UNIVERSITY OF PLYMOUTH

Mathematics

MATH3628

Gravitational Lensing

Author:
Thomas WOODLAND-SCOTT
{*10588263*}

Supervisor:
Dr. Thomas HEINZL

Figure 1: 12 images within the four arcs are visible, these are copies of the same galaxy called PSZ1 G311.65-18.48, which is almost 11 billion light-years away, [4].

UNIVERSITY OF PLYMOUTH

Acknowledgements

I would like to thank my supervisor Tom Heinzl, whose knowledge, guidance and patience has been greatly appreciated throughout and has greatly improved this project as a whole.

In addition, I would like thank Susie Scott for proof reading.

Abstract

Gravitational lensing occurs when light rays, emanating from a distant source, pass a massive object, gravity from this object causes the light rays to bend. This project presents the basics of gravitational lensing, deriving some key equations needed to understand the phenomenon. Sticking with the chronological development of gravitational lensing, we start by using Newtonian mechanics to derive a rudimentary approximation for a deflection angle of a light ray. To yield a better approximation, we make use of two main topics, optics and general relativity. In doing so, we introduce lensing by a point mass, a spherically symmetric mass and a general mass distribution, calculating the deflection angle of a passing light ray for each. Finally, we will discuss multiple imaged objects including the formation of Einstein rings and arcs.
The aim of this project it to give the reader a solid foundation in the mathematics of gravitational lensing as well an understanding of its importance in history. We assume that the reader is familiar with basic algebra, geometry and calculus.

Contents

1 Introduction

1.1 Motivation

Space is fascinating: how dots in the sky are actually extremely massive stars; the Sun, which is millions of miles away, has a significant effect on the planet we live on; and stars, planets and other matter in space move with shocking speeds. For example, to us, the earth appears stationary, however, it orbits the Sun at a speed close to 67,000 mph, [12]. Gravity permeates the universe and connects all masses: the moon orbits the earth; the earth orbits the Sun; and the Sun is one of an estimated 100 billion stars in the milky way galaxy alone. Eignstein's general theory of relativity allows us to use mathematics to explore many of the phenomena in space, including gravitational lensing. Gravitational lensing has had a critical part to play in accepting Einstein's theory. In addition, gravitational lensing has optical effects that are very interesting.

1.2 Gravitational Lensing

Gravitational lensing is a consequence of general relativity, where the gravitational force field produced by a massive object in space bends the path of light [3]. This project will explore how light propagates through space when perturbed by these gravitational fields and the effects this phenomenon produces. We consider the light path in three zones. The first zone starts at the source and corresponds to unperturbed spacetime, where the light ray travels in a straight line towards the lens. The second, and most interesting zone, where the light rays relatively close proximity to the lens, causes deflection. In the third zone, light rays are, again, in unperturbed spacetime, and thus are straight lines to the observer [10]. As we will see, the strength of the gravitational field is dependent on the distance from its source, hence the light ray that is deflected the most, is the light ray that passes at the minimum distance from the lens.

1.3 Optics

We will start chronologically, first making use of Newtonian mechanics (assuming that light takes the form of particles) to calculate a rudimentary approximation for the deflection angle caused by a mass. We shall confirm this approximation using Kibble's relation, tying the impact parameter of the light particle with the scattering angle, see Figure 3.
A major issue with the above derivations is clear. The assumption, that light behaves as a particle, we now know not to be completely true. We can more accurately describe light as light rays, where these rays are electromagnetic waves that obey Maxwell's equations, [13]. Hence, working from Maxwell's famous equations describing electromagnetism, we can derive the *eikonal equation* of geometrical optics, providing us with an important relationship between the physical wave optics and geometrical ray optics. This enabled us to form a ray equation. Using this equation along side the *refractive index* (a scalar quantity describing how fast light travels in a medium) of a gravitational field, we can calculate a solution to the eikonal equation and form an equation describing the motion of a light ray.
This result is supported by Fermat's principle. In addition, importantly, Fermat's principle leads

to the concept of *Geodesics*, the shortest connection between two points.
The simplest case of gravitational lensing is the lensing caused by a point mass. Einstein's general theory of relativity implies a specific refractive index of the gravitational field and gravitational potential, produced by a point mass lens. Knowing these values allows us to apply our geometrical optics formula to this specific case. This leads us to Einstein's historically important point mass deflection angle formula which is twice that of the Newtonian approximation.
In reality, the point mass approximation is limited as often lensing occurs by large clumps of matter such as galaxies. To increase our domain of validity we use Poisson's equation to generalise the light deflection formula of a point mass to that of a mass distribution. Poisson's equation is a vital equation in understanding gravity. Later, we will make a link directly between general relativity and the Poisson's equation.

1.4 General Relativity

To fill in some of the holes in logic, the refractive index and gravitational potential of a point mass, we must introduce general relativity. To introduce some key concepts such as the manifold and metric tensor, we stick with the historical time line and, first, briefly introduce Minkowski space-time.
The aim of introducing Einstein's theory is to re-derive Schwarzschild's solution, a solution to the Einstein vacuum field equations that describes spacetime when perturbed by a massive, non-rotating, spherically symmetric body. Starting with the special relativity metric, which we take as spherically symmetric, in spherical polar coordinates. We form a general spherically symmetric metric and calculate its respective Ricci tensor field and Ricci scalar field, thus we can find the famous spherically symmetric solution. We also make links between Newton's second law of motion with the geodesic equation in general relativity and Poisson's equation with the field equations.
From Schwarzschild's solution, we use the geodesic equation and formulate an expression for the deflection angle of a light ray in such a field. We go on to show that, under certain constraints, Newton's theory can still hold, and, under these constraints, we do in fact show the desired refractive index and Newtonian potential which we called upon when calculating the deflection angle with the use of optics.
To tie general relativity and optics together, we go back to the Schwarzschild solution with the aim of recovering Einstein's deflection angle once more.

1.5 Lensing Effects

One application of the mathematics discussed in the Optics and the General Relativity sections is to look at the images an observer would see as light is bent by gravitational fields.
Imposing conditions on the geometry of a system, the positioning of the source, lens and observer relative to each other, determines the image. Choosing a point mass or spherically symmetric source, which is radially omitting light, to be directly behind the lens, produces a ring shaped image, commonly know as an Einstein ring. Determining the radius of this ring is straight froward, assuming we know the positioning of the source, lens and observer.

1.6 History

In 1687, I. Newton unified Keplar's laws of planetary motion with Galilei's theory of falling bodies in *The Mathematical Principles of Natural Philosophy*, allowing for a first deflection angle approximation.

In 1905, A. Einstein invented the special theory of relativity, a theory defining the relationship between space and time. Special relativity is based on two postulates: (i) the special relativity principle stating that the laws of physics are invariant in all inertial frames of reference; and (ii) the principle of the constancy of the speed of light in vacuum. Amongst other things, special relativity explains the invariance of Maxwell's equations with respect to a change in inertial frame.

Two years later, A. Einstein formed the equivalence principle, showing a gravitational field is equivalent to an acceleration. For Newton's equation of motion we see, $(\textit{inertial mass}) \cdot (\textit{acceleration}) = (\textit{gravitational mass}) \cdot (\textit{Intensity of the gravitational field})$.

In 1908, H. Minkowski introduced *Minkowski space*, a four-dimensional formulation of special relativity.

In 1915, A. Einstein presented the formulation of general relativity with his gravitational field equation, generalising the special relativity theory and refining Newton's law of universal gravitation. One consequence of this formulation was a prediction of a light ray's angle of deflection when passing by the Sun (for a light ray grazing the surface of the Sun, this angle is 1.7 arcseconds, [11]). Later that same year, K. Schwarzschild derived the spherically symmetric static solution to Einstein's field equation in vacuum. This specific solution is known as the *Schwarzschild solution.*

In May 1919, an expedition headed by Sir A. Eddington was sent to the island of Príncipe, situated just off the west cost of Africa, to observe the position of stars in the Haydes cluster. At the time, this cluster was located near to the Sun during a solar eclipse. This expedition showed the image of these stars was off its expected position by an angle of 1.9 arcseconds. This was sufficiently accurate to help verify Einstein's claim. Today, the deflection formula has been verified to a relative accuracy of 0.02%[3] [11].

2 Newtonian Mechanics

2.1 Rudimentary Approximation of Deflection Angle

An approximation of the light deflection can be calculated using Newtonian theory. If we assume that light behaves like a continuous stream of massless particles, deflection can be calculated using Newton's theory of gravitation [1].

Definition: The *deflection angle*, α, is the angle between the trajectory of the unperturbed light ray and the trajectory of the light ray after lensing.

Example: Let us consider a light ray travelling along the x-axis that is perturbed by a spherically symmetric mass, i.e. the Sun, with mass M and radius R centred as seen in Figure 2 below. In this example, the x-axis is the unperturbed light ray.

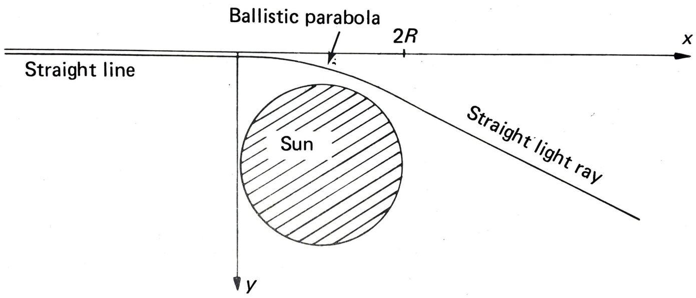

Figure 2: Deflection of a light ray grazing the surface of the Sun, [15].

On the light ray's path through the Sun's gravitational field, maximum deflection of the light ray will occur at the minimum distance from the Sun. We make the simplifying assumption that all lensing occurs in the range $0 < x < 2R$. Here we have two equations of motion: $\ddot{y} = g$, where $g = \frac{GM}{R^2}$ is the acceleration due to gravity; and $\ddot{x} = 0$, since any change will be negligibly small. The solution to the equation of motion along x:

$$x = v_0 t + x_0 = ct \Rightarrow t = \frac{x}{c}. \tag{1}$$

The solution to the equation of motion along y:

$$y = y(t) = \frac{1}{2}gt^2 = \frac{1}{2}\frac{GM}{R^2}t^2. \tag{2}$$

Now, substituting (1) to get $y = y(x)$ we find,

$$y(x) = \frac{1}{2}\frac{MG}{R^2}\frac{x^2}{c^2}. \tag{3}$$

To find an equation describing the deflection angle α, assume $\alpha \ll 1$, which will be true for a large domain of validity. Then, the Taylor expansion of $\tan(\alpha)$ is:

$$\tan(\alpha) = \alpha + \frac{\alpha^3}{3} + O(\alpha^5); \tag{4}$$

therefore, in our case, we can neglect higher order terms. Hence, $\tan(\alpha) \approx \alpha$. Now, we evaluate the change in y by differentiating (3) with respect to x, evaluated at $x = 2R$,

$$\tan(\alpha) \approx \frac{\partial y}{\partial x}\bigg|_{x=2R} = \frac{MGx}{R^2c^2}\bigg|_{x=2R} = \frac{2MG}{Rc^2} \Rightarrow \alpha \approx \frac{2MG}{Rc^2}. \tag{5}$$

This is a quick, informal, derivation approximating the deflection of a light ray caused by a spherically symmetric mass M.

A spherically symmetric mass affects external objects gravitationally as though all of its mass were concentrated at a point at its centre [17]. As the example above meets this requirement, the result can be applied to a point mass. We make the further simplification that, the radius R is roughly equal to the impact parameter b and so we have,

$$\alpha \approx \frac{2MG}{bc^2}. \tag{6}$$

Definition: The *impact parameter*, b, is the perpendicular distance between the unperturbed light ray and the point mass.

2.2 Kibble's Relation

To confirm this result with an improved derivation, we introduce Kibble's relation for the impact parameter b and the *scattering angle*, Θ.

$$b = \frac{|k|}{2E}\cot\left(\frac{\Theta}{2}\right). \tag{7}$$

See [6].

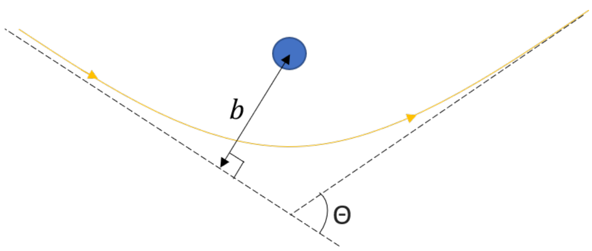

Figure 3: Example of an elliptical orbit, showing the scattering angle Θ [6].

When we apply gravitational lensing to Kibble's relation we have: k, the *inverse square law constant* $k = mMG$; the kinetic energy, $E = \frac{1}{2}mV_0^2 = \frac{1}{2}mc^2$; and Θ will be small, $\Theta \ll 1$. Therefore, we have,

$$b \simeq \frac{mMG}{mc^2}\frac{2}{\Theta} = \frac{2MG}{c^2\Theta}. \tag{8}$$

We notice in equation (8), the lighter mass, m, drops out of the equation, hinting we may be able to describe the motion of even a massless particle, like a photon, using this argument. Hence,

$$\Theta \simeq \frac{2MG}{c^2 b}. \tag{9}$$

Now we introduce $\hat{\boldsymbol{\alpha}}$ as the difference between our initial unit vector at the time lensing begins, $\hat{\mathbf{e}}_i$ and our final unit vector at the time the light ray proceeds in a straight line $\hat{\mathbf{e}}_f$.

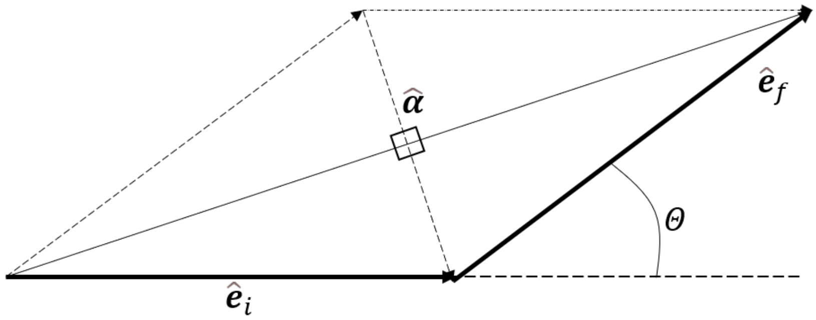

Figure 4: Geometric visualisation of the relationship between $\hat{\boldsymbol{\alpha}}$ and angle Θ.

Figure 4 shows that the vector $\hat{\boldsymbol{\alpha}}$ has the magnitude Θ, giving us an approximate deflection angle,

$$|\hat{\boldsymbol{\alpha}}| \simeq \frac{2MG}{c^2 b}. \tag{10}$$

3 Optics - Short-Wave Asymptotics

3.1 Maxwell's Equations

The complete classical theory of electromagnetic fields are encoded into Maxwell's equations. These describe phenomena caused by the interaction between electromagnetic fields and matter, neglecting quantum effects [14]. In a general form, Maxwell's equations are:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \tag{11}$$

$$\boldsymbol{\nabla} \wedge \mathbf{E} = -\kappa \frac{\partial \mathbf{B}}{\partial t}, \tag{12}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0, \tag{13}$$

$$\boldsymbol{\nabla} \wedge \mathbf{B} = \kappa \mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right). \tag{14}$$

Where $\mathbf{E}$ denotes the *electric field strength*, $\mathbf{B}$ is the *magnetic induction*, ρ is the *electric charge density* and $\mathbf{j}$ is the *electric current density.* The constants κ, μ_0 and ϵ_0 can be chosen depending on the system of units used. If we use the *Gaussian system*, we have,

$$\kappa = \tfrac{1}{c}, \quad \mu_0 = 4\pi, \quad \epsilon_0 = \tfrac{1}{4\pi}. \tag{15}$$

Many optical phenomena are often a result of the fact that the wavelength of light is small compared to the dimensions of the optical system, and in these situations the wave-like features of the light are often hidden. Since the wavelengths of light are tiny with respect to the optical system, in our case gravitational lensing, the relevant solutions of the wave equation are locally given by plane waves. From this, we see that the trajectory and amplitude of a wave will not change noticeably over a small distance (i.e. will not change over several wavelengths). Generalising the expression for a plane wave with fixed frequency ω and a wave vector of absolute value, $k = \omega/c$, we can write

$$u_k(t, \mathbf{x}) = \phi_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{x})} = \phi_0 e^{-ik(ct - \mathbf{n} \cdot \mathbf{x})}. \tag{16}$$

We re-write in the form of our wave ansatz,

$$u_k(t, \mathbf{x}) = \phi_k(\mathbf{x}) e^{ik[S(\mathbf{x}) - ct]}, \tag{17}$$

where the real function $S(\mathbf{x})$ determines the *phase* of the wave at $\mathbf{x}$ and the function $\phi_k(\mathbf{x})$ fixes the amplitude of the wave. $S(\mathbf{x})$ is called the *eikonal function* and the surfaces are $S(\mathbf{x})$ = constant, the shape of which depends on the given problem. In a locally isotropic medium, Maxwell's equations for a field with a fixed frequency ω and vanishing charge and current densities are:

$$\epsilon_0 \boldsymbol{\nabla} \cdot (\epsilon \mathbf{E}) = 0, \tag{18}$$

$$\mu_0 \boldsymbol{\nabla} \cdot (\mu \mathbf{H}) = 0, \tag{19}$$

$$\boldsymbol{\nabla} \wedge \mathbf{H} = -i\omega\kappa\epsilon\epsilon_0\mathbf{H}, \tag{20}$$

and

$$\boldsymbol{\nabla} \wedge \mathbf{E} = -i\omega\kappa\mu\mu_0\mathbf{H}, \tag{21}$$

where $\mathbf{B} = \mu\mu_0\mathbf{H}$, ϵ and μ are functions of wave frequency, ω, and position, $\mathbf{x}$. Inserting our wave ansatz (**??**), in other words, letting

$$\mathbf{E}(t, \mathbf{x}) = \mathbf{e}(\mathbf{x})e^{ik[S(\mathbf{x})-ct]}, \tag{22}$$

and

$$\mathbf{H}(t, \mathbf{x}) = \mathbf{h}(\mathbf{x})e^{ik[S(\mathbf{x})-ct]}. \tag{23}$$

Recall, $k = \omega/c$, therefore, we have

$$\boldsymbol{\nabla} \wedge \mathbf{E} = (\boldsymbol{\nabla} \wedge \mathbf{e} + ik\boldsymbol{\nabla}S \wedge \mathbf{e})e^{ik[S(\mathbf{x})-ct]}, \tag{24}$$

$$\boldsymbol{\nabla} \cdot \mathbf{E} = (\boldsymbol{\nabla} \cdot \mathbf{e} + ik\boldsymbol{\nabla}S \cdot \mathbf{e})e^{ik[S(\mathbf{x})-ct]}, \tag{25}$$

and so we obtain:

$$\begin{aligned}
\mathbf{e} \cdot \boldsymbol{\nabla}S &= -\frac{1}{ik}\left(\mathbf{e} \cdot \frac{\boldsymbol{\nabla}\epsilon}{\epsilon} + \boldsymbol{\nabla} \cdot \mathbf{e}\right), \\
\mathbf{h} \cdot \boldsymbol{\nabla}S &= -\frac{1}{ik}\left(\mathbf{h} \cdot \frac{\boldsymbol{\nabla}\mu}{\mu} + \boldsymbol{\nabla} \cdot \mathbf{h}\right), \\
(\boldsymbol{\nabla}S \wedge \mathbf{e} - \mu\mu_0\kappa c\mathbf{h}) &= -\frac{1}{ik}\boldsymbol{\nabla} \wedge \mathbf{e}, \\
(\boldsymbol{\nabla}S \wedge \mathbf{h} + \epsilon\epsilon\kappa c\mathbf{e}) &= -\frac{1}{ik}\boldsymbol{\nabla} \wedge \mathbf{h}.
\end{aligned} \tag{26}$$

For the situations we are interested in, we can assume $k \to \infty$ and thus the right-hand side of these equations are negligible in leading order, we have:

$$\begin{aligned}
\mathbf{e} \cdot \boldsymbol{\nabla}S = 0, &\qquad \mathbf{h} \cdot \boldsymbol{\nabla}S = 0, \\
\boldsymbol{\nabla}S \wedge \mathbf{e} = \mu\mu_0\kappa c\mathbf{h}, &\quad \boldsymbol{\nabla}S \wedge \mathbf{h} = -\epsilon\epsilon_0\kappa c\mathbf{e}.
\end{aligned} \tag{27}$$

Eliminating $\mathbf{e}$ or $\mathbf{h}$ leads to the *eikonal equation*

$$(\boldsymbol{\nabla}S)^2 = n^2, \tag{28}$$

where the *refractive index*, n, is such that $n^2(\omega, \mathbf{x}) = \epsilon\mu$. Hence, if we know the refractive index, we can calculate the eikonal function S.
Definition: The *refractive index* is a dimensionless number that describes how fast light travels through a medium.

3.2 Eikonal Equation

The eikonal equation is a partial differential equation (PDE) for the eikonal function $S(\mathbf{x})$,

$$(\boldsymbol{\nabla}S(\mathbf{x}))^2 = n^2(\mathbf{x}), \tag{29}$$

where $\mathbf{x} = \mathbf{x}(x, y, z)$ is a position vector, $S = S(\mathbf{x})$. The specific solution for this equation can vary depending on the shape of the wave field. In gravitational lensing, the refractive index is a consequence of gravity. In vacuum, light is not "slowed"; therefore, vacuum has a refractive index of 1, thus light travels at speed c. Generally, $n = n(\mathbf{x}) \geq 1$ and the speed of light is $c/n \leq c$.

Defitition: *Light rays* are the trajectories taken by light over time, obtained such that the curve's tangent at each point is proportional to $\boldsymbol{\nabla}(S)$.
Succinctly, light rays are the integral curves of the vector field $\boldsymbol{\nabla} S$ and are *null geodesics* (geodesics will be explained later), [16]. Hence, we have the ray equation

$$\frac{d\mathbf{x}(\tau)}{d\tau} = \boldsymbol{\nabla} S, \tag{30}$$

where

$$|\boldsymbol{\nabla} S| = n. \tag{31}$$

We can see the level surfaces, or wave fronts, $S = C_i$, are orthogonal to the propagating light rays, $\mathbf{x}(\tau)$, as shown in Figure 5.

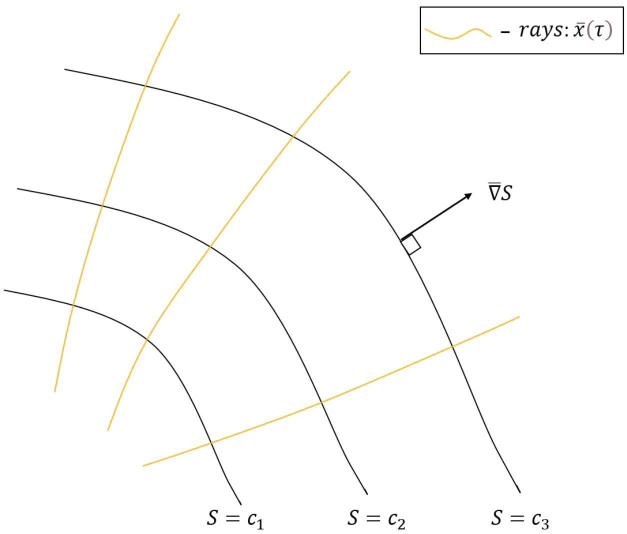

Figure 5: Light rays propagating across wave fronts.

As the eikonal equation is a PDE, to get a unique solution, we will have to impose boundary conditions on the field.
Example 1: Assuming $n(\mathbf{x}) = 1$, that is the refractive index in vacuum.
We find a solution to the eikonal equation is

$$S(\mathbf{x}) = |\mathbf{x}|. \tag{32}$$

For this solution, we see the wave fronts, $S = const$ are spheres and, hence, at any two points that are the same distance from the origin will have the same value for the eikonal function.

The gradient of the eikonal function follow easily,

$$\nabla S = \frac{\mathbf{x}}{|\mathbf{x}|} = \dot{\mathbf{x}}. \tag{33}$$

Integrating with respect to τ gives us the light rays

$$\mathbf{x}(\tau) = \tau \frac{\mathbf{x}(0)}{|\mathbf{x}(0)|}. \tag{34}$$

For $\mathbf{x}(0) \neq \mathbf{0}$, we see

$$|\mathbf{x}(\tau)| = \tau. \tag{35}$$

Hence, as τ increases, the length of the light ray also increases.

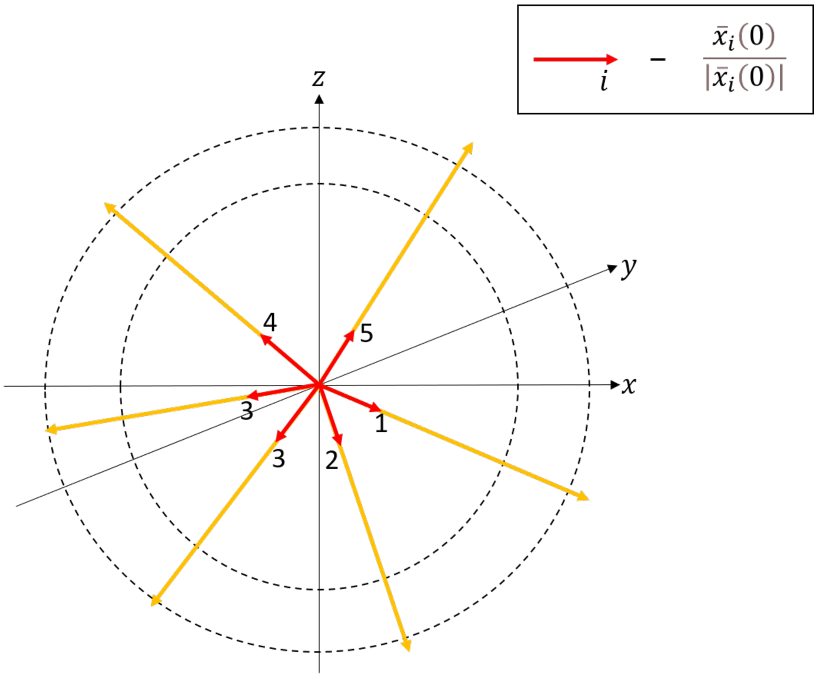

Figure 6: Light rays propagating through spherically symmetric wave fronts.

Example 2: $n(\mathbf{x}) = constant$. Then, $S = \mathbf{n} \cdot \mathbf{x}$ solves the eikonal equation where $\mathbf{n}$ is a constant vector with length n, $|\mathbf{n}| = n$, as the gradient of S equals n. From the ray equation,

$$\dot{\mathbf{x}} = \nabla S = \mathbf{n}. \tag{36}$$

The general solution of which is

$$\mathbf{x}(\tau) = \mathbf{n}\tau + \mathbf{x}_0. \tag{37}$$

From boundary conditions, when $\tau = 0$, $\mathbf{x} = \mathbf{x}_0$.
Compare $S = \mathbf{n} \cdot \mathbf{x}$ with $S = |\mathbf{x}|$, these obey different boundary conditions. One gives us a spherically symmetric solution with light rays emanating radially. The other, gives us plane waves, where the light rays emanate parallel to $\mathbf{n}$.

3.3 Normalising the Ray Equation

To normalise our ray equation, we parametrise light rays using arc length; therefore, $\mathbf{x} = \mathbf{x}(s)$. The length L_γ of curve γ is given by:

$$L_\gamma = \int_\gamma ds. \tag{38}$$

This implies:

$$\left| \frac{d\mathbf{x}}{ds} \right| = 1. \tag{39}$$

Whence, $d\mathbf{x}/ds$ is the unit vector tangent to the curve $\mathbf{x} = \mathbf{x}(s)$.

Re-arranging the Eikonal equation, (29), gives:

$$\left| \frac{\boldsymbol{\nabla} S}{n} \right| \equiv 1. \tag{40}$$

Let,

$$\hat{\mathbf{e}} \equiv \frac{\boldsymbol{\nabla} S}{n}. \tag{41}$$

This is a unit vector which we identify with the unit tangent vector along light rays,

$$\frac{d\mathbf{x}}{ds} = (1/n)\boldsymbol{\nabla} S = \hat{\mathbf{e}}. \tag{42}$$

Multiplying through by n and taking the derivative with respect to arc length gives

$$\frac{d}{ds}\left(n\frac{d\mathbf{x}}{ds} \right) = \frac{d}{ds}(\boldsymbol{\nabla} S). \tag{43}$$

Using the chain rule and Einstein's summation convention we

$$\frac{d\boldsymbol{\nabla} S}{ds} = \frac{\partial \boldsymbol{\nabla} S}{\partial x_i}\frac{dx_i}{ds} = \frac{\partial \boldsymbol{\nabla} S}{\partial x_i}\hat{e}_i. \tag{44}$$

Using substitution, into (43), yields

$$\frac{d}{ds}\left(n\frac{d\mathbf{x}}{ds} \right) = \hat{\mathbf{e}} \cdot \boldsymbol{\nabla}(\boldsymbol{\nabla} S) = \frac{1}{n}(\boldsymbol{\nabla} S \cdot \boldsymbol{\nabla})\boldsymbol{\nabla} S. \tag{45}$$

Alternatively, in components summed over j,

$$\frac{d}{ds}\left(n\frac{dx_i}{ds} \right) = \frac{1}{n}(\partial_j S)\partial_j\partial_i S = \frac{1}{n}(\partial_j S)\partial_i\partial_j S. \tag{46}$$

To simplify further, notice $\partial_i(\partial_j S \partial_j S) = 2(\partial_j S)\partial_i\partial_j S$. We correct by dividing by 2 to obtain,

$$\frac{d}{ds}\left(n\frac{dx_i}{ds} \right) = \frac{1}{2n}\partial_i(\boldsymbol{\nabla} S)^2. \tag{47}$$

Substituting from (9) gives

$$\frac{d}{ds}\left(n\frac{dx_i}{ds}\right) = \frac{1}{2n}\partial_i n^2 = \frac{1}{.}2n(2n\partial_i n) = \partial_i n. \tag{48}$$

In vector notation, we find the normalised ray equation:

$$\frac{d}{ds}\left(n\frac{d\mathbf{x}}{ds}\right) = \boldsymbol{\nabla} n. \tag{49}$$

3.4 Fermat's Principle

Lets evaluate the integral,

$$\int_{\gamma\mathbf{x}_0}^{\mathbf{x}} \boldsymbol{\nabla} S \cdot d\mathbf{x} = \int_{\gamma\mathbf{x}_0}^{\mathbf{x}} n\hat{\mathbf{e}} \cdot d\mathbf{x} = \int_{\gamma\mathbf{x}_0}^{\mathbf{x}} n\hat{\mathbf{e}} \cdot \frac{d\mathbf{x}}{ds} ds = S(\mathbf{x}) - S(\mathbf{x}_0). \tag{50}$$

The path γ that we choose to connects the two points $\mathbf{x}_0$ and $\mathbf{x}$ will not change the integral, hence, the integral is path independent. Making use of the inequality

$$\hat{\mathbf{e}} \cdot \frac{d\mathbf{x}}{ds} \leq |\hat{\mathbf{e}}| \left|\frac{d\mathbf{x}}{ds}\right| = 1. \tag{51}$$

We obtain

$$\int_{\gamma\mathbf{x}_0}^{\mathbf{x}} n\hat{\mathbf{e}} \cdot d\mathbf{x} \leq \int_{\gamma\mathbf{x}_0}^{\mathbf{x}} nds. \tag{52}$$

The inequality is dependent on the path, for all paths γ that are not the shortest path (the path of light rays) the less than sign holds. For the shortest path we have an inequality. As, in our case, we are dealing with light rays, we know the direction of the light ray points in the direction of $\hat{\mathbf{e}}$. Therefore, we have $\hat{\mathbf{e}} \cdot (d\mathbf{x}/ds) = 1$. We now acknowledge *Fermat's principle.*
Definition: Fermat's principle states that among all trajectories γ running from $\mathbf{x}_0$ to $\mathbf{x}$, the trajectories given by light rays yield the smallest possible values, $S(\mathbf{x}) - S(\mathbf{x}_0)$, for the integral $\int_{\gamma\mathbf{x}_0}^{\mathbf{x}} n\ ds$, [13].
We define a new *light-length* element:

$$dl^2 = n^2(\mathbf{x})\dot{x}_i\dot{x}_i d\tau^2 = n^2 ds^2. \tag{53}$$

We see that the arc length of a curve $\mathbf{x}(t)$ with respect to the *natural-length* element, ds, from t_1 to t_2 is

$$s_{21} = \int ds = \int_{t_1}^{t_2} \sqrt{\dot{\mathbf{x}}^2} d\tau. \tag{54}$$

Expressing the arc-length of this curve in terms of the light-length element dl, we have

$$l_{21} = \int dl = \int_{t_1}^{t_2} n(\mathbf{x}(\tau))\sqrt{\dot{\mathbf{x}}} d\tau. \tag{55}$$

Definition: *Geodesics* are the shortest connections between two points.
With respect to the light-length element, geodesics are light rays! Recall $n(x) = c/c(x)$. Therefore, (55) implies

$$l = c\int_\gamma \frac{ds}{c(x)}. \tag{56}$$

As c is constant, it is clear l is proportional to the time that the light ray takes to traverse the path γ. We, therefore, have the equivalent expression for Fermat's principle, a light ray runs from $\mathbf{x}_1$ to $\mathbf{x}_2$ along the path that is traversed in the *least* time.

Example: Fermat's principle and ray equation.
The equation for geodesics with respect to the light-like element seen in (53) can be expressed as

$$\delta\int nds = \delta\int n\sqrt{\left(\frac{d\mathbf{x}}{dt}\right)^2}dt = 0. \tag{57}$$

From the product rule,

$$0 = \int\left[\boldsymbol{\nabla}n\sqrt{\dot{\mathbf{x}}^2}\cdot\delta\mathbf{x} + n\frac{\dot{\mathbf{x}}}{\sqrt{\dot{\mathbf{x}}}^2}\cdot\frac{d\delta\mathbf{x}}{dt}\right]dt = \int\left[\boldsymbol{\nabla}n\cdot\delta\mathbf{x} + n\frac{d\mathbf{x}}{ds}\cdot\frac{d\delta\mathbf{x}}{ds}\right]ds, \tag{58}$$

recall, from (53), it is clear $ds^2 = \dot{x}_i\dot{x}_i d\tau$. Next, using the chain rule yields

$$0 = \int\left[\boldsymbol{\nabla}n - \frac{d}{ds}\left(n\frac{d\mathbf{x}}{ds}\right)\right]\cdot\delta\mathbf{x}ds. \tag{59}$$

The right-hand side vanishes for all $\delta\mathbf{x}$ and hence this implies

$$\boldsymbol{\nabla}n - \frac{d}{ds}\left(n\frac{d\mathbf{x}}{ds}\right) = 0. \tag{60}$$

We arrive back at the normalised ray equation, see (49)!!!

3.5 Deflection of Light from a Point Mass

The simplest example of gravitational lensing is lensing due to a point mass, i.e. a massive body with zero spatial extent. A point mass has a gravitational field that acts as a medium and, as such, perturbs the light rays. Generally, a medium will slow light and we shall quantify the slowing effect over that of a vacuum, as δ. So, we have the refractive index n as:

$$n(\mathbf{x}) = 1 + \delta(\mathbf{x}). \tag{61}$$

In a relativity weak gravitational field, the refractive index will almost always be small $\delta \ll 1$. The refractive index of a gravitational field can be derived from the general relativistic Maxwell equations by inserting the eikonal ansatz for the Maxwell field [16],

$$\mathbb{F}_{\mu\nu} = \mathrm{Re}(f_{\mu\nu}e^{iS}), \tag{62}$$

where $f_{\mu\nu}$ is the varying amplitude and S is the eikonal. This yields the eikonal equation, $(\boldsymbol{\nabla} S)^2 = n^2$, with the effective index of refraction,

$$\delta = \frac{-2}{c^2}\phi(\mathbf{x}), \tag{63}$$

with ϕ known as the Newtonian gravitational potential. For spherical objects of mass M, $\phi = -GM/r$ and $r \equiv |\mathbf{x}|$, see 71.
From equation (49) we have

$$\boldsymbol{\nabla} n = \frac{d}{ds}(n\hat{\mathbf{e}}) = \frac{dn}{ds}\hat{\mathbf{e}} + n\frac{d\hat{\mathbf{e}}}{ds}. \tag{64}$$

Rearranging for change in the unit tangent vector with respect to arclength gives:

$$\begin{aligned}\frac{d\hat{\mathbf{e}}}{ds} &= \frac{1}{n}(\boldsymbol{\nabla} n - \frac{\partial n}{\partial x_i}\frac{\partial x_i}{\partial s}\hat{\mathbf{e}}) = \frac{1}{n}(\boldsymbol{\nabla} n - \hat{\mathbf{e}} \cdot (\boldsymbol{\nabla} n)\hat{\mathbf{e}}) = \frac{1}{n}(\boldsymbol{\nabla} - \hat{\mathbf{e}}(\hat{\mathbf{e}} \cdot \boldsymbol{\nabla}))n \\ &\equiv \frac{1}{n}\mathbb{P}\boldsymbol{\nabla} n,\end{aligned} \tag{65}$$

where $\mathbb{P}_{ij} \equiv \delta_{ij} - e_i e_j$, for i and j summed over the three spatial axis x, y and z. Therefore,

$$\frac{d\hat{e}}{ds} = \frac{1}{n}\boldsymbol{\nabla}_{\perp} n, \tag{66}$$

where $\boldsymbol{\nabla}_{\perp}$ is the differential operator perpendicular to the light ray, i.e. if the light ray is along the z-axis $\boldsymbol{\nabla}_{\perp} = (\partial_x, \partial_y, 0)$. Substituting yields,

$$\frac{d\hat{e}}{ds} = \frac{1}{1+\delta(\mathbf{x})}\boldsymbol{\nabla}_{\perp}(1+\delta(\mathbf{x})). \tag{67}$$

Recalling $\delta \ll 1$, we may use the Taylor expansion

$$\frac{1}{1+\delta} = 1 - \delta + 0(\delta^2) \simeq 1 - \delta. \tag{68}$$

Giving us

$$\frac{d\hat{e}}{ds} \simeq (1-\delta)\boldsymbol{\nabla}_{\perp}\delta \simeq \boldsymbol{\nabla}_{\perp}\delta = \boldsymbol{\nabla}_{\perp} n. \tag{69}$$

Gravity from a point mass will manifest as radial attraction. It is proportional to the mass M of the lens and the gravitational constant G, and inversely proportional to the radius r. As,

$$\mathbf{x} = r = \sqrt{x^2 + y^2 + z^2}, \tag{70}$$

the Newtonian potential is

$$\phi(\mathbf{x}) = \phi(r) = -\frac{GM}{r} = -\frac{GM}{\sqrt{x^2 + y^2 + z^2}}. \tag{71}$$

As $\delta = -(2/c^2)\phi$, (69) becomes,

$$
\begin{aligned}
\frac{d^2\mathbf{x}}{ds^2} = \frac{d\hat{e}}{ds} \simeq -\frac{2}{c^2}\boldsymbol{\nabla}_{\perp}\phi(r) = \frac{2GM}{c^2}\boldsymbol{\nabla}_{\perp}\frac{1}{r} = \frac{2GM}{c^2}\left(-\frac{1}{r^2}\right)\boldsymbol{\nabla}_{\perp}r \\
= -\frac{2GM}{c^2}\frac{1}{r^2}\frac{\mathbf{x}_{\perp}}{r} = -\frac{2GM}{c^2 r^3}\mathbf{x}_{\perp}.
\end{aligned}
\tag{72}
$$

This is the second derivative of the light ray position with respect to arclength; therefore, integrating with respect to arclength will formalise how the light ray changes with regard to change in arclength. (72) is a non-linear system of ordinary differential equations for trajectory ('light ray') $\mathbf{x} = \mathbf{x}(s)$.

In general, this system can be difficult to solve. We will only focus on finding the deflection $\hat{\boldsymbol{\alpha}}$, recall $\hat{\boldsymbol{\alpha}}$ is the vector describing the change in direction from the initial unit vector (the unit vector when deflection begins) and the final unit vector (the unit vector when the light ray travels once again in a straight trajectory after being deflected). Note, $\hat{\boldsymbol{\alpha}}$ is not a unit vector, the hat is simply convention [16].

Let our unperturbed light ray be along the z-axis, i.e. $ds = dz$. Then the deflection is

$$
\hat{\boldsymbol{\alpha}} = -\int \frac{d\hat{\mathbf{e}}}{ds}\ ds = -\int \boldsymbol{\nabla}_{\perp}n\ dz = \frac{2}{c^2}\int \boldsymbol{\nabla}_{\perp}\phi(z)\ dz. \tag{73}
$$

Substituting in our Newtonian potential from equation (71) we have

$$
\hat{\boldsymbol{\alpha}} = -\frac{2GM}{c^2}\int \boldsymbol{\nabla}_{\perp}\frac{1}{\sqrt{x_{\perp}^2 + z^2}}dz. \tag{74}
$$

$|x_{\perp}|$ is the perpendicular distance between the unperturbed light ray and the point mass, this is known as the *impact parameter* and is commonly denoted with the letter b, see Figure(3), [10].
Since z is orthogonal to $\boldsymbol{\nabla}_{\perp} = (\partial_x,\partial_y,0)$, we have $\boldsymbol{\nabla}_{\perp}z = 0$, notice

$$
\boldsymbol{\nabla}_{\perp}\frac{1}{\sqrt{x_{\perp}^2 + z^2}} = -\mathbf{x}_{\perp}(x_{\perp}^2 + z^2)^{-\frac{3}{2}}; \tag{75}
$$

therefore,

$$
\int -\mathbf{x}_{\perp}(x_{\perp}^2 + z^2)^{-\frac{3}{2}}\ dz = -\mathbf{x}_{\perp}\int (x_{\perp}^2 + z^2)^{-\frac{3}{2}}\ dz. \tag{76}
$$

Due to the fact that most lensing will occur in the range $-x_{\perp}^2 < z < x_{\perp}^2$, which relative to the distance from source to observer is tiny, we generalise the bounds from source to observer, to the bounds $-\infty < z < +\infty$ without introducing a sizeable error.

By substituting $z = b\tan(u)$, we integrate and find

$$
\int_{-\infty}^{+\infty} (x_{\perp}^2 + z^2)^{-\frac{3}{2}}\ dz = \left[\frac{z}{b^2\sqrt{b^2 + z^2}}\right]_{-\infty}^{+\infty} = \frac{2}{|\mathbf{x}_{\perp}|^2}, \tag{77}
$$

resulting in the deflection vector,

$$
\hat{\boldsymbol{\alpha}} = \mathbf{x}_{\perp}\frac{4GM}{c^2|\mathbf{x}_{\perp}|^2}, \tag{78}
$$

with modulus,

$$|\hat{\boldsymbol{\alpha}}| \equiv \alpha = \frac{4GM}{c^2|\mathbf{x}_\perp|} \equiv \frac{4GM}{c^2 b}. \tag{79}$$

This equation is Einstein's famous result for the angle of deflection for a point mass lens. Notice, this is twice the previous Newtonian result (10).

3.6 Generalising Light Deflection to a Mass Distribution

The point mass deflection angle is not always a good approximation. When light travels through space it encounters many deflections due to the fluctuating gravitational potential along its path [3]. We, therefore, generalise our previous discussion by considering deflection caused by a mass distribution. Similar to the point mass example, we shall consider only one lens. However, we allow this lens to extend in the transverse direction and thus, this solution will be applicable to a wider range of circumstances.

Poisson's equation relates a given mass density distribution $\rho(\mathbf{x})$ to the gravitational potential ϕ [6]:

$$\triangle\, \phi(\mathbf{x}) = 4\pi G\rho(\mathbf{x}). \tag{80}$$

As seen above Poisson's equation is an inhomogeneous Laplace PDE. Knowing the mass density distribution allows us solve for ϕ. Hence, we can substitute the solution, for ϕ in (80), into (73):

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\int \boldsymbol{\nabla}_\perp \phi(s)\; ds. \tag{81}$$

Alternatively, in individual components we have

$$\hat{\alpha}_i = \frac{2}{c^2}\int \mathbb{P}_{ij}\partial_j\phi\; ds. \tag{82}$$

As before, we shall choose our co-ordinates such that light will travel down the z-axis and; therefore, we have a projection transverse to the z direction:

$$\mathbb{P}_\perp = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Leftrightarrow \mathbb{I} - \mathbb{P}_z. \tag{83}$$

Therefore, we can simplify from $\mathbb{R}^3$ to $\mathbb{R}^2$ as our z component will always be zero,

$$\hat{\alpha}_a = \frac{2}{c^2}\int \partial_a\phi dz, \tag{84}$$

with $a = x,\, y$. To make use of Poisson's equation, we take partial derivatives,

$$\partial_a\hat{\alpha}_a = \frac{2}{c^2}\int \partial_a\partial_a\phi dz = \frac{2}{c^2}\int \triangle_\perp\, \phi\;\; dz. \tag{85}$$

From (83),

$$\partial_a \hat{\alpha} = \frac{2}{c^2} \int_{-\infty}^{+\infty} (\triangle\ \phi - \partial_z^2 \phi) dz = \frac{2}{c^2} \int \triangle\ \phi - \frac{2}{c^2} \partial_z \phi \Big|_{-\infty}^{+\infty}. \tag{86}$$

However,

$$\frac{2}{c^2} \partial_z \phi \Big|_{-\infty}^{+\infty} = 0, \tag{87}$$

therefore, we have

$$\partial_a \hat{\alpha} = \frac{2}{c^2} \int \triangle\ \phi dz. \tag{88}$$

Substituting in the Poisson equation, yields

$$\partial_a \hat{\alpha}_a = \frac{8\pi G}{c^2} \int \rho(\mathbf{x})\ dz \equiv \frac{8\pi G}{c^2} \Sigma(\mathbf{x}_\perp). \tag{89}$$

We have introduced Σ, the *surface mass density.* As the integral of the mass density along the the light ray, in our case the z-axis,

$$\Sigma(\mathbf{x}_\perp) \equiv \int \rho(\mathbf{x}_\perp, z) dz. \tag{90}$$

We now introduce the potential $\hat{\psi}$, such that

$$\hat{\alpha} =: \frac{2}{c^2} \partial_a \hat{\psi}. \tag{91}$$

Here again, the hat is just convention and not related to unit length.

$$\frac{2}{c^2} \int \partial_a \phi dz \equiv \frac{2}{c^2} . \partial_a \hat{\psi} \tag{92}$$

Therefore, (89) becomes

$$\partial_a \hat{\alpha}_a = \frac{8\pi G}{c^2} \Sigma = \frac{2}{c^2}\ \triangle_\perp\ \hat{\psi}, \tag{93}$$

which leads to the two-dimensional Poisson equation:

$$\triangle_\perp\ \hat{\psi} = 4\pi G \Sigma. \tag{94}$$

A formal solution for the potential is given by

$$\hat{\psi} = 4\pi G\ \triangle_\perp^{-1}\ \Sigma, \tag{95}$$

where $\triangle_\perp^{-1}$ denotes the Green function of the two-dimensional Laplacian. Which is explicitly given by,

$$\mathbb{G}(\mathbf{x}_\perp) \equiv \mathbb{G}(x, y) = \frac{1}{2\pi} \log |\mathbf{x}_\perp|. \tag{96}$$

[5]. Hence, (95) becomes:

$$\hat{\psi}(\mathbf{x}_\perp) = 2G \int \log |\mathbf{x}_\perp - \mathbf{y}_\perp| \Sigma(\mathbf{y}_\perp) d^2 y_\perp. \tag{97}$$

Substituting this into (92), we obtain

$$\hat{\alpha} = \frac{2}{c^2}\boldsymbol{\nabla}_\perp \psi(\mathbf{x}_\perp) = \frac{4G}{c^2}\boldsymbol{\nabla}_\perp \int \log|\mathbf{x}_\perp - \mathbf{y}_\perp|\Sigma(\mathbf{y}_\perp) d^2 y_\perp. \tag{98}$$

Evaluating the transverse gradient finally yields the transverse vector

$$\hat{\boldsymbol{\alpha}} = \frac{4G}{c^2}\int \frac{\mathbf{x}_\perp - \mathbf{y}_\perp}{|\mathbf{x}_\perp - \mathbf{y}_\perp|^2}\Sigma(\mathbf{y}_\perp) d^2 y_\perp. \tag{99}$$

Hence, we have derived an expression for the deflection vector describing a light ray when perturbed by a general mass distribution, with arbitrary projected mass density (equally know as *surface density*) Σ. Taking the modulus of this expression will lead to the deflection angle.

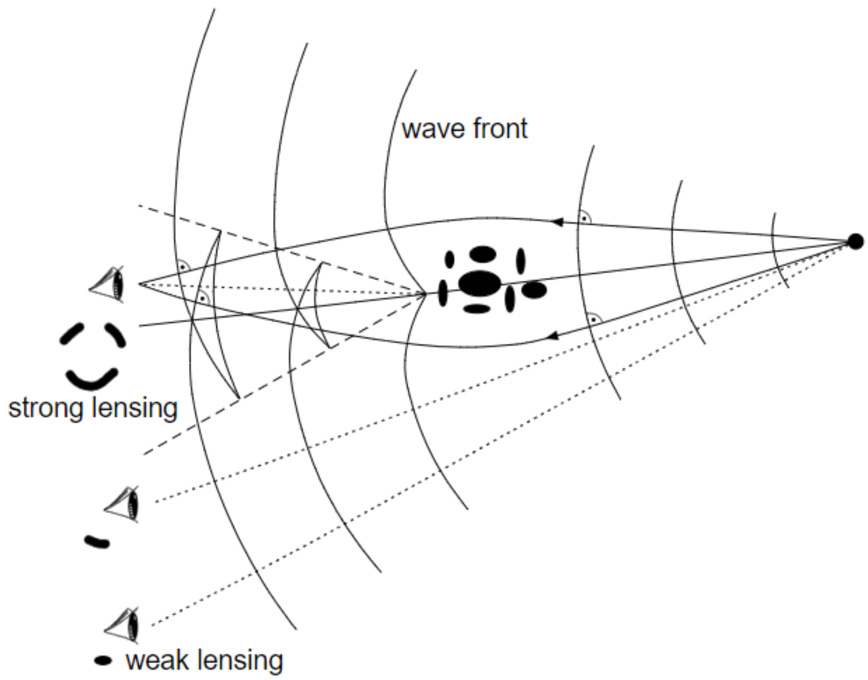

Figure 7: Diagram showing multiple light rays being perturbed by mass distributions, together with hinting at the lensing images received by the observer [16].

In Figure 7, we can see visually how the distance that the light passes the lens effects the magnitude of the bending. This is a result of the potential being anti-proportional to r, which is seen clearly during the point mass example (71). The figure also shows how the lensing and its magnitude affects the observer's view of the source, more on this later.

Example: Mass distribution: point mass.
To check the deflection angle, α, for the general result seen in (99), we revisit the case of a point mass, i.e. a massive object with zero spatial extent. Therefore, its mass density is

$$\rho(\mathbf{x}) = M\delta(x)\delta(y)\delta(z) = M\delta^3(\mathbf{x}). \tag{100}$$

Implying a surface density,

$$\Sigma(x, y) = \int \rho dz = M\delta(x)\delta(y) \int \delta(z)dz = M\delta(x)\delta(y) \equiv M\delta^2(\mathbf{x}_\perp). \tag{101}$$

Therefore, equation (99) leads to

$$\hat{\boldsymbol{\alpha}} = \frac{4G}{c^2} \int_{\mathrm{Re}^2} \frac{\mathbf{x}_\perp - \mathbf{y}_\perp}{|\mathbf{x}_\perp - \mathbf{y}_\perp|^2} M\delta^2(\mathbf{y}_\perp) d^2 y_\perp = \frac{4GM}{c^2} \frac{\mathbf{x}_\perp}{|\mathbf{x}_\perp|^2}. \tag{102}$$

Taking the modulus, we re-obtain Einstein's original result for the deflection angle due to a point mass M

$$\alpha = \frac{4GM}{c^2 |\mathbf{x}_\perp|}, \tag{103}$$

see (79).

4 General Relativity

This section will briefly introduce general relativity with the aim of retrieving a solution that describes the gravitational field produced by a spherically symmetric, non-rotating, body, i.e. a star [11]. This originally was solved by K. Schwarzschild in 1916, hence, it is known as the *Schwarzschild Solution.* To achieve our aim, we shall omit much of the differential geometry.
In general relativity, Einstein replaced the flat Minkowski spacetime with a curved spacetime, this curvature is a result of energy and momentum produced by masses [11]. Curved spacetime forms a four-dimensional Lorentzian Manifold M. In accordance with literature, the coordinates we use to describe spacetime will be denoted by letters with Greek indices, these indices will take the values $0, 1, 2, 3$, with zero being the time coordinate, i.e.

$$q^\alpha = (x^0, x^1, x^2, x^3) = (ct, x, y, z). \tag{104}$$

Latin indices will separately describe the spatial coordinates only.

4.1 Spherically Symmetric Metric

To start our derivation of Schwarzschild's solution, we will need to introduce the concept of a metric tensor.
Definition: A *pseudo-Riemannian metric* on a manifold M is a tensor field g that is symmetric and non-degenerate, i.e. $\forall p \in M$ and $X_p, Y_p \in T_pM$ we have $g_p(X_p, Y_p) = g_p(Y_p, X_p)$ and $g_p(X_p, \cdot) = 0 \Rightarrow X_p = 0$.
Representing a pseudo-Riemannian metric in a coordinate system such that $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$ implies $g_{\mu\nu} = g_{\nu\mu}$ from the symmetry condition, and from the non-degenerate condition $\det(g_{\mu\nu}) \neq 0$ implies we will always have the existence of inverse matrix $g^{\mu\nu}$ such that $g^{\mu\nu} g_{\mu\tau} = \delta^\tau_\mu$.
As a consequence of being a symmetric metric, for any point $p \in M$ we can diagonalise the metric by choosing suitable coordinates. Also, from the second condition $det(g_{\mu\nu}) \neq 0$, all of these diagonal elements must be non-zero. Hence, we can always compress or stretch the axis in such a way that we make the diagonal elements equal to ± 1. Therefore, a pseudo-Riemannian metric can be written in the form,

$$g_{\mu\nu} = diag(-1, ..., -1, 1..., 1). \tag{105}$$

In differential geometry we would acknowledge $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$ as a 2-form, we omit this for simplicity. In practical terms, the metric tensor allows us to define lines and lengths of curves on the manifold.
We instead will use the line element squared which, for simplicity, we shall just refer to as the *line element*, given by:

$$ds^2 = d\mathbf{s} \cdot d\mathbf{s} = dx^\mu \partial_\mu \cdot dx^\nu \partial_\mu = (\partial_\mu \cdot \partial_\nu) dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu. \tag{106}$$

Alternatively, we will call ds^2 the metric.

Definition: A *general relativistic spacetime* is a four-dimensional manifold with a pseudo-Riemannian metric of signature (-1,+1,+1,+1), [2].

A *spherically symmetric metric* is one such that the field described is the same at all points located at the same distance from the centre.

We will begin by introducing a special case, Minkowski spacetime. As well as acknowledging an important historical model of spacetime, this will allow us to introduce some key concepts before adding curvature.

Minkowski spacetime is the spacetime model implied by the postulates of special relativity, (i) and (ii) seen section 1.6.

Minkowski's spacetime is a four-dimensional manifold that is flat, which consists of three spatial dimensions and one dimension of time, [19]. Points in this space correspond to events in space-time and, to each event, there is a past and a future light-cone that allows us to differentiate between *timelike*, *spacelike* and *null/lightlike* vectors. If the vector lies on a cone, it is lightlike. If the vector lies in the space enclosed by a cone, it is timelike. If the vector is outside the cones, it is spacelike. This flat spacetime can be described in Cartesian coordinates by the metric tensor of form:

$$\eta_{\mu\nu} = \pm diag(-1,+1,+1,+1). \tag{107}$$

In Cartesian coordinates, the line element is:

$$ds^2 = \eta_{\alpha\beta}dq^{\alpha}dq^{\beta} = -(ct)^2 + (dx)^2 + (dy)^2 + (dz)^2, \tag{108}$$

where c is the speed of light, and x, y, z are spatial coordinates.

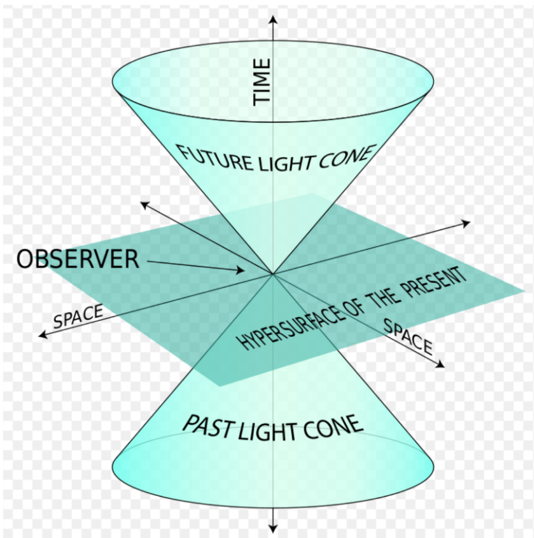

Figure 8: Minkowski spacetime [19].

To make use of Minkowski spacetime for our goal of re-forming Schwarzschild's solution, we make the transformation to spherical polar coordinates. Let the position vector

$$\mathbf{r}(r,\theta,\phi)=\begin{bmatrix}x\\y\\z\end{bmatrix}=r\begin{bmatrix}\sin(\theta)\cos(\phi)\\\sin(\theta)\sin(\phi)\\\cos(\theta)\end{bmatrix}. \tag{109}$$

Coordinates transform from Cartesian to curvilinear: $q^\alpha \mapsto x^\mu$, where $x^\mu=(ct,r,\theta,\phi)$. Then, using the chain rule, the line element is

$$ds^2=\eta_{\alpha\beta}dq^\alpha dq^\beta=\eta_{\alpha\beta}\frac{\partial q^\alpha}{\partial x^\mu}\frac{\partial q^\beta}{\partial x^\nu}dx^\mu dx^\nu\equiv g_{\mu\nu}(x)dx^\mu dx^\nu. \tag{110}$$

It follows, that the metric tensor is

$$g_{\mu\nu}=\begin{bmatrix}-c^2&0&0&0\\0&1&0&0\\0&0&r^2&0\\0&0&0&r^2\sin^2(\theta)\end{bmatrix}. \tag{111}$$

Hence, in spherical polar coordinates, we find the line element:

$$ds^2=-c^2dt^2+dr^2+r^2(\sin^2\theta d\phi^2+d\theta^2), \tag{112}$$

for $t\in\,]-\infty,\infty[$,$r\in\,]r_*,\infty[$, where $r_*=r_*(t)$ is the radius of the spherically symmetric body, $\theta\in[0,\pi]$ and $\phi\in[0,2\pi[$.

Any rotation of this coordinate system will not change the overall geometry of the metric; therefore, under rotations, the special relativity metric is invariant and hence spherically symmetric. We now generalise this whilst preserving the spherical symmetry. Multiplying each term by functions parametrised by t and/or r will not effect its spherical symmetry. In addition, the only mixed terms that do not break the spherical symmetry are $dtdr$ and $drdt$. Therefore, we can write for the most general spherically symmetric metric:

$$ds^2=l(r,t)dt^2+h(r,t)dr^2+k(r,t)(\sin^2\theta d\phi^2+d\theta^2)+a(r,t)(drdt+dtdr), \tag{113}$$

where a, h, k and l are functions parametrised by the radius and time [8]. From here, we may transform the coordinates because of the arbitrary choice of reference system, provided the spherical symmetry is maintained.

If for the transformation $(t,r,\theta,\phi)\longmapsto(\tilde{t},\tilde{r},\theta,\phi)$, where $\tilde{r}=f(r,t)$ and $\tilde{t}=g(r,t)$. Then, clearly both transformations are independent of any angular change. Hence, they do not break the spherical symmetry.

To maintain similar structure to (112), we choose coordinates $\tilde{r}$ and $\tilde{t}$ such that $a(\tilde{r},\tilde{t})=0$, also $k(\tilde{r},\tilde{t})=\tilde{r}^2$, [8]. Next, we choose $h(\tilde{r},\tilde{t})=e^{\lambda(\tilde{r},\tilde{t})}$ and $l(\tilde{r},\tilde{t})=-c^2e^{\nu(\tilde{r},\tilde{t})}$. Dropping the tilde for clarity, we now arrive at the line element:

$$ds^2=-e^{\nu(t,r)}c^2dt^2+e^{\lambda(t,r)}dr^2+r^2(d\theta^2+\sin^2\theta d\phi^2). \tag{114}$$

Any spherically symmetric metric can, therefore, be described in the form of (114), providing we can do the coordinate transformation $(t, r, \theta, \phi) \longmapsto (\tilde{t}, \tilde{r}, \theta, \phi)$ and the factors have signs such that they can be written as exponentials [11].

4.2 Christoffel Symbols

In order to calculate the Schwarzschild solution we have introduced the idea of a manifold, M, described by a metric, g. We have shown the most general case for the spherically symmetric metric through manipulation of coordinates whilst maintaining the spherical symmetry (114). However, we still have unknown functions $\nu(t, r)$ and $\lambda(t, r)$ to find.

We now introduce Christoffel Symbols, $\Gamma^{\mu}_{\nu\sigma}$. Christoffel symbols are arrays of numbers describing the difference between the flat and the curved metric. For our argument, the Christoffel symbols are used to represent the difference in unit vectors from Cartesian to spherical polar coordinates [3], and are defined by:

$$\Gamma^{\tau}_{\mu\nu} = \frac{g^{\sigma\tau}}{2}(\partial_{\mu} g_{\nu\sigma} + \partial_{\nu} g_{\sigma\mu} - \partial_{\sigma} g_{\mu\nu}). \tag{115}$$

Thus, they can be calculated by working out the derivatives of the metric. An alternative approach is to look at the geodesic equation.

Definition: A *geodesic* is the shortest curve on a manifold between two given points.

In general relativity, the geodesic equation describes motion of a *freely falling* particle, that is, a particle under the influence of gravity and absent of all other forces. To derive the geodesic equation, we start with the general equation describing particles in a gravitational field, [11].

$$L(x, \dot{x}) = \frac{1}{2} g_{\mu\nu}(x) \dot{x}_{\mu} \dot{x}_{\nu}. \tag{116}$$

Then the curve defined by $x^{\mu}(s)$ is a geodesic if, and only if, it is a stationary point of the action functional: $W = \int_a^b L(x(s), \dot{x}(s)) ds$, $\delta W = 0$, [11]. $\delta W = 0$ this is known as Hamilton's principle or, it is sometimes referred to, as the principle of least action, analogous to the Euler-Lagrange equations from Lagrangian mechanics

$$0 = \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}_{\mu}}\right) - \frac{\partial L}{\partial q_{\mu}}, \tag{117}$$

[7], where the over-dot signifies differentiation with respect to the curve parameter s, also from Lagrangian mechanics recall velocities and position vectors are independent. Using our Lagrangian, (116), we quickly arrive at,

$$g_{\sigma\mu} \ddot{x}^{\mu} + \frac{1}{2}(\partial_{\mu} g_{\sigma\nu} + \partial_{\nu} g_{\sigma\mu}) \dot{x}^{\mu} \dot{x}^{\nu} - \frac{1}{2}(\partial_{\sigma} g_{\mu\nu}) \dot{x}^{\mu} \dot{x}^{\nu} = 0. \tag{118}$$

Multiplying through by $g^{\sigma\tau}$ and comparing with (115) the standard form of the geodesic equation easily follows:

$$g^{\sigma\tau} g_{\sigma\mu} \ddot{x}^{\mu} + \frac{1}{2} g^{\sigma\tau}(\partial_{\mu} g_{\sigma\nu} + \partial_{\nu} g_{\sigma\mu}) \dot{x}^{\mu} \dot{x}^{\nu} - \frac{1}{2} g^{\sigma\tau}(\partial_{\sigma} g_{\mu\nu}) \dot{x}^{\mu} \dot{x}^{\nu} = \ddot{x}^{\tau} + \Gamma^{\tau}_{\mu\nu} \dot{x}^{\mu} \dot{x}^{\nu} = 0. \tag{119}$$

We are now in a position to calculate the Christoffel Symbols of the spherically symmetric metric tensor. The non-zero components of (114) are

$$g_{00} = e^{\nu}, \quad g_{11} = e^{\lambda}, \quad g_{22} = -r^2, \quad g_{33} = -r^2 \sin^2\theta. \tag{120}$$

The inverses are denoted by the superscript indices of the respective component, i.e $(g^{00})^{-1} \equiv g_{00}$; therefore, we have

$$g^{00} = e^{-\nu}, \quad g^{11} = e^{-\lambda}, \quad g^{22} = -r^{-2}, \quad g^{33} = -r^{-2}\sin^{-2}\theta. \tag{121}$$

For the general spherically symmetric metric shown in (114), we have, in the Lagrangian form,

$$L(x,\dot{x}) = \frac{1}{2}(-e^{\nu}c^2\dot{t}^2 + e^{\lambda}\dot{r}^2 + r^2(\sin^2\theta\dot{\phi}^2 + \dot{\theta})^2)). \tag{122}$$

Again recalling the Euler-Lagrange equation from (117), where q_μ are generalised coordinates, in our case $(q_0, q_1, q_2, q_3) = (t, r, \theta, \phi)$. Plugging (122) into (117) and equating to the corresponding components of the geodesic equation (119) leads to the set of equations:

$$\begin{aligned}
\ddot{t} &= -\frac{\partial\nu}{\partial r}\dot{r}\dot{t} - \frac{1}{2}\frac{\partial\nu}{\partial t}\dot{t}^2 - \frac{1}{2c^2}e^{\lambda-\nu}\frac{\partial\lambda}{\partial t}\dot{r}^2 = -\Gamma^t_{\mu\nu}\dot{x}^\mu\dot{x}^\nu,\\
\ddot{r} &= -\frac{1}{2}\frac{\partial\lambda}{\partial r}\dot{r}^2 - \frac{\partial\lambda}{\partial t}\dot{t}\dot{r} - \frac{1}{2}e^{\nu-\lambda}\frac{\partial\nu}{\partial r}c^2\dot{t}^2 + re^{-\lambda}(\sin^2\theta\dot{\phi}^2 + \dot{\theta}^2) = -\Gamma^r_{\mu\nu}\dot{x}^\mu\dot{x}^\nu,\\
\ddot{\theta} &= -\frac{2}{r}\dot{r}\dot{\theta} + \sin\theta\cos\theta\dot{\phi}^2 = -\Gamma^\theta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu,\\
\ddot{\phi} &= -2\cot\theta\dot{\theta}\dot{\phi} - \frac{2}{r}\dot{r}\dot{\phi} = -\Gamma^\phi_{\mu\nu}\dot{x}^\mu\dot{x}^\nu.
\end{aligned} \tag{123}$$

Comparing the right-hand side with terms in the centre column clearly shows the Christoffel symbols as:

$$\begin{array}{lll}
\Gamma^t_{rt} = \frac{1}{2}\frac{\partial\nu}{dr}, & \Gamma^t_{tr} = \frac{1}{2}\frac{\partial\nu}{dr}, & \Gamma^r_{tr} = \frac{1}{2}\frac{\partial\lambda}{\partial t},\\
\Gamma^r_{rt} = \frac{1}{2}\frac{\partial\lambda}{\partial t}, & \Gamma^\theta_{r\theta} = \frac{1}{r}, & \Gamma^\theta_{\theta r} = \frac{1}{r},\\
\Gamma^\phi_{\theta\phi} = \cot\theta, & \Gamma^\phi_{\phi\theta} = \cot\theta, & \Gamma^\phi_{\phi r} = \frac{1}{r},\\
\Gamma^\phi_{r\phi} = \frac{1}{r}, & \Gamma^t_{tt} = \frac{1}{2}\frac{\partial\nu}{dt}, & \Gamma^t_{rr} = \frac{1}{2c^2}e^{\lambda-\nu}\frac{\partial\lambda}{\partial t},\\
\Gamma^r_{rr} = \frac{1}{2}\frac{\partial\lambda}{\partial r}, & \Gamma^r_{tt} = \frac{c^2}{2}e^{\nu-\lambda}\frac{\partial\nu}{\partial r}, & \Gamma^r_{\phi\phi} = -re^{-\lambda}\sin^2\theta,\\
\Gamma^r_{\theta\theta} = -re^{-\lambda}, & \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta. &
\end{array} \tag{124}$$

All components that do no appear in (124) are zero.

4.3 Ricci Tensor

From the Riemann tensor we can find the Ricci tensor. Specifically, the Ricci tensor is formed by contracting the upper index with one of the lower indices. We can visualise the Ricci tensor as a tensor that quantifies the difference between the local geometry of a given metric and pseudo-Euclidean space, given by:

$$R_{\mu\nu} = R^\rho_{\rho\mu\nu} = \partial_\rho\Gamma^\rho_{\mu\nu} - \partial_\mu\Gamma^\rho_{\rho\nu} + \Gamma^\sigma_{\mu\nu}\Gamma^\rho_{\mu\nu} - \Gamma^\sigma_{\rho\nu}\Gamma^\rho_{\mu\sigma}. \tag{125}$$

With the help of the metric, we can further contract the Ricci tensor resulting in the Ricci scalar:

$$R = g_{\mu\nu} R_{\mu\nu}. \tag{126}$$

Both the Ricci tensor and Ricci scalar are key components of Einstein's field equation.
As we have already found, the Christoffel symbols (124) can be used to calculate the components of the Ricci tensor,

$$\begin{aligned}
R_{rt} = R_{tr} &= \partial_\rho \Gamma^\rho_{tr} - \partial_t \Gamma^\rho_{\rho r} + \Gamma^\sigma_{tr}\Gamma^\rho_{\rho\sigma} - \Gamma^\sigma_{\rho r}\Gamma^\rho_{t\sigma} \\
&= \partial_t \Gamma^t_{tr} + \partial_r \Gamma^r_{tr} - \partial_t(\Gamma^t_{tr} + \Gamma^r_{rr} + \Gamma^\theta_{\theta r} + \Gamma^\phi_{\phi r}) + \Gamma^t_{rt}(\Gamma^t_{tt} + \Gamma^r_{rt}) + \Gamma^r_{rt}(\Gamma^t_{tr} + \Gamma^r_{rr} + \Gamma^\theta_{\theta r} + \Gamma^\phi_{\phi r}) \\
&\quad - \Gamma^t_{tr}\Gamma^t_{tt} - \Gamma^r_{tr}\Gamma^t_{tr} - \Gamma^t_{rr}\Gamma^r_{tt} - \Gamma^r_{rr}\Gamma^r_{tr} \\
&= \partial_r \left(\frac{1}{2}\frac{\partial \lambda}{\partial t}\right) - \partial_t \left(\frac{1}{2}\frac{\partial \lambda}{\partial r} + \frac{2}{r}\right) + \frac{1}{2}\frac{\partial \nu}{\partial r}\frac{1}{2}\frac{\partial \lambda}{\partial t} + \frac{1}{2}\frac{\partial \lambda}{\partial t}\left(\frac{2}{r}\right) - \frac{1}{2c^2} e^{\lambda-\nu}\frac{\partial \lambda}{\partial t}\frac{c^2}{2} e^{\nu-\lambda}\frac{\partial \nu}{\partial r} \\
&= \frac{1}{2}\frac{\partial \lambda}{\partial t}.
\end{aligned} \tag{127}$$

We omit the analogous calculations to above for the remaining components. The components are:

$$\begin{aligned}
R_{\theta t} &= R_{t\theta} = 0, \quad R_{\phi t} = R_{t\phi} = 0, \quad R_{\theta t} = R_{t\theta} = 0, \\
R_{\theta r} &= R_{r\theta} = 0, \quad R_{\phi r} = R_{r\phi} = 0, \quad R_{\phi\theta} = R_{\theta\phi} = 0,
\end{aligned} \tag{128}$$

with the non-zero components are:

$$\begin{aligned}
R_{tt} &= c^2 e^{\nu-\lambda}\left(\frac{1}{2}\frac{\partial^2 \nu}{\partial r^2} + \frac{1}{4}\left(\frac{\partial \nu}{\partial r}\right)^2 - \frac{1}{4}\frac{\partial \lambda}{\partial r}\frac{\partial \nu}{\partial r} + \frac{1}{r}\frac{\partial \nu}{\partial r}\right) - \frac{1}{2}\frac{\partial^2 \lambda}{\partial t^2} + \frac{1}{4}\frac{\partial \nu}{\partial t}\frac{\partial \lambda}{\partial t} - \frac{1}{4}\left(\frac{\partial \lambda}{\partial t}\right)^2, \\
R_{rr} &= -\frac{1}{2}\frac{\partial^2 \nu}{\partial r^2} - \frac{1}{4}\left(\frac{\partial \nu}{\partial r}\right)^2 + \frac{1}{4}\frac{\partial \lambda}{\partial r}\frac{\partial \nu}{\partial r} + \frac{1}{r}\frac{\partial \lambda}{\partial r} + \frac{1}{c^2} e^{\lambda-\nu}\left(\frac{1}{2}\frac{\partial^2 \lambda}{\partial t^2} + \frac{1}{4}\left(\frac{\partial \lambda}{\partial t}\right)^2 - \frac{1}{4}\frac{\partial \nu}{\partial t}\frac{\partial \lambda}{\partial t}\right), \\
R_{\theta\theta} &= 1 - e^{-\lambda} - \frac{r}{2} e^{-\lambda}\left(\frac{\partial \nu}{\partial r} - \frac{\partial \lambda}{\partial r}\right),
\end{aligned} \tag{129}$$

and, finally,

$$R_{\phi\phi} = \sin^2\left(1 - e^{-\lambda} - \frac{r}{2} e^{-\lambda}\left(\frac{\partial \nu}{\partial r} - \frac{\partial \lambda}{\partial r}\right)\right) = \sin^2\theta R_{\theta\theta}. \tag{130}$$

4.4 Einstein's Vacuum Field Equation

We have introduced a geometric framework for general relativity: the idea of a space-time manifold, whilst omitting much of the differential geometry; the geodesic equation, how a freely falling particle traverses this manifold. However, we still need an equation to describe the gravitational fields, Einstein's vacuum field equation. This is analogous to the Poisson equation in Newtonian theory of gravity

$$\Delta\phi = 4\pi G\rho. \tag{131}$$

We have already shown the relationship between the geodesic equation in general relativity and

the Newtonian equation of motion.
Recall, the geodesic equation:

$$\ddot{x}^{\mu} + \Gamma^{\mu}_{\nu\sigma}\dot{x}^{\nu}\dot{x}^{\sigma} = 0, \tag{132}$$

where the over-dot represents differentiation by proper time τ.
The relevant Newtonian equation of motion can be written as:

$$\frac{d^2}{dt^2}\mathbf{x} - \boldsymbol{\nabla}\phi = 0. \tag{133}$$

Comparing the Newtonian equation of motion for a particle in a gravitational field with the geodesic equation, we see the gradient of the Newtonian potential ϕ is analogous to the Christoffel symbols. Recall the Christoffel symbols are formed from first-order derivatives of the metric g, (115). This suggests that the metric is analogous to ϕ which supports the fundamental idea that information about the gravitational field is deeply encoded in the metric tensor field. ρ, the mass density, is replaced by the energy-momentum tensor, T. From (131), the only variables with no information are $1/4\pi G$ which is just a differential operator we choose to label D. Hence, the field equation we are formulating should be of the form

$$Dg = T. \tag{134}$$

We require D to satisfy two conditions. First, motivated by comparison with Newtonian theory, Dg contains derivatives of g up to second order. Second, Dg is a tensor field of rank (0,2), i.e.

$$Dg = (Dg)_{\mu\nu}dx^{\mu} \otimes dx^{\nu} \in \tau^{0}_{2}M, \tag{135}$$

that satisfies

$$\nabla^{\mu}(Dg)_{\mu\nu} = 0, \tag{136}$$

[11].
These two properties uniquely fix the operator D into a form given by David Lovelock, shown over two papers [J. Math. Phys. 12, 498 (1971), J. Math. Phys. 13, 874 (1972)].
Theorem, Lovelock: Both axioms above are satisfied if, and only if, Dg has the following form:

$$(Dg)_{\mu\nu} = \frac{1}{\kappa}\left(R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} + \Lambda g_{\mu\nu}\right). \tag{137}$$

Here, $R_{\mu\nu}$ is the Ricci tensor field, R is the Ricci scalar and both κ and Λ are constants.
Einstein's field equation quickly follows,

$$R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} + \Lambda g_{\mu\nu} = -\kappa T_{\mu\nu}. \tag{138}$$

Λ is known as the *cosmological constant*, and κ is *Einstein's gravitational constant.* Also, notice that there is not an imposed specific coordinate system for this equation. Therefore, Einstein's field equation will hold the same form in any coordinate system. This is in contrast to special relativity.
Interestingly, we see Lovelock's thoerem was proved many years after Einstein's field vacuum

equations. Originally, Einstein arrived at this equation only after some strong assumptions. Also, in the first version, Einstein's field equation had no cosmological constant at all, this was added later. Einstein, according to G. Gamow, called the introduction of the cosmological constant his "biggest blunder", [11]!
The Einstein field equations formalises the distribution of energy and momentum on the spacetime manifold in a second-order, non-linear, PDE for the metric tensor.
Einstein's field equations for a vacuum, $T_{\mu\nu} = 0$, are

$$R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} + \Lambda g_{\mu\nu} = 0. \tag{139}$$

Contracting with $g^{\mu\nu}$ gives

$$R - 4\frac{R}{2} + 4\Lambda = 0. \tag{140}$$

Rearranging for the Ricci scalar, we find $R = 4\Lambda$. Inserting this into the vacuum field equation, we have

$$R_{\mu\nu} - 2\Lambda g_{\mu\nu} + \Lambda g_{\mu\nu} = 0. \tag{141}$$

Hence,

$$R_{\mu\nu} = \Lambda g_{\mu\nu}. \tag{142}$$

An *Einstein manifold* is the label given to any spacetime manifold whose Ricci tensor is in this form!
Critically for us, for determining the gravitational field near an isolated celestial body, we may consider Einstein's vacuum field equation without a cosmological constant, $\Lambda = 0$, [11].

$$R_{\mu\nu} = 0. \tag{143}$$

We are now in a position to solve for the functions $\mu(t, r)$ and $\nu(t, r)$. From our Ricci tensor, we see that twelve, out of the possible eighteen, Ricci tensor components are zero due to the spherical symmetry we imposed, also, we can omit (130) as it is dependent on $R_{\theta\theta}$. We are left with a system of four non-trivial, independent, PDEs:

$$R_{tr} = 0 \Rightarrow \frac{\partial \lambda}{\partial t} = 0, \tag{144}$$

$$\begin{aligned} &R_{tt} = 0 \\ &\Rightarrow 0 = c^2 e^{\nu-\lambda} \left(\frac{1}{2}\frac{\partial^2 \nu}{\partial r^2} + \frac{1}{4}\left(\frac{\partial \nu}{\partial r}\right)^2 - \frac{1}{4}\frac{\partial \lambda}{\partial r}\frac{\partial \nu}{\partial r} + \frac{1}{r}\frac{\partial \nu}{\partial r} \right) - \frac{1}{2}\frac{\partial^2 \lambda}{\partial t^2} + \frac{1}{4}\frac{\partial \nu}{\partial t}\frac{\partial \lambda}{\partial t} - \frac{1}{4}\left(\frac{\partial \lambda}{\partial t}\right)^2, \end{aligned} \tag{145}$$

$R_{rr} = 0$

$$\Rightarrow 0 = -\frac{1}{2}\frac{\partial^2\nu}{\partial r^2} - \frac{1}{4}\left(\frac{\partial\nu}{\partial r}\right)^2 + \frac{1}{4}\frac{\partial\lambda}{\partial r}\frac{\partial\nu}{\partial r} + \frac{1}{r}\frac{\partial\lambda}{\partial r} + \frac{1}{c^2}e^{\lambda-\nu}\left(\frac{1}{2}\frac{\partial^2\lambda}{\partial t^2} + \frac{1}{4}\left(\frac{\partial\lambda}{\partial t}\right)^2 - \frac{1}{4}\frac{\partial\nu}{\partial t}\frac{\partial\lambda}{\partial t}\right), \tag{146}$$

$$R_{\theta\theta} = 0 \Rightarrow 1 - e^{-\lambda} - \frac{r}{2}e^{-\lambda}\left(\frac{\partial\nu}{\partial r} - \frac{\partial\lambda}{\partial r}\right) = 0. \tag{147}$$

From (144), we see all terms with $\partial_t\lambda$ are zero. Hence,

$$\frac{\partial^2\nu}{\partial r^2} + \frac{1}{2}\left(\frac{\partial\nu}{\partial r}\right)^2 - \frac{1}{2}\frac{\partial\nu}{\partial r}\frac{\partial\lambda}{\partial r} + \frac{2}{r}\frac{\partial\nu}{\partial r} = 0, \tag{148}$$

$$\frac{\partial^2\nu}{\partial r^2} + \frac{1}{2}\left(\frac{\partial\nu}{\partial r}\right)^2 - \frac{1}{2}\frac{\partial\nu}{\partial r}\frac{\partial\lambda}{\partial r} - \frac{2}{r}\frac{\partial\lambda}{\partial r} = 0. \tag{149}$$

Subtracting (149) from (148), we have

$$\frac{\partial\nu}{\partial r} + \frac{\partial\lambda}{\partial r} = 0. \tag{150}$$

Then, differentiating with respect to t,

$$\frac{\partial^2\nu}{\partial t\partial r} + \frac{\partial^2\lambda}{\partial t\partial r} = 0. \tag{151}$$

Recall (144) implies the λ term drops out allowing us to directly integrate with respect to both parameters t and r,

$$\nu(r,t) = \tilde{\nu}(r) + f(t). \tag{152}$$

Inserting back into (150), we obtain

$$\frac{d}{dr}\left(\tilde{\nu}(r) + f(x) + \lambda(r)\right) = 0. \tag{153}$$

Therefore,

$$\tilde{v}(r) + \lambda(r) = k = const. \tag{154}$$

Hence,

$$\nu(t,r) = -\lambda(r) + f(t) + k. \tag{155}$$

Recalling the general spherically symmetric metric, from (114),

$$ds^2 = e^{-\lambda(r)}r^k e^{f(t)}c^2dt^2 - e^{\lambda(t,r)}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{156}$$

Transforming the time coordinate $\tilde{t} = e^{k/2}\int e^{f(t)/2}dt$, we can take the derivative with respect to t and separate giving $d\tilde{t} = e^{k/2}e^{f(t)/2}dt$.
And so we have,

$$e^{\nu(t,r)}dt^2 = e^{-\lambda(r)}d\tilde{t}^2. \tag{157}$$

Renaming $\tilde{t}$ with t and ds with g, we arrive at our spherically symmetric metric with time-independent components, with $\lambda = \lambda(r)$ and $\nu = -\lambda(r)$, we have

$$g = -e^{-\lambda(r)}c^2dt^2 + e^{\lambda(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{158}$$

Now we can reform equations: (144), (148), (149) and (147) with the aim of solving for λ. Clearly (144) is satisfied, both (148) and (149) give

$$-\frac{\partial^2\lambda}{\partial r^2} + \left(\frac{\partial\lambda}{\partial r}\right)^2 - \frac{2}{r}\frac{\partial\lambda}{\partial r} = 0, \tag{159}$$

and from (147),

$$1 - e^{-\lambda} + re^{-\lambda}\frac{d\lambda}{dr} = 0. \tag{160}$$

(160) is an inhomogeneous ordinary differential equation of first order with variable coefficients and can be solved directly with integration. Let,

$$-u = e^{-\lambda} \Rightarrow \frac{du}{dr} = -e^{-\lambda}\left(\frac{d\lambda}{dr}\right). \tag{161}$$

Therefore,

$$1 - u - r\frac{du}{dr} = 0, \tag{162}$$

rearranging, we can form

$$\int \frac{du}{1-u} = \int \frac{dr}{r}, \tag{163}$$

integrating gives

$$-\log(1-u) = \log(r) - log(r_s), \tag{164}$$

then, rearranging for u,

$$u = e^{-\lambda} = 1 - \frac{r_s}{r}. \tag{165}$$

Hence, $\lambda = -\log(1 - r_s/r)$. To accept this solution for λ, first we must check it satisfies (159), using substitution to rewrite the left-hand side we have

$$-\frac{\partial^2\lambda}{\partial r^2} + \left(\frac{\partial\lambda}{\partial r}\right)^2 - \frac{2}{r}\frac{\partial\lambda}{\partial r} = \frac{-rr_s(2r - r_s) + rr^2 + 2r_s(r^2 - r_sr))}{r(r^2 - r_sr)^2} = 0, \tag{166}$$

as required. Inserting this solution into the metric (156) we finally arrive at the Schwarzschild solution:

$$g = -\left(1 - \frac{r_s}{r}\right)c^2dt^2 + \frac{dr^2}{(1 - \frac{r_s}{r})} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{167}$$

Where r is the radius from the centre of the body causing the perturbation of space-time and r_s is some arbitrary length known as the *Schwarzschild radius* and c is the speed of light in vacuum.

4.5 Lightlike Geodesics in Spherically Symmetric Spacetime

We can use the Schwarzchild solution, as seen in (167), to calculate how light propagates spherically symmetric spacetime. Recall this spacetime geometry forms around a point mass and thus we aim to re-derive (79). Lightlike geodesics are the "straight" lines a *freely falling* massless particle will take and, as seen above, we can use Lagrangian mechanics to analyse the geodesic equation (116). We are working with a spherically symmetric spacetime which produced because of a mass, M. Thus, any light rays being effected by the curvature will only be perturbed towards the mass, critically, staying on the same plane. We can therefore, without breaking generality, arrange our coordinate system such that we work in the equatorial plane, $\theta = \pi/2$. Hence, our Lagrange function is simplified,

$$L(x, \dot{x}) = \frac{1}{2} g_{\mu\nu}(x) \dot{x}^\mu \dot{x}^\nu = \frac{1}{2} \left(- \left(1 - \frac{r_s}{r}\right) c^2 \dot{t}^2 + \frac{1}{\left(1 - \frac{r_s}{r}\right)} \dot{r}^2 + r^2 \dot{\phi}^2 \right), \tag{168}$$

recall $\dot{x}_i \equiv dx_i/ds$. By definition, for lightlike geodesics

$$\frac{1}{2} g_{\mu\nu}(x) \dot{x}^\mu \dot{x}^\nu = 0, \tag{169}$$

[11]. Computing the Euler-Lagrange equations of motion, (117), firstly with respect to t and secondly with respect to ϕ, will give us two more equations describing lightlike geodesics.
t-component:

$$0 = -c^2 \frac{d}{ds} \left(\left(1 - \frac{r_s}{r}\right) \dot{t} \right) \Rightarrow E = \left(1 - \frac{r_s}{r}\right) \dot{t} \rightarrow \dot{t} = \frac{E}{\left(1 - \frac{r_s}{r}\right)}, \tag{170}$$

ϕ-component:

$$0 = \frac{d}{ds}(r^2 \dot{\phi}) \Rightarrow L = r^2 \dot{\phi} \rightarrow \dot{\phi} = \frac{L}{r^2}, \tag{171}$$

where E and L are constants.
To get interpretable information, first divide (169) by $\dot{\phi}^2$,

$$- \left(1 - \frac{r_s}{r}\right) \frac{c^2 \dot{t}^2}{\dot{\phi}^2} + \frac{1}{\left(1 - \frac{r_s}{r}\right)} \frac{\dot{r}^2}{\dot{\phi}^2} + r^2 \tag{172}$$

Next, substituting (170) and (171) yields,

$$\frac{-c^2 E^2 r^4}{\left(1 - \frac{r_s}{r}\right) L^2} + \frac{1}{\left(1 - \frac{r_s}{r}\right)} \left(\frac{dr}{d\phi}\right)^2 + r^2 = 0. \tag{173}$$

Rearranging for $d\phi$,

$$d\phi = dr \left(\frac{c^2 E^2 r^4}{L^2} - r^2 + r_s r \right)^{-1/2}. \tag{174}$$

A light ray coming from a distant source, $r \simeq \infty$, will have a minimum radius $r = r_m$. As the lens is spherically symmetric, the magnitude of the lensing when the light ray is approaching the lens will be equal to the magnitude of the lensing when the light ray is travelling away from the lens, with r_m occurring at the half way point, see figure 9. Therefore, we can decompose the integral into two equal halves,

$$\int d\phi = 2\int_{r_m}^{\infty}\left(\frac{c^2E^2r^4}{L^2} - r^2 + r_s r\right)^{-1/2} dr. \tag{175}$$

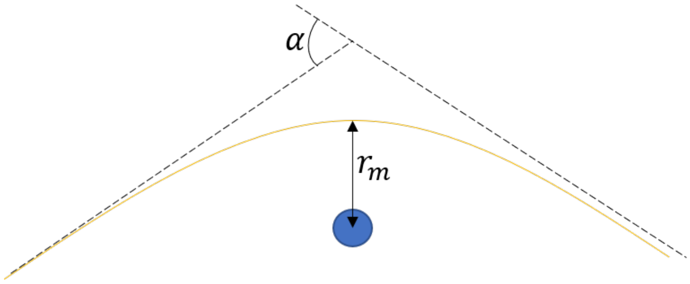

Figure 9: *Lightlike* geodesic propagating a spherically symmetric spacetime perturbation, i.e. a star, [11] .

From Figure 9, it is clear the change in ϕ,

$$d\phi = \pi + \alpha, \tag{176}$$

where α is the deflection angle.
We can further work with r_m, we see r_m is a turning point, see Figure 9; therefore, differentiating r with respect to ϕ and evaluating at r_m will be zero,

$$0 = \left(\frac{dr}{d\phi}\right)^2 |_{r_m}. \tag{177}$$

(173) evaluated at r_m gives

$$0 = \frac{c^2E^2r_{r_m}^4}{L^2} - r_m^2 + r_s r_m, \tag{178}$$

it follows,

$$\frac{c^2E^2}{L^2} = \frac{1}{r_m^2} = \frac{r_s}{r_m^3}. \tag{179}$$

Relating (175) and (176) and substituting (179). Yields

$$|\alpha| + \pi = 2\int_{r_m}^{\infty}\left(\left(\frac{1}{r_m^2} - \frac{r_s}{r_m^3}\right)r^4 - r^2 + r_s r\right)^{-1/2} dr. \tag{180}$$

We have derived a formula allowing the calculation of the deflection angle of a *lightlike* geodesic when perturbed by a spherically symmetric gravitational field, assuming that the Schwarzchild radius is small relative to the physical radius from centre of mass, r_*. Furthermore, the closest the light ray reaches radially, r_m, is clearly bigger then the physical radius. We have $r_s \ll r_* \leq r_m$.

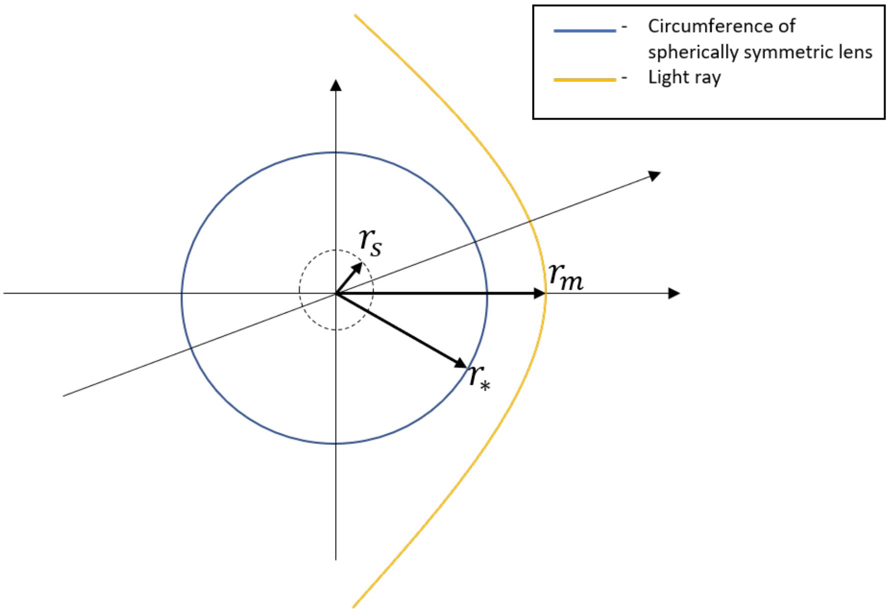

Figure 10: Picture to help visualise r_m, r_s and r_*

Therefore, r_s/r_m is small, and so we omit terms.
Exception: This breaks down with black holes. If a star is massive enough, the gravitational force on itself could cause it to collapse causing the matter to become so dense that the distance r_s approaches r_m and maybe even exceeds it. In a situation such as this one, our assumption would not be valid. In many cases, matter is not nearly as dense as required to cause problems. For example the Sun, our closest star, has a radius $r_m \approx 700,000\,\text{km}$ and a schwarzschild radius $r_s \approx 3\,\text{km}$.

4.6 Newtonian Limit

We now apply Einstein's theory of gravitation with constraints to show we can reproduce Newton's gravitational theory. Our aim is to show that Newton's gravitational theory is a viable theory for a large, but constrained, domain of validity. Newton's theory relies on the equation of motion for a particle in a gravitational potential and the field equation for the gravitational potential. First recall the geodesic equation, (119)

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\sigma}\dot{x}^\nu\dot{x}^\sigma = 0, \tag{181}$$

where the over-dot represents differentiation by proper time τ. The related statement in Newtonian theory of gravity is

$$\frac{d^2}{dt^2}\mathbf{x} - \boldsymbol{\nabla}\phi = 0, \tag{182}$$

and, second, Einstein's field equations

$$R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} + \Lambda g_{\mu\nu} = -\kappa T_{\mu\nu}, \tag{183}$$

with the analogous statement in Newtonian theory of gravity, Poisson equation (80):

$$\Delta\phi = 4\pi G\rho. \tag{184}$$

First we list the assumption necessary for the Newtonian limit, [11].
(a) The gravitational field is weak, i.e. the metric differs only a little from the metric of special relativity, $\eta_{\mu\nu}$, allowing us to express the metric as a sum of the special relativity metric, and $h_{\mu\nu}$ is sufficiently small such that we can omit terms above first order in $h_{\mu\nu}$ and $\partial_\sigma h_\mu \nu$,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \tag{185}$$

Making an ansatz, that the inverse metric up to order h is given by: $g^{\nu\rho} = (\eta^{\mu\rho} - \eta^{\nu\tau}\eta^{\rho\lambda}h_{\tau\lambda})$
Check:

$$\begin{aligned} g_{\mu\nu}g^{\mu\rho} &= (\eta_{mu\nu} + h_{\mu\nu})(\eta^{\mu\rho} - \eta^{\nu\tau}\eta^{\rho\lambda}h_{\tau\lambda}) = \eta_{\mu\nu}\eta^{\nu\rho} + h_{\mu\nu}\eta_{\nu\rho} - \eta_{\mu\nu}\eta^{\nu\tau}\eta^{\rho\lambda}h_{\tau\lambda} \\ &= \delta^\rho_\mu + h_{\mu\lambda}\eta^{\lambda\rho} - \delta^\tau_\mu\eta^{\rho\lambda}h_{\tau\lambda} = \delta^\rho_\lambda \end{aligned} \tag{186}$$

From the assumption, we have omitted terms greater than first order. As required, $g^{\mu\nu}$ is our inverse metric
(b) The gravitational field varies so slowly that it can be approximated as being time independent,

$$\partial_0 h_{\mu\nu} = 0. \tag{187}$$

(c) The particle velocity is small in comparison to the speed of light,

$$\left|\frac{dx^0}{d\tau}\right| = \left|\frac{cdt}{d\tau}\right| \approx c, \tag{188}$$

and

$$\left|\frac{dx^i}{d\tau}\right| \approx \left|\frac{dx^i}{dt}\right| \ll c, \tag{189}$$

where τ denotes proper time. *Proper time* along a world line is defined as time as measured by a clock following that line, [11]. The second equation just shows that if the local spacetime geometry is close to flat, the length of a world line does not differ noticeably from that of a similar world line on a flat space time.
(d) Matter moves so slowly is can be approximate as being at rest, leaving only the mass density, μ, as a source of gravity $T_{00} = c^2\mu$, $T_{0i} = 0$ and $T_{ik} = 0$.
First, to calculate the Newtonian limit, sum over the spatial components of the geodesic equation for a freely falling particle, as seen in (181),

$$\frac{d^2x^i}{d\tau^2} + \Gamma^i_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} + \Gamma^i_{j0}\frac{dx^j}{d\tau}\frac{dx^0}{d\tau} + \Gamma^i_{0j}\frac{dx^0}{d\tau}\frac{dx^j}{d\tau} + \Gamma^i_{jk}\frac{dx^j}{d\tau}\frac{dx^k}{d\tau} = 0. \tag{190}$$

From (c), omit the third, fourth and fifth term,

$$\frac{d^2x^i}{d\tau^2} = -\Gamma^i_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} = -\Gamma^i_{00}c^2. \tag{191}$$

From (115) and using (a), we can calculate Γ^i_{00}:

$$\Gamma^i_{00} = \frac{1}{2} g^{i\mu}(\partial_0 g_{\mu 0} + \partial_0 g_{\mu 0} + \partial_\mu g_{00}) = \frac{1}{2}\eta^{i\mu}(2\partial_0 h_{\mu 0} - \partial_\mu h_{00}). \tag{192}$$

Then, the time independent condition (b) and the fact that the spacial components of $\eta^{\mu\sigma}$ are the identity matrix yields

$$\Gamma^i_{00} = -\frac{1}{2}\delta^{ij}\partial_j h_{00}. \tag{193}$$

Hence, we can write the equation of motion in the following form:

$$\frac{d^2 x^i}{dt^2} = \frac{1}{2}\delta^{ij}\partial_j h_{00}. \tag{194}$$

Writing the gravitational potential as $-\partial_i \phi = \ddot{x}_i$, gives

$$\ddot{x}_i = -\partial_i \phi = \frac{1}{2}\delta^{ij}\partial_j h_{00}. \tag{195}$$

Therefore,

$$h_{00} = -\frac{2\phi}{c^2}, \tag{196}$$

knowing $\eta_{00} = \pm 1$ and using (a), the curvature component easily follows:

$$g_{00} = -\left(1 + \frac{2\phi}{c^2}\right). \tag{197}$$

This is an important result, recall from section 3.5 we start by introducing the refractive index $n(\mathbf{x}) = 1 + \delta(\mathbf{x})$ where $\delta = -2\phi(\mathbf{x})/c^2$!
Now consider Einstein's field equation, (183), and contract $g^{\mu\nu}$. Then we have,

$$R - \frac{R}{2}4 + \Lambda 4 = \kappa T_{\mu\nu} g^{\mu\nu}, \tag{198}$$

therefore,

$$R = 4\Lambda - \kappa T_{\mu\nu} g^{\mu\nu}. \tag{199}$$

Hence, we can re-write the field equation

$$R_{\mu\nu} = \Lambda g_{\mu\nu} + \kappa\left(T_{\mu\nu} - \frac{1}{2}T_{\rho\sigma}g^{\rho\sigma}g_{\mu\nu}\right). \tag{200}$$

Consider the 00-component of the field equation,

$$R_{00} = \Lambda g_{00} + \kappa\left(T_{00} - \frac{1}{2}T_{00}g^{00}g_{00}\right), \tag{201}$$

from (d), we can simplify, giving

$$R_{00} = \Lambda g_{00} + \kappa c^2 \rho\left(1 - \frac{1}{2}g^{00}g_{00}\right), \tag{202}$$

from (a), we have $g_{00} = \eta_{00} + h_{00} \approx -1$ and $g_{00}g^{00} = 1$, giving us

$$R_{00} = -\Lambda + \frac{1}{2}\kappa c^2 \rho. \tag{203}$$

Recall (125), we can calculate R_{00}

$$R_{00} = R^{\rho}_{\rho 00} = \partial_\rho \Gamma^{\rho}_{00} - \partial_0 \Gamma^{\rho}_{\rho 0} + \Gamma^{\sigma}_{00}\Gamma^{\rho}_{00} - \Gamma^{\sigma}_{\rho 0}\Gamma^{\rho}_{0\sigma}. \tag{204}$$

However, from (a), we can neglect the second two terms, and from (b) we know $\partial_0 \Gamma^{\mu}_{\rho\sigma} = 0$. Hence,

$$R_{00} = \partial_\mu \Gamma^{\mu}_{00} = \partial_i \Gamma^{i}_{00}, \tag{205}$$

together with our result that $\Gamma^{i}_{00} = -\frac{1}{2}\delta^{ij}\partial_j h_{00}$, we come to

$$R_{00} = -\frac{1}{2}\delta^{ij}\partial_i\partial_j h_{00} = -\frac{1}{2}\Delta h_{00}. \tag{206}$$

Substituting this into (203), we find

$$\Delta h_{00} = 2\Lambda - \kappa c^2 \rho. \tag{207}$$

However, we have previously calculated $h_{00} = -2\phi/c^2$; therefore,

$$\Delta \phi = -\Lambda c^2 + \frac{1}{2}\kappa c^4 \rho. \tag{208}$$

If $\Lambda = 0$ and $\kappa = 8\pi G/c^4$ we have Poisson's equation of Newton's gravitational theory:

$$\Delta \phi = 4\pi G \rho. \tag{209}$$

This then implies, if Λ and κ are as required. Newton's gravitational theory can replace Einstein's gravitational theory for a large range of validity.
Example: Let the mass density $\rho = 0$, using spherical coordinates

$$\Delta \phi(r) = \frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{d}{dr}\phi(r)\right) = 0. \tag{210}$$

Therefore, by directly integrating we find the general solution for $r \neq 0$,

$$\phi(r) = \frac{C_1}{r} + C_2. \tag{211}$$

To take this further, we can use Gauss's theorem,

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{F} d^3V = \oint_{\partial V} \mathbf{F} \cdot \mathbf{da}. \tag{212}$$

For this equation, the left-hand side is a volume integral, and the right-hand side is a closed surface integral. We choose to integrate over a large sphere radius R, which has volume parametrised by

$$V(r) = r \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}. \tag{213}$$

For $\theta \in \ [0, \pi]$, $\phi \in \ [0, 2\pi[$ and $r \in \ [0, R]$. The surface element $\mathbf{da} = R^2 \sin(\theta)\hat{\mathbf{e}}_r d\theta\phi$, also $\boldsymbol{\nabla} \cdot \mathbf{F} = 4\pi G\rho$. For a central potential we have $\mathbf{F} = F_r(r)\hat{\mathbf{e}}_r$, and the right-hand side of (212) becomes,

$$\oint_{\partial V} \mathbf{F} \cdot \mathbf{da} = F_r(r) 4\pi R^2. \tag{214}$$

The left-hand side of (212) is,

$$\int_V 4\pi G\rho dV = 4\pi G \int_V \rho dV = 4\pi GM, \tag{215}$$

rearranging gives Newton's inverse square law:

$$F_r(r) = \frac{GM}{r^2}. \tag{216}$$

Differentiating, yields the Newtonian potential

$$\phi(r) = \frac{dF_r(r)}{dr} = -\frac{GM}{r}. \tag{217}$$

This is an implication of Newton's force law $F = GMm/r^2$.

4.7 Schwarzschild Radius

We now wish to further discuss the physical interpretation of the integration constant r_s from (167). Recall that the Newtonian limit was valid if four conditions were met: if r is sufficiently large, we see the first condition, (a) is satisfied, i.e. the gravitational field is sufficiently weak; (b) the gravitational field varies so slowly it can be approximated as being time independent, is satisfied everywhere; (d) matter moves so slowly it can be approximated at being at rest, is also always satisfied; and (c) gives no restrictions on the metric. Hence, if r is sufficiently large the Newtonian limit holds for the Schwarzschild metric. Therefore, from (197),

$$-\left(1 + \frac{2\phi}{c^2}\right) = -\left(1 - \frac{r_s}{r}\right). \tag{218}$$

A spherically symmetric Newtonian field implies

$$\phi = \phi(r) = -\frac{GM}{r}, \tag{219}$$

as seen in (71). Therefore,

$$-\left(1 - \frac{r_s}{r}\right) = -\left(1 + \frac{2GM}{c^2 r}\right). \tag{220}$$

Hence, the integration constant

$$r_s = \frac{2GM}{c^2}. \tag{221}$$

It is clear from this that the Schwarzschild radius, r_s, is determined only by the mass M of the body causing the spherically symmetric metric, i.e. the mass of a star or black hole. Thus, in some situations, we can measure an unknown mass M using Newtonian theory!!
It is clear for positive M that the Schwarzschild radius will be positive. For the vast majority of cases, the radius of the central body, the body causing the curvature, r_* is much bigger then r_s. Hence, the Schwarzschild metric will be regular in its entire range of validity.

4.8 Einstein's Deflection Angle

Let us solve the elliptical integral seen in (180):

$$\int_{r_m}^{\infty} \left(\left(\frac{1}{r_m^2} - \frac{r_s}{r_m^3} \right) r^4 - r^2 + r_s r \right)^{-1/2} dr \equiv I \tag{222}$$

let $\epsilon \equiv r_s/r_m$, recall in the large domain of validity $r_s \ll r_m$. Hence, $\epsilon \ll 1$ and so whilst keeping an acceptable degree of accuracy we can omit terms of order ϵ^2 and above. To do this, first substitute$u = r/r_m$

$$\begin{aligned} I &= \int_1^{\infty} \left((1-\epsilon)u^4 - u^2 + \epsilon u \right)^{-1/2} du, \\ &= \int_1^{\infty} (u^4 - u^2)^{-1/2} + \left(1 + \epsilon \frac{u - u^4}{u^4 - u^2} \right)^{-1/2} du, \\ &\approx \int_1^{\infty} (u^4 - u^2)^{-1/2} + \left(1 + \frac{\epsilon}{2} \frac{u - u^4}{u^4 - u^2} \right) du, \\ &= \int_1^{\infty} (u^{-1}(u^2-1))^{-1/2} du + \frac{\epsilon}{2} \int_1^{\infty} \frac{1}{u^2} \frac{u^2+u+1}{(u+1)\sqrt{u^2-1}} du. \end{aligned} \tag{223}$$

With the aid of computer algebra, see appendix A, we have

$$I = \frac{\pi}{2} + \epsilon = \frac{\pi}{2} + \frac{r_s}{r_m}. \tag{224}$$

Now, substituting this into (180) yields,

$$|\alpha| + \pi = 2(\frac{\pi}{2} + \frac{r_s}{r_m}). \tag{225}$$

Therefore,

$$|\alpha| = 2\frac{r_s}{r_m} = \frac{4GM}{c^2 r_m}. \tag{226}$$

Now, for a light ray grazing the surface of the Sun, we have $r_m \approx 7 \times 10^5$km and $r_s \approx 3$km. Hence,

$$\delta = 1.75''. \tag{227}$$

5 Lensing Effects

5.1 Multiple Images

We have discussed lensing by point masses, spherically symmetric mass and a general mass distribution, however, not how these phenomena are observed.
Gravity is an attractive force that's strength decays as the radius increases. When light rays pass a gravitational lens, the distance that any ray passes the centre of the lens will dictate the magnitude of its deflection. Light rays passing at a greater distance are deflected less then light rays passing closer to the centre of the lens. Because of this, we get *focal lines* instead of a single focal point. For this section, we omit any distortion in the shape of the image.

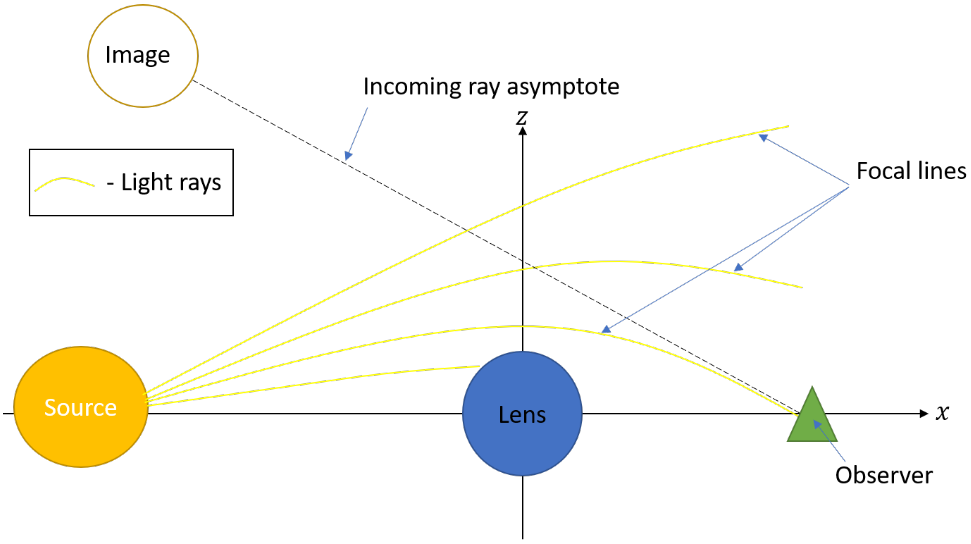

Figure 11: Light rays from a Source being deflected by a Lens. The image that the Observer sees has an apparent position which is different from the Source's actual position.

The Source, which is radially omitting light rays, is located roughly behind a lens from the Observer's point of view. The Observer receives an image, this implies they are located in a range between the maximal deflected light ray and minimal deflected light ray. The *image* of the Source is in an *apparent position* which lies on the incoming ray's asymptote. How much this image differs in position from the actual Source depends on the magnitude of the deflection angle.
If two light rays emanating from the same source travelled to an observer taking two different paths, the observer would see two images of the source at two different apparent positions, this phenomenon is known as *multiple imaging*, see Figure 1.
The picture that the observer sees is dependent on the geometry of the system in question. If the source, radially omitting light, is directly behind an appropriate spherically symmetric (or point mass) lens, the observer will see an Einstein ring.
Definition: An *Einstein ring* is the ring shaped image formed when the source, lens and observer all lie in perfect alignment on the same axis.
Definition: The *Einstein radius* is the radius of an Einstein ring when measured at the lens.

If the system, the position of the Source, lens and observer, is not perfectly aligned then a single arc or multiple arcs could form, see Figure 12.

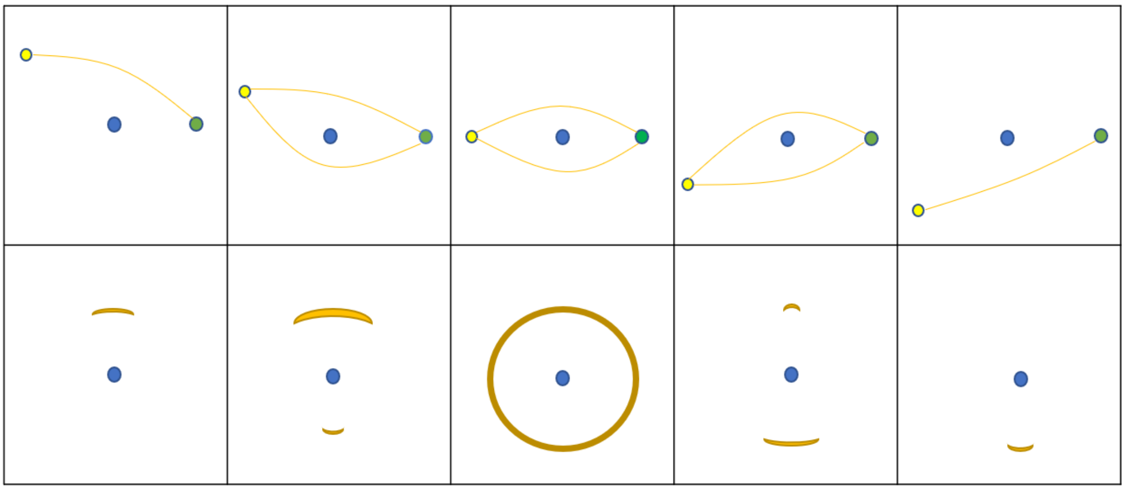

Figure 12: The top row shows the geometry of the system and the bottom row shows the image the observer will see. The yellow, blue and green circles represent the Source, Lens and Observer respectively.

Example: Deflection angle for a point mass lens.
Light from a source would be observed at at angle θ_2 if it were not deflected by a lens. We choose the x-axis to be on the line connecting the Observer and the Source. The distance in x between: the Source plane and Lens plane is D_{SL}; the Lens plane and Observer plane is D_L; and the Source plane and Observer plane is D_S.

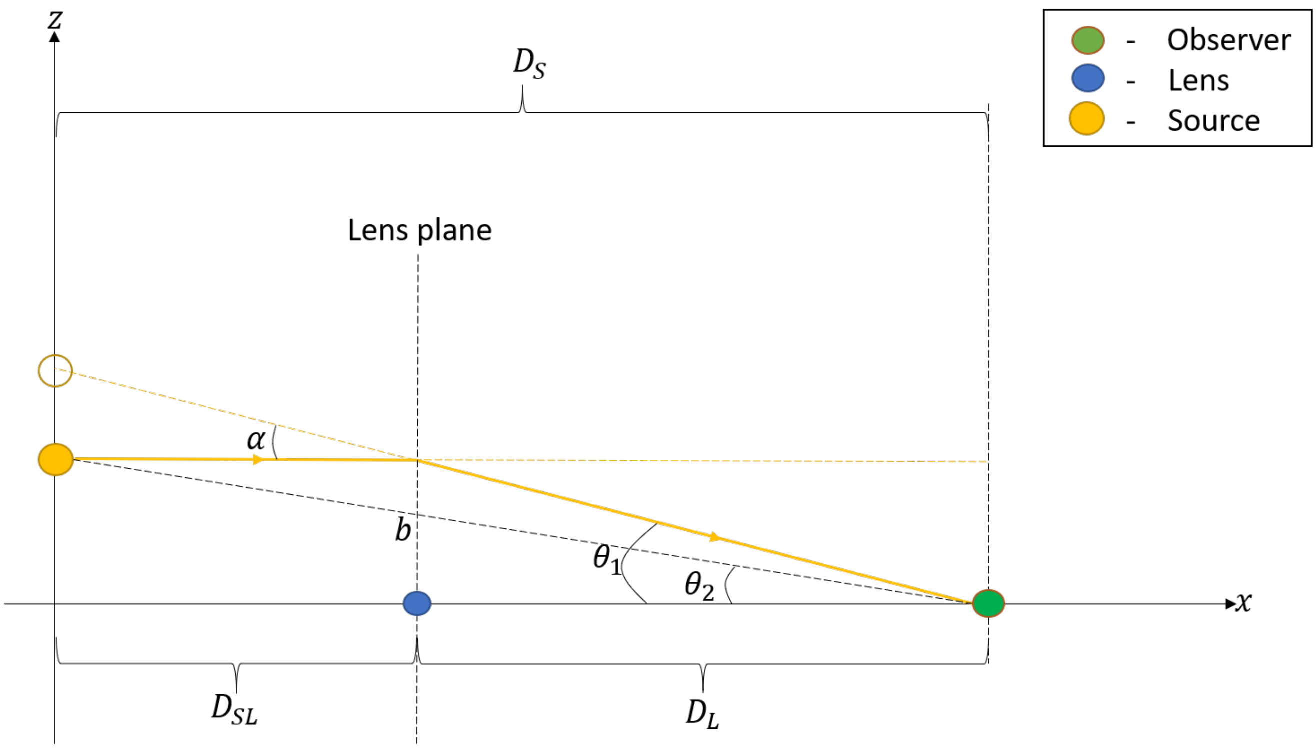

Figure 13: An example of one light ray being deflected by a point mass. With deflection angle α, impact parameter b and the angle between the apparent position and the x-axis θ_1.

As we are working with a small deflection angle, $\theta_1 \ll 1$, $\theta_2 \ll 1$ and $\alpha \ll 1$, we can approximate the impact parameter as

$$b \approx D_L \tan(\theta_1) \approx D_L \theta_1. \tag{228}$$

Recalling Einstein's angle of deflection, from a point mass, seen in (79), we have

$$\alpha = \frac{4GM}{c^2 b} \Rightarrow \alpha(\theta_1) = \frac{4GM}{c^2 D_L \theta_1}. \tag{229}$$

The vertical distance in the Source plane spanned by θ_1 is equal to the sum of the vertical distances spanned by θ_2 and α in the Source plane,

$$D_S \theta_1 \approx D_S \theta_2 + D_{SL} \alpha. \tag{230}$$

Re-arranging for α, we find

$$\alpha \approx \frac{D_S}{D_{SL}}(\theta_1 - \theta_2). \tag{231}$$

Example: Einstein ring.
If the system is axially symmetric about the x-axis, then the formation of an Einstein ring is possible. We assume the deflection angle is small, hence, $r_m \approx b$. Using this approximation, we similarly arrive at (231) for a spherically symmetric lens.

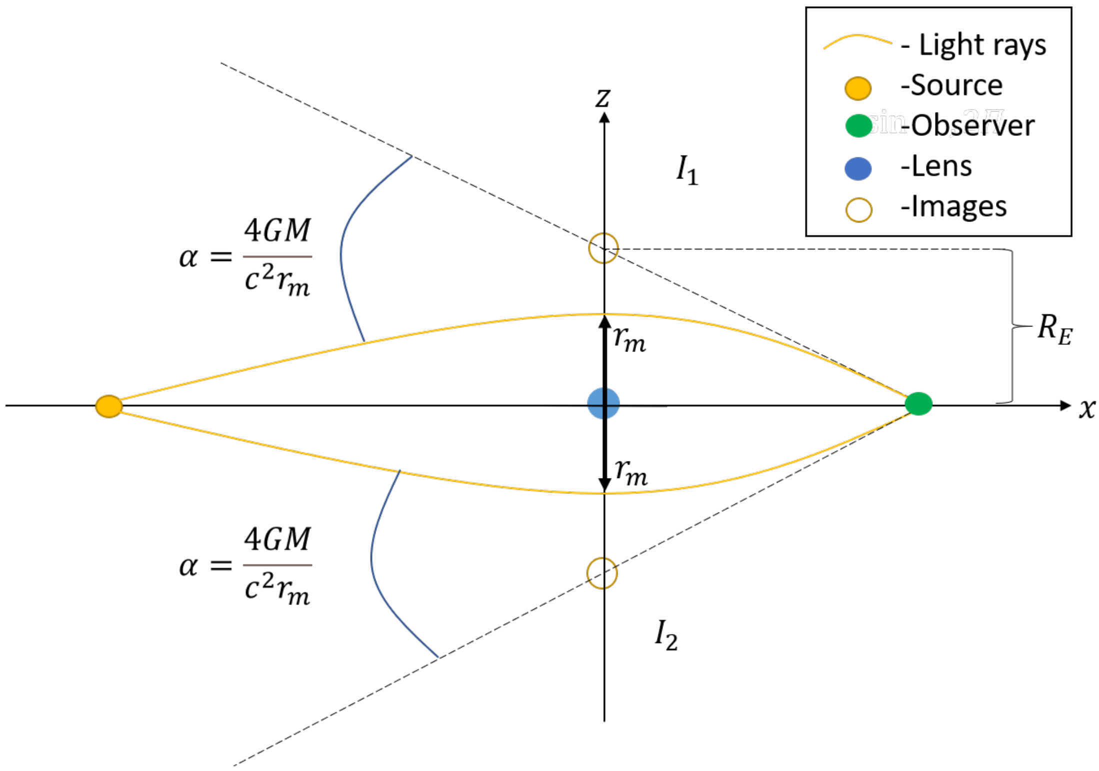

Figure 14: The geometry of a system capable of producing an Einstein ring. The Source, Lens and Observer lie on a similar axis.

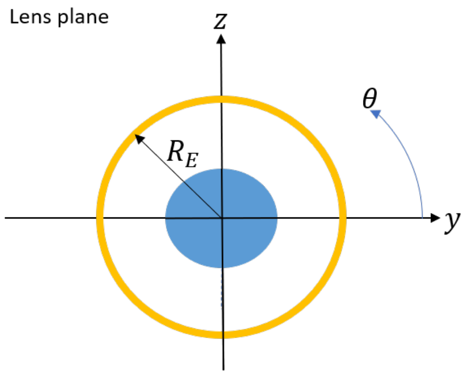

Figure 15: Cross section of Figure 14 at the Lens.

In Figure 14, $\theta_2 = 0$. Hence,

$$\alpha = \alpha(\theta_1) = \frac{D_S \theta_1}{D_{SL}}. \tag{232}$$

By substituting (229) into (232), we find

$$\theta_1^2 = \frac{4GM}{c^2} \frac{D_{SL}}{D_L D_S}, \tag{233}$$

therefore,

$$\theta_1 = \sqrt{\frac{4GM}{c^2} \frac{D_{SL}}{D_L D_S}} \equiv \theta_E. \tag{234}$$

Using the small angle approximation, the Einstein radius is the angle that the light from the Source hits the Observer with respect to the common axis, multiplied by the distance to the Lens,

$$R_E \approx \theta_E D_L. \tag{235}$$

Example: Let a person on earth be the Observer, the Sun is the spherically symmetric lens and there is a point source directly behind the Sun which is the same distance again, i.e. $D_S = 2D_{SL} = 2D_L$. The distance from the earth to the Sun is approximately 150 million kilometres, [18]. Knowing the Sun's Schwartzschild radius,

$$r_s = \frac{2GM}{c^2} \approx 3\,\text{km}, \tag{236}$$

[11]. We easily calculate the theoretical Einstein radius,

$$R_E \approx \sqrt{6 \frac{D_{SL}}{D_L D_S}} D_L \approx 21,000\,\text{km}. \tag{237}$$

Clearly, in this case, the Einstein radius is much larger then the Schwartzschild radius; however, it is not as large as the physical radius of the Sun. This implies only extremely dense clumps of matter will be suitable to form a visible Einstein ring.

6 Summary

6.1 Discussion

We introduce the propagation of light through inhomogeneous media in the 'Optics - Short-Wave Asymptotics' section. Using Maxwell's equations, we input a wave ansatz and form the wave equation, relating the eikonal function with the refractive index of the media through which the light rays are propagating. In general, optical systems are large in comparison to the wave length of light; therefore, light ray solutions to the wave equation are plane waves. This is useful because we can explore different types of lensing, including gravitational lensing, by imposing boundary conditions on the eikonal function as well as the specific refractive index of the lens. We solve the eikonal equation for a point mass, giving us Einstein's deflection angle for a point mass. One of the key positives for using this method to explore gravitational lensing is that it is easily adaptable for general mass distributions, assuming we know the relevant density distribution. However, using this method does not give us much understanding of the physics behind these results.

In the 'General Relativity' section, we omit lots of the underlining differential geometry of manifolds and metrics, this is a clear limitation; however, not particularly enlightening content. We solve Einstein's vacuum field equations for a spherically symmetric case, known as Schwarzschild's solution. The limitations of this solution arise because of the constraints: the metric is asymptotically flat, meaning that for large radii the metric approaches the Minkowski metric; r_s/r_m is small, which is not the case for black holes; the Schwarzschild metric is spherically symmetric and static; and, finally, the mass body causing the spacetime curvature is not rotating. A natural progression for what we have covered in this section would be to study Schwarzschild black holes.

The 'Lensing Effects' section discusses the idea of multiple imaging and how, under specific circumstances, this leads to the formation of arcs and rings. We have not, however, explored many other phenomena arising from gravitational lensing including, magnification, microlensing and image shapes, see [3]. The image shape of an extended source (e.g. a galaxy) could be distorted; light from opposite ends of such a source will be deflected by different amounts; therefore, the image appears squished.

Due to time constraints, I was unable to explore the implication of dark matter. Using gravitational lensing, we can estimate the amount of mass in a section of space. In some cases, manually estimating the mass in a certain section, using stars and gas only, falls short, i.e. we expect a certain mass but lensing tells us there is much more mass that we cannot see. This unseen mass is known as dark matter, see [9].

6.2 Conclusion

Gravitational lensing is a consequence of general relativity and is a useful tool that we can use to learn more about the universe we live in. This project has explored the basic mathematics, through optical short-wave asymptotics and general relativity, that form a foundation needed in order to delve deeper into the topic. We have discussed its historical relevance as well as some of its manifestations including the formation of Einstein rings.

References

[1] M. Bartelmann and P. Schneider. Weak gravitational lensing. *Physics Reports*, 340(4-5):291–472, Jan 2001.

[2] S. M. Carroll. Lecture notes on general relativity, 1997. University of Chicago.

[3] S. Dodelson. *Gravitational Lensing*. Cambridge University Press, 2017.

[4] Rivera-Thorsen et al. ESA/Hubble, NASA. The sunburst arc.

[5] T. Heinzl. Partial differential equations. University of Plymouth lecture notes, 2020-21.

[6] T.W.B. Kibble and F.H. Berkshire. *Classical Mechanics*. Imperial College Press, london, 2004.

[7] B. King. Classical mechanics. University of Plymouth lecture notes, 2020-21.

[8] L. D. Landau and E. M. Lifshitz. *The classical theory of fields*. Revised second edition. Course of Theoretical Physics, Vol. 2. Translated from the Russian by Morton Hamermesh. Pergamon Press, Oxford-London-Paris-Frankfurt; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1962.

[9] R. Massey, T. Kitching, and J. Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, Jul 2010.

[10] R. Narayan and M. Bartelmann. Lectures on gravitational lensing. *arXiv preprint astro-ph/9606001*, 1996.

[11] V. Perlik. General relativity. *Universität Bremen*, 2012.

[12] H. Rhett. How fast is the earth moving? https://www.scientificamerican.com/article/how-fast-is-the-earth-mov/, 1998. [Online; accessed 04-May-2021].

[13] H. Römer. *Theoretical Optics*, chapter 7 - Short-Wave Asymptotics, pages 107–119. John Wiley & Sons, Ltd, 2004. Weinheim, Germany.

[14] H. Römer. *Theoretical Optics*, chapter 2 - The Electrodynamics of Continuous Media, pages 15–43. John Wiley & Sons, Ltd, 2004. Weinheim, Germany.

[15] R. Sexl and H. Sexl. White dwarfs-black holes. an introduction to relativistic astrophysics. 1979. Academic Press, New York.

[16] N. Straumann. Weak and strong lensing statistics. *Space Science Reviews*, 100:29–38, 01 2002.

[17] Wikipedia. Shell theorem — Wikipedia, 2020. [Online; accessed 4-December-2020].

[18] Wikipedia. Astronomical unit — Wikipedia, the free encyclopedia, 2021. [Online; accessed 29-March-2021].

[19] Wikipedia. Minkowski space. `https://en.wikipedia.org/w/index.php?title=Minkowski_space&oldid=1011762927`, 2021. [Online; accessed 23-March-2021].

A Elliptic integral

Evaluating Perlick's elliptic integral using maple algebra.
The integral may be written as:

$$E = \int_1^{\infty} \frac{1}{\sqrt{u^4 - u^2 - \epsilon(-u^4 + u)}} \tag{238}$$

where $\epsilon \equiv \frac{r_s}{r_m}$.
Expanding, we find the leading term is $\pi/2$. Whilst the $O(\epsilon)$ term is 2.

We can evaluate the elliptic integral seen in (238), Maple can do this for all values of ϵ, however this gives three ugly, non illuminating, equations. The relevant equation is $\epsilon \ll 1$, and substitute $v = \frac{1}{u}$.

```
> assume(epsilon < 2/3);
```

The elliptic integral becomes:

```
> int((1-v^2-epsilon*(1-v^3))^(-1/2), v=0..1);
```

$$\frac{1}{\sqrt{1-\epsilon}\,\epsilon}\Big(\big(3\epsilon - 1 + \sqrt{-3\epsilon^2 + 2\epsilon + 1}\big)\sqrt{2}\sqrt{\frac{\epsilon}{3\epsilon - 1 + \sqrt{-3\epsilon^2+2\epsilon+1}}}\sqrt{\frac{1-\epsilon+\sqrt{-3\epsilon^2+2\epsilon+1}}{-3\epsilon+1+\sqrt{-3\epsilon^2+2\epsilon+1}}}\sqrt{\frac{-1+\epsilon+\sqrt{-3\epsilon^2+2\epsilon+1}}{3\epsilon-1+\sqrt{-3\epsilon^2+2\epsilon+1}}}\,\mathrm{EllipticF}\Big(\sqrt{2}\sqrt{\frac{\epsilon}{3\epsilon-1+\sqrt{-3\epsilon^2+2\epsilon+1}}}, \sqrt{-\frac{3\epsilon-1+\sqrt{-3\epsilon^2+2\epsilon+1}}{-3\epsilon+1+\sqrt{-3\epsilon^2+2\epsilon+1}}}\Big)\Big) \tag{6}$$

This does not look quite so bad. The ugly square roots presumamby are related to the roots of the fourth-order polynomials in the integrand. Expanding around $\varepsilon = 0$ yields the series

```
> series(%,epsilon=0);
```

$$\frac{\pi}{2} + \epsilon + \left(\frac{15\pi}{32} - \frac{1}{2}\right)\epsilon^2 + \left(-\frac{15\pi}{32} + \frac{61}{24}\right)\epsilon^3 + \left(\frac{3465\pi}{2048} - \frac{65}{16}\right)\epsilon^4 + \mathrm{O}(\epsilon^5) \tag{7}$$

Figure 16: Maple algebra used to evaluate the elliptical integral