be

$$\pi = \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}\,. \tag{9.31}$$

For the free scalar field this is

$$\pi = \dot{\phi}\,. \tag{9.32}$$

Of course since we are using time derivative we have assumed a particular inertial frame and therefore the Hamiltonian procedure necessarily violates manifest Lorentz invariance. With care however, the observables remain Lorentz invariant. The Hamiltoian is represented as the integral of a Hamiltonian density over the spatial directions directions. The Hamiltonian density is related to the Lagrangian by a Legendre transformation,

$$\begin{aligned} \mathcal{H}(\phi,\pi) &= \pi\dot{\phi} - \mathcal{L}(\phi, \partial_\mu\phi) \\ &= \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{m^2}{2}\phi^2\,, \end{aligned} \tag{9.33}$$

with $(\nabla\phi)^2 = \delta^{ij}\partial_i\phi\partial_j\phi$. In comparison to the harmonic oscillator the field $\phi(x)$ plays the role of the coordinate x and the momentum field $\pi(x)$ plays the role of p. Instead of a state being specified by two number (x,p) at some fixed time, the initial conditions are values of the field over all of the spatial directions at a fixed time.

Note that $\phi(x^\mu)$ is not a wave function; it is a dynamical variable generalising the single degree of freedom x in the case of the harmonic oscillator. We will use a Heisenberg picture of time evolution where we promote ϕ to an operator.

First we need to solve the classical theory. The solutions of the Klein–Gordon equation include the plane wave solution

$$\phi(x^\mu) = \phi_0 \mathrm{e}^{\mathrm{i}p_\mu x^\mu} = \phi_0 \mathrm{e}^{-\mathrm{i}p^0 t + \mathrm{i}\vec{p}\cdot\vec{x}}\,, \tag{9.34}$$

where the wave vector has components

$$p^\mu = (p^0, \vec{p})\,, \tag{9.35}$$

and the frequency must satisfy

$$(p^0)^2 = \vec{p}^{\,2} + m^2\,, \qquad p^0 > 0\,. \tag{9.36}$$

The latter condition is in order to consider the positive frequency modes only.

We can write down the most general solution by constructing a complete orthonormal set of modes in terms of which any solution may be expressed. We need to first define an inner product on the space of solutions. To inner product is an integral over a constant time hypersurface Σ_t and is

$$(f,g) = \mathrm{i}\int_{\Sigma_t} (f^* \overleftrightarrow{\partial_t} g)\mathrm{d}^{n-1}x\,, \qquad f^* \overleftrightarrow{\partial_t} g = f^*\partial_t g - \partial_t f^* g\,. \tag{9.37}$$

By using Stoke's theorem and the equation of motion one can check that this is independent of the chosen hypersurface. Let us define

$$\psi_p = N_p \mathrm{e}^{\mathrm{i}p_\mu x^\mu}\,, \tag{9.38}$$

with $p^2 + m^2 = 0$. Then $\{\psi_p, \psi_p^*\}$ form a basis of solutions and any field configuration can be expanded as

$$\phi(x) = \int \mathrm{d}^3p \big(a_p\psi_p(x) + a_p^*\psi_p^*(x)\big)\,, \tag{9.39}$$

with a_p and a_p^* are complex constants. In order for the basis to be orthonormal we take

$$N_p = \frac{1}{\sqrt{2p^0}(2\pi)^{3/2}},\,. \tag{9.40}$$

We quantise the theory by promoting ϕ and π to be operators and impose the standard commutation relations:

$$[\phi(t,\vec{x}), \pi(t,\vec{y})] = \mathrm{i}\delta^{(3)}(\vec{x}-\vec{y})\,, \quad [\phi(t,\vec{x}), \phi(t,\vec{y})] = 0\,, \quad [\pi(t,\vec{x}), \pi(t,\vec{y})] = 0\,. \tag{9.41}$$

This may then be translated into commutation relations for the a's, with

$$[a_p, a_q^\dagger] = \delta(\vec{p} - \vec{q})\,, \quad [a_p, a_q] = 0\,, \quad [a_p^\dagger, a_q^\dagger] = 0\,. \tag{9.42}$$

We may then define a vacuum state by

$$a_p|0\rangle = 0\,, \quad \forall p\,. \tag{9.43}$$

It may seem that the definition of the vacuum state depends on the initial choice of inertial frame, however this is not the case. Consider a different inertial frame $\tilde{x}^\mu$ related by a Lorentz transformation $\tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu$. In this new frame the positive frequency mode functions are

$$\tilde{\psi}_p = N_p \mathrm{e}^{\mathrm{i}p_\mu \tilde{x}^\mu}\,, \tag{9.44}$$

and the field expansion is

$$\phi(\tilde{x}) = \int \mathrm{d}^3p \Big(\tilde{a}_p \tilde{\psi}_p + \tilde{a}_p^\dagger \tilde{\psi}_p^*\Big)\,, \tag{9.45}$$

and in terms of these modes the new vacuum state satisfies $\tilde{a}_p|\tilde{0}\rangle = 0, \forall p$. We need to show that

$$a_p|0\rangle = 0 \;\; \forall p \quad \Rightarrow \quad \tilde{a}_p|0\rangle \;\; \forall p\,. \tag{9.46}$$

We have

$$\tilde{\psi}_p = \frac{1}{\sqrt{2p_0}(2\pi)^{3/2}} \mathrm{e}^{\mathrm{i}p_\mu \tilde{x}^\mu} = \left(\frac{\tilde{p}^0}{p^0}\right)^{1/2} \frac{1}{\sqrt{2\tilde{p}^0}(2\pi)^{3/2}} \mathrm{e}^{\mathrm{i}\tilde{p}_\mu x^\mu} = \left(\frac{\tilde{p}^0}{p^0}\right)^{1/2} \psi_{\tilde{p}}\,. \tag{9.47}$$

More over since we restrict to the orthochronous subgroup of the Lorentz group, i.e. $\Lambda^0{}_0 > 0$ we have $p^0 > 0 \Rightarrow \tilde{p}^0 > 0$. Therefore we have

$$a_p|0\rangle = 0 \;\; \forall p \quad \Rightarrow \quad \tilde{a}_p|0\rangle \;\; \forall p\,, \tag{9.48}$$

and the converse follows by symmetry and the vacuum state is independent of the choice of frame.

9.3 QFT in curved spacetime

We now want to consider what changes when we try to quantise a field theory on curved spacetime. We fix a background (M, g) and assume that it is globally hyperbolic. Recall that this means that the spacetime admits a Cauchy surface and from initial conditions on the Cauchy surface we can solve the equations of motion on all of spacetime. We perform minimal coupling of the theory so that $\eta^{\mu\nu} \to g^{\mu\nu}$ and $\partial_\mu \to \nabla_\mu$. The Klein–Gordon equation becomes

$$\nabla^2\phi \equiv g^{\mu\nu}\nabla_\mu\partial_\nu\phi = m^2\phi\,, \tag{9.49}$$

while the inner product is modified to

$$(f_1, f_2) = \mathrm{i}\int_\Sigma \mathrm{d}^3x\sqrt{\gamma}n^\mu(f_1^*\partial_\mu f_2 - \partial_\mu f_1 f_2^*)\,, \tag{9.50}$$

with Σ a spacelike hypersurface and n^μ a unit normal vector and γ the determinant of the induced metric. Let the background admit a Killing vector, K, then on functions we have

$$[K, \nabla^2]f = 0\,. \tag{9.51}$$

Since ∇^2 and $\mathrm{i}K$ are both self-adjoint and commuting they admit a complete set of common eigenfunctions

$$\nabla^2 f = m^2 f\,, \quad \mathrm{i}K^\mu\partial_\mu f = \omega f\,. \tag{9.52}$$

If K is timelike we are entitled to call the eigenvalue the frequency. Indeed this is how it works in Minkowski space where $K = \partial_t$. If f is an eigenfunction with positive frequency ω then f^* is an eigenfunction of negative frequency $-\omega$. We can then without loss of generality expand our fields in terms of positive and negative frequency eigenfunctions of the Laplacian in a basis $\{\psi_i\}$ of positive frequency modes and $\{\psi_i^*\}$ of negative frequency modes. We expand our field as

$$\phi = \sum_i (a_i\psi_i + a_i^\dagger\psi_i^*)\,, \tag{9.53}$$

with

$$[a_i, a_{i'}^\dagger] = \delta_{ij}\,. \tag{9.54}$$

Consider a sandwich spacetime (M, g) made up of three regions, region B bottom, region C for centre and region T for top, and assume the Klein–Gordon equation holds throughout spacetime. Region B is stationary and admits a timelike Killing vector K^B, region C is not stationary and all sorts of dynamical processes might take place so long as it remains globally hyperbolic, and finally region T is once again stationary with a new timelike Killing vector K^T. If we quantise in region B we pick a set of modes $\{f_i, f_i^*\}$ that satisfy $\mathrm{i}K^B f_i = \omega_i f_i$ with $\omega_i > 0$. On the other hand in region T we choose another set of modes $\{g_i, g_i^*\}$ that satisfy $\mathrm{i}K^T g_i = \tilde{\omega}_i g_i$ with $\tilde{\omega}_i > 0$. Note that even though the positive-frequency conditions are imposed using the Killing vectors in specific regions the modes extend throughout the whole of spacetime. In the two cases the respective expansion is then

$$\phi(x) = \sum_i \left(a_i f_i + a_i^\dagger f_i^*\right) = \sum_i \left(b_i g_i + b_i^\dagger g_i^*\right)\,, \tag{9.55}$$

where the modes have been normalised with respect to the Klein–Gordon inner product so that the commutation relations are

$$[a_i, a_j^\dagger] = \delta_{ij} \,, \qquad [b_i, b_j^\dagger] = \delta_{ij} \,. \tag{9.56}$$

Since $\{f_i\}$ forms a basis we can also expand any function in terms of it, we have

$$g_i = \sum_i A_{ij} f_j + B_{ij} f_j^* \,. \tag{9.57}$$

The coefficients A_{ij} and B_{ij} are called the *Bogoliubov coefficients* and the transformation between the different bases is called a *Bogoliubov transformation.* Using the normalisation conditions it can be shown that they satisfy

$$\begin{aligned} \sum_k A_{ik} A_{jk}^* - B_{ik} B_{jk}^* &= \delta_{ij} \,, \\ \sum_k A_{ik} B_{jk} - B_{ik} A_{jk} &= 0 \,. \end{aligned} \tag{9.58}$$

Or in matrix notation

$$AA^\dagger - BB^\dagger = 1 \,, \qquad AB^T = BA^T \,. \tag{9.59}$$

We can also relate the different operator coefficients to each other

$$b_i = \sum_j A_{ij}^* a_j - B_{ij}^* a_j^\dagger \,. \tag{9.60}$$

The procedure above defines a vacuum state associated with the modes $\{f_i, f_i^*\}$ called the *in*-vacuum as the states satisfy $a_i|0\rangle_{in} = 0 \ \forall i$. In a stationary reference frame in region B (i.e. an integral curve of K^B this will appear empty. What about in region T? What is the expected number of particles of the state $|0\rangle_{in}$ with momentum i. It is given by the expectation value

$$\langle N_i \rangle = {}_{in}\langle 0|b_i^\dagger b_i|0\rangle_{in} = \sum_j B_{ij} B_{ij}^* \qquad \text{no summation over i} \,. \tag{9.61}$$

If this is non-zero there is pair production. Alternatively one can see this as the in-vacuum and out-vacuum are different. Hence a changing spacetime geometry generically causes particle production.

9.4 Unruh effect

Even though we have made an effort above to understand QFT in curved space we will first consider a phenomenon that uses the above ideas but manifests in flat space. This is the Unruh effect, which states that an accelerating observer in the Minkowski vacuum will observe a thermal spectrum of particles.

The basic idea is very simple, observers with different notions of positive and negative frequency modes will disagree on the particle content of a given state. A uniformly accelerated observe in Minkowski moves along an orbit of a time-like Killing vector, however this is not the usual time-translation Killing vector. We can therefore expand the field in terms of modes appropriate for the accelerated observer and calculate the number operator in the ordinary Minkowski vacuum. We will find that this leads to a thermal spectrum of particles.

To simplify things as much as possible let us consider a massless scalar field in two dimensions. The wave equation is

$$\Box\phi = 0\,. \tag{9.62}$$

Before trying to quantise the theory consider a uniformly accelerating observer, we have seen this earlier in section 4.1.3, but let us review the details. In inertial coordinates the metric can be written as

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2\,. \tag{9.63}$$

An observer moving at a uniform acceleration of magnitude α follows the trajectory

$$t(\tau) = \frac{1}{\alpha}\sinh(\alpha\tau)\,, \qquad x(\tau) = \frac{1}{\alpha}\cosh(\alpha\tau)\,, \tag{9.64}$$

note that

$$x^2 = t^2 + \alpha^2\,. \tag{9.65}$$

We can choose new coordinates on two-dimensional Minkowski space that are adapted to uniformly accelerated motion as

$$t = \frac{1}{a}\mathrm{e}^{a\xi}\sinh(a\eta)\,, \qquad x = \frac{1}{a}\mathrm{e}^{a\xi}\cosh(a\eta)\,, \quad (x > |t|)\,. \tag{9.66}$$

The new coordinates have ranges

$$-\infty < \eta, \xi < \infty\,, \tag{9.67}$$

and cover the wedge $x > |t|$ *Rindler space* corresponds to the right wedge $x > |t|$ foliated by the worldlines of the accelerated observers and labelled by region I in figure 32. In these

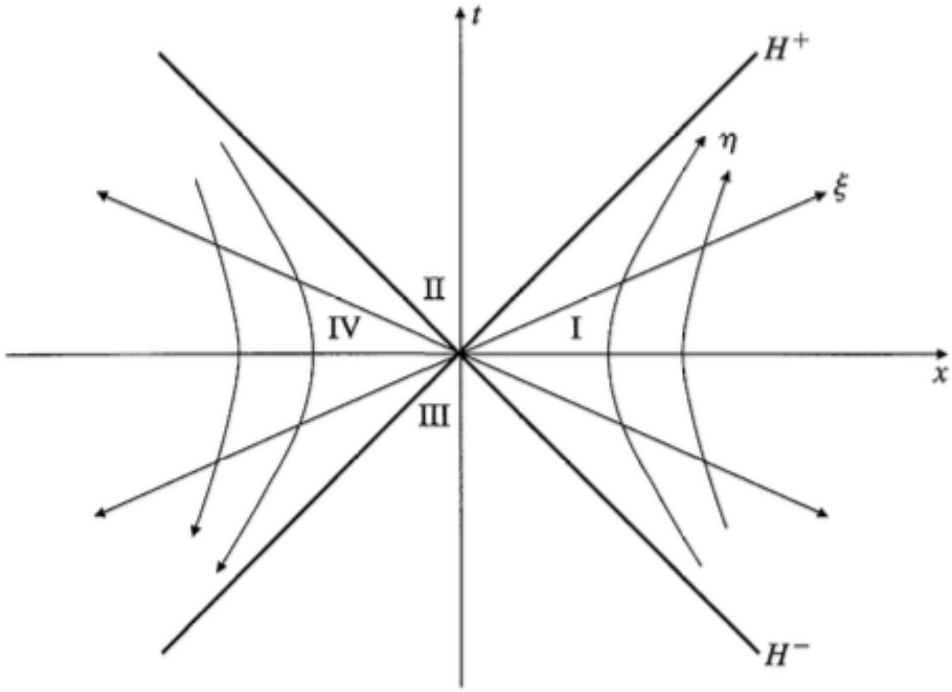

Figure 32: Minkowski spacetime in Rindler coordinates. Region I is the region accessible to an observer undergoing constant acceleration in the $+x$-direction. The coordinates (η, ξ) can be used in region I or region IV, where they point in the opposite direction. The vector filed ∂_η corresponds to the generator of Lorentz boosts and the horizons $H^{\pm}$ are Killing horizons for this vector field, which represent the boundaries of the past and future as witnessed by the Rindler observer.

coordinates the constant acceleration path is

$$\eta(\tau) = \frac{\alpha}{a}\tau\,, \qquad \xi(\tau) = \frac{1}{a}\log\frac{a}{\alpha}\,, \tag{9.68}$$

and we see that the proper time is proportional to η and the spatial constant ξ is constant. Then an observer with acceleration $\alpha = a$ moves along the path

$$\eta = \tau\,, \quad \xi = 0\,. \tag{9.69}$$

The metric in these coordinates takes the form

$$\mathrm{d}s^2 = \mathrm{e}^{2a\xi}\big(-\mathrm{d}\eta^2 + \mathrm{d}\xi^2\big)\,. \tag{9.70}$$

The null line $t = x$ labelled by H^+ is a future Cauchy horizon for any $\eta =$ constant spacelike hypersurface in region I. Similarly H^- is a past Cauchy horizon.

The metric is independent of η and therefore ∂_η is a Killing vector, however since this is Minkowski spacetime there are more of course. Indeed if we express ∂_η in the (t,x) coordinates we have

$$\partial_\eta = a(x\partial_t + t\partial_x)\,. \tag{9.71}$$

This is the Killing vector which generates a boost in the x-direction. It is clear that this Killing vector naturally extends throughout the spacetime. This extends naturally throughout the spacetime, in regions II and III it is spacelike while in region IV it is timelike but past-directed. The horizons are Killing horizons for ∂_η.

We can define coordinates (η,ξ) in region IV by flipping the signs in (9.66),

$$t = -\frac{1}{a}\mathrm{e}^{a\xi}\sinh(a\eta)\,, \qquad x = -\frac{1}{a}\mathrm{e}^{a\xi}\cosh(a\eta)\,, \quad (x<|t|)\,. \tag{9.72}$$

The sign guarantees that ∂_η and ∂_t point in opposite directions. Strictly speaking we cannot use the (η,ξ) simultaneously in regions I and IV since the ranges are the same in each region, we must explicitly indicate to which region the coordinate belongs to. We add labels to distinguish so that the metric takes the same form in both regions.

Along the surface $t=0$ the Killing vector ∂_η is a hypersurface-orthogonal timelike Killing vector except for the single point $x=0$ where it vanishes. We can therefore use it to define a set of positive and negative frequency modes on which we can build a Fock space for the scalar-field Hilbert space. The massless Klein–Gordon equation in Rindler coordinates takes the form

$$\Box\phi = \mathrm{e}^{-2a\xi}(-\partial_\eta^2 + \partial_\xi^2)\phi = 0\,. \tag{9.73}$$

Therefore a normalised plane wave

$$g_k = \frac{1}{\sqrt{4\pi\omega}}\mathrm{e}^{-\mathrm{i}\omega\eta+\mathrm{i}k\xi}\,, \qquad \omega = |k|\,, \tag{9.74}$$

solves the equation and has positive frequency with respect to ∂_η since

$$\mathcal{L}_{\partial_\eta} g_k = -\mathrm{i}\omega g_k\,. \tag{9.75}$$

However this is only true in region I since we need our modes to be positive frequency with respect to a future directed Killing vector, in region IV the relevant Killing vector is $\partial_{-\eta} = -\partial_\eta$. To remove this problem of defining the modes we introduce two sets of modes one with support in region I and one with support in region IV:

$$g_k^{(1)} = \begin{cases} \frac{1}{\sqrt{4\pi\omega}} \mathrm{e}^{-\mathrm{i}\omega\eta + \mathrm{i}k\xi} & I \\ 0 & IV \end{cases}$$
$$g_k^{(2)} = \begin{cases} 0 & I \\ \frac{1}{\sqrt{4\pi\omega}} \mathrm{e}^{\mathrm{i}\omega\eta + \mathrm{i}k\xi} & IV \end{cases} \tag{9.76}$$

with $\omega = |k|$ in each region. These then define the positive frequency with respect to the relevant future directed timelike Killing vector. The two sets with their conjugates form a complete set of modes for any solution to the wave equation throughout the spacetime. Both sets are non-vanishing in regions II and III however this is obscured by the choice of (η, ξ) coordinates. Denoting the associated annihilation and creation operators as $b_k^{(i)}$ and $b_k^{(i)\dagger}$, we can write

$$\phi = \int \mathrm{d}k \Big(b_k^{(1)} g_k^{(1)} + b_k^{(1)\dagger} g_k^{(1)*} + b_k^{(2)} g_k^{(2)} + b_k^{(2)\dagger} g_k^{(2)*} \Big) \,. \tag{9.77}$$

This gives an alternative expansion to the original Minkowski modes:

$$\phi = \int \mathrm{d}k \big(a_k f_k + a_k^\dagger f_k^* \big) \,. \tag{9.78}$$

The inner product of the Rindler modes gives

$$(g_{k_1}^{(i)}, g_{k_2}^{(j)}) = \delta^{ij} \delta(k_1 - k_2) \,, \tag{9.79}$$

and similarly for the conjugate modes. There are two sets of modes, Minkowski and Rindler, that we can expand the solution of the Klein–Gordon equation in. Although the Hilbert spaces are the same the Fock spaces are different, in particular the definition of the vacuum. The Minkowski vacuum $|0_M\rangle$ satisfies

$$a_k |0_M\rangle = 0 \,, \tag{9.80}$$

while the Rindler vacuum satisfies

$$b_k^{(1)} |0_R\rangle = b_k^{(2)} |0_R\rangle = 0 \,. \tag{9.81}$$

We see that because an individual Rindler mode cannot be written in terms of positive frequency Minkowski modes, the Rindler annihilation modes are a superposition of both the Minkowski creation and annihilation operators.

A Rindler observer will be static with respect to orbits of the boost Killing vector ∂_η. Such an observer in region I will describe particles in terms of the Rindler modes $g_k^{(1)}$ and will observer a state in the Rindler vacuum to be devoid of particles, a state $b_k^{(1)\dagger}|0_R\rangle$ to contain a single particle of frequency $\omega = |k|$ and so forth. Conversely a Rindler observer travelling through the Minkowski vacuum state will detect a background of particles, even though to the inertial observer the vacuum is completely empty.

We would like to know what kind of particles does the Rindler observer detect? We know how to answer this, we need to compute the Bogolubov coefficients relating the Minkowski modes to the Rindler modes, and then use this to compute the expectation values. Unruh found a shortcut to this somewhat tedious computation. His idea was to find a set of modes that share the same vacuum as the Minkowski modes but for which the overlap with the Rindler modes is more direct. We start with the Rindler modes and extend them to all of spacetime, and then express the extension in terms of the original Rindler modes.

We have

$$\begin{aligned} \mathrm{e}^{-a(\eta-\xi)} &= \begin{cases} a(x-t) & I \\ a(t-x) & IV \end{cases} \\ \mathrm{e}^{a(\eta+\xi)} &= \begin{cases} a(t+x) & I \\ -a(t+x) & IV \end{cases} \end{aligned} \tag{9.82}$$

We can express the spacetime dependence of a mode $g_k^{(1)}$ with $k > 0$ in terms of the Minkowski coordinates in region I as

$$\sqrt{4\pi\omega}\, g_k^{(1)} = a^{\mathrm{i}\omega/a}(x-t)^{\mathrm{i}\omega/a}\,. \tag{9.83}$$

The analytic continuation of this throughout all of spacetime is then obvious, we just use this final expression for all (t, x). We want to express the result in terms of the Rindler modes everywhere and so we need to bring the $g_k^{(2)}$ modes into the game. We have

$$\sqrt{4\pi\omega}\, g_k^{(2)} = a^{-\mathrm{i}\omega/a}(-t-x)^{-\mathrm{i}\omega/a}\,. \tag{9.84}$$

This doesn't match the behaviour of our analytically extended mode, however if we take the complex conjugate and reverse the wave number we find

$$\sqrt{4\pi\omega}\, g_{-k}^{(2)*} = a^{\mathrm{i}\omega/a}\mathrm{e}^{\pi\omega/a}(-t+x)^{\mathrm{i}\omega/a}\,, \tag{9.85}$$

and therefore

$$\sqrt{4\pi\omega}\Big(g_k^{(1)} + \mathrm{e}^{-\pi\omega/a} g_{-k}^{(2)*}\Big) = a^{\mathrm{i}\omega/a}(-t+x)^{\mathrm{i}\omega/a}\,. \tag{9.86}$$

An identical result holds for the $k < 0$ modes. The properly normalised mode is

$$h_k^{(1)} = \frac{1}{\sqrt{2\sinh\frac{\pi\omega}{a}}}\Big(\mathrm{e}^{\pi\omega/(2a)} g_k^{(1)} + \mathrm{e}^{-\pi\omega/(2a)} g_{-k}^{(2)*}\Big)\,. \tag{9.87}$$

This is the appropriate analytic extension of the $g_k^{(1)}$ modes, the extension of the $g_k^{(2)}$ modes is

$$h_k^{(2)} = \frac{1}{\sqrt{2\sinh\frac{\pi\omega}{a}}}\left(\mathrm{e}^{\pi\omega/(2a)} g_k^{(2)} + \mathrm{e}^{-\pi\omega/(2a)} g_{-k}^{(1)*}\right). \tag{9.88}$$

One can check that these are correctly normalised. We can now expand in these modes as

$$\phi = \int \mathrm{d}k \left(c_k^{(1)} h_k^{(1)} + c_k^{(1)\dagger} h_k^{(1)*} + c_k^{(2)} h_k^{(2)} + c_k^{(2)\dagger} h_k^{(2)*}\right). \tag{9.89}$$

The modes $h_k^{(i)}$ can be expressed purely in terms of positive frequency Minkowski modes f_k and therefore they share the same vacuum state $|0_M\rangle$ so that

$$c_k^{(i)}|0_M\rangle = 0\,. \tag{9.90}$$

In the Minkowski vacuum an observer in region I will observe particles defined by the operators $b_k^{(1)}$; the expected number of such particle of frequency ω is

$$\begin{aligned}\langle 0_M|n_R^{(1)}(k)|0_M\rangle &= \langle 0_M|b_k^{(1)\dagger} b_k^{(1)}|0_M\rangle \\ &= \frac{1}{2\sinh\frac{\pi\omega}{a}}\langle 0_M|\mathrm{e}^{-\pi\omega/a} c_{-k}^{(1)} c_{-k}^{(1)\dagger}|0_M\rangle \\ &= \frac{1}{\mathrm{e}^{2\pi\omega/a}-1}\delta(0)\,.\end{aligned} \tag{9.91}$$

Planck's law describes the spectral density of electromagnetic radiation emitted by a black body in thermal equilibrium at a give temperature T. It says that the spectral radiance of a body for frequency ω at temperature T is given by

$$B(\omega, T) = \frac{\hbar\omega^3}{4\pi^2 c^2}\frac{1}{\mathrm{e}^{\hbar\omega/(K_B T)}-1}\,. \tag{9.92}$$

We conclude that an observer moving with uniform acceleration through the Minkowski vacuum observes a thermal spectrum of particles. (There is more to saying this is a thermal spectrum than just the above, one needs to check that there are no hidden correlations in the observed particles, this has indeed been shown and therefore the radiation detected by a Rindler observer is truly thermal.)

The temperature $T = \frac{a}{2\pi}$ is what would be measured by an observer moving along the path $\xi = 0$, which feels the acceleration $a = \alpha$. Any other path with $\xi =$ constant feels an acceleration

$$\alpha = a\mathrm{e}^{-a\xi}\,, \tag{9.93}$$

and thus should measure thermal radiation with a temperature $T = \frac{\alpha}{2\pi}$. As $\xi \to \infty$ the temperature goes to 0, which is consistent with the fact that near ∞ the Rindler observer is nearly inertial.

The Unruh effect tells us that an accelerated observer will detect particles in the Minkowski vacuum state. An inertial observer would say that the same state is completely empty, the expectation value of the energy momentum tensor $\langle T_{\mu\nu}\rangle = 0$. If there is no energy momentum how can the Rindler observer detect particles? If the Rindler observer is to detect background particles, they must carry a detector. This must be coupled to the particle being detected. However if a detector is being maintained at constant acceleration, energy is not conserved. From the point of view of the Minkowski observer the Rindler detector emits as well as absorbs particles, once the coupling is introduced the possibility of emission is unavoidable. When the detector registers a particle the inertial observer would say that it had emitted a particle and felt a radiation-reaction force in response. Ultimately the energy needed to excite the Rindler detector does not come from the background energy momentum tensor but from the energy we put into the detector to keep it accelerating.

9.5 Hawking temperature

We may now use a very quick argument following the above to conclude that a black hole has a temperature. Consider a static observer at radius $r_1 > R_S$ outside the Schwarzschild black hole. Such an observer moves along orbits of the time-like Killing vector $K = \partial_t$. The red-shift factor is given by

$$V = \sqrt{1 - \frac{2G_N M}{r}}\,, \tag{9.94}$$

and the magnitude of the acceleration is given by

$$a = \frac{G_N M}{r\sqrt{r - 2G_N M}}\,. \tag{9.95}$$

For observed close to the event horizon $r_1 - 2G_N M \ll 2G_N M$ this acceleration becomes very large compared to the scale set by the Schwarzschild radius

$$a_1 \gg \frac{1}{2G_N M}\,. \tag{9.96}$$

Let us assume that the quantum state of some scalar field ϕ looks like the Minkowski vacuum as seen by a freely falling observer near the black hole. The static observer looks just like a constant acceleration observer in flat spacetime and will detect Unruh radiation at a temperature $T_1 = a_2/(2\pi)$.

Now consider a static observer at infinity. The radiation will propagate to infinity with an appropriate red-shift factor. We find

$$T_\infty = \frac{V_1}{V_\infty}\frac{a}{2\pi}\,. \tag{9.97}$$

At infinity we have $V_\infty = 1$ so the observed temperature is

$$T_\infty = \lim_{r_1 \to 2G_N M} \frac{V_1 a_1}{2\pi} = \frac{\kappa}{2\pi}\,. \tag{9.98}$$

This is the Hawking effect and the radiation is known as Hawking radiation.

We can be more rigorous in the derivation of the Hawking temperature. Consider a spacetime that corresponds to a spherically symmetric collapsing star which forms a black hole, recall that the Penrose diagram is given in 11. This is a curved spacetime which is globally hyperbolic, for instance $\mathscr{I}^-$ is a Cauchy surface. Even though the Schwarzschild black hole solution is a static spacetime the collapsing star is not, and involves complicated dynamics. However the spacetime is approximately stationary in the far asymptotic past (near $\mathscr{I}^+$) and the far asymptotic future (near $\mathscr{I}^+$). We can therefore perform second quantisation with respect to stationary observers near $\mathscr{I}^-$ which give us "in"-modes and the "in"-vacuum and also a second quantisation associated with stationary observers at $\mathscr{I}^+$ leading to the "out"-vacuum. We have a sandwich spacetime and we can ask will observes in the far future see particles in the *in*-vacuum.

The field expansion defining the *in*-vacuum can be constructed by specifying a complete set of positive frequency modes on $\mathscr{I}^-$. For the quantisation in the far future $\mathscr{I}^+$ is not a Cauchy surface for the spacetime, one must take $\mathscr{I}^+ \cup \mathcal{H}^+$. We may therefore quantise the field in the far future by specifying a complete set on it. There are three sets of modes:

$$\begin{aligned}
&f_i : \text{positive frequency on } \mathscr{I}^- \\
&g_i : \text{positive frequency on } \mathscr{I}^+ \text{ and zero on } \mathcal{H}^+ \\
&h_i : \text{positive frequency on } \mathcal{H}^+ \text{ and zero on } \mathscr{I}^+
\end{aligned} \tag{9.99}$$

Strictly speaking there is no timelike Killing vector on $\mathcal{H}$ so the term positive frequency is somewhat misleading, however the choice of modes h_i does not affect the outcome of the calculation. We can choose an arbitrary set and call them positive frequency modes and attach them to annihilation operators in the field expansion, we only require that the set $\{g, h\}$ give a basis of modes. We can therefore expand

$$\phi(x) = \sum_i a_i f_i(x) + \text{h.c.} = \sum_I b_I g_I(x) + \sum_\alpha c_\alpha h_\alpha(x) + \text{h.c.}\,. \tag{9.100}$$

The Bogoliubov coefficients in the expansion satisfy

$$g_i = \sum_j \left(A_{ij} f_j + B_{ij} f_j^*\right). \tag{9.101}$$

We now want to look at the analytic solutions of the Klein–Gordon equation in the Schwarzschild black hole background. This is hard. Instead we can ask if we impose boundary condition to the solution at $\mathscr{I}^+$ and investigate what its corresponding form must be on $\mathscr{I}^-$. This amounts to tracing back in time the solution from $\mathscr{I}^+$ to $\mathscr{I}^-$.

The metric of the Schwarzschild black hole spacetime with coordinates (t, r_*, θ, ϕ) reads

$$\mathrm{d}s^2 = \left(1 - \frac{2M}{r}\right)\left(- \mathrm{d}t^2 + \mathrm{d}r_*^2\right) + r^2 \mathrm{d}s^2(S^2)\,. \tag{9.102}$$

We will also use the light-cone coordinates $u = t - r_*$ and $v = t + r_*$. We can find the Klein–Gordon equation for the field $\phi(t, r_*, \theta, \phi)$. Expanding in spherical harmonics

$$\phi(t, r_*, \theta, \phi) = \chi_l(t, r_*) Y_{lm}(\theta, \phi)\,, \tag{9.103}$$

we find

$$\left[\partial_t^2 - \partial_{r_*}^2 + V_l(r_*)\right]\chi_l = 0\,, \tag{9.104}$$

where

$$V_l(r_*) = \left(1 - \frac{2M}{r}\right)\left[\frac{l(l+1)}{r^2} + \frac{2M}{r^3}\right]. \tag{9.105}$$

We set

$$\chi_l(t, r_*) = \mathrm{e}^{-\mathrm{i}\omega t} R_{l\omega}(r_*)\,, \tag{9.106}$$

so that

$$(\partial_{r_*}^2 + \omega^2) R_{\omega l} = V_l R_{\omega l}\,. \tag{9.107}$$

We can get some intuition by looking at the potential. Both near the horizon $\mathcal{H}^+$ $(r_* \to -\infty)$ and near $\mathscr{I}^\pm$ $(r_* \to \infty)$ the potential tends to zero. It takes the for of a potential barrier. If we consider how any solution to the above evolves in time, it will be partly transmitted and partly reflected as it comes in from $r_* = \infty$.

Near $\mathscr{I}^\pm$ the solutions are just plane waves. We define outgoing and ingoing as those which correspond to r_* increasing or decreasing with time. We define the *early modes*

$$\begin{aligned} f_{lm\omega+} &= \frac{1}{\sqrt{2\pi\omega}}\mathrm{e}^{-\mathrm{i}\omega u}\frac{Y_{lm}}{r}\,, \quad \text{outgoing} \\ f_{lm\omega-} &= \frac{1}{\sqrt{2\pi\omega}}\mathrm{e}^{-\mathrm{i}\omega v}\frac{Y_{lm}}{r}\,, \quad \text{ingoing} \end{aligned} \tag{9.108}$$

at $\mathscr{I}^-$ and *late modes*

$$\begin{aligned} g_{lm\omega+} &= \frac{1}{\sqrt{2\pi\omega}}\mathrm{e}^{-\mathrm{i}\omega u}\frac{Y_{lm}}{r}\,, \quad \text{outgoing} \\ g_{lm\omega-} &= \frac{1}{\sqrt{2\pi\omega}}\mathrm{e}^{-\mathrm{i}\omega v}\frac{Y_{lm}}{r}\,, \quad \text{ingoing} \end{aligned} \tag{9.109}$$

at $\mathscr{I}^+$. We will be interested mainly in ingoing early modes and outgoing late modes, so we will use the shorthand notation:

$$f_\omega \sim f_{lm\omega-}\,, \qquad g_\omega \sim g_{lm\omega+}\,. \tag{9.110}$$

We need to express g_ω in terms of $f_{\omega'}$ and $f^*_{\omega'}$ on $\mathscr{I}^-$. First note that plane waves such as g_ω are in fact completely delocalised since they have support everywhere on $\mathscr{I}^+$.

We want to trace the solution of the late modes back in time in terms of the early modes. As the wave travels inwards from $\mathscr{I}^+$ toward decreasing values of r_*, it will encounter the potential barrier. One part of the wave, $g^{(r)}_\omega$ will be reflected and end up on $\mathscr{I}^-$ with the same frequency ω. This will correspond to a term of the form $A_{\omega\omega'} \propto \delta(\omega - \omega')$ in the expansion in (9.101). The remaining part $g^{(t)}_\omega$ will be transmitted through the barrier and will enter the collapsing matter. In that region the precise geometry of spacetime is unknown. However since we are interested in a packet peaked at late times and at some finite frequency ω_0 we know that the packet will be peaked at a very high frequency as it enters the collapsing matter due to the gravitational blueshift. This allows us to assume that the packet will obey the *geometric optics approximation* which means that g_ω takes the form $A(x)\mathrm{e}^{\mathrm{i}S(x)}$ where $A(x)$ is slowly varying compared to S. Substituting into the Klein–Gordon equation we find $\nabla_\mu S \nabla^\mu S = 0$, which means that surfaces of constant phase are null. Given a wave we can trace its surfaces of constant phase back in time by following null geodesics.

Consider tracing back the wave along a particular null geodesic γ which starts off at some $u = u_0$ at $\mathscr{I}^+$ and hits $\mathscr{I}^-$ at $v = v_0$. Denote by γ_H a null generator of the horizon $\mathcal{H}^+$ which has been extended into the past until it hits $\mathscr{I}^-$ at some value of v. We may set this value to $v = 0$ without loss of generality since the spacetime is invariant under shifts $v \to v + c$. We therefore have $v_0 < 0$ for the geodesic γ. Let n be a connecting vector between the two curves and fix its normalisation by requiring $n \cdot \xi = -1$ with ξ the generator of the Killing horizon $\mathcal{H}^+$. Near the horizon the Kruskal coordinate $U = -\mathrm{e}^{-\kappa u}$ is an affine distance along n and we can use it to measure the distance between γ and γ_H. In order to find the form of the wave at $\mathscr{I}^-$ we need to understand how the affine distance along the connecting vector n will change by the time γ reaches $\mathscr{I}^-$. At $\mathscr{I}^-$ the cooridnate v is an affine parameter aong the null geodesic integral curves of n. If $U_0 = 0$ then the affine distance is zero at $\mathscr{I}^-$. Hence we can expand the affine distance between γ and γ_H at $\mathscr{I}^-$ in powers of U_0: $v = cU_0 + \mathcal{O}(U_0^2)$ for some constant $c > 0$. Using $u = -\kappa^{-1}\log(-U) = -\kappa^{-1}\log(-cv)$ we can conclude that if a mode takes the form $g_\omega \sim \mathrm{e}^{-\mathrm{i}\omega u}$ on $\mathcal{I}^+$, the transmitted part $g^{(t)}_\omega$ on $\mathscr{I}^-$ will take the form

$$g^{(t)}_\omega \sim \begin{cases} \mathrm{e}^{\mathrm{i}\omega/\kappa \log(-v)} & \text{for} \quad v < 0 \\ 0 & \text{for} \quad v > 0 \end{cases} \tag{9.111}$$

up to a constant phase. This is exactly analogous to the Rindler modes in the previous section with $\kappa \leftrightarrow a$. We have $A_{\omega\omega'} = \mathrm{e}^{-\pi\omega/\kappa} B_{\omega\omega'}$ and therefore

$$\langle N_\omega \rangle \propto \frac{1}{\mathrm{e}^{\hbar\omega/(k_B T)} - 1}\,, \tag{9.112}$$

where the *Hawking temperature* is given by

$$T = \frac{\hbar\kappa}{2\pi k_B}\,. \tag{9.113}$$

Since the temperature is inversely proportional to the mass, the black hole hets up as it evaporates.

9.6 Black hole evaporation

If a black hole has a temperature it must evaporate. This leads to a serious problem with unitarity. We can compute the rate of mass loss due to the Hawking radiation. Stefan's law for the rate of energy loss by a blackbody:

$$\frac{\mathrm{d}E}{\mathrm{d}t} \sim -\alpha A T^4\,, \tag{9.114}$$

Plugging in $E = M$ and $A \propto M^2$ and $T \propto M^{-1}$ we have

$$\frac{\mathrm{d}M}{\mathrm{d}T} \propto -\frac{1}{M^2}\,, \tag{9.115}$$

and hence the black hole evaporates away completely in a time

$$\tau \sim \frac{G_N^2}{\hbar c^4} M^3\,, \tag{9.116}$$

note that the calculation of Hawking radiation assumed no backreaction, M was taken to be constant. This is good when $\frac{\mathrm{d}M}{\mathrm{d}t} \ll M$ but fails in the final stages of evaporation.

Consider a black hole which forms from collapsing matter and then evaporates away completely, leaving just thermal radiation. It should be possible to arrange that the collapsing matter is in a definite quantum state $|\psi\rangle$, the associated density matrix would be the one of a pure state, $\rho = |\psi\rangle\langle\psi|$. When the black hole is formed the Hilbert space naturally splits into the tensor product of Hilbert spaces, one with support in the interior of the black hole and the other with support on the exterior of the black hole: $\mathcal{H} = \mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{out}}$. An outside

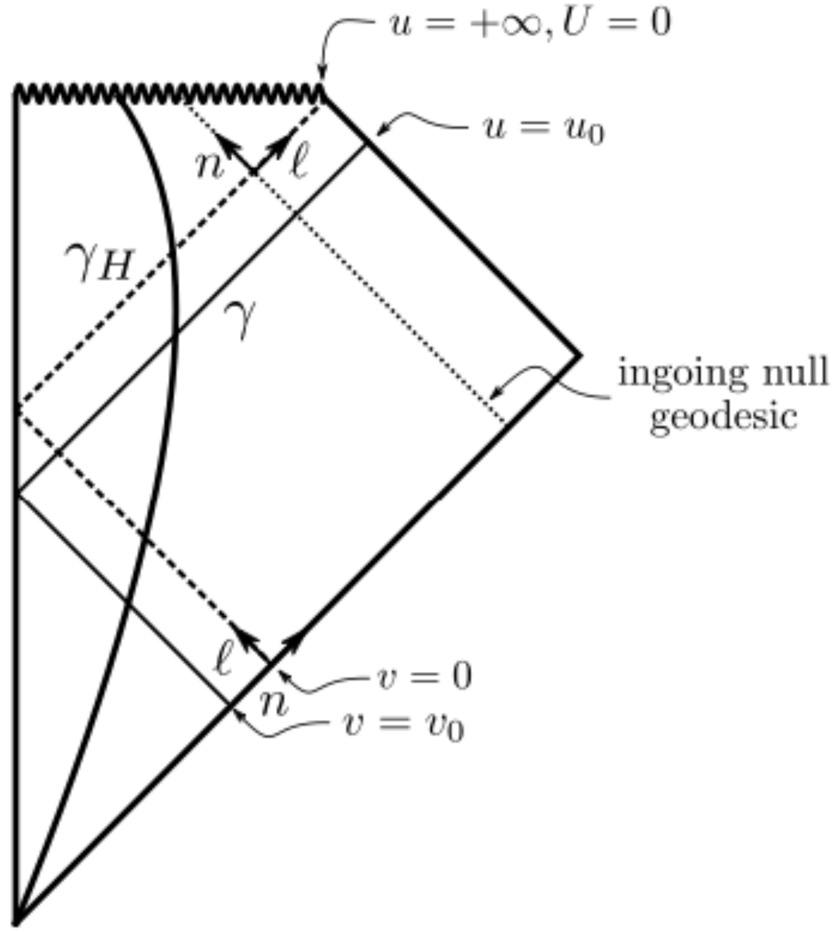

Figure 33: The evolution of the modes.

observer does not have access to $\mathcal{H}_{\text{in}}$ so their description of the black hole state is necessarily incomplete. They will describe the state outside the horizon as a reduced density matrix obtained by tracing over $\mathcal{H}_{\text{in}}$: $\rho_{\text{out}} = \text{tr}_{\text{in}}\rho$.

Since it described by a non-trivial density matrix the outside state is mixed. This is consistent with the fact that it contains thermal radiation, so there is no issue so far. The external state is entangled with the interior and the reduced density matrix is just a way in which the outside observer parametrises their ignorance of the interior. If we assume that the black hole has completely evaporated nothing is left in the interior and the exterior reduced density matrix will describe the full state, which is therefore a mixed state. However evolution from a pure state to a mixed state is forbidden by unitarity in quantum mechanics.

This is the black hole information paradox. It is important to emphasise the difference between thermal radiation produced in ordinary processes which do not violate unitarity. If we burn a printed copy of these lecture notes, thermal radiation is produced, however the process is unitary and in principle one could reconstruct all the information contained in the notes by studying the radiation and ashes. The early radiation is entangled with excitations inside the burning body, however the excitations inside the burning body can still transmit information to the radiation emitted later on which will thus contain non-trivial information. On the other hand, throwing the notes into a black hole, the information appears to be really lost once the black hole has fully evaporated because the final radiation is exactly thermal. The internal excitations are shielded by the horizon and by causality cannot influence the later outgoing radiation.

Nearly half a century after Hawking formulated the black hole information paradox it is still and open and active area of research. Our analysis has been in a funny hybrid theory of quantum field theory coupled to classical general relativity. General relativity predicts a singularity at the centre of a black hole, this is a regime where quantum effects will dramatically alter our classical expectations. **We need a quantum theory of gravity.**

Acknowledgments

It is a pleasure to thank all the students who have attended these lectures. You have made this an enjoyable course to teach and all your questions are much appreciated. I hope that you have enjoyed the course and learnt some of the amazing things that GR has to offer.

I am indebted to Alice Lüscher for spotting many typos in the lecture notes, of which many probably still remain.

A A short review of GR1

To keep our conventions in order we will briefly review the essential material from GR1. For those who have done a GR course but not studied manifolds I recommend consulting the GR1 notes as manifolds will appear at times in the lectures.

A.1 Manifolds

The underlying structure of General relativity is differential geometry. This is the study of manifolds.

Definition Let X be any set and $\mathcal{T} = \{U_i | i \in I\}$ denote a certain collection of subsets of X. The pair $(X, \mathcal{T})$ is called a *topological space* if $\mathcal{T}$ satisfies

1. Both the set X and the empty set $\emptyset$ are open subsets: $M \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$.
2. If $\mathcal{T}$ is any, possibly infinite, subcollection of I, then the family $\{U_j | j \in J\}$ satisfies $\cup_{j \in J} U_j \in \mathcal{T}$.
3. If K is any finite subcollection of I then the set $\{U_k | k \in K\}$ satisfies $\cap_{k \in K} U_k \in \mathcal{T}$.

Definition M is an n-dimensional *differentiable manifold* if satisfies:

1. M is a Hausdorff topological space,
2. M is provided with a family of pairs $\{(U_i, \varphi_i)\}$;
3. $\{U_i\}$ is a family of open sets which covers M: $\cup_i U_i = M$.
4. φ_i is a homeomorphism from U_i onto an open subset U_i' of $\mathbb{R}^n$,
5. Given U_i and U_j such that $U_i \cap U_j \neq \emptyset$, then the map $\psi_{ij} = \varphi_i \circ \varphi_j^{-1}$ from $\varphi_j(U_i \cap U_j)$ to $\varphi_i(U_i \cup U_j)$ is infinitely differentiable. ψ_{ij} is known as a *transition function.*

Differentiable maps Let $f : M \to N$ be a map from an m-dimensional manifold M to an n-dimensional manifold N. A point $p \in M$ is mapped to a point $f(p) \in N$. We may take a chart (U, φ) on M and a chart (V, ψ) in N where for all $p \in U$, $f(p) \in V$. Then f has the following coordinate presentation:

$$\psi \circ f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}^n \,. \tag{A.1}$$

If we write $\varphi(p) = \{x^\mu\}$ and $\psi\big(f(p)\big) = \{y^\alpha\}$ then, $\psi \circ f \circ \varphi^{-1}$ is just the usual vector-valued function $y = \psi \circ f \circ \varphi^{-1}(x)$ of m variables. Sometimes it is useful to abuse notation and write $y = f(x)$ or $y^\alpha = f^\alpha(x^\mu)$ when we know the coordinate systems on M and N that are in use.

Definition We say that a function $f : M \to \mathbb{R}$ is *smooth* if the map $f \circ \varphi^{-1} : U \to \mathbb{R}$ is smooth for all charts. We let the set of all small functions on M be denoted by $\mathcal{F}(M)$.

Definition We say that a map $f : M \to N$ between two manifolds is smooth if the map $\psi \circ f \circ \varphi^{-1} : U \to V$ is smooth for all charts $\varphi : M \to \mathbb{R}^m$ and $\psi : N \to \mathbb{R}^n$. If $y = \psi \circ f \circ \varphi^{-1}(x)$ is C^∞ then we say that f is *differentiable* at p. This is actually independent of the coordinate system.

Definition Let $f : M \to N$ be a homeomorphism and ψ and φ coordinate functions. If $\psi \circ f \circ \varphi^{-1}$ is invertible, f is called a *diffeomorphism* and M is said to be *diffeomorphic* to N and vice-versa. This is denoted by $M \equiv N$.

Since the map is invertible it follows that if $M \equiv N$ then $\dim M = \dim N$. Homeomorphisms classify spaces according to whether it is possible to deform one space into another *continuously*. *Diffeomorphisms* classify spaces into equivalence classes according to whether it is possible to deform one space into the other *smoothly*. As such a diffeomorphism is stronger than a homeomorphism, it requires that both the map and its inverse are smooth. Two diffeomorphic manifolds are viewed as the same manifold.

Tangent vectors We can define curves on our manifold, $\gamma : (a, b) \to M$ and the tangent to such a curve. If we collect all curves passing through the point p and find all tangent vectors to the point p, this defines the *tangent space* at p: $T_p(M)$ which is a vector space. A basis of the tangent space is given by

$$\{e_\mu\} = \left\{ \frac{\partial}{\partial x^\mu} \right\}, \tag{A.2}$$

and any vector field X may be expanded in terms of this basis as

$$X = X^\mu \frac{\partial}{\partial x^\mu} . \tag{A.3}$$

When we are looking at vector fields in $T_p(M)$ the X^μ are just numbers, however we can equally consider the tangent bundle which is the union of all tangent spaces in M. Then a vector field in the tangent bundle has X^μ which are functions on M.

Let U_i, j be two coordinate patches with coordinates $x = \varphi_i(p)$ and $y = \varphi_j(p)$ respectively and let $p \in U_i \cup U_j$. Then we can give the vector field X in both sets of coordinates and we have that

$$\frac{\partial}{\partial x^\mu} = \frac{\partial y^\nu}{\partial x^\mu} \frac{\partial}{\partial y^\nu} , \tag{A.4}$$

and therefore the components of the vector field X transform as

$$X = X^\mu \frac{\partial}{\partial x^\mu} = \tilde{X}^\mu \frac{\partial}{\partial y^\mu} \quad \Rightarrow \quad \tilde{X}^\mu = X^\nu \frac{y^\mu}{x^\nu} . \tag{A.5}$$

One-forms Since $T_p(M)$ is a vector space there exists a dual vector space whose element is a linear function $T_p(M) \to \mathbb{R}$. The dual space is called the *cotangent space* at p, and denoted $T_p^*(M)$. An element $\omega \in T_p^*(M)$ is a linear map $T_p(M) \to \mathbb{R}$ and is called a *cotangent vector, dual vector* or *one-form.*

The natural basis of the cotangent space is given by the differential of the coordinates: $\{\mathrm{d}x^\mu\}$. Using the bilinear map arising from the tangent and cotangent spaces being dual vector spaces, one takes

$$\left\langle \mathrm{d}x^\mu, \frac{\partial}{\partial x^\nu} \right\rangle = \delta^\mu_\nu \,. \tag{A.6}$$

An arbitrary one-form can then be expanded out in this basis as $\omega = \omega_\mu \mathrm{d}x^\mu$. Let us take $p \in U_i \cup U_j$ as before, then for $\omega \in T_p^*(M)$ we have

$$\omega = \omega_\mu \mathrm{d}x^\mu = \tilde{\omega}_\mu \mathrm{d}y^\mu \quad \Rightarrow \quad \tilde{\omega}_\nu = \omega_\mu \frac{\partial x^\mu}{\partial y^\nu} \,. \tag{A.7}$$

Tensors We can now define tensors of type (q, r) to be a multilinear object which maps q elements of $T_p^*(M)$ and r elements of $T_p(M)$ to $\mathbb{R}$. We denote the set of (q, r) tensors at p to be $\mathcal{T}_p^{(q,r)}(M)$. An element of $\mathcal{T}^{(q,r)}(M)$ can be written in terms of the bases described above as

$$T = T^{\mu_1 \dots \mu_q}{}_{\nu_1 \dots \nu_r} \frac{\partial}{\partial x^{\mu_1}} \dots \frac{\partial}{\partial x^{\mu_q}} \mathrm{d}x^{\nu_1} \dots \mathrm{d}x^{\nu_r} \,. \tag{A.8}$$

T is a linear function

$$T : \otimes^q T_p^*(M) \otimes^r T_p(M) \to \mathbb{R} \,. \tag{A.9}$$

Let $V_i = V_i^\mu \frac{\partial}{\partial x^\mu}$ with $1 \leq i \leq r$ and $\omega_j = \omega_{j\mu} \mathrm{d}x^\mu$ with $1 \leq j \leq q$ then the action of T is

$$T(\omega_1, ..., \omega_q; V_1, ...V_r) = T^{\mu_1 \dots \mu_q}{}_{\nu_1 \dots \nu_r} \, \omega_{1\mu_1} .... \omega_{q\mu_q} V_1^{\mu_1} .... V_r^{\mu_r} \,. \tag{A.10}$$

Tensor fields So far we have defined vectors, one-forms and tensors at a particular point $p \in M$. We want to be able to smoothly assign such an object to every point of M. For a vector we call such an object a *vector field.* In other words if V is a vector field then for every $f \in \mathcal{F}(M)$ then $V[f] \in \mathcal{F}(M)$. We will denote the set of all vector fields on M as $\mathcal{X}(M)$. A vector field X at $p \in M$ is denoted by $X|_p$ which is an element of $T_p(M)$. Similarly we may define a *tensor field* of type (q, r) by a smooth assignment of an element of $\mathcal{T}^q_{r,p}(M)$ at each point $p \in M$. The set of tensor fields of type (q, r) on M is denoted by $\mathcal{T}^q_r(M)$.

Differential forms A *differential form* of order r, or more succinctly an r-form, is a totally anti-symmetric tensor of type $(0, r)$.

The *Wedge product* $\wedge$ of r one-forms is defined to be the totally anti-symmetric tensor product of the one-forms

$$\mathrm{d}x^{\mu_1} \wedge \mathrm{d}x^{\mu_2} \wedge ...\mathrm{d}x^{\mu_r} \equiv \sum_{P \in S_r} \mathrm{sgn}(P) \mathrm{d}x^{\mu_{P(1)}} \otimes \mathrm{d}x^{\mu_{P(2)}} \otimes .... \otimes \mathrm{d}x^{\mu_{P(r)}} \,. \tag{A.11}$$

Thus

$$\mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} = \mathrm{d}x^{\mu} \otimes \mathrm{d}x^{\nu} - \mathrm{d}x^{\nu} \otimes \mathrm{d}x^{\mu} \,. \tag{A.12}$$

The wedge product satisfies the following conditions

- $\mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_r} = 0$ if some index is repeated.
- $\mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_r} = \mathrm{sgn}(P) \mathrm{d}x^{\mu_{P(1)}} \wedge ... \wedge \mathrm{d}x^{\mu_{P(r)}}$.
- $\mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_r}$ is linear in each $\mathrm{d}x^{\mu}$.

We will denote the vector space of r-forms at the point $p \in M$ by $\Omega^r_p(M)$, a basis is provided by the set of all wedge products in (A.11). We can then expand an element of $\Omega^r_p(M)$ as

$$\omega = \frac{1}{r!} \omega_{\mu_1 \cdots \mu_r} \mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_r} \,, \tag{A.13}$$

where $\omega_{\mu_1 \cdots \mu_r}$ are taken to be totally anti-symmetric.

We may define the *exterior product* to be the map $\wedge : \Omega^q_p(M) \times \Omega^r_p(M) \to \Omega^{q+r}_p(M)$. Its action follows by trivial extension of the wedge product defined above. Let $\omega \in \Omega^q_p(M)$ and $\xi \in \Omega^r_p(M)$ be an q-form and and r-form respectively. The action of the $(q+r)$-form $\omega \wedge \xi$ on $q+r$ vectors V_i is

$$(\omega \wedge \xi)(V_1, ..., V_{q+r}) = \frac{1}{q!r!} \sum_{P \in S_{q+r}} \mathrm{sgn}(P) \omega\big(V_{P(1)}, ..., V_{P(q)}\big) \xi\big(V_{P(q+1)}, ..., V_{P(q+r)}\big) \,. \tag{A.14}$$

The *exterior deriavtive* d_r is a map $\Omega^r(M) \to \Omega^{r+1}(M)$, whose action on an r-form

$$\omega = \frac{1}{r!} \omega_{\mu_1 \cdots \mu_r} \mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_r} \,, \tag{A.15}$$

is

$$\mathrm{d}_r \omega = \frac{1}{r!} \left(\frac{\partial}{\partial x^{\nu}} \omega_{\mu_1 \cdots \mu_r} \right) \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_r} \,. \tag{A.16}$$

It is common to drop the r subscript and simply write d. The wedge product automatically anti-symmetrises the coefficient so it is indeed a $(r+1)$-form that we obtain. It follows that for $\xi \in \Omega^q_p(M)$, $\eta \in \Omega^r_p(M)$ we have

$$\mathrm{d}(\xi \wedge \eta) = \mathrm{d}\xi \wedge \eta + (-1)^q \xi \wedge \mathrm{d}\eta \,. \tag{A.17}$$

The exterior derivative satisfies $\mathrm{d}^2 = 0$.

Let X be a vector field and $\omega \in \Omega^r(M)$ then the *interior product* of the r-form ω with respect to the vector X is

$$\mathrm{i}_X\omega(X_1, ..., X_{r-1}) \equiv \omega(X, X_1, ..., X_{r-1})\,. \tag{A.18}$$

If we introduce coordinates: $X = X^\mu \frac{\partial}{\partial x^\mu}$ then

$$\mathrm{i}_X\omega = \frac{1}{(r-1)!} X^\nu \omega_{\nu\mu_1\ldots\mu_{r-1}} \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_{r-1}}\,. \tag{A.19}$$

A.2 Riemannian geometry

Definition: Let M be a differentiable manifold. A *Riemannian metric* g on M is a type $(0,2)$ tensor field on M which at each point $p \in M$ satisfies

- Symmetric: $g_p(X, Y) = g_p(Y, X)$,
- $g_p(X, X) \geq 0$ with equality iff $X = 0$

with $X, Y \in T_p(M)$. A tensor field g of type $(0,2)$ is a *pseudo-Riemannian metric* if it satisfies the first condition and

- Non-degenerate. If for any $p \in M$ $g_p(X, Y) = 0$ for all $Y \in T_p(M)$ then $X_p = 0$,

We may extend the tensor g_p over the full manifold. With a choice of coordinates we can write the metric as

$$g = g_{\mu\nu}(x)\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu\,. \tag{A.20}$$

We will often write this as the line elements $\mathrm{d}s^2$,

$$\mathrm{d}s^2 = g_{\mu\nu}(x)\mathrm{d}x^\mu\mathrm{d}x^\nu\,. \tag{A.21}$$

We may view $g_{\mu\nu}$ as a matrix, which by the symmetry property above is symmetric. This implies that the matrix is diagonalisable, with real eigenvalues. If there are i positive eigenvalues and j negative eigenvalues the pair (i, j) is called the *index* of the metric. If $j = 1$ the metric is called a *Lorentz metric*, for $j = 0$ we have a *Euclidean* metric. The number of negative entries is called the *signature* and by Sylvester's law of inertia[26], this is independent of the choice of basis.

[26] This has nothing to do with inertia, Sylvester just wanted a law of inertia like Newton.

Lorentzian manifolds For our purposes Riemannian manifolds are not what we want to consider, instead we want to consider Lorentzian manifolds. The simplest example is Minkowski space. This is $\mathbb{R}^{1,m-1}$ equipped with the metric

$$\eta = -\mathrm{d}x^0 \otimes \mathrm{d}x^0 + \mathrm{d}x^1 \otimes \mathrm{d}x^1 + ... + \mathrm{d}x^{m-1} \otimes \mathrm{d}x^{m-1}\,, \tag{A.22}$$

which has components $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, ..., 1)$. Note that on a Lorentzian manifold we take the index to run over $0, 1, .., m-1$.

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

$$g_{\mu\nu}|_p = \eta_{\mu\nu}\,. \tag{A.23}$$

This is closely related to the equivalence principle (see later).

The fact that locally the metric looks locally like Minkowski space allows us to import some of the ideas of special relativity, namely we can classify the elements of $T_p(M)$ into three classes

- $g(X, X) > 0 \longrightarrow X$ is *spacelike* ,
- $g(X, X) = 0 \longrightarrow X$ is *lightlike* or *null* ,
- $g(X, X) < 0 \longrightarrow X$ is *timelike* .

At each point on M we can then draw light cones which are the null tangent vectors at that point. The novelty is that the directions of these light cones can vary smoothly as we move around the manifold. This specifies the causal structure of spacetime which determines which regions of spacetime can interact together.

We can use the metric to determine the length of curves. The nature of a curve is inherited from the nature of its tangent vector. A curve is called *timelike* if its tangent vector is everywhere timelike. We then measure the proper time

$$\tau = \int_a^b \mathrm{d}t \sqrt{-g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}}\,. \tag{A.24}$$

The existence of a metric comes with a large number of benefits.

The metric as an isomorphism The metric gives a natural isomorphism between vectors and covectors, $g : T_p(M) \to T_p^*(M)$ for each p. In a coordinate basis we can write $X = X^\mu \partial_\mu$, and map it to a one-form $X = X_\mu \mathrm{d}x^\mu$, as

$$X_\mu = g_{\mu\nu} X^\nu\,. \tag{A.25}$$

We will usually say that we use the metric to lower (or raise) an index. What we really mean is that the metric provides and isomorphism between a vector space and its dual. Since g is non-degenerate and is thus invertible we also have the inverse map. We take the inverse of $g_{\mu\nu}$ to be $g^{\mu\nu}$ so that $g^{\mu\nu}g_{\nu\rho} = \delta^{\mu}_{\rho}$. This can then be thought of as the components of a symmetric $(2,0)$ tensor

$$\hat{g} = g^{\mu\nu}\partial_{\mu} \otimes \partial_{\nu} \,. \tag{A.26}$$

Then

$$X^{\mu} = g^{\mu\nu}X_{\nu} \,. \tag{A.27}$$

The Volume form The metric also gives a natural volume form on the manifold M. On a Riemannian manifold we take the volume form to be

$$\mathrm{vol}(M) = \sqrt{\det(g_{\mu\nu})}\mathrm{d}x^1 \wedge ...\mathrm{d}x^m \,, \tag{A.28}$$

and we use the shorthand $\sqrt{\det(g_{\mu\nu})} = \sqrt{g}$. On a Lorentzian manifold the determinant is negative and therefore we take the volume form to be

$$\mathrm{vol}(M) = \sqrt{-g}\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^{n-1} \,. \tag{A.29}$$

As it is written it looks coordinate dependent however it is not.

Hodge dual On an oriented manifold M we can use the totally anti-symmetric tensor density to define a map which takes a p-form $\omega \in \Omega^p(M)$ to a $(m-p)$-form $\star\omega \in \Omega^{m-p}(M)$. We define this map to be

$$(\star\omega)_{\mu_1...\mu_{m-p}} = \frac{1}{p!}\sqrt{|g|}\epsilon_{\mu_1...\mu_{m-p}\nu_1...\nu_p}\omega^{\nu_1..\nu_p} \,, \tag{A.30}$$

where $\epsilon_{\mu_1...\mu_m}$ is the totally anti-symmetric tensor, with $\epsilon_{123...m} = 1$ and for even permutations, -1 for odd permutations and 0 otherwise.

This is called the *Hodge dual* and is independent of coordinates. One can see that it satisfies

$$\star(\star\omega) = \pm(-1)^{p(m-p)}\omega \,, \tag{A.31}$$

with $+$ for a Riemannian metric and $-$ for a Lorentzian.

Connections An *affine connections* ∇ is a map $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$, $(X,Y) \mapsto \nabla_X Y$ which satisfies

$$\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z \,, \tag{A.32}$$

$$\nabla_{(fX+gY)}Z = f\nabla_X Z + g\nabla_Y Z\,, \tag{A.33}$$
$$\nabla_X(fY) = X[f]Y + f\nabla_X Y\,, \tag{A.34}$$

for vector fields $X, Y, Z \in \mathcal{X}(M)$ and functions $f, g \in \mathcal{F}(M)$.

We may introduce connection coefficients so that the connection acts on an arbitrary tensor of rank (q, r) as

$$\begin{aligned}\nabla_\mu T^{\nu_1\ldots\nu_q}{}_{\rho_1\ldots\rho_r} =& \frac{\partial}{\partial x^\mu} T^{\nu_1\ldots\nu_q}{}_{\rho_1\ldots\rho_r} + \Gamma^{\nu_1}{}_{\mu\sigma}T^{\sigma\ldots\nu_q}{}_{\rho_1\ldots\rho_r} + ... + \Gamma^{\nu_q}{}_{\mu\sigma}T^{\nu_1\ldots\nu_{q-1}\sigma}{}_{\rho_1\ldots\rho_r} \\ &- \Gamma^{\sigma}{}_{\mu\rho_1}T^{\nu_1\ldots\nu_q}{}_{\sigma\ldots\rho_r} - \ldots - \Gamma^{\sigma}{}_{\mu\rho_r}T^{\nu_1\ldots\nu_q}{}_{\rho_1\ldots\rho_{r-1}\sigma}\;.\end{aligned} \tag{A.35}$$

In words, you first differentiate the tensor and then for each upper index you add in a $+\Gamma T$ and for every down index a $-\Gamma T$. The connection takes tensors to tensors, the (q, r) tensor gets mapped to a $(q, r + 1)$ tensor.

The connection coefficients are not tensors themselves, but transform as

$$\tilde{\Gamma}^\mu{}_{\nu\rho} = (\Lambda^{-1})^\mu{}_\kappa \Lambda^\sigma{}_\rho \Lambda^\tau{}_\nu \Gamma^\kappa{}_{\sigma\tau} + (\Lambda^{-1})^\mu{}_\kappa \Lambda^\sigma{}_\rho \partial_\sigma \Lambda^\kappa{}_\nu\,, \qquad \text{with} \qquad \Lambda^\mu{}_\nu = \frac{\partial y^\mu}{\partial x^\nu}\,. \tag{A.36}$$

The difference

$$T^\kappa{}_{\sigma\tau} = \Gamma^\kappa{}_{\sigma\tau} - \Gamma^\kappa{}_{\tau\sigma}\,, \tag{A.37}$$

is called the torsion tensor, and is indeed a tensor. If the torsion tensor vanishes we say that the connection is *torsion free*.

Levi–Civita connection Given a metric there we have:

Theorem There exists a unique, torsion free, connection that is compatible with the metric g:

$$\nabla_X g = 0\,, \tag{A.38}$$

for all vector fields X.

The connection compatible with the metric is called the *Levi–Civita connection* while the components of the Levi–Civita connection are called the *Christoffel symbols* and are given by:

$$\Gamma^\lambda{}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}\big(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu} - \partial_\rho g_{\mu\nu}\big)\,. \tag{A.39}$$

Given a vector field X which is tangent to the curve γ with coordinates x^μ, we say that a tensor field T is parallel transported along γ if

$$\nabla_X T = 0\,. \tag{A.40}$$

Let γ connect two points $p, q \in M$. The condition (A.40) provides a map from the vector space defined at p to the vector space defined at q. Consider a second vector field Y. In components (A.40) reads

$$X^\nu\big(\partial_\nu Y^\mu + \Gamma^\mu{}_{\nu\rho} Y^\rho\big) = 0\,. \tag{A.41}$$

If we evaluate it on the curve γ, we can write $Y^\mu = Y^\mu(x(\lambda))$ and therefore the condition is

$$\frac{\mathrm{d}Y^\mu}{\mathrm{d}\lambda} + X^\nu \Gamma^\mu{}_{\nu\rho} Y^\rho\,. \tag{A.42}$$

A *geodesic* is a curve tangent to a vector field X that obeys

$$\nabla_X X = 0\,. \tag{A.43}$$

Along the curve γ with coordinates x^μ and tangent vector X this implies

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu{}_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda} = 0\,. \tag{A.44}$$

This is the same geodesic equation one obtains by varying the action

$$S = \int \mathrm{d}\lambda \sqrt{-g_{\mu\nu}(x)\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}\,, \tag{A.45}$$

and picking an affine parameter.

Using the Levi–Civita connection we can define the curvature and torsion tensors. In components the Riemann tensor is

$$R^\sigma{}_{\rho\mu\nu} = \partial_\mu \Gamma^\sigma{}_{\nu\rho} - \partial_\nu \Gamma^\sigma{}_{\mu\rho} + \Gamma^\lambda{}_{\nu\rho}\Gamma^\sigma{}_{\mu\lambda} - \Gamma^\lambda{}_{\mu\rho}\Gamma^\sigma{}_{\nu\lambda}\,. \tag{A.46}$$

It has the following symmetries and properties

$$R^\sigma{}_{\rho\mu\nu} = -R^\sigma{}_{\rho\nu\mu}\,, \tag{A.47}$$

$$R_{\mu\nu\rho\sigma} = R_{\sigma\rho\mu\nu}\,, \tag{A.48}$$

$$R_{\mu[\nu\rho\sigma]} = 0\,, \tag{A.49}$$

$$\nabla_{[\mu} R_{\sigma\rho]\tau\nu} = 0\,. \tag{A.50}$$

Given a rank $(1,3)$ tensor we can construct a rank $(0,2)$ tensor by contraction, for the Riemann tensor the resultant $(0,2)$-rank tensor is called the *Ricci* tensor and is defined by

$$R_{\mu\nu} = R^\rho{}_{\mu\rho\nu}\,. \tag{A.51}$$

It inherits symmetry in its indices from the properties of the Riemann tensor

$$R_{\mu\nu} = R_{\nu\mu}\,. \tag{A.52}$$

We can create a scalar by contracting the indices again

$$R = g^{\mu\nu} R_{\mu\nu}\,. \tag{A.53}$$

A.3 Einstein's equation

The Einstein–Hilbert action is

$$S_{\mathrm{EH}} = \int \mathrm{d}^4x\sqrt{-g}R\,. \tag{A.54}$$

Variation with respect to the metric gives Einstein's field equations

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0\,. \tag{A.55}$$

A cosmological constant term may be added to the action

$$S = \frac{1}{16\pi G_N}\int \mathrm{d}^4x\sqrt{-g}(R-2\Lambda)\,. \tag{A.56}$$

Varying the action as before yields the Einstein equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = -\Lambda g_{\mu\nu}\,. \tag{A.57}$$

Coupling to matter We can couple gravity to matter. We do this via minimal coupling. We replace covariant derivatives with the connection, add in the correct volume measure and insert a metric for summed space-time indices.

We need to consider the combined action

$$S = \frac{1}{16\pi G_N}\int \mathrm{d}^4x\sqrt{-g}(R-2\Lambda) + S_{\mathrm{Matter}}\,, \tag{A.58}$$

where S_{Matter} is the action for any matter fields in the theory minimally coupled to gravity. The *Energy-Momentum tensor* is defined to be

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_{\mathrm{Matter}}}{\delta g^{\mu\nu}}\,. \tag{A.59}$$

A.4 Schwarzschild solution

The Scwarzschild solution is

$$\mathrm{d}s^2 = -\left(1-\frac{2G_NM}{r}\right)\mathrm{d}t^2 + \left(1-\frac{2G_NM}{r}\right)^{-1}\mathrm{d}r^2 + r^2\big(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\big)\,. \tag{A.60}$$

This solves Einstein's equations in a vacuum, $R_{\mu\nu} = 0$.

Birkhoff's theorem The Schwarzschild solution is the unique spherically symmetric asymptotically flat solution to the vacuum Einstein equations.

New coordinates The Schwarzschild solution in Schwarzschild coordinates has a coordinate singularity at $r = R_s = 2G_N M$. This surface is called the *event horizon*. In GR no signals can come out from within the event-horizon, once you fall past the event horizon you are lost to the outside world.

The apparent singularity at $r = R_s$ is only a coordinate singularity and can be removed by a coordinate transformation. First introduce the *tortoise coordinate* r_*

$$r_* = r + 2G_N M \log\left(\frac{r - 2G_N M}{2G_N M}\right), \tag{A.61}$$

then in these coordinates the null radial in-going/out-going geodesics are particularly simple:

$$t = \pm r_* + \text{constant}\,. \tag{A.62}$$

Next introduce a pair of null coordinates further adapted to the null geodesics:

$$v = t + r_*\,, \qquad u = t - r_*\,. \tag{A.63}$$

Ingoing Eddington–Finkelstein coordinates Eliminating t via $t = v - r_*(r)$, known as ingoing Eddington–Finkelstein coordinates, we find

$$\mathrm{d}s^2 = -\left(1 - \frac{2G_N M}{r}\right)\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r + r^2\mathrm{d}s^2(S^2)\,. \tag{A.64}$$

Even though the metric coefficient g_{vv} vanishes at $r = 2G_N M$ there is no real degeneracy there and the metric is well-defined as one can see by computing the determinant.

There is also the complementary *outgoing Eddington–Finkelstein coordinates* where we eliminate t using u above. With Eddington–Finkelstein coordinates we are able to continue the Schwarzschild solution beyond the horizon to $r > 0$. In fact there are two ways to do this with either the ingoing or outgoing Eddington–Finkelstein coordinates. In fact we can do better and write a metric which captures both of these regions simultaneously.

To begin write the Schwarzschild metric using both null (u, v)-coordinates, the metric is

$$\mathrm{d}s^2 = -\left(1 - \frac{2G_N M}{r}\right)\mathrm{d}u\mathrm{d}v + r^2\mathrm{d}s^2(S^2)\,, \tag{A.65}$$

where r is a function of $u - v$. In these coordinates the metric is again degenerate at $r = R_s$ so we need to perform another change of coordinates. We introduce the *Kruskal-Szekeres coordinates*,

$$U = -\exp\left(-\frac{u}{4G_N M}\right), \qquad V = \exp\left(\frac{v}{4G_N M}\right), \tag{A.66}$$

both are null coordinates. The original Schwarzschild black hole is parametrised by $U < 0$ and $V > 0$. Outside the horizon they satisfy

$$UV = -\exp\left(\frac{r_*}{2G_N M}\right) = \frac{2G_N M - r}{2G_N M}\exp\left(\frac{r}{2G_N M}\right), \tag{A.67}$$

and similarly

$$\frac{U}{V} = -\exp\left(-\frac{t}{2G_N M}\right). \tag{A.68}$$

The metric is then

$$\mathrm{d}s^2 = -\frac{32(G_N M)^3}{r}\mathrm{e}^{-\frac{r}{2G_N M}}\mathrm{d}U\mathrm{d}V + r^2\mathrm{d}s^2(S^2), \tag{A.69}$$

with $r(U,V)$ defined by inverting (A.67). The original Schwarzschild metric covers just $U < 0$ and $V > 0$ however there is no obstruction to extending $U, V \in \mathbb{R}$. Nothing bad happens at $r = 2G_N M$, the metric is smooth and non-degenerate and now we have a metric which covers all regions. The Kruskal spacetime is the maximal extension of the Schwarzschild solution.

References

[1] S. M. Carroll, *Spacetime and Geometry*. Cambridge University Press, 7, 2019.

[2] B. F. Schutz, *A FIRST COURSE IN GENERAL RELATIVITY*. Cambridge Univ. Pr., Cambridge, UK, 1985, 10.1017/CBO9780511984181.

[3] R. M. Wald, *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984, 10.7208/chicago/9780226870373.001.0001.

[4] A. Buonanno, *Gravitational waves*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime*, 9, 2007. 0709.4682.

[5] N. Straumann, *General Relativity*. Graduate Texts in Physics. Springer, Dordrecht, 2013, 10.1007/978-94-007-5410-2.

[6] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford University Press, 2007, 10.1093/acprof:oso/9780198570745.001.0001.

[7] M. Maggiore, *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*. Oxford University Press, 3, 2018.

[8] S. M. Christensen and M. J. Duff, *Quantizing Gravity with a Cosmological Constant*, *Nucl. Phys. B* **170** (1980) 480–506.

[9] P. C. Peters and J. Mathews, *Gravitational radiation from point masses in a Keplerian orbit*, *Phys. Rev.* **131** (1963) 435–439.

[10] P. C. Peters, *Gravitational Radiation and the Motion of Two Point Masses*, *Phys. Rev.* **136** (1964) B1224–B1232.

[11] C. Caprini and D. G. Figueroa, *Cosmological Backgrounds of Gravitational Waves*, *Class. Quant. Grav.* **35** (2018) 163001, [1801.04268].

[12] J. Arrechea, C. Barceló, G. García-Moreno and J. Polo-Gómez, *Dissecting Buchdahl's limit: A surgeon's guide to compact objects*, 2411.14018.

[13] E. Annala, T. Gorda, A. Kurkela and A. Vuorinen, *Gravitational-wave constraints on the neutron-star-matter Equation of State*, *Phys. Rev. Lett.* **120** (2018) 172703, [1711.02644].

[14] C. Hoyos, N. Jokela and A. Vuorinen, *Holographic approach to compact stars and their binary mergers*, *Prog. Part. Nucl. Phys.* **126** (2022) 103972, [2112.08422].

[15] J. M. Bardeen, B. Carter and S. W. Hawking, *The Four laws of black hole mechanics*, *Commun. Math. Phys.* **31** (1973) 161–170.