Light-Like Trajectories Around Neutral Gravitational Sources

Andrew Chang

April 12, 2021

SID: 3032009553

Start Date: December 1st, 2020

Date Completed: February 10, 2021

0 Abstract

Ever since its conception by Albert Einstein, General Relativity has been known to be a highly nonlinear theory. However, there are a class of important phenomena, for massless force carriers specifically, in which we can consider a linearized theory of gravity. In this exposition, we will focus on one of these phenomena, the deflection of a photon by a neutral (uncharged, stationary) gravitational source. Namely, we compute the angle of deflection of a photon with impact parameter b about a massive body with mass M: $\alpha = \frac{4MG}{bc^2}$.

Furthermore, we derive a closed-form expression (in terms of elliptic integrals) to the preceding result by considering light-like trajectories performing an orbit about a Schwarzschild Black Hole. We conclude this by performing an expansion to linear order in r_s/b, showing that this closed-form expression reduces to the initial answer of $\frac{4MG}{bc^2}$ in the large-b limit.

Note: Various times throughout this exposition, we will use natural units where $c = 1$ for notational simplicity. We implore the reader to work out the units themselves, or to refer to other textbooks or writings where they include c as a constant.

1 Introduction

1.1 What is Photon Deflection?

Simply stated, it is the deflection of a photon, the quanta of electromagnetic radiation, by a massive body.

For a more comprehensive answer: in General Relativity, the basic idea is that gravity is observable through a smooth deformation of spacetime. Photon deflection is one instance of phenomena in which this is observed. According to Fermat's Principle, light propagates in a trajectory that minimizes the time elapsed in traversing between two points.

For the mathematically minded reader, this statement comes as minimizing the following action [9]:

$$L = \int \mathcal{L} dx$$

Where $\mathcal{L} = (1 + y'(x))^{1/2}$ for $y = f(x)$. With the appropriate generalization for higher-dimensional spaces.

The Euler-Lagrange Equations, with appropriate boundary conditions, will give us that the arc length is minimized when $f(x) = mx + b$.

This mathematical statement that the trajectory that minimizes the distance/time, *naively*[1], describes something called a **geodesic**. As demonstrated, at any local point on Earth, this geodesic is a straight line between two points[2]. This is an excellent approximation on Earth (in a consistent refractive medium) as there is a uniform gravitational field[3], namely $g \approx 9.81$ ms^{-2}.

Later computations will show that, the straight line is only a zeroeth-order contribution to the true trajectory of light. In the presence of more exotic gravitational fields, first-order corrections to the trajectory of light become more apparent, and this is precisely what Albert Einstein predicted with the Theory of General Relativity.

With more context now, we can confidently define an answer to our initial question: **The deflection angle is the angle between a higher-order and zeroeth-order trajectory of the photon in the presence of a gravitational field.**

1.2 Goals

There are three goals we want to achieve in this exposition:

1. Summarize everything we need to compute our deflection angle in a theory of Gravitation. Our emphasis will be on linearization and perturbation theory.

2. Compute the angle of deflection of photons by massive bodies.

[1] This is indeed a very naive view of geodesics. Spheres are the simplest example of a manifold in which the geodesics are not necessarily the points of shortest distance. In the case of flat space, our straight line is, indeed, a geodesic, so we will not worry too much about this distinction.

[2] We can approximately describe Earth as a smooth manifold, which locally "looks like" the plane $\mathbb{R}^2$.

[3] the difference in the gravitational field at various points on the Earth is miniscule, and definitely not enough so that the deflection of light is noticeable by any instrument

3. Generalize (2) and compute the deflection angle of photons by a Schwarzschild Black Hole.

2 Physics Background: Black Holes and Einstein's Equations

We will explore the necessary background needed to model our relevant problem. This will be a brief discussion about the Schwarzschild Solution, Einstein's equations, and First-Order Perturbation Theory.

2.1 Schwarzschild Solution

Author's Note: We follow the derivation given in [3].This derivation is not original, although we may claim that our computations are our own.

Definition 2.1. Schwarzschild Metric

The Schwarzschild Metric is a spherically symmetric solution to Einstein's Equation given explicitly by:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right)dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{1}$$

Where

$$d\Omega^2 = d\theta^2 + \sin^2(\theta)d\phi^2$$

The short story to this can be given by:

Theorem 1. ***Birkhoff's Theorem***

The Schwarzschild Metric is the unique, most general spherically symmetric and asymptotically flat solution to Einstein's Equations in vacua.

The following, longer, story is significantly more insightful to us, and constitutes a "good enough" proof for us. In order to enforce positive definiteness, we give the most general form of the metric:

$$ds^2 = -e^{2\alpha(r)}dt^2 + e^{2\beta(r)}dr^2 + e^{2\gamma(r)}d\Omega^2$$

Enforcing spatial symmetry, we set $\gamma(r) = \ln(r)$ to obtain:

$$ds^2 = -e^{2\alpha(r)}dt^2 + e^{2\beta(r)}dr^2 + r^2 d\Omega^2$$

We set the following components of the metric according to the first fundamental form:

$$g_{tt} = -e^{2\alpha(r)} \qquad g_{rr} = e^{2\beta(r)} \qquad g_{\theta\theta} = r^2 \qquad g_{\phi\phi} = r^2\sin^2(\theta)$$

Using **Proposition 2** gives us 9 non-trivial Christoffel Symbols:

$$\Gamma^t_{tr} = \partial_r\alpha \qquad \Gamma^r_{tt} = e^{2(\alpha-\beta)}\partial_r\alpha \qquad \Gamma^r_{rr} = \partial_r\beta$$

$$\Gamma^{\theta}_{r\theta} = \frac{1}{r} \qquad \Gamma^{r}_{\theta\theta} = -re^{-2\beta} \qquad \Gamma^{\phi}_{r\phi} = \frac{1}{r}$$
$$\Gamma^{r}_{\phi\phi} = -re^{-2\beta}\sin^2(\theta) \qquad \Gamma^{\theta}_{\phi\phi} = -\sin(\theta)\cos(\theta) \qquad \Gamma^{\phi}_{\theta\phi} = \cot(\theta)$$

We can then compute the Ricci Tensor by using **Definition A.24**. Note that in the computation, the equality of the Ricci Tensor is an assignment, and not a true equality[4].

$$R_{\mu\nu} \equiv R^{\rho}_{\mu\rho\nu} = \partial_\rho \Gamma^{\rho}_{\mu\nu} - \partial_\nu \Gamma^{\rho}_{\rho\mu} + \Gamma^{\rho}_{\rho\lambda}\Gamma^{\lambda}_{\mu\nu} - \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\rho\nu}$$

It turns out that $\mu = \nu = \sigma$, as any nondiagonal terms of the Ricci Tensor will vanish. Computation of the Ricci Tensor yields:

$$R_{tt} = R^{r}_{trt} + R^{\theta}_{t\theta t} + R^{\phi}_{t\phi t}$$
$$R_{rr} = R^{t}_{rtr} + R^{\theta}_{r\theta r} + R^{\phi}_{r\phi r}$$
$$R_{\theta\theta} = R^{t}_{\theta t\theta} + R^{r}_{\theta r\theta} + R^{\phi}_{\theta\phi\theta}$$
$$R_{\phi\phi} = R^{t}_{\phi t\phi} + R^{r}_{\phi r\phi} + R^{\theta}_{\phi\theta\phi}$$

Clearly, any $R^{\mu}_{\mu\mu\mu} = 0$. Computing this results in:

$$R_{tt} = e^{2(\alpha-\beta)}\left[\partial_r^2\alpha + (\partial_r\alpha)^2 - \partial_r\alpha\partial_r\beta + \frac{2}{r}\partial_r\alpha\right]$$
$$R_{rr} = -\partial_r^2\alpha - (\partial_r\alpha)^2 + \partial_r\alpha\partial_r\beta + \frac{2}{r}\partial_r\beta$$
$$R_{\theta\theta} = e^{-2\beta}\left[r\left(\partial_r\beta - \partial_r\alpha\right)\right] + 1$$
$$R_{\phi\phi} = \sin^2(\theta)R_{\theta\theta}$$

Therefore, since we are solving for the metric in vacua, we may now simply solve $R_{\mu\nu} = 0$: As R_{tt} and R_{rr} independently vanish, we may set:

$$e^{2(\alpha-\beta)}R_{tt} + R_{rr} = 0$$

This gives us that $\alpha = -\beta + c$, with boundary conditions making the constant $c = 0$. Then, we also solve for $R_{\theta\theta} = 0$, which gives us:

$$\partial_r\left(re^{2\alpha}\right) = 1 \Longrightarrow e^{2\alpha} = 1 - \frac{R_s}{r}$$

We take it as a result from the weak-field approximation of a metric about a point mass, that:

$$g_{tt} = -\left(1 - \frac{2GM}{r}\right)$$

This is exactly the form of the metric right now, thus, $R_s = 2GM$. This is known as the **Schwarzschild Radius**, and it effectively defines the mass M. Our metric is therefore:

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

[4] The real Ricci Tensor, in terms of the Riemann Curvature tensor, is defined through a series of contractions.

2.2 Einstein's Equations

There are many formulations of Einstein's Equations, and we will lay them out here:

Einstein's Equation in Vacua

$$R_{\mu\nu} = 0 \tag{2}$$

Einstein's Equation (Ricci Tensor)[5]

$$R_{\mu\nu} = 8\pi G\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right) \qquad T = T^{\mu\nu}T_{\mu\nu} \tag{3}$$

Einstein's Equation (Einstein Tensor)

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \qquad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \tag{4}$$

Note that we define the Einstein Tensor purely for convenience.

2.3 Linearized Gravity

Note from the Author: This section follows closely from [4]. It is, by no means, original work of mine. However, I thought it useful to include because we will blindly refer to this in later computations.

The essence of **linearized gravity** is to decompose the metric into the following:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{5}$$

Where $\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$ is the Minkwoski Metric and $h_{\mu\nu}$ is a metric such that $|h_{\mu\nu}| \ll 1$.

To make more sense of this perturbed metric tensor, we give the following mathematical setting.

2.3.1 Perturbations and Physical Spacetime

Let M_b be the manifold on which the Minkowski Metric is defined[6]. We call this the **Background Space**. Likewise, let M_p be a manifold on which $h_{\mu\nu}$ is defined. We call this the **Physical Space**, for reasons that will be apparent after a brief walkthrough. We note that there exists a local diffeomorphism between M_b and M_p. Pictorially, we envision something like this:

$$\phi : M_b \longrightarrow M_p$$

[5] There is some discrepancy in notation here. What we defined as T here is actually known as T^2 in some literature (we are thinking of **General Relativity** by Robert M. Wald). We will stick with defining $T^{\mu\nu}T_{\mu\nu}$ as T, but please note the discrepancy.

[6] Clearly, this implies that M_b must be a submanifold of $\mathbb{R}^4$, but we do not concern ourselves with too much mathematical detail.

Note that this induces a pullback:

$$\phi^* : M_p \longrightarrow M_b$$

Let us look exclusively at the perturbation metric, $h_{\mu\nu}$.

$$h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$$

If we consider the local diffeomorphism between the background and physical space, we see that:

$$h_{\mu\nu} = (\phi^* g)_{\mu\nu} - \eta_{\mu\nu}$$

Hence the pullback allows us to analyze the perturbation in physical space, in our background space[7]. Moreover, this means that whenever we perform a contraction of indices (”raising” and ”lowering”), we will be allowed to use the Minkowski Metric, $\eta^{\mu\nu}$.

On our background space M_b, consider a vector field $\xi^\mu \in T_p M_b$. Without loss of generality, we may regard ξ^μ as the generator of a 1-parameter family of diffeomorphisms on M_b:

$$\psi_\varepsilon : M_b \longrightarrow M_b$$

We are, again, interested in the pullback of this one-parameter family:

$$h^\varepsilon_{\mu\nu} = (\phi \psi_\varepsilon g)^*_{\mu\nu} - \eta_{\mu\nu} \Longrightarrow h^\varepsilon_{\mu\nu} = (\psi^*_\varepsilon \phi^* g)_{\mu\nu} - \eta_{\mu\nu}$$

This implies that:

$$h^\varepsilon_{\mu\nu} = \psi^*_\varepsilon (\eta_{\mu\nu} + h_{\mu\nu}) - \eta_{\mu\nu} = (\psi^*_\varepsilon h)_{\mu\nu} + (\psi^*_\varepsilon \eta_{\mu\nu} - \eta_{\mu\nu})$$

Which further implies that:

$$h^\varepsilon_{\mu\nu} = (\psi^*_\varepsilon h)_{\mu\nu} + \varepsilon \left[\frac{\psi^*_\varepsilon \eta_{\mu\nu} - \eta_{\mu\nu}}{\varepsilon} \right]$$

Note that the term in square brackets above is the definition of a **Lie Derivative along the vector field** ξ. Hence, we write this as:

$$h^\varepsilon_{\mu\nu} = (\psi^*_\varepsilon h)_{\mu\nu} + \varepsilon \mathcal{L}_\xi \eta_{\mu\nu}$$

This expression highlights what we call a gauge transformation in linearized gravity. The reason why **physical space**, M_b, is the domain of our tensor field $h_{\mu\nu}$ is because an infinitesimal translation of $h_{\mu\nu}$ corresponds to a first order expansion in the Lie Derivative. According to the following expression:

$$h^\varepsilon_{\mu\nu} - (\psi^*_\varepsilon h)_{\mu\nu} = \varepsilon \mathcal{L}_\xi \eta_{\mu\nu}$$

[7]This is essentially the idea of a pullback. A pullback, without too much regard for rigor, is really a reparametrization, as it is inherently a composition of a smooth function with a local diffeomorphism.

If the perturbation, $h_{\mu\nu}$, is arbitrarily close to our family of perturbations $\forall \varepsilon > 0$, this means that $\mathcal{L}_\xi \eta_{\mu\nu} = 0$. This is known as **Killing's Equation**, and it identifies ξ^μ as a **Killing Vector**. A Killing Vector always corresponds to a constant of motion along a suitable geodesic (in fact, the constant of motion corresponding to a Killing Vector, ξ_α, is precisely $u^\alpha \xi_\alpha$, where u^α is a tangent vector along the geodesic), [8]. Therefore, M_p is the physical space as it is the domain of a gauge-invariant family of tensor fields.

2.3.2 Einstein's Equations in Coordinates

Author's Note: We follow the derivation in [4]. Again, the computation is useful to highlight the methods underlying our results.

Moving away from the abstract, coordinate-free formalism, let us introduce coordinates into our discussion through fixing an inertial frame. Our ultimate goal here is to show the gauge freedom in General Relativity and exhibit a few useful gauges, one of which our computations will take place in. With gravitation in mind, we identify the following components of $h_{\mu\nu}$:

$$h_{00} = -2\Phi \qquad h_{0i} = \omega_i \qquad h_{ij} = 2s_{ij} - 2\Psi\delta_{ij}$$

Here, s_{ij} denotes the **strain tensor**, and it is inherently the trace-free portion of a rotationally-invariant (0,2)-tensor, h_{ij}. $\Psi = \mathrm{Tr}(h_{ij}) = -\frac{1}{6}\delta^{ij}h_{ij}$, $\mathrm{Tr}(s_{ij}) = 0$. This gives us an explicit form for s_{ij}:

$$s_{ij} = \frac{1}{2}\left(h_{ij} - \frac{1}{3}\delta^{kl}h_{kl}\delta_{ij}\right)$$

We may now write out the full metric $g_{\mu\nu}$ with the first fundamental form:

$$ds^2 = -(1+2\Phi)dt^2 + \omega_i\left(dx^i dt + dx^i dt\right) + \left[(1-2\Psi)\delta_{ij} + 2s_{ij}\right]dx^i dx^j \tag{6}$$

From here, it is simple enough to compute the Christoffel Symbols, which will directly give us the Riemann Tensor, and consequently, either the Ricci Tensor or the Einstein Tensor, with which we may solve Einstein's Equations.

The non-trivial Christoffel Symbols are:

$$\Gamma^0_{00} = \partial_0\Phi \qquad \Gamma^i_{00} = \partial_i\Phi + \partial_0\omega_i \qquad \Gamma^0_{j0} = \partial_j\Phi \qquad \Gamma^i_{j0} = \partial_{[j\omega_i]} + \frac{1}{2}\partial_0 h_{ij}$$

$$\Gamma^0_{jk} = -\partial_{(j\omega_k)} + \frac{1}{2}\partial_0 h_{jk} \qquad \Gamma^i_{jk} = \partial_{(jh_k)i} - \frac{1}{2}\partial_i h_{jk}$$

Where $\partial_{[]}$ and $\partial_{()}$ denote the antisymmetrization, symmetrization of indices, respectively.

[8]The mathematical nuance here is that the vector field ξ^μ is not only a tangent vector, but it is actually the generator of the tangent space, T_pM_b. If our background space, M_b, is a **Lie Group**, then this implies that our Killing Vector is actually a generator of the **Lie Algebra**. In flat spacetime, $\mathbb{R}^4$, this is precisely the case (as Euclidean Space can be embedded into a known Lie Group, $GL_n(\mathbb{R})$).

In the interest of moving onto the more important aspects of our theory, we will skip the computation of the Riemann Tensor. The nontrivial physics was all in the Christoffel Symbols, which effecively gives us the Riemann, Ricci, Einstein Tensors. We will bluntly state the components of the Einstein Tensor:

$$G_{00} = 2\nabla^2\Psi + \partial_k\partial_l s^{kl} \qquad G_{0j} = -\frac{1}{2}\nabla^2\omega_j + \frac{1}{2}\partial_i\partial_k\omega^k + 2\partial_0\partial_j\Psi + \partial_0\partial_k s_j^k$$

$$G_{ij} = \left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)(\Psi - \Phi) + \delta_{ij}\partial_0\partial_k\omega^k - \partial_0\partial_{(i\omega_j)} + 2\delta_{ij}\partial_0^2\Psi - \Box s_{ij} + 2\partial_k\partial_{(is_j)^k} - \delta_{ij}\partial_k\partial_l s^{kl}$$

Now we give the explicit form of Einstein's Equations:

$$G_{00} = 8\pi G T_{00}$$

In our metric, this becomes:

$$\nabla^2\Psi = 4\pi G T_{00} - \frac{1}{2}\partial_k\partial_l s^{kl} \tag{7}$$

Then for

$$G_{0j} = 8\pi G T_{0j}$$

In our metric, this becomes:

$$\left(\delta_{jk}\nabla^2 - \partial_j\partial_k\right)\omega^k = -16\pi G T_{0j} + 4\partial_0\partial_j\Psi + 2\partial_0\partial_k s_j^k \tag{8}$$

Then for

$$G_{ij} = 8\pi G T_{ij}$$

In our metric, this becomes:

$$\left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)\Phi = 8\pi G T_{ij} + \left(\delta_{ij}\nabla^2 - \partial_i\partial_j - 2\delta_{ij}\partial_0^2\right)\Psi - \delta_{ij}\partial_0\partial_k\omega^k + \partial_0\partial_{(i\omega_j)} + \Box s_{ij} - 2\partial_k\partial_{(is_j)^k} - \delta_{ij}\partial_k\partial_l s^{kl} \tag{9}$$

Now, under the gauge transformations that our first-order perturbation metric underwent, our perturbed fields transform in the following manner:

$$\Phi \longmapsto \Phi + \partial_0\xi^0 \qquad \omega_i \longmapsto \omega_i + \partial_0\xi^i - \partial_i\xi^0$$

$$\Psi \longmapsto \Psi - \frac{1}{3}\partial_i\xi^i \qquad s_{ij} \longmapsto s_{ij} + \partial_{(i\xi_j)} - \frac{1}{3}\partial_k\xi^k\delta_{ij}$$

2.3.3 Gauge Choice

Here, we will choose the gauge that will aid us in the goal of this paper, the **Transverse Gauge**. The physical motivation behind this gauge is to regard the strain as spatially transverse, i.e. $\partial_i s^{ij} = 0$

We can accomplish this by choosing ξ^j to satisfy:

$$\nabla^2\xi^j + \frac{1}{3}\partial_j\partial_i\xi^i = -2\partial_i s^{ij}$$

Furthermore, we regard the vector fields as spatially transverse, i.e. $\partial_i \omega^i = 0$

We may accomplish this by choosing ξ^0 to satisfy:

$$\nabla^2 \xi^0 = \partial_i \omega^i + \partial_0 \partial_i \xi^i$$

Plugging in our choice of gauge, we obtain the following forms of Einstein's Equations for our perturbed metric:

$$G_{00} = 8\pi G T_{00}$$

Which is explicitly:

$$\nabla^2 \Psi = 4\pi G T_{00} \tag{10}$$

Then

$$G_{0j} = 8\pi G T_{0j}$$

Which is explicitly:

$$\delta_{jk} \nabla^2 \omega^k = -16\pi G T_{0j} + 4\partial_0 \partial_j \Psi \tag{11}$$

And finally

$$G_{ij} = 8\pi G T_{ij}$$

Which explicitly becomes:

$$\left(\delta_{ij} \nabla^2 - \partial_i \partial_j\right) (\Phi - \Psi) - \partial_0 \partial_{(i\omega_j)} + 2\delta_{ij} \partial_0^2 \Psi - \Box s_{ij} = 8\pi G T_{ij} \tag{12}$$

Note that, in conclusion, there are 10 degrees of freedom corresponding to 10 equations of motion, if we consider our equations coordinate-wise.

The work we did here will pay off as the Transverse Gauge for our linearized metric will be exactly what we need for the computation of the deflection of a photon.

3 Photon Deflection

The central problem we seek to address is the seemingly paradoxical phenomenon of the deflection of the photon, a massless particle, on a macroscale. We will look at this phenomenon through (i) dimensional analysis (ii) approximately with first-order perturbation theory (iii) exactly through solving for the Schwarzschild Orbit.

3.1 Dimensional Analysis

In the Newtonian regime, dimensional analysis is able to give us a quick, back-of-the-envelope way to compute dimensionless quantities. It turns out, that dimensional analysis gets us *most* of our true answer.

3.1.1 Dimensional Analysis: The Theory

Dimensional analysis concerns the relationship between physical quantities and a set of "basis" units. For example, a Newton is a unit of force that can be decomposed into a time unit, a mass unit, and a distance unit. In fact, **the SI unit system** is a **vector space of physical units spanned by** a *time unit* (seconds), *mass unit* (kilograms), and *unit of length* (meters)[9].

Definition 3.1. Vector Space of Physical Units

Let $\beta = \{v_1, \dots, v_m\}$ be a basis of physical units, we call it the **Unit System**. Then β spans V, which we call the **vector space of physical units** [1], over $\mathbb{Q}$,[10] whose elements are called **physical units**[11]

Definition 3.2. Dimension Matrix

For a function f that depends on variables with physical units $\{q_1, \dots, q_m\}$, we define the **Dimension Matrix**, A, as the following $m \times n$ matrix:

$$A = \begin{pmatrix} \vdots & & \vdots \\ \vec{q}_1 & \cdots & \vec{q}_n \\ \vdots & & \vdots \end{pmatrix} \tag{13}$$

Where n is the **number of variables**, and $m = \dim(V)$ is the **number of basis units**.

For a variable with physical units of the following form:

$$q_i = v_1^{a_{1i}} \times \cdots \times v_m^{a_{mi}} \tag{14}$$

We define the **ith physical unit**, $\vec{q}_i$ as:

$$\vec{q}_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \tag{15}$$

Furthemore, we define $\ker(A)$ as the **space of dimensionless parameters** corresponding to f. Hence, $\text{Nullity}(A) = n - m$.

[9] A reference that everyone in the sciences should read is [1]. It formalizes the idea of physical units and unit systems as a linear algebra problem, which is exactly what we will do here.

[10] The requirement that the space of physical units have rational coefficients is a choice more than anything. We want to avoid basis units with transcendental powers, like $\mathrm{m}^{\ln(\pi)}$, which are horribly defined in the physical sense.

[11] Don't confuse *physical units* with a **unit** in ring theory, the latter are elements of a ring with an inverse. "Physical unit" is just terminology that english-speaking scientists have adopted to describe the physical quantity.

Theorem 2. ***Buckingham Pi Theorem***

*Let π_j denote the jth dimensionless parameter. Without loss of generality, take $j = 1$. Given variables with physical units q_i, f is a function of physical units of the form in **Definition 3.2**, then the relation:*

$$f(q_1, \ldots, q_m) = 0$$

Can be turned into the relation:

$$\pi_1 = F(\pi_2, \ldots, \pi_k)$$

Where $k = n - m$.

Furthermore, A is the matrix representation of the map f, and:

$$\pi_1 \in \ker(A) \qquad \text{Nullity}(A) = k$$

The precise statement of this theorem and the definitions does not seem to make anything easier, but we will use it directly in our computation below.

3.1.2 Method 1: Application of the Theory

Now let us consider the 4 most fundamental variables in our current situation: G, M, b, c. And let our unit system be $\{[t], [m], [l]\}$, corresponding to time, mass, length, respectively. And now consider the following relation:

$$f(G, M, b, c) = 0$$

Then by the Buckingham Pi Theorem, we can represent a dimensionless quantity, π, as an element of the kernel of f:

$$\pi = G^{c_1} M^{c_2} b^{c_3} c^{c_4} \tag{16}$$

Now let us write our variables in terms of the basis units:

$$G = [t]^{-2}[m]^{-1}[l]^3 \qquad M = [m] \qquad b = [l] \qquad c = [l][t]^{-1}$$

Therefore, our dimension matrix, A is a matrix representation of f, consisting of all the powers of each basis unit for each variable:

$$\begin{pmatrix} -2 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 1 \end{pmatrix}$$

Now, let us find a solution for the following equation:

$$\begin{pmatrix} -2 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \vec{0}$$

Our system is underdetermined (more variables than equations), hence we will have a free parameter, c_1. Let us make a choice of c_1, and choose $c_1 = 1$.[12] When we choose $c_1 = 1$, it follows that:

$$c_2 = 1 \qquad c_3 = -1 \qquad c_4 = -2$$

So that:

$$\vec{\pi} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -2 \end{pmatrix} \Longrightarrow \pi = GMb^{-1}c^{-2} = \frac{GM}{bc^2}$$

Remark: We know that $\vec{\pi}$ is not unique as if we scale any vector in a linear subspace by a constant, it will still be in the subspace. Hence, the most general dimensionless parameter corresponds to something of the form $K\vec{\pi}$, which corresponds to the quantity:

$$\left(\frac{GM}{bc^2}\right)^K \tag{17}$$

The simpliest choice of constant is $K = 1$, hence, we conclude that our dimensionless parameter is:

$$\frac{GM}{bc^2}$$

We know that the deflection angle, α, is a **dimensionless quantity**, and given that the Nullity$(A) = 1$, this indicates that $\alpha = GMb^{-1}c^{-2}$ is the unique dimensionless quantity that is derivable through dimensional analysis. Hence, we conclude, from dimensional analysis, that our deflection angle is:

$$\alpha = \frac{GM}{bc^2} \tag{18}$$

If we convert to natural units, assuming the appropriate basis change ($c = 1$), we obtain:

$$\alpha = \frac{GM}{b}$$

3.1.3 Method 2: Educated Guess

Finding the deflection angle becomes a lot more simple if we recognize that it is a dimensionless quantity. Note that we use units where $\hbar = c = 1$ for the sake of simplicity. It becomes a game to construct the simplest dimensionally trivial quantity with the fundamental constants.

Let $[m]$ denote the fundamental mass unit. In our natural units[13], the mass

[12] We find this to be appropriate as G appears as a single power, G in the expression for the gravitational potential obtained from Newton's Law.

[13] Recall that all quantities in natural units are expressible as eV and some fundamental constant.

$$\text{Length} = \hbar c\text{eV}^{-1}$$
$$\text{Mass} = \text{eV}c^{-2}$$
$$\text{Time} = \hbar\text{eV}^{-1}$$

Once we set $\hbar = c = 1$, we obtain the necessary conversions to obtain our results.

dimensions of the relevant fundamental constants are the following:

$$G = [m]^{-2} \qquad M = [m] \qquad r = [m]^{-1}$$

Where r denotes the magnitude of some vector displacement. It is very apparent now, that the most fundamental dimensionless quantity with the fundamental constants that appear in a purely gravitational source is:

$$\frac{GM}{r}$$

If we let b denote the distance from the gravitational source (of mass M) to our photon[14]. Technically, we have that the deflection angle is something of the form:

$$\alpha \propto \frac{GM}{b}$$

The simplest constant of proportionality here is 1. Hence, we have that, through educated guesses alone:

$$\alpha = \frac{GM}{b}$$

3.2 First-Order Perturbations about a "Classical" Source

In the previous section, we derived an explicit dimensionless parameter from three (technically four) fundamental constants and parameters in our scenario. The rest of this section is dedicated to finding the true proportionality constant.

Note: By a "classical" source, we really mean a mass M such that the volume of the mass M constitutes a physical radius larger than its Schwarzschild Radius, i.e. something that's not a black hole. For example, any planetary body will have an actual radius larger than its Schwarzschild Radius.[15]

3.2.1 Modeling Spacetime

We model the region of spacetime in which we are considering the light propagating as **dust**. Dust is defined as a perfect fluid with no pressure [4].

In the rest frame of the dust, we see that the energy-momentum tensor is:

$$T_{\mu\nu} = \rho U_\mu U_\nu = \mathrm{diag}(\rho, 0, 0, 0) \tag{19}$$

The background manifold, M_b is just Minkowski Space, so we may boost to another frame as our Energy-Momentum Tensor is Lorentz-Invariant.

[14] Technically, looking at this as a "hard sphere" scattering problem, b is really the impact parameter of our photon with a really, really large mass. This is how we will treat it in our semi-classical computation in the next section.

[15] For our purposes, a black hole is a gravitational source such that there exists a non-zero radius at which null trajectories cannot escape to spatial infinity. This expresses the idea that all black holes have an event horizon, the threshold at which "even light cannot escape".

We now work in the transverse gauge. Since only our 00 component is nontrivial, we obtain a simple form of Einstein's Equations.

$$\nabla^2 \Psi = 4\pi G\rho \tag{20}$$

$$\nabla^2 \omega_j = 0 \tag{21}$$

$$\left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)(\Phi - \Psi) - \nabla^2 s_{ij} = 0 \tag{22}$$

For nonsingular solutions, we enforce that $\omega_i = 0$, so that when we take the trace of the third of Einstein's Equations, we obtain that:

$$\nabla^2 (\Phi - \Psi) = 0$$

Therefore, $\Phi = \Psi$ up to a constant. This implies that:

$$\nabla^2 s_{ij} = 0$$

And for a traceless s_{ij}, the only possible solution is when $s_{ij} = 0$. Therefore, we have that a well-behaved solution of Einstein's Equations is given when $\omega_i = 0$ and $s_{ij} = 0$. Therefore, we have that our perturbed metric is given as:

$$ds^2 = -(1 + 2\Phi)\, dt^2 + (1 - 2\Phi)\left(dx^2 + dy^2 + dz^2\right) \tag{23}$$

It is important to note that our perturbing metric, $h_{\mu\nu}$ takes on the following form:

$$h_{\mu\nu} = -2\Phi I_{4\times 4} \tag{24}$$

3.2.2 Perturbations and Geodesics

We decompose the geodesic for our photon into:

$$x^\mu(\lambda) = x^{(0)\mu}(\lambda) + x^{(1)\mu}(\lambda) \tag{25}$$

Then we define the background and physical wave vectors, respectively:

$$k^\mu \equiv \frac{dx^{(0)\mu}}{d\lambda} \tag{26}$$

$$l^\mu = \frac{dx^{(1)\mu}}{d\lambda} \tag{27}$$

Because our photon obviously follows a null geodesic, we set:

$$ds^2 = 0 \Longrightarrow g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0 \tag{28}$$

Then, noting our first-order expansion of the geodesic for our photon and the first-order perturbation of the Minkowski metric, we plug it in and obtain:

$$(\eta_{\mu\nu} + h_{\mu\nu})\frac{d}{d\lambda}\left(x^{(0)\mu} + x^{(1)\mu}\right)\frac{d}{d\lambda}\left(x^{(0)\nu} + x^{(1)\nu}\right) = 0$$

Work this out further to obtain that:

$$(\eta_{\mu\nu}+h_{\mu\nu})\left[\frac{dx^{(0)\mu}}{d\lambda}\frac{dx^{(0)\nu}}{d\lambda}+\frac{dx^{(0)\mu}}{d\lambda}\frac{dx^{(1)\nu}}{d\lambda}+\frac{dx^{(0)\nu}}{d\lambda}\frac{dx^{(1)\mu}}{d\lambda}+\frac{dx^{(1)\mu}}{d\lambda}\frac{dx^{(1)\nu}}{d\lambda}\right]=0$$

At the 0th order in this expansion, we see that:

$$\eta_{\mu\nu}k^{\mu}k^{\nu}=0 \Longrightarrow k^2 \equiv -\left(k^0\right)^2+\vec{k}^2=0$$

Denote $\vec{k}^2=k^2$.

At first order, we obtain:

$$\eta_{\mu\nu}k^{\mu}l^{\nu}+\eta_{\mu\nu}l^{\mu}k^{\nu}+h_{\mu\nu}k^{\mu}k^{\nu}=0$$

By symmetry of the metric, we obtain:

$$2\eta_{\mu\nu}k^{\mu}l^{\nu}+h_{\mu\nu}k^{\mu}k^{\nu}=0$$

Which becomes:

$$\vec{k}\cdot\vec{l}=k^0l^0+2\Phi k^2$$

We now compute the Christoffel Symbols. Using the formula in one of our propositions above, we see that:

$$\Gamma^0_{0i}=\Gamma^i_{00}=\partial_i\Phi \qquad \Gamma^i_{jk}=\delta_{jk}\partial_i\Phi-\delta_{ik}\partial_j-\delta_{ij}\partial_k\Phi$$

Then, we may expand the geodesic equation

$$\frac{d^2x^{\mu}}{d\lambda^2}+\Gamma^{\mu}_{\rho\sigma}\frac{dx^{\rho}}{d\lambda}\frac{dx^{\sigma}}{d\lambda}=0$$

by plugging in our perturbed geodesic

$$\frac{d^2x^{(0)\mu}}{d\lambda^2}+\frac{d^2x^{(1)\mu}}{d\lambda^2}+\Gamma^{\mu}_{\rho\sigma}\left(\frac{dx^{(0)\rho}}{d\lambda}\frac{dx^{(0)\sigma}}{d\lambda}+2\frac{dx^{(0)\rho}}{d\lambda}\frac{dx^{(1)\sigma}}{d\lambda}+\frac{dx^{(1)\rho}}{d\lambda}\frac{dx^{(1)\sigma}}{d\lambda}\right)=0$$

What we recognize is that our christoffel symbols $\Gamma^{\mu}_{\rho\sigma}$ are small quantities on the order of $|h_{\mu\nu}|$,[16] therefore, the two latter terms resolve as follows:

$$\Gamma^{\mu}_{\rho\sigma}\frac{dx^{(0)\rho}}{d\lambda}\frac{dx^{(1)\sigma}}{d\lambda}\sim O\left(|h_{\mu\nu}|^2\right) \qquad \Gamma^{\mu}_{\rho\sigma}\frac{dx^{(1)\rho}}{d\lambda}\frac{dx^{(1)\sigma}}{d\lambda}\sim O\left(|h_{\mu\nu}|^3\right)$$

The terms $\leq O\left(|h_{\mu\nu}|\right)$ are:

$$\frac{d^2x^{(0)\mu}}{d\lambda^2}+\frac{d^2x^{(1)\mu}}{d\lambda^2}+\Gamma^{\mu}_{\rho\sigma}\frac{dx^{(0)\rho}}{d\lambda}\frac{dx^{(0)\sigma}}{d\lambda}=0 \tag{29}$$

We may read it off order-by-order:

[16] This is not obvious, but notice the form of our perturbing metric $h_{\mu\nu}$, it is $2\Phi I_{4\times4}$. Because $|h_{\mu\nu}|\ll 1$ necessarily, it follows that $\Phi \ll 1$, hence, any Φ or $\partial_i\Phi$ is a small quantity.

Zeroeth Order:
$$\frac{d^2x^{(0)\mu}}{d\lambda^2} = 0 \tag{30}$$
This is the most trivial solution, it is the Newtonian result that light travels in a straight line.

First Order:
$$\frac{d^2x^{(1)\mu}}{d\lambda^2} + \Gamma^{\mu}_{\rho\sigma}\frac{dx^{(0)\rho}}{d\lambda}\frac{dx^{(0)\sigma}}{d\lambda} = 0 \tag{31}$$
This easily resolves to:
$$\frac{dl^{\mu}}{d\lambda} = -\Gamma^{\mu}_{\rho\sigma}k^{\rho}k^{\sigma}$$
Plug in $\mu = 0$ and $\mu = i$ separately.
$$\frac{dl^0}{d\lambda} = -\left(\Gamma^0_{0i}k^0k^i + \Gamma^0_{i0}k^ik^0\right) = -2\Gamma^0_{0i} = -2k^0\left(\vec{k}\cdot\nabla\Phi\right)$$
But we know that, because the photon follows a null geodesic, that $k^0 = k$, hence:
$$\frac{dl^0}{d\lambda} = -2k\left(\vec{k}\cdot\nabla\Phi\right)$$
Likewise,
$$\frac{dl^i}{d\lambda} = -\left(\Gamma^i_{00}k^0k^0 + \Gamma^i_{jk}k^jk^k\right)$$
Thus, this implies that
$$\frac{d\vec{l}}{d\lambda} = -2k^2\nabla\Phi + 2\left(\vec{k}\cdot\nabla\Phi\right)\vec{k}$$
Remark: Note that we may define the transverse gradient:
$$\nabla_{\perp} = \nabla\Phi - \frac{\vec{k}}{k^2}\left(\vec{k}\cdot\nabla\Phi\right)$$
It is aptly named so as the following observation is true:
$$l^0 = \int\frac{dl^0}{d\lambda}d\lambda = -2k\int\left(\vec{k}\cdot\nabla\Phi\right)d\lambda = -2k\int\frac{d\vec{x}^{(0)}}{d\lambda}\cdot\nabla\Phi d\lambda = -2k\int\nabla\Phi\cdot d\vec{x}^{(0)} = -2k\Phi$$
Noting that $\vec{l}\cdot\vec{k} = kl^0 + 2k^2\Phi$, we see that:
$$\vec{k}\cdot\vec{l} = 0$$
Therefore, we see that $\vec{l}$ and $\vec{k}$ are orthogonal up to first order, and hence, this gradient is the sum of the original gradient and the gradient along the path perpendicular to the original path.

Our spatial equation of motion becomes:
$$\frac{d\vec{l}}{d\lambda} = -2k^2\nabla_{\perp}\Phi \tag{32}$$

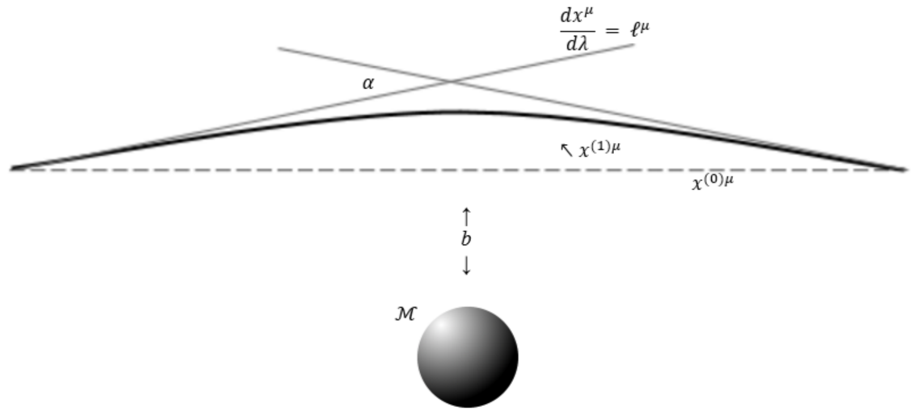

Figure 1: A figure of a massive body M perturbing the trajectory of the photon path.

3.2.3 The Deflection Angle

For the rest of this section, we refer to **Figure 1**[17]. Note that the angle α is effectively the ratio of the change in the perturbed wave vector and the background geodesic.

$$\alpha = -\frac{\Delta\vec{l}}{k} \tag{33}$$

We can simplify this by plugging in:

$$\Delta\vec{l} = \int \frac{d\vec{l}}{d\lambda} d\lambda = -2k^2 \int \nabla_\perp \Phi d\lambda$$

Make a suitable change of variables to make the integration measure in terms of a spatial variable:

$$x = k\lambda \Longrightarrow dx = kd\lambda$$

This will give us the following expression for the first-order deflection angle:

$$\alpha = 2\int \nabla_\perp \Phi dx \tag{34}$$

For a Newtonian body (such as a sun, planet, whatever it is), we can use the following Newtonian potential:

$$\Phi = -\frac{GM}{\left|\vec{r} - \vec{b}\right|} \tag{35}$$

[17]My wonderful partner, Kelly Lam, graciously drew this for me, as I am not the most competent artist.

Without loss of generality, we can pick an inertial frame in which the system is constrained to 2-dimensional plane. Hence, $\vec{r} = x\hat{x}$ and $\vec{b} = b\hat{y}$, so that $\left|\vec{r} = \vec{b}\right|^{1/2} = \left(x^2 + b^2\right)^{1/2}$. So our potential in this simplified model is:

$$\Phi = -\frac{GM}{\left(x^2 + b^2\right)^{1/2}}$$

If we compute $\nabla_\perp \Phi$, we see that:

$$\nabla \Phi = \frac{GM}{\left(x^2 + b^2\right)^{3/2}} \vec{b}$$

Where we take the gradient with respect to b. Then, we see that:

$$\nabla_\perp \Phi = \nabla \Phi - \frac{\vec{k}}{k^2} \left(\vec{k} \cdot \nabla \Phi\right) = \nabla \Phi$$

Since $\vec{b} \cdot \vec{k} = 0$.

Integrate to infinity,

$$\alpha = 2GMb \int_{-\infty}^{\infty} \frac{1}{(x^2 + b^2)^{3/2}} dx$$

The integration gives us a factor of $2/b^2$. Hence, we obtain:

$$\alpha = \frac{4GM}{b} \tag{36}$$

3.3 Closed-Form of the Schwarzschild Orbit

In this section, we will analyze the deflection of a photon, considering it as a null trajectory following a Schwarzschild Orbit.

3.3.1 Physical Intuition

We consider a stationary, spherically symmetric black hole of mass M. Where $r_s = 2GM$, we have the **event horizon**[18]. Informally, we know the Event Horizon to be the *point of no return*. While the validity of this identification may be disputed, it is accurate enough of a characterization for our purposes.

We first identify that the effective potential of the Schwarzschild Orbit is the following:

$$V(r) = \left(\frac{r^2}{1 - r_s}\right)^{1/2}$$

[18]Note that the Schwarzschild Radius and the Event Horizon only coincide for a *Schwarzschild* Black Hole. The Kerr Metric, for example, may have multiple event horizons. It is not a formal reference, but try the wikipedia page for the Kerr Metric for more information on this. This is beyond the scope of our paper.

We can square this potential:

$$V^2(r) = \frac{r^2}{1 - r_s}$$

Then we can minimize this expression:

$$\frac{dV^2}{dr} = -\frac{r_s}{\left(1 - \frac{r_s}{r}\right)^2} + \frac{2r}{1 - \frac{r_s}{r}} = 0$$

Then the critical value of this orbit is:

$$r_{crit} = \frac{3}{2} r_s$$

At $R = \frac{3}{2} r_S$, we obtain that:

$$b^2 = \frac{27}{2} r_s^2 \Longrightarrow b_{crit} = 3\sqrt{3}GM$$

At this impact parameter, we have two cases:

1. At $b < b_{crit}$, the photon will be attracted to the Schwarzschild radius r_s, and the null trajectory will never fly off to spatial infinity (i.e. the photon "falls" into the black hole). To see why this occurs, consider that the impact parameter, b_{crit}, corresponds to the distance of closest approach, $R = \frac{3}{2} r_s$, (given by **Equation 49**), and is called the **Photon Sphere**[19]. It is the radial distance from the center of a massive body such that the null geodesics are forced to traverse in an orbit. Clearly, if we perturb this radial distance to be closer to the gravitational source, the orbit will degenerate and the photon will be attracted to the **Event Horizon**.

2. For some choice of $\varepsilon > 0$, at $b > b_{crit} + \varepsilon$, the photon is deflected by an angle α and flies off to spatial infinity. To see why, let us take the distance of closest approach, R,[20] and consider the impact parameter:

   $$b = \sqrt{\frac{R^3}{R - r_s}}$$

   And expand it in a Taylor Series:

   $$b \approx R + \frac{r_s}{2}$$

   In the instance that our impact parameter $b \gg M$, we obtain that $r_s \longmapsto 0$ and:

   $$b \approx R$$

   This means that the impact parameter and the distance of closest approach coincide the limit where our impact parameter is sufficiently large.

   We further see the following in this case:

[19] See **Remark 2.1** and **Remark 2.2** to see the significance of this in our computation.

[20] To see how it is defined, see Equation 49 in the later section.

(a) Take $\varepsilon \longmapsto \infty$, $b \gg b_{crit}$, then the effects of the black hole on the trajectory of the photon is miniscule (hence it is appropriate for linear-order perturbations, which assume that the metric perturbation in a gravitational source with a Minkowski background is miniscule), but deflects, at first-order, by an angle $\alpha = \frac{4GM}{b}$.

(b) Take $0 < \varepsilon < \infty$, the the photon will actually orbit the black hole for multiple periods, and fly off to spatial infinity[21].

3.3.2 Exact Form for the Deflection Angle of a Photon

Note: The result obtained in this section will be exactly what is reported in [2]. Our intention is to go more in-depth into the computation in our convention[22]. However, the calculation to obtain these results were done entirely originally.

Recall the Schwarzschild Metric:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2$$

Where $d\Omega^2 = d\theta^2 + \sin^2(\theta) d\phi^2$ is the metric for S^2.

We want to compute the deflection angle of a photon that is in **orbit** around a massive body (of mass M). The keyword is *orbit*. In orbital motion, we recognize that the particle is constrained to a plane. Namely, our ds^2 is symmetric about the plane $\theta = \pi/2$. Therefore, we may consider the motion of the particle only on this plane, which is a **totally geodesic submanifold** of spacetime[23].

Let λ be an affine parameter. Then we see that our metric, in terms of the affine parameter is:

$$d\lambda^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\phi^2$$

We will now define a scalar quantity, T, assuming a space-like trajectory, as follows:

$$2T = 1 = -\left(1 - \frac{2GM}{r}\right)\left(\frac{dt}{d\lambda}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2 + r^2\left(\frac{d\phi}{d\lambda}\right)^2 \quad (37)$$

Furthermore, we will use the shorthand notation, $\dot{x}^\mu = \frac{dx^\mu}{d\lambda}$. Hence, T becomes:

$$2T = 1 = -\left(1 - \frac{2GM}{r}\right)\left(\dot{t}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\dot{r}\right)^2 + r^2\left(\dot{\phi}\right)^2$$

We will now use variational methods to compute the equations of motion. Equivalently, direct computation of Christoffel Symbols or even Hamiltonian Methods

[21] We do not have a derivation of this fact, it is given in the reference [6], on pages 189-191.
[22] Said authors use geometerized units, which we will discuss in the Appendix, **Section C**.
[23] This means that any geodesic on the submanifold is also a geodesic on the larger manifold.

will work as well. We define:

$$L = \int \sqrt{g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}d\lambda = \int \sqrt{2T}d\lambda \tag{38}$$

Then we will minimize this action L by taking the functional derivative with respect to T and setting it equal to 0:

$$\delta L \equiv \frac{\delta L}{\delta T} = \int \frac{\delta T}{\sqrt{2T}}d\lambda = \int \delta T d\lambda = \delta \int T\lambda = 0$$

Clearly, minimizing δL is equivalent to finding the **Euler-Lagrange Equations** for our scalar T:

$$\frac{d}{d\lambda}\left(\frac{\partial T}{\partial \dot{x}^\mu}\right) = \frac{\partial T}{\partial x^\mu} \tag{39}$$

Clearly, since there is only explicit dependence on r, our three Euler-Lagrange Equations are:

$$\frac{d}{d\lambda}\left(\frac{\partial T}{\partial \dot{t}}\right) = 0 \tag{40}$$

$$\frac{d}{d\lambda}\left(\frac{\partial T}{\partial \dot{\phi}}\right) = 0 \tag{41}$$

$$\frac{d}{d\lambda}\left(\frac{\partial T}{\partial \dot{r}}\right) = \frac{\partial T}{\partial r} \tag{42}$$

The former two equation of motions are trivial:

$$\frac{\partial T}{\partial \dot{t}} = \text{some constant} \qquad \frac{\partial T}{\partial \dot{t}} = \text{another constant}$$

We may now set[24]:

$$\frac{\partial T}{\partial \dot{t}} = \frac{a}{b} \qquad \frac{\partial T}{\partial \dot{\phi}} = a$$

Thus, our equation of motions for t and ϕ become:

$$\dot{t} = ab^{-1}\left(1 - \frac{2GM}{r}\right)^{-1} \tag{43}$$

$$\dot{\phi} = ar^{-2} \tag{44}$$

We have three cases to consider here:

Note: The notion of space-like, time-like, and null geodesics are highly dependent upon our choice of metric signature. Our conventions here reflect the use of the $(-, +, +, +)$ signature.

[24]Our choice of these constants is what we call *proof by hindsight*. This will give us the more direct form of our answer that is consistent with the dimensional analysis we did in the earlier sections.

1. **Time-Like Geodesic**: For a time-like geodesic, we may consider the case where $2T = -1$.

$$2T = -\left(1 - \frac{2GM}{r}\right)\dot{t}^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\phi}^2 = -1$$

Rearrange the $\dot{\phi}$:

$$-\left(1 - \frac{2GM}{r}\right)\left(\frac{dt}{d\phi}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\phi}\right)^2 + r^2 = -\dot{\phi}^{-2}$$

Plugging in the explicit expressions for $\dot{\phi}$ and $\dot{t}$, we have that:

$$\frac{dt}{d\phi} = \frac{r^2}{b}\left(1 - \frac{2GM}{r}\right)^{-1} \qquad \dot{\phi}^2 = a^{-2}r^4$$

Then, after plugging these in and rearranging the expression, we obtain:

$$\left(\frac{dr}{d\phi}\right)^2 = \frac{r^4}{b^2} - \left(1 - \frac{2GM}{r}\right)\left(\frac{r^4}{a^2} + r^2\right)$$

Then, casting it into integral form, we obtain:

$$\phi = \int \frac{dr}{r^2\sqrt{\frac{1}{b^2} - \left(1 - \frac{2GM}{r}\right)\left(1 + \left(\frac{r}{a}\right)^2\right)r^{-2}}} \tag{45}$$

2. **Space-Like Geodesic**: For a space-like geodesic, we may consider the case where $2T = 1$.

The calculation is almost identical to the previous case, except there is a negative sign in one of the factors.

$$\left(\frac{dr}{d\phi}\right)^2 = \frac{r^4}{b^2} + \left(1 - \frac{2GM}{r}\right)\left(\frac{r^4}{a^2} - r^2\right)$$

This results in the following integral:

$$\phi = \int \frac{dr}{r^2\sqrt{\frac{1}{b^2} - \left(1 - \frac{2GM}{r}\right)\left(1 - \left(\frac{r}{a}\right)^2\right)r^{-2}}} \tag{46}$$

3. **Light-Like (Null) Geodesic**: Note that, with the assumption that we are solving for the deflection of a photon, we will now consider the case where either (a) *the geodesic is light-like* (b) *the orbiting particle is assumed to be massless.* We will find that both formulations are equivalent.

(a) Set $T = 0$ and $a = 1$.

$$-\left(1 - \frac{2GM}{r}\right)\left(\dot{t}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\dot{r}\right)^2 + r^2\left(\dot{\phi}\right)^2 = 0$$

We may use the Chain Rule to see that $\dot{t}/\dot{\phi} = \frac{dt}{d\phi}$ and $\dot{r}/\dot{\phi} = \frac{dr}{d\phi}$. Making the rearrangements necessary, we obtain:

$$-\left(1 - \frac{2GM}{r}\right)\left(\frac{dt}{d\phi}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\phi}\right)^2 + r^2 = 0$$

Plugging in $\dot{t}/\dot{\phi}$ and simplifying, we obtain:

$$\left(\frac{dr}{d\phi}\right)^2 = r^4\left[\frac{1}{b^2} - \left(1 - \frac{2GM}{r}\right)r^{-2}\right]$$

Further simplifying this results in:

$$\frac{d\phi}{dr} = \frac{1}{r^2\sqrt{b^{-2} - \left(1 - \frac{2GM}{r}\right)r^{-2}}}$$

Integrate this, and we obtain:

$$\phi(r) = \int \frac{dr}{r^2\sqrt{b^{-2} - \left(1 - \frac{2GM}{r}\right)r^{-2}}} \tag{47}$$

(b) Take either $2T = \pm 1$ (so when the trajectory is space-like or time-like), and send $a \longrightarrow \infty$ (so when the particle is massless). This will give us an identical integral:

$$\phi(r) = \int \frac{dr}{r^2\sqrt{b^{-2} - \left(1 - \frac{2GM}{r}\right)r^{-2}}}$$

We are interested in **Case 3** since our **orbiting particle is a photon**. However, for an exact solution, we are most interested in the time-like case[25], as we can always take the appropriate limit to reduce to the case of our photon orbit.

We will use the coordinate transformation:

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

Then, through the use of the chain rule, we obtain that:

$$\left(\frac{dr}{d\phi}\right)^2 = \left(\frac{du}{d\phi}\right)^2\frac{1}{u^4} - \frac{1}{u^4b^2} - (1 - 2GMu)\left(\frac{1}{u^2} + \frac{1}{u^4a^2}\right)$$

[25]You could easily use space-like too. It's not a huge deal.

Simplifying gives us:

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{b^2} - (1 - 2GMu)\left(u^2 + \frac{1}{a^2}\right)$$

Let us reduce to the case of a photon, which gives us:

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{b^2} - (1 - 2GMu)\, u^2$$

We can replace the impact parameter b with a more fundamental quantity, **the distance of closest approach**:

$$\frac{dr}{d\lambda} = \frac{dr}{d\phi}\dot{\phi} = 0 \tag{48}$$

Solving this results in:

$$b^2 = \frac{r^3}{r - 2GM}$$

Define any r that satisfies this relation as R, then we have:

$$b^2 = \frac{R^3}{R - 2GM} \tag{49}$$

We will substitute this into the expression for $(du/d\phi)^2$, then we obtain:

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{R^3}\left(2GMR^3u^3 - R^3u^2 + R - 2GM\right) \equiv P(u) \tag{50}$$

This gives us:

$$\phi = \int \frac{R^{3/2}du}{\left(2GMR^3u^3 - R^3u^2 + R - 2GM\right)^{1/2}}$$

Our original integration bounds were given by:

$$\int_R^\infty$$

However, with our coordinate transformation, we obtain:

$$\phi = \int_0^{1/R} \frac{R^{3/2}du}{\left(2GMR^3u^3 - R^3u^2 + R - 2GM\right)^{1/2}} \tag{51}$$

We must factor the polynomial $P(u)$:

$$P(u) = 2GMR^3u^3 - R^3u^2 + R - 2GM \tag{52}$$

Notice that $u = 1/R$ is a root. Therefore, we obtain a factorization:

$$P(u) = \left(u - \frac{1}{R}\right) S(u) \tag{53}$$

Where we can determine $S(u)$ by long division (or just plug into software[26]):

$$S(u) = 2GMR^3u^2 + [2GM - R]\,[Ru + 1]\,R$$

Therefore,

$$P(u) = \left(u - \frac{1}{R}\right)\left(2GMR^3u^2 + [2GM - R]\,[Ru + 1]\,R\right)$$

Completing the square on $S(u)$ will yield the following:

$$S(u) = 2GMR^3\left(\left(u + \frac{1}{2}\left[\frac{1}{R} - \frac{1}{2GM}\right]\right)^2 + \left(\frac{1}{R} - \frac{1}{2GM}\right)\left(\frac{1}{R} - \frac{1}{4}\left[\frac{1}{R} - \frac{1}{2GM}\right]\right)\right) \tag{54}$$

Define the constant $\gamma = \frac{1}{R} - \frac{1}{2GM}$, then we can write our full polynomial, $P(u)$, as:

$$P(u) = 2GMR^3\left(u - \frac{1}{R}\right)\left(\left[u + \frac{\gamma}{2}\right]^2 + \gamma\left[\frac{1}{R} - \frac{\gamma}{4}\right]\right) \tag{55}$$

We can obtain a factorization into roots of $P(u)$ by just finding the roots of $S(u)$:

$$P(u) = 2GMR^3\,(u - X)\,(u - Y)\,(u - Z)$$

Where:

$$X = -\frac{\gamma}{2} + \sqrt{\left(\frac{\gamma}{4} - \frac{1}{R}\right)\gamma} \qquad Y = \frac{1}{R} \qquad Z = -\frac{\gamma}{2} - \sqrt{\left(\frac{\gamma}{4} - \frac{1}{R}\right)\gamma}$$

Plug in the integral, then we obtain:

$$\phi = \frac{1}{\sqrt{2GM}}\int_0^Y \frac{du}{[(u - X)(u - Y)(u - Z)]^{1/2}} \tag{56}$$

Now, we define the following constant:

$$Q^2 = (6GM + R)\,(R - 2GM) \tag{57}$$

Then, our roots X, Y, Z become:[27]:

$$X = \frac{R - 2 + Q}{4R} \qquad Y = \frac{1}{R} \qquad Z = \frac{R - 2 - Q}{4R} < 0$$

$$X = \frac{1}{4GMR}\left[(R - 2GM) + Q\right] \qquad Y = \frac{1}{R} \qquad Z = \frac{1}{4GMR}\left[(R - 2GM) - Q\right]$$

[26] Mathematica is my favorite.

[27] If we use geometerized units and set $GM = 1$, we obtain something consistent with the reference [2].

Remark 2.1. We may prove the following fact: $X > Y > Z$ with $X > Y$ holding for certain R (that depends on the value of GM).

Remark 2.2. This value of R is special, it is exactly the **Photon Sphere**, which we briefly defined in **Section 3.3.1**. In terms of the consequences on our computation, we see that the expression for our deflection angle is ill-defined if $R < \frac{3}{2}r_s$, as this is exactly the radial distance at which the orbit of the photon will degenerate and be unable to escape to infinity. We give a "graphical proof" of this for the case of $GM = 1$ in **Figure 2**.

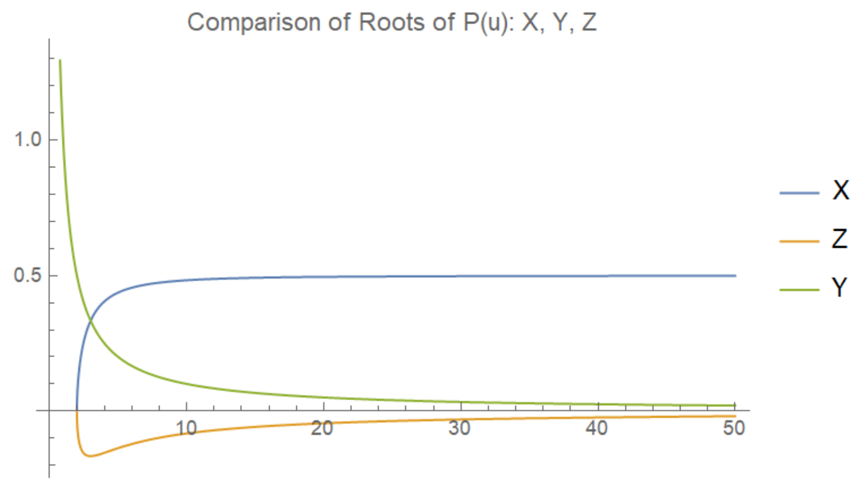

Figure 2: A graphical comparison of our roots X, Y, Z. Clearly, we see that $X > Z$ for any $R > 0$, while $X > Y$ for sufficiently large values of R. Note, solving for when the $X = Y$ shows that our expressions are only valid for $R > 3$. This suffices as a "graphical proof".

We use geometrized units, $G = c = 1$ and set the mass $M = \frac{1}{G}$ (see **Appendix C**).

Source: Wolfram Mathematica 12.2

We now use the following identity from [7][28].

$$\int_v^Y \frac{du}{\sqrt{(u-X)(u-Y)(u-Z)}} = \frac{2}{\sqrt{X-Z}}F(\delta, q) \qquad X > Y > v \geq Z \tag{58}$$

Where

$$\delta = \arcsin\left[\sqrt{\frac{(X-Z)(Y-v)}{(Y-Z)(X-v)}}\right] \qquad q = \sqrt{\frac{Y-Z}{X-Z}}$$

If we explicitly calculate these, we obtain:

$$X - Z = \frac{Q}{2GMR} \qquad Y - Z = \frac{6GM + Q - R}{4GMR} \qquad \frac{Y}{X} = \frac{4GM}{R - 2GM + Q}$$

[28]In this reference, it is on page 254, and is Formula 4 on section 3.131.

Hence

$$\delta = \arcsin\left[\sqrt{\frac{(X-Z)Y}{(Y-Z)X}}\right] = \arcsin\left[\left(\frac{8GMQ}{(R-2GM+Q)\,(6GM+Q-R)}\right)^{1/2}\right]$$

And for convenience, set:

$$\beta = \left(\frac{8GMQ}{(R-2GM+Q)\,(6GM+Q-R)}\right)^{1/2}$$

$$q = \left(\frac{6GM+Q-R}{2Q}\right)^{1/2}$$

Now, the **elliptic integral of the first kind** is defined in [2][29] as:

$$F(\phi,k) = \int_0^{\phi} \frac{d\alpha}{\sqrt{1-k^2\sin^2(\alpha)}} = F(\sin(\phi),k) = \int_0^{\sin(\phi)} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \tag{59}$$

Therefore, we finally obtain that:

$$\phi = \frac{1}{\sqrt{2GM}}\int_0^{1/R} \frac{du}{(u-X)(u-Y)(u-Z)} = 2\sqrt{\frac{R}{GMQ}}F\left[\beta, \left(\frac{6GM+Q-R}{2Q}\right)^{1/2}\right]$$

We can get our true deflection angle α, by:

$$\alpha = 2\phi - \pi$$

Hence, this becomes:

$$\alpha = 4\sqrt{\frac{R}{GMQ}}F\left[\beta, \left(\frac{6GM+Q-R}{2Q}\right)^{1/2}\right] - \pi \tag{60}$$

We will now use a reduction formula to reduce this elliptic integral into smaller parts.

Proposition 3. ***Reduction Formula for the Elliptic Integral*** $F(\delta,k)$

The following identity for an Elliptic Integral of the First Kind:

$$F(1,k) = F(x_1,k) + F(x_2,k)$$

holds if and only if:

$$\sqrt{1-k^2}\frac{x_1x_2}{\sqrt{(1-x_1^2)(1-x_2^2)}} = 1$$

[29]This is the first expression on the top of page 860 of said reference.

If we choose suitable x_1, x_2 and k, we may apply this formula immediately.

$$x_1 = \sqrt{\frac{8GMQ}{(R-2GM+Q)(6GM+Q-R)}} \qquad x_2 = \sqrt{\frac{2+Q-R}{6+Q-R}} \qquad k = \sqrt{\frac{6GM-R+Q}{2Q}}$$

We ask the reader to take it on faith, or to verify themselves, that the conditions of **Proposition 3** hold. We take it for granted and proceed with the computation.

$$F(x_1, k) = F(1, k) - F(x_2, k)$$

This gives us a representation of our deflection angle α:

$$\alpha = 4\sqrt{\frac{R}{GMQ}}\left(F\left(1, k\right) - F(x_2, k)\right) - \pi \tag{61}$$

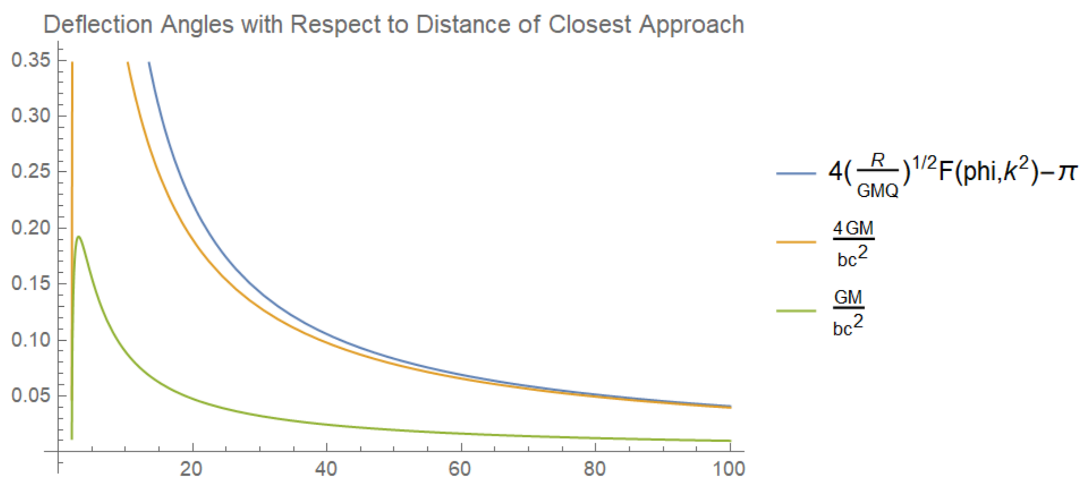

Figure 3: **Blue**: Exact Result. **Orange**: First-Order Deflection Angle from Perturbation Theory. **Green**: Result from Dimensional Analysis.

The deflection angle, α, is plotted with respect to the parameter R, defined by **Equation 49**. We use geometrized units, $G = c = 1$ and set the mass $M = \frac{1}{G}$ (see **Appendix C**).

Source: Wolfram Mathematica 12.2

3.3.3 The First-Order Limit

Our goal here is to recover the original computation we performed in section **3.2.3**. Let us analyze the limits of our original integral:

$$\phi = \int_R^\infty \frac{dr}{r^2\left(\frac{1}{b^2} - \left(1 - \frac{2GM}{r}\right)\frac{1}{r^2}\right)^{1/2}}$$

Let $r_s = 2GM$, and write our integrand in terms of a dimensionless parameter, $x = r_s/b$ and then make the change of variables $r = b/u$. We note that this parameter, r_s/b, should be extremely small.

$$\phi = \int_0^U \frac{du}{\sqrt{1 - u^2 (1 - xu)}} \tag{62}$$

The trick we use here is due to [8].

We express the integral for ϕ as such:

$$\phi = \int_0^U \frac{1}{\sqrt{1 - ux}} \frac{1}{\sqrt{(1 - ux)^{-1} - u^2}} du$$

We expand all the terms that have $(1 - ux)$, to the linear term, this results in the following:

$$\phi = \int_0^U \frac{(1 + \frac{ux}{2})}{\sqrt{1 + ux - u^2}} du \tag{63}$$

Now, there is, quite remarkably, a closed-form anti-derivative for the integrand. We state it here[30]:

$$\int \frac{(1 + \frac{ux}{2})}{\sqrt{1 + ux - u^2}} du = \frac{1}{2} \left(-x\sqrt{1 - u^2 + ux} - \frac{1}{2} \left(4 + x^2\right) \arctan \left[\frac{-2u + x}{2\sqrt{1 - u^2 + ux}} \right] \right)$$

Now plug in the limits of integration[31]:

$$\phi = \int_0^U \frac{1}{\sqrt{1 - ux}} \frac{1}{\sqrt{(1 - ux)^{-1} - u^2}} du = -\left(-1 - \frac{x^2}{4} \right) \frac{\pi}{2} + \frac{x}{2} + \left(1 + \frac{x^2}{4} \right) \arctan \left[\frac{x}{2} \right]$$

Note the sign of the first term in the evaluated integral[32]. Now we simplify this:

$$\phi = \frac{\pi}{2} + \frac{\pi x^2}{8} + \frac{x}{2} + \left(1 + \frac{x^2}{4} \right) \arctan \left[\frac{x}{2} \right] \tag{64}$$

Then we may now find our true deflection angle:

$$\alpha = 2\phi - \pi \tag{65}$$

This results in:

$$\alpha = \frac{\pi x^2}{4} + x + \left(2 + \frac{x^2}{2} \right) \arctan \left[\frac{x}{2} \right] \tag{66}$$

[30] We used mathematica for it.

[31] Note here that U is actually a root of the polynomial $1 + ux - u^2$. This is a result of the work we did in the previous section, where we analyze the roots of the polynomial in the denominator of the integral for ϕ. This tremendously simplifies our computation.

[32] The arctan in the antiderivative actually goes to $-\frac{\pi}{2}$ instead of $\frac{\pi}{2}$, with the assumption that $U \gg x$. Note however, even if the sign is incorrect, it will not change the answer when we truncate all terms that are non-linear. However, if we want to compute the corrections to this, it will be important to keep track of the sign.

Now recognize that our parameter $x \ll 1$, hence we may ignore any terms that are non-linear in x.

$$\alpha = x + 2\left(\frac{x}{2} + \frac{x^3}{24} + \mathcal{O}(x^5)\right) \overset{\text{Truncate to Linear Order}}{\Longrightarrow} \alpha = 2x \tag{67}$$

Therefore, we obtain that:

$$\alpha = 2x$$

If we put the answer in our original constants, then we get:

$$\alpha = \frac{4GM}{b} \tag{68}$$

Or if we convert back to SI units:

$$\alpha = \frac{4GM}{bc^2} \tag{69}$$

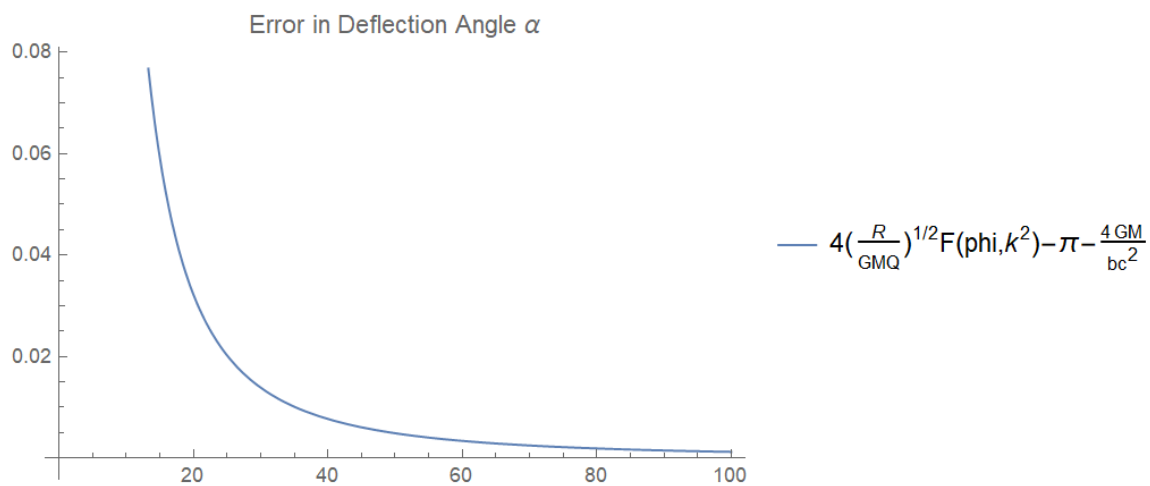

Figure 4: Note how the error in the deflection angle approaches 0 as we take $R \longmapsto \infty$, which means $b \longmapsto \infty$. Hence, this justifies the use of perturbation theory (**Equation 5**), which assumes that $|h_{\mu\nu}| \ll 1$ and, **Equation 24** and **Equation 35** imply that perturbation theory works when the impact parameter $b \longmapsto \infty$.

Source: Wolfram Mathematica 12.2

This justifies the work that we did in **Section 3.3.2** (Equation (36)) and the work that we did in **Section 3.1.2** (Equation (18)). The Elliptic Integral that we obtained, in the appropriate limit and with careful approximations, reduces to the first-order deflection obtained using first-order perturbation theory.

Note: This expansion of our analytical result in the previous section relies on the fact that our impact parameter is sufficiently larger than the Schwarzschild Radius (and hence, the mass), i.e. $b \gg r_s$. For a Schwarzschild Black Hole, as the Schwarzschild Radius and the Event Horizon are synonymous, this is expected. The error in the deflection angle decreases for large b as shown in **Figure 4**.

4 Conclusion

Throughout this paper, we focused on photon deflection (a form of gravitational lensing), one of the first phenomena that solidified General Relativity's validity as a physical theory. We used three different techniques that coincide to a correct answer (up to a constant I suppose):

1. Using a mathematically rigorous formulation of dimensional analysis, utilizing the **Buckingham Pi Theorem**. We found (up to a constant) our deflection angle to be the kernel of a linear transformation of dimensionless parameters, and deduced that our quantity is the correct deflection angle by arguing that it was the unique quantity up to an overall power. The advantage of this method is its sheer simplicity: it took an otherwise nonlinear problem into the realm of elementary linear algebra.

2. Using **first-order perturbation theory**, we are led to the fact that our **linearized theory is actually gauge invariant**, indicating that we are working with a physical theory. We then pick a convenient gauge called the **transverse gauge** in which our strain is spatially transverse, which leads to a convenient form of the metric tensor, by which we are able to compute the photon deflection angle between the perturbed geodesic and the background geodesic, by expanding our geodesic equation to first-order. The advantage of this result is that it is simple and highlights the use of perturbation theory to observe physical quantities.

3. **Using calculus of variations to find an equation of motion for the Light-Like Trajectory of a Schwarzschild Orbit**. Whereby, we are able to compute the deflection angle between the true trajectory of the photon, and the background geodesic, as an **Elliptic Integral of the First Kind**. Further analysis shows that the Elliptic Integral we obtain reduces to the result using method (2) and closely relates to the result found by using method (1). We find that the advantage of this argument is that we can obtain arbitrarily high precision by expanding the analytical result past linear order.

In every instance, we found the well-known result for the deflection angle of a photon, about a gravitational source with mass M and impact parameter b: $\frac{4GM}{c^2b}$. And furthermore, we found that these results are first-order, and rely on the assumption that our photon has an impact parameter much greater than the Schwarzschild Radius (hence the mass) of the gravitating body. Hence, this first-order theory suffices for the case when a photon orbits well beyond the Schwarzschild Radius of a gravitational source.

If we were to continue this work, we would look at more exotic gravitational sources, namely, the Reissner-Nordstrom Black Hole, or the Kerr Black Hole, both being examples of closed-form solutions to Einstein's Equations for **charged black holes**[33]. We expect that, in these regimes where assumptions that the source be stationary and have trivial charge are relaxed, that the first-order result for the deflection angle remains unchanged; again, with the assumption that our photon propagates sufficiently far away from the event horizon of such black holes.

Furthermore, we hypothesize that such black holes would require more advanced techniques to account for other fundamental parameters, such as charge or angular dependence on θ.

5 Acknowledgements

I would like to thank my advisor, Andrés Franco Valiente, for his consistent guidance. There are no words to describe the immense amount of knowledge and wisdom and tricks he has bestowed upon me over the years. Even coming to a consensus on the topic of this paper was an arduous task. I am indebted to Andrés for his patience, as what started as a project on String Theory, turned into a project on Renormalization in Quantum Field Theory, before, in the last month, turning into a project on General Relativity. I will likely never find someone, friend nor academic peer, that will ever be as patient as Andrés has been.

I would like to thank my partner, Kelly Lam, for always being present for me, and being one of the smartest and most patient individuals I've met. She edited and drew **Figure 1**, once again proving that she is immensely relevant and capable of uplifting me in all aspects of my life. She may not know much about General Relativity, but she knows everything about me, and that is an infinitely harder topic.

I would like to thank countless other friends for the discussions we have had over the years that slowly added up to the little knowledge I have today, that I have used to write this paper. I will avoid naming names as I will most definitely mistakenly exclude someone that I shouldn't have, but they know who they are. This is the first time I worked independently on a project, so it is an experience I will cherish for a while.

This paper was written as a project, satisfying requirements for the Physics Directed Reading Program of Fall 2020.

[33]Note that this was not in the scope of the paper given that we were regarding *Neutral* Gravitational Sources.

A Math Background: Geometry and Topology of Spacetime

This section is meant to be an introduction to some relevant concepts in General Relativity for the mathematically-minded readers. This is in no way necessary to understand, at least in this depth and generality. I encourage readers to not spend too much time dwelling on this.

A.1 Topological Considerations

Definition A.1. When we say "Manifold", we really mean a C^∞, Paracompact, Hausdorff Manifold, M, a disjoint union of C^∞ charts, (ϕ_α, U_α).

Proposition 4. *Such a smooth manifold necessarily has a globally-defined **Pseudo-Riemann Metric**.*

Proof. The proof of this is available in more mathematically advanced GR texts. As a sketch, it involves taking the defined pseudo-Riemannian Metric in each neighborhood of the manifold (which, by definition, is diffeomorphic to Euclidean Space), recognizing that the Manifold is a disjoint union of such neighborhoods (and the associated local diffeomorphisms), and using the assumption of paracompactness and Hausdorffness, we obtain a partition of unity, $\{\theta_i\}$, subordinate to an open cover, $\{U_i\}$. By using the definition of a partitiion of unity, we may create a global pseudo-Riemannian Metric out of stitching together all pseudo-Riemannian Metrics in each neighborhood of the manifold. Therefore, we conclude the existence of a globally-defined pseudo-Riemannian Metric. □

Definition A.2. Spacetime is a four-dimensional, Simply-Connected **Lorentzian** Manifold with a signature $(-, +, +, +)$.

Definition A.3. A **path** is a smooth, regular curve, $\gamma : [0, 1] \longrightarrow M$, mapping to our manifold M.

A.2 Algebraic Objects

The primary algebraic object of interest in General Relativity[34] is the tensor. We will give three definitions of the tensor, all in different terms familiar to different people.

Definition A.4. Algebraically, a Tensor is an element of $F(M \times N)/H$, where $F(M \times N)$ is the free group generated by the module $M \times N$, quotiented by the submodule, H. Given a bilinear projection map $T : M \times N \longrightarrow F(M \times N)/H$, H is the submodule of $F(M \times N)$ generated by the ordered pairs:

$$(m_1 + m_2, n) - (m_1, n) - (m_2, n)$$

$$(m, n_1 + n_2) - (m, n_1) - (m, n_2)$$

[34] And really all of physics if you think hard enough.

$$(rm, n) - (m, rn) \qquad r \in R$$

These are the undesirable elements of $F(M \times N)$ that would violate the bilinearity of T.

This R-module $F(M \times N)/H \cong M \otimes N$, and is called the **tensor product of M and N**. All decomposable elements (images of T), $m \otimes n \in M \otimes N$ is actually defined via:

$$m \otimes n = (m, n) + H$$

Therefore, with the machinery defined, a **Tensor** is the **image of a bilinear map**[35] $T : M \times N \longrightarrow M \otimes N$. Consider M and N to be vector spaces[36], and allow T to vary as a function of coordinates, then we obtain a **Tensor Field**.

Definition A.5. Under a local diffeomorphism $\phi : U \longrightarrow U'$, a **vector field** is an function, $u(x)$, that transforms as:

$$u' = d\phi_x u$$

Or, in terms of explicit coordinates:

$$u^{\alpha'}(x) = \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} u^{\alpha}(x)$$

Likewise, a **covector field** is a function, $p(x)$, that transforms as:

$$p' = d\phi_x^{-1} p$$

Or in coordinates:

$$p_{\alpha'}(x) = \frac{\partial x^{\alpha}}{\partial x^{\alpha'}} p_{\alpha}(x)$$

Now we get to the definition we really care about:

Definition A.6. A **Tensor Field** is a smooth[37], multilinear function that takes in, as its inputs, smooth vector fields and smooth covector fields, both of which vary as a function of coordinates.

The **Rank of a Tensor (Field)** is an ordered pair (n, k) describing the fact that T takes in, as its arguments, n vector fields, and k covector fields[38].

Now, for another popular definition of a tensor.

Definition A.7. A **Rank-(n, k) Tensor Field** is a function that transforms like a vector field in its first n components, and transforms like a covector field in its last k components[39]. This is commonly referred to as transforming *tensorially* or transforming *like a tensor*.

[35] Mathematicians call this bilinear map itself, a tensor.

[36] free modules over a field

[37] We are implicitly assumiing that we are working over a smooth manifold. We can weaken it to C^k manifolds and there are no limitations except for the level of differentiability

[38] Some people have the opposite convention, where n is the covector fields and k is the vector fields. Any permutation of the arguments is a diffeomorphism, so this is only convention.

[39] Refer to Definition A.4

Definition A.8. "Transforming *like a tensor*" entails the following transformation behavior for a set of local diffeomorphisms ϕ^i:

$$T' = \left(\bigotimes_{i=1}^{n} d\phi^i_{x_i} \right) \otimes \left(\bigotimes_{j=1}^{k} \left(d\phi^j\right)^{-1}_{x_j} \right) T$$

Definition A.9. Component-wise, we often represent an (n, k)-Tensor Field T as:

$$T^{\mu_1 \dots \mu_n}_{\sigma_1 \dots \sigma_k}$$

Whose transformation behavior (equivalent to **Definition A.8**) is described as follows:

$$T^{\mu'_1 \dots \mu'_n}_{\sigma'_1 \dots \sigma'_k} = \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \cdots \frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{\sigma_1}}{\partial x^{\sigma'_1}} \cdots \frac{\partial x^{\sigma_k}}{\partial x^{\sigma'_k}} T^{\mu_1 \dots \mu_n}_{\sigma_1 \dots \sigma_k}$$

Now that we have repeatedly beaten the definition of a tensor into the ground, we now discuss tangent vector fields and cotangent vector fields.

Definition A.10. The **tangent space of a manifold** M **at a point** $p \in M$, also denoted T_pM, is the set of all vectors tangent to p on the manifold M.

Seeking a more concrete characterization of T_pM, let U be a neighborhood of a point p, given a parametrization, $\phi : U \subset \mathbb{R}^k \longrightarrow V \subset M$, such that $\phi(0) = x$. The tangent space is actually defined as the image of the derivative:

$$T_pM = d\phi_0 \left[U\right]$$

Hence, given that the image of a linear map is defined exclusively by its action on the basis, we do the following:

$$d\phi_0 e_\mu = \begin{pmatrix} \nabla \phi_1^T \\ \vdots \\ \nabla \phi_k^T \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \partial_\mu \phi_1 \\ \vdots \\ \partial_\mu \phi_k \end{pmatrix} = \partial_\mu \phi$$

It is not obvious why this is meaningful, but let us consider a (smooth) real-valued map $f : M \longrightarrow \mathbb{R}$. By the chain rule:

$$df_{\phi(0)} \circ d\phi_0 = d(f \circ \phi)_0 = \partial_\mu f(\phi(0))$$

It becomes very clear here that we may identify $\partial_\mu \phi$ with ∂_μ as the function f we use is arbitrary. This gives us the following definition:

Definition A.11. The **Tangent Space of** M **at the point** p is the real vector space spanned by the elements:

$$\{\partial_1, \dots, \partial_k\}$$

These are partial derivatives in each principal direction, and is often referred to as the **coordinate basis**.

For the definition most relevant to us, we use the following proposition:

Proposition 5. *Every tangent vector is a velocity vector of some curve* $\gamma : [0,1] \longrightarrow M$.

Proof. Let $f : M \longrightarrow \mathbb{R}$ be a smooth, real-valued map. Let (U, ϕ) be the chart mentioned in the above demonstration. Let γ be a path, as defined in **Definition A.3**. Let λ parametrize γ. By assumption, $p \in \gamma$. We define the tangent vector to γ at p as:

$$\frac{df}{d\lambda} = \frac{d}{d\lambda}(f \circ \gamma) = \frac{d}{d\lambda} f \circ (\phi \circ \phi^{-1}) \circ \gamma = \frac{d}{d\lambda}\left[(f \circ \phi) \circ (\phi^{-1} \circ \gamma)\right]$$

We can use the chain rule to obtain that:

$$\frac{d(\phi^{-1} \circ \gamma)}{d\lambda} \frac{\partial (f \circ \phi)}{\partial x^\mu} = \frac{dx^\mu}{d\lambda} \frac{\partial f}{\partial \mu} \equiv \frac{dx^\mu}{d\lambda} \partial_\mu f$$

We conclude that:

$$\frac{df}{d\lambda} = \frac{\partial x^\mu}{d\lambda} \partial_\mu f$$

As f is arbitrary, we conclude that:

$$\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda} \partial_\mu$$

Hence, every tangent vector is the velocity vector of some curve, γ. **Note, we may also take this as the definition of a tangent vector**. □

Definition A.12. The **Tangent Space of** M **at the point** p is the vector space of velocity vectors (i.e. directional derivatives), $\frac{d}{d\lambda}$. The proposition preceding this definition shows that the coordinate basis $\{\partial_1, \ldots, \partial k\}$ generates all directional derivatives of arbitrary scalar functions f.

Definition A.13. The **exterior derivative of a scalar function** is defined as:

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

Definition A.14. Let f be a real-valued function on M. Let d be the exterior derivative. A **cotangent vector** is defined through its action on tangent vectors:

$$df\left(\frac{d}{d\lambda}\right) = \frac{df}{d\lambda} = \frac{dx^\mu}{d\lambda} \partial_\mu f$$

Definition A.15. The **cotangent space of a manifold** M **at a point** p, also denoted T_p^*M, is the dual vector space of T_pM. The cotangent space is spanned by the exterior derivative of the coordinate projection map, π_i:

$$d\pi_\mu(x) = dx^\mu$$

By **Definition A.13** and **A.14**, we may completely define the action of T_p^*M on T_pM, through the coordinate bases, as:

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

There is a meaningful connection between tangent/cotangent vectors and tangent/cotangent bundles, but we will not go into too much detail here. We have presented more than enough rigorous mathematics to work through the computations of General Relativity.

A.3 Geometric Objects

Let $V = V^\mu X_\mu$ and $W = W^\nu X_\nu$ where $V, W \in T_pM$ and X_μ spans T_pM and is the smooth cross section of the Tangent Bundle, TM.

Definition A.16. Tangent/Cotangent Bundles

Assume that for M, that we pick a suitable chart (U, ϕ) such that $\phi(0) = x$ and $\phi : U \subset \mathbb{R}^k \longrightarrow V \subset M$. Let $p \in V$ where V is an open neighborhood about p in M. The local coordinate system $\phi^{-1}(p) = (x^0(p), \ldots, x^k(p))$ induces a local frame for the cotangent bundle T^*M and the tangent bundle TM. **Elements in the local frame of T^*M** are exterior derivatives of projections of the coordinate system, i.e. 1-forms dx^μ. Likewise, **elements in the local frame of TM** are tangent vectors in the principal directions indicated by the coordinate system, ∂_μ.

The following is tautology, given **Definition A.13-A.15**:

$$dx^\mu \partial_\nu = \delta^\mu_\nu$$

$$\partial_\nu x^\mu = \delta^\mu_\nu$$

This highlights the duality of the tangent and cotangent vector. In fact, as one could tell by the section we put this into, **we take this to be the definition of the basis 1-form**.

Definition A.17. Intrinsically, the **metric tensor** is a nondegenerate, symmetric, bilinear, quadratic form on M of the form:

$$g : T_pM \times T_pM \longrightarrow \mathbb{R}$$

$$g(V, W) = \sum_{\mu,\nu} g(X_\mu, X_\nu) V^\mu \otimes W^\nu$$

Where $g(X_\mu, X_\nu) \equiv g_{\mu\nu}$ are metric coefficients for the matrix $d\phi_0^T d\phi_0$. Quick inspection shows us that we may compute these components quickly by using:

$$g_{\mu\nu} = X_\mu \cdot X_\nu$$

Now pick a chart (U, ϕ) such that $\phi(0) = x$. Then we have a local coordinate system about some neighborhood of $p \in M$ dictated by $\phi^{-1}(p) = (x^0(p), x^1(p), \ldots, x^k(p))$. Using the local frame on the cotangent bundle induced by our choice of chart, we may plug in the cotangent vector $dx^\mu \partial_\mu$:

$$g(dx^\mu \partial_\mu, dx^\nu \partial_\nu) = ds^2 = \sum_{\mu,\nu} g_{\mu\nu} dx^\mu \otimes dx^\nu$$

Definition A.18. Causality

Note that this is with respect to our choice of metric signature and will be different for the other canonical choice.

A tangent vector $u^\mu \in T_pM$ is **Space-Like Separated** if:

$$u^\mu u_\mu > 0$$

A tangent vector $u^\mu \in T_pM$ is **Time-Like Separated** if:

$$u^\mu u_\mu < 0$$

A tangent vector $u^\mu \in T_pM$ is **Null** if:

$$u^\mu u_\mu = 0$$

A path γ is any of these characterizations if the corresponding tangent vector $u^\mu \in T_pM$ for $p \in \gamma$ is any of these characterizations.

Definition A.19. For our purposes, a **connection** on a manifold is a correction term that makes the covariant derivative transform tensorially.

Proposition 6. ***Existence and Uniqueness of the Levi-Civita Connection***

There exists a unique metric-compatible, torsion-free connection on any smooth pseudo-riemmanian manifold (M, g). It is of the explicit form:

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} \left(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}\right)$$

Proof. The existence and uniqueness of this is given by the explicit form. It is defined through a universal property as a parallel transport functor on the category of vector bundles. $\square$

Definition A.20. The **covariant derivative** is a **derivation**[40] **of a tensor field** that transforms tensorially. For a general (k, l)-tensor field T, we have:

$$\nabla^\alpha T^{\mu_1\mu_2\ldots\mu_k}_{\nu_1\nu_2\ldots\nu_l} = \partial_\alpha T^{\mu_1\mu_2\ldots\mu_k}_{\nu_1\nu_2\ldots\nu_l} + \left(\Gamma^{\mu_1}_{\alpha\lambda} T^{\lambda\mu_2\ldots\mu_k}_{\nu_1\nu_2\ldots\nu_l} + \Gamma^{\mu_2}_{\alpha\lambda} T^{\mu_1\lambda\ldots\mu_k}_{\nu_1\nu_2\ldots\nu_l} + \cdots\right)$$
$$- \left(\Gamma^{\lambda}_{\alpha\nu_1} T^{\mu_1\mu_2\ldots\mu_k}_{\lambda\nu_2\ldots\nu_l} + \Gamma^{\lambda}_{\alpha\nu_2} T^{\mu_1\mu_2\ldots\mu_k}_{\nu_1\lambda\ldots\nu_l} + \cdots\right)$$

From this, it is easy to deduce that the covariant derivative of a vector field and one-form are, respectively:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda$$

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda$$

[40]Do not be scared by big words. A derivation is a map that has all the properties of a derivative.

Definition A.21. The **directional covariant derivative** is simply the chain rule applied to the covariant derivative with respect to an affine parameter:

$$\frac{D}{d\lambda} \equiv U^\mu \nabla_\mu$$

Where $U \in T_pM$ is the tangent vector to the path x^μ at some point $p \in x^\mu$. λ affinely parametrizes the path x^μ.

Definition A.22. Parallel Transport

The equation of parallel transport for a tensor field T is:

$$\frac{DT}{d\lambda} = \frac{dx^\mu}{d\lambda} \nabla_\mu T^{\mu_1\mu_2\ldots\mu_k}_{\nu_1\nu_2\ldots\nu_l} = 0$$

Intuitively, this means that a tensor field defined on a manifold remains parallel according to the connection, which holds information about the intrinsic geometry of the manifold.

As a special case, the parallel transport of a vector field V is of the form:

$$\frac{dV^\mu}{d\lambda} + \Gamma^\mu_{\alpha\rho} \frac{dx^\alpha}{d\lambda} V^\rho = 0$$

Definition A.23. Geodesic Equation

The parallel transport of a tangent vector to a path is known as the **geodesic equation**.

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\sigma} \frac{dx^\alpha}{d\lambda} \frac{dx^\sigma}{d\lambda} = 0$$

Author's Note: The following discussions of the Riemann Curvature Tensor and its derived objects are given in [5].

Definition A.24. Riemann Curvature Tensor

The **Riemann Curvature Tensor** is a multilinear map whose components are defined uniquely by the following:

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\mu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}$$

We further define the **Ricci Tensor** and **Ricci Scalar** as the following:

$$R_{\mu\nu} \equiv R^\lambda_{\mu\lambda\nu} \qquad R = g^{\mu\nu} R_{\mu\nu}$$

Proposition 7. ***Properties of the Curvature Tensor***

The following are fundamental identities of the Riemann Curvature Tensor and its associated quantities:

Symmetry in the Last Two Components:

$$R^\mu_{\nu\sigma\rho} = R^\mu_{\nu\rho\sigma}$$

Furthermore, if we lower the upper index, we obtain symmetry in the first two components, and the last two components:

$$R_{\mu\nu\sigma\rho} = R_{\nu\mu\sigma\rho} = R_{\mu\nu\rho\sigma}$$

***Cyclic Property**:*

$$R_{\rho\sigma\mu\nu} + R_{\rho\nu\sigma\mu} + R_{\rho\mu\nu\sigma} = 0$$

This can also be expressed compactly as:

$$R_{\rho[\sigma\mu\nu]} = 0$$

The Ricci Tensor has a few properties of its own:

***Symmetry**:*

$$R_{\mu\nu} = R_{\nu\mu}$$

Definition A.25. Stress Energy Tensor

The **Stress-Energy Tensor** is a multlinear map that describes the mass density of spacetime. We have various forms for various types of matter in spacetime:

Dust:

$$T_{\mu\nu} = \rho U^\mu U^\nu$$

Perfect Fluid:

$$T^{\mu\nu} = (\rho + p)\, U^\mu U^\nu + p\eta^{\mu\nu}$$

Where ρ and p denote the rest energy density and the pressure, respectively.

Now we get to the primary object of our study.

Definition A.26. Einstein Field Equations

The Einstein Field Equations are a system of equations of motion for a spacetime with matter described by the stess-energy tensor T:

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

In **a perfect vacuum**, we see that the field equations are given as the homogenous equation:

$$R_{\mu\nu} = 0$$

B Geometrized Units

The widely-accepted units are the SI-unit system. As discussed in **Section 3.1.1**, we see that the choice of unit system can be seen as a choice of basis, hence it is relative. Different unit systems serve different purposes. For example, the natural unit system is one in which two fundamental constants, $\hbar$ and c, are normalized, i.e. $\hbar = c = 1$. This serves two purposes: (i) we achieve mass-energy equivalence (ii) planck constant is on the order of 10^{-34}, leading to extremely small values that may not be computationally tractable. Setting $\hbar = 1$ "magnifies" the scale of the numbers we work with by about $1/\hbar$.

In General Relativity, that the scale we work with is typically large (the mass of the Solar System's Sun is on the order of 10^{30} kg). However, unlike in high energy physics, we have various constants that serve to suppress this scale, namely G and c. Normalizing these often serves to show that sizeable corrections occur when relativistic effects are taking place.

The aforementioned system is a popular unit system in General Relativity known as the *Geometrized Unit System*. This is an informal unit system where velocities are dimensionless, and the gravitational coupling is dimensionless (so $G = c = 1$). Various authors will be using this convention, however, in calculation, we only set $c = 1$ for the purpose of mass-energy equivalence.

The biggest utility of this is that we can forego writing constants as long as our expressions are consistent in dimension and units.

References

[1] Gerrit Ansmann. Natural units and the vector space of physical values. *European Journal of Physics*, 36(3):035008, Mar 2015.

[2] G. S. Bisnovatyi-Kogan and O. Yu. Tsupko. Strong gravitational lensing by schwarzschild black holes. *Astrophysics*, 51(1):99–111, Jan 2008.

[3] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*, pages 193–197. Pearson Education Inc. Addison Wesley, 2004.

[4] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*, pages 274–293. Pearson Education Inc. Addison Wesley, 2004.

[5] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*, pages 93–139. Pearson Education Inc. Addison Wesley, 2004.

[6] Charles Galton Darwin. *The Gravity Field of a Particle*, volume 245, pages 189–191. Royal Society, January 1959.

[7] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*, pages 254, 860. Elsevier Academic Press, 2007.

[8] James B. Hartle. *Gravity: An Introduction to Einstein's General Relativity*, pages 210–212. Pearson Education Inc. Addison Wesley, 2003.

[9] John R. Taylor. *Classical Mechanics*, pages 215–228. University Science Books, 2005.

Any errors and typos that are present in this paper are my own.