Tests of General Relativity

James Fielder

Abstract

In this report we will examine the experimental evidence for Einstein's theory of General Relativity (GR). First, the basic ideas of GR will be developed, starting with the equivalence principle and leading to the Einstein equation. Then the theory of gravitational waves will be introduced. Experimental verifications for GR will be reviewed and the use of GR in the Global Positioning System presented.

This piece of work is a result of my own work except where it forms an assessment based on group project work. In the case of a group project, the work has been prepared in collaboration with other members of the group. Material from the work of others not involved in the project has been acknowledged and quotations and paraphrases suitably indicated.

Contents

List of Figures

1 | Introduction

In 1905 a young Albert Einstein published a paper entitled "On the Electrodynamics of Moving Bodies". This paper was the beginning of the revolution that relativistic physics would bring to our understanding of nature. Einstein abandoned the notion that time and space were separate entities, and instead intimately tied them together. However, the origins of this significant change to our view of nature were founded earlier, with the advent of classical electrodynamics.

Around 1862 James Clerk Maxwell published his equations which united electric and magnetic phenomena into one unified theory. One of the most striking results which can be derived from these equations is that light is an electromagnetic phenomenon. However, there was a problem: physicists of the time were unable to pinpoint the medium through which the light waves propagated. The solution to this problem was suggested by Maxwell; light propagated through the "luminiferous æther", a substance pervading the whole universe which light waves would perturb as they travelled. This idea would plague physics for the next 40 or so years, with successive experiments finding evidence which would both support and contradict the concept of the æther.

Einstein's great insight in 1905 was to do away with the æther and to suggest that the speed of light is the same in all inertial frames. This was a radical idea, going against centuries of established science at the time. In doing so Einstein vastly simplified the calculations required to fit with all observational data at that point [5].

On coming up with this idea Einstein had to rethink commonly held notions of time and space. Due to the constant nature of the speed of light in all frames, time had to proceed differentially for observers moving at different speeds. From here the absolute notions of length had to be discarded, and instead lengths would be observed differently by different observers. These were the main changes brought by the special theory of relativity [5].

Special relativity had a major flaw unfortunately: it did not account for gravity and thus could not be a full explanation of the mechanics of moving bodies. Once Einstein had published special relativity he began to work on the problem of integrating gravity with his new views of physics. One of the most important features of special relativity was that the physical laws of electromagnetism and mechanics had to be invariant under transformation from one inertial frame to another. Einstein realised this coordinate free description of physical laws was how any physical law should be written, so as to not prefer any frame of reference. He referred to this as the principle of general covariance and used it as a guide while inventing general relativity [17].

Guided by his desire for the physical laws to be coordinate independent, Einstein developed his theory of gravity using the mathematical language of tensors. This theory would relate the gravitational field at a point to the curvature of the spacetime as caused by nearby mass. Describing the fundamentals of this theory will be the topic of chapter 2 of this document.

The claims made by Einstein's theories are drastic and counterintuitive. Hence the scientific community would only accept them with great supporting evidence. Much of this document will be exploring that evidence, and seeing how we can produce experiments to verify Einstein's bold claims.

Finally, the applications to technology that relativistic physics have produced are fairly substantial, mostly in the implementation of the Global Positioning System. Without our understanding of relativity the GPS would not function to the level of accuracy that it does. The effects of relativity on the system will be examined and some calculations of the corrections due to relativistic effects will be demonstrated.

Throughout this text we will take the speed of light $c = 1$ in order to make formulas slightly clearer. If terms involving the speed of light are introduced they will be explicitly mentioned in the text.

2 | Basics of General Relativity

The following discussion relies primarily on the book by Cheng [5] and the book by Carroll [4] although some other references have been used, and they are referenced in the text.

2.1 The Equivalence Principle

The fundamental idea which caused Einstein to suggest that relativistic physics occurs on a 4 dimensional differentiable manifold is the equivalence principle. The equivalence principle states that physical laws in a freely falling frame of reference must be the same as those in an inertial frame without gravity. Thus, any point of spacetime must locally look like the flat spacetime of special relativity [5, 4].

The principle of equivalence had already existed in a simpler form before Einstein, as a statement that the inertial mass of an object and the gravitational mass were equivalent. This statement had already received experimental verification in various forms, first by Galileo (the story of dropping two balls of different mass from the Leaning Tower of Pisa seems to be apocryphal) and Newton, and then later in a more sophisticated form of the Eötvös experiment [5].

Einstein extended this principle from just motion in inertial frames to any frame of reference using any coordinate system. In order to capture this coordinate free description of physics Einstein turned to differential geometry and tensors. This mathematical formalism was chosen due to a differentiable manifold locally looking like flat space, but globally potentially having curvature. This condition replicates that of the equivalence principle: in small regions of space we can easily transform the metric to look like the flat space metric, and thus return to special relativity, but globally this is not possible [5, 4, 20].

2.2 Manifolds, Vectors and Tensors

The fundamental object in differential geometry is a differentiable manifold M. A differentiable manifold is an n-dimensional space which locally looks like the equivalent flat $\mathbb{R}^n$ space but may have global curvature, unlike flat space. In the case of relativity we use a 4-dimensional manifold which we call spacetime. Points on this manifold are called events and are a combination of both the time coordinate and the space coordinates. We then have local coordinates x^μ on this manifold. In the relativistic case greek indices will run from 0 to 3, with the 0 index being the time coordinate and 1 to 3 being spatial.

At all points p on the manifold we have the tangent space $T_p(M)$. The tangent space is the vector space of all possible tangent vectors at a point p on the manifold. The basis vectors of the tangent space can be taken to be the partial derivative operators at that point and thus are tangential to the manifold at that point. That is

$$\hat{e}_\mu = \partial_\mu = \frac{\partial}{\partial x^\mu}. \tag{2.1}$$

This is the basis we will use while studying General Relativity and we call this the coordinate basis. As an example in 2D we could visualise this as a plane intersecting a manifold at a point [1].
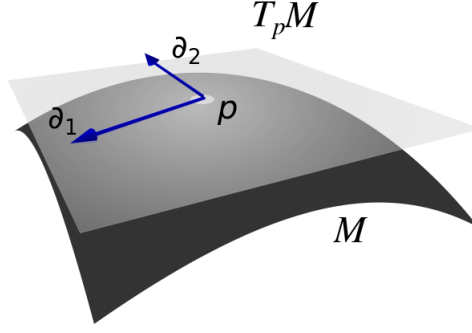

Figure 2.1: A visualisation of the tangent plane $T_p(M)$ to a manifold M in 2 dimensions.

The vectors in the tangent space are what we will identify as contravariant vectors [12, 4]. Due to using the coordinate basis we can think of such vectors as being differential operators on the tangent plane. Thus we see that a vector can be represented as a sum of its coordinates A^μ and the differential operators at the point ∂_ν. That is

$$\vec{A} = A^\mu \partial_\mu = A^0 \partial_0 + A^1 \partial_1 + A^2 \partial_2 + A^3 \partial_3. \tag{2.2}$$

We can now introduce another set of vectors at a point p on the manifold. If we consider the set of linear maps from the tangent space to the real numbers, $\omega : T_p(M) \to \mathbb{R}$, we can construct the dual vector space of the tangent space to the manifold at a point p, $T_p^*(M)$ [4]. We choose the basis of this vector space, $\hat{e}^\nu$ such that [4, 12]

$$\hat{e}^\nu \hat{e}_\mu = \mathrm{d}x^\nu(\partial_\mu) = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu. \tag{2.3}$$

From this basis, $\hat{e}^\mu = \mathrm{d}x^\mu$, we can construct any general linear map from vectors in the tangent space to the real numbers. Thus we can give an expression for the covectors much like (2.2) [4]

$$\vec{\omega} = \omega_\mu \mathrm{d}x^\mu = \omega_0 \mathrm{d}x^0 + \omega_1 \mathrm{d}x^1 + \omega_2 \mathrm{d}x^2 + \omega_3 \mathrm{d}x^3. \tag{2.4}$$

Explicitly, by choosing the bases as we have, the linear mapping $\omega : T_p(M) \to \mathbb{R}$ must be the same as the standard euclidean inner product [12]. That is, if we take the contraction of a vector and a covector $A^\mu \omega_\mu$ we obtain the sum of the product of the components of the vector and covector [12]

$$\begin{aligned}
\omega_\mu \mathrm{d}x^\mu(A^\nu \partial_\nu) &= \omega_\mu A^\nu (\mathrm{d}x^\mu \partial_\nu) = \omega_\mu A^\nu \delta^\mu_\nu = \omega_\mu A^\mu \\
&= \omega_0 A^0 + \omega_1 A^1 + \omega_2 A^2 + \omega_3 A^3 \in \mathbb{R}.
\end{aligned} \tag{2.5}$$

---

[1]Edited from original file at http://commons.wikimedia.org/wiki/File:Tangentialvektor.png

From here on in we will refer to a vectors components, A^μ, as the vector itself, and the components of a covector as the covector itself. The implicit understanding here is that we are using the coordinate basis, and that therefore vectors at any point in the manifold will always be defined in the way given above.

We are now prepared to define the objects which we will use most throughout our discussion of General Relativity, tensors. First we recall a theorem from linear algebra.

Theorem 2.2.1. *In a finite dimensional vector space V with a dual space V^* the dual of the dual space V^{**} is isomorphic to the original space V.* [20]

Thus, due to 2.2.1, we see that in addition to a covector acting on a vector as a linear operator we can consider the converse to be true: a vector can act on a covector as a linear operator. This is generally the case with a tensor: a tensor will take as an input vectors and covectors and map them to the real numbers. Thus we can define a tensor T as follows [2] [4, 20]

$$T : \underbrace{T_p^*(M) \times \cdots \times T_p^*(M)}_{p \text{ times}} \times \underbrace{T_p(M) \times \cdots \times T_p(M)}_{q \text{ times}} \to \mathbb{R}. \tag{2.6}$$

In the above expression we take $\times$ to be the Cartesian product: that is the above expression takes as an input ordered pairs of p covectors and q vectors.

We now can define the type of a tensor: the tensor above is of type $\begin{bmatrix} p \\ q \end{bmatrix}$. We will call the p covectors we input as p "up" indices and the q vectors "down" indices. That is, the tensor above would be given as $T^{\sigma_1 \cdots \sigma_p}{}_{\alpha_1 \cdots \alpha_q}$ [4]. In this way we can see that vectors are a type $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ tensor and covectors are of type $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We will define scalars to be of type $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ [20].

Finally we can introduce the tensor product, $\otimes$. The tensor product is a linear product of tensors, that is if we add a tensor T of type $\begin{bmatrix} p \\ q \end{bmatrix}$ and a tensor S of type $\begin{bmatrix} r \\ s \end{bmatrix}$ we form $T \otimes S$ of type $\begin{bmatrix} p+r \\ q+s \end{bmatrix}$. In terms of indices we have that the tensor product is like multiplication between tensors: if $T = T^{\lambda\sigma}$ and $S = S_{\mu\nu}$ then $T \otimes S = T^{\lambda\sigma} S_{\mu\nu}$ [4, 20].

2.2.1 Index Rules

Before we continue it is worth summarizing the rules for index calculus with tensors. We use the Einstein summation convention throughout the text, where any repeated index (one upper and one lower) is summed over. That is, in an n dimensional space we would have

$$v^i w_i = \sum_{i=0}^{n-1} v^i w_i \tag{2.7}$$

assuming that we are using zero indexes. We can now introduce several other operations on tensors.

- Contraction is an operation which maps a $\begin{bmatrix} p \\ q \end{bmatrix}$ type tensor to a $\begin{bmatrix} p-1 \\ q-1 \end{bmatrix}$ tensor by setting two of the indices equal to each other and summing them. That is if we have a tensor $K^{\mu\lambda\sigma}{}_{\alpha\gamma\varphi}$ and we contract the φ and σ indices we will obtain a new tensor, of type $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ [4]

$$K^{\mu\lambda\sigma}{}_{\alpha\gamma\sigma} = L^{\mu\lambda}{}_{\alpha\gamma}. \tag{2.8}$$

---

[2]In this definition we state that tensors map to the real numbers, however to be more correct we would have to say that tensors map to the field that the manifold is constructed over. Here we will always take this to be $\mathbb{R}$.

In the case of a matrix this operation is the same as the trace. In order to contract over two upper or two lower indices we will need the metric tensor, which we will introduce in the next section.

- A tensor is considered to be symmetric in some of its indices if we can exchange their order and still obtain the same tensor. As an example, if $S^{\mu\nu}{}_{\sigma\alpha}$ is symmetric in the upper indices then $S^{\mu\nu}{}_{\sigma\alpha} = S^{\nu\mu}{}_{\sigma\alpha}$ [4].

- Similarly a tensor is antisymmetric if when we exchange two indices we obtain the same tensor, but the tensor also changes sign. That is $S^{\mu\nu}{}_{\sigma\alpha} = -S^{\nu\mu}{}_{\sigma\alpha}$.

- Given a tensor which is not symmetric we can create a symmetric tensor from it in any indices we wish. To do so we sum all permutations of the indices we are symmetrizing and divide by the number of terms in the sum. We will make clear which indices we are symmetrizing by surrounding them with round brackets [4]

$$T_{(\mu_1\mu_2\cdots\mu_n)\rho}{}^{\sigma} = \frac{1}{n!}(T_{\mu_1\mu_2\cdots\mu_n\rho}{}^{\sigma} + \text{sum over permutations of indices}). \tag{2.9}$$

- In a similar way we can antisymmetrize a tensor in some of its indices using the same procedure, but alternating the sign of the terms in the sum over the permutations [4]. We denote this by surrounding the relevant indices with square brackets.

$$T_{[\mu_1\mu_2\cdots\mu_n]\rho}{}^{\sigma} = \frac{1}{n!}(T_{\mu_1\mu_2\cdots\mu_n\rho}{}^{\sigma} + \text{alternating sum over permutations of indices}). \tag{2.10}$$

2.3 The Metric Tensor

We now can define a new object on a manifold: the metric tensor $g_{\mu\nu}$. This tensor is the "star" of General Relativity, as we will see in the subsequent discussion. From allowing us to perform calculations on the indices to allowing us to define the curvature of spacetime and eventually appearing in the all important Einstein Equation, the metric is omnipresent [4, 20, 5].

We define the metric as being a type $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ tensor which takes two vectors as an input and gives back a separation between them

$$ds^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu. \tag{2.11}$$

Thus the metric supplies the idea of an inner product between two vectors, as opposed to a vector and a covector [12]

$$g(\vec{A},\vec{B}) = \vec{A}\cdot\vec{B} = g_{\mu\nu}A^\mu B^\nu. \tag{2.12}$$

We can now see how the metric must be used to raise and lower indices: in order for the inner product to work like (2.5) we must have that the metric allows us to raise an index

$$A_\mu = g_{\mu\nu}A^\nu. \tag{2.13}$$

If we now introduce the inverse metric $g^{\mu\nu}$ which is defined to be the tensor such that [4]:

$$g^{\mu\nu}g_{\nu\sigma} = g_{\lambda\sigma}g^{\lambda\mu} = \delta^\mu_\sigma , \tag{2.14}$$

we can then use the inverse metric to lower an index

$$A^\mu = g^{\mu\nu} A_\nu. \tag{2.15}$$

One of the most important features of relativistic theory is that ds^2 must be a relativistic invariant, that is, no matter what choice of coordinates ds^2 must always take the same value for the same spacetime events. We will see this is due to how tensors transform, with vectors transforming in such a way as to cancel the transformation of the matrix.

In the 4 dimensional spacetime of general relativity the metric takes the form of a position dependent symmetric 4 by 4 matrix with one negative and three positive eigenvalues [3]. Explicitly, this is

$$g_{\mu\nu}(x^\sigma) \equiv \begin{pmatrix} g_{00}(x^\sigma) & g_{01}(x^\sigma) & g_{02}(x^\sigma) & g_{03}(x^\sigma) \\ g_{01}(x^\sigma) & g_{11}(x^\sigma) & g_{12}(x^\sigma) & g_{13}(x^\sigma) \\ g_{02}(x^\sigma) & g_{12}(x^\sigma) & g_{22}(x^\sigma) & g_{23}(x^\sigma) \\ g_{03}(x^\sigma) & g_{13}(x^\sigma) & g_{23}(x^\sigma) & g_{33}(x^\sigma) \end{pmatrix}. \tag{2.16}$$

In the special relativistic case where spacetime is flat this reduces to the Minkowski metric

$$\eta_{\mu\nu}(x^\sigma) \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.17}$$

In keeping with the equivalence principle it is also possible to transform the metric at any point such that is it the same as the Minkowski metric $\eta_{\mu\nu}$ and that the first order partial derivatives vanish [4].

2.4   Covariant Differentiation and the Connection

In order to understand how to perform differentiation on the manifold we first need to understand how tensors transform from one coordinate system to another in general.

If we wish to switch from one set of coordinates x^μ to a new set of coordinates $x^{\mu'}$. We define the matrix of transformation between the coordinates for a vector as [5]

$$A^\mu \to A^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} A^\mu. \tag{2.18}$$

For a covector we have [5]

$$A_\mu \to A_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} A_\mu. \tag{2.19}$$

Now we can define the transformation for an arbitrary tensor [4] [4]

$$T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} \to T^{\mu'_1\cdots\mu'_k}{}_{\nu'_1\cdots\nu'_l} = \frac{\partial x^{\mu'_1}}{\partial x^\mu} \cdots \frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} \cdots \frac{\partial x^{\nu_l}}{\partial x^{\nu'_l}} T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}. \tag{2.20}$$

---

[3] We can also take this to be one positive and three negative eigenvalues. The choice between the two is just convention.
[4] We can consider this tensor transformation as another way of defining tensors, that is that they are objects which transform linearly under coordinate transformations.

Now we can consider partial derivatives. The partial derivative operator is a covector, as shown in (2.1), and transforms in the same way as any other covector given in (2.19). However, when we apply the derivative to a vector and then attempt to change coordinates the result does not follow the rules. To see this consider $\partial_{\nu'} A^{\mu'}$ the derivative of a vector field. Use the definition (2.18) and differentiate on both sides [5]

$$
\begin{aligned}
\partial_{\nu'} A^{\mu'} &= \frac{\partial}{\partial x^{\nu'}} \left(\frac{\partial x^{\mu'}}{\partial x^{\rho}} A^{\rho} \right) \\
&= \frac{\partial x^{\lambda}}{\partial x^{\nu'}} \frac{\partial x^{\mu'}}{\partial x^{\rho}} (\partial_{\lambda} A^{\rho}) + \frac{\partial^2 x^{\mu'}}{\partial x^{\nu'} \partial x^{\rho}} A^{\rho}
\end{aligned}
\tag{2.21}
$$

which is not the same as the tensor transformation law given in (2.20).

This discrepancy is due to the position dependent transformation of tensors. Since on an arbitrary manifold the basis vectors of the tangent plane at a point are dependent on the choice of coordinate system, a tensor defined at one point will not necessarily be the same at another point [5].

In order to solve this problem we must introduce a new type of derivative which will transform as we expect under coordinate transformations. This is the covariant derivative. In order to find the covariant derivative we must first list several of the properties we will expect it to have [4].

- Linearity: $\nabla(T + S) = \nabla T + \nabla S$.

- Obeys the product rule: $\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$.

- Commutes with the contraction operation: $\nabla_{\mu}(S^{\lambda}{}_{\lambda\rho}) = (\nabla S)_{\mu}{}^{\lambda}{}_{\lambda\rho}$.

- Like the partial derivative it maps a $\begin{bmatrix} r \\ s \end{bmatrix}$ type tensor to a $\begin{bmatrix} r \\ s+1 \end{bmatrix}$ type tensor.

- For a scalar, the covariant derivative is the same as a the partial derivative $\nabla_{\mu}\varphi = \partial_{\mu}\varphi$.

In order for the operation to follow the product rule we will require it to be the same as the partial derivative plus an additional linear correction term [4]. In order to do this we introduce a new linear term in the derivative, called the connection coefficients, which make the transformation of a derivative occur in the tensorial way. These coefficients are viewed as a matrix for each of the coordinate functions, that is $(\Gamma_{\mu})^{\rho}{}_{\sigma}$ [4]. We will drop the use of the brackets and understand that $\Gamma^{\nu}{}_{\mu\sigma}$ has the same meaning. The covariant derivative of a vector will now be

$$
\nabla_{\mu} V^{\nu} = \partial_{\mu} V^{\nu} + \Gamma^{\nu}{}_{\mu\sigma} V^{\sigma}.
\tag{2.22}
$$

The transformation of this expression is designed to be tensorial, and as such the connection coefficients will not transform in a tensorial way but will transform so as to make the overall transformation tensorial. The transformation of the coefficients will be as follows [4]:

$$
\Gamma^{\nu}{}_{\mu\sigma} \to \Gamma^{\nu'}{}_{\mu'\sigma'} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\sigma}}{\partial x^{\sigma'}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} \Gamma^{\nu}{}_{\mu\sigma} + \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\sigma}}{\partial x^{\sigma'}} \frac{\partial^2 x^{\nu'}}{\partial x^{\mu} \partial x^{\sigma}}.
\tag{2.23}
$$

This will now correct for the second term in the expansion of the partial derivative of a vector, making it tensorial. In order to find the covariant derivative of a covector we first consider the derivative of the scalar $\omega_{\mu} A^{\mu}$ [4]. The definition of the covariant derivative given above is only for vectors, however it is reasonable to assume that the covariant derivative for a covector will be similar.

Thus we introduce a new coefficient, which we will call $\xi^{\sigma}_{\mu\lambda}$ for now, and covariantly differentiate $\omega_{\mu}A^{\mu}$:

$$
\begin{aligned}
\nabla_{\mu}(\omega_{\lambda}A^{\lambda}) &= (\nabla_{\mu}\omega_{\lambda})A^{\lambda} + \omega_{\lambda}(\nabla_{\mu}A^{\lambda}) \\
&= (\partial_{\mu}\omega_{\lambda})A^{\lambda} + \xi^{\sigma}_{\mu\lambda}\omega_{\sigma}A^{\lambda} + \omega_{\lambda}(\partial_{\mu}A^{\lambda}) + \omega_{\lambda}\Gamma^{\lambda}_{\mu\rho}A^{\lambda}.
\end{aligned}
\tag{2.24}
$$

Since this is the covariant derivative of a scalar, we must have that it is equal to the partial derivative. Thus we can also write:

$$
\begin{aligned}
\nabla_{\mu}(\omega_{\lambda}A^{\lambda}) &= \partial_{\mu}(\omega_{\lambda}A^{\lambda}) \\
&= (\partial_{\mu}\omega_{\lambda})A^{\lambda} + \omega_{\lambda}(\partial_{\mu}A^{\mu}).
\end{aligned}
\tag{2.25}
$$

From here we can obviously see that $\xi^{\sigma}_{\mu\lambda} = -\Gamma^{\lambda}_{\mu\rho}$ and thus for a covector the covariant derivative is given by [4]

$$
\nabla_{\mu}\omega_{\nu} = \partial_{\mu}\omega_{\nu} - \Gamma^{\lambda}_{\mu\nu}\omega_{\lambda}.
\tag{2.26}
$$

We can now define the covariant derivative for any tensor, once again by considering it as a product of vectors and covectors, thus we see [4]

$$
\begin{aligned}
\nabla_{\sigma}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} &= \partial_{\sigma}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} + \Gamma^{\mu_1}_{\sigma\lambda}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} + \cdots + \Gamma^{\mu_k}_{\sigma\lambda}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} \\
&\quad - \Gamma^{\lambda}_{\sigma\nu_1}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} - \cdots - \Gamma^{\lambda}_{\sigma\nu_l}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}.
\end{aligned}
\tag{2.27}
$$

Now we can introduce the specific connection we will use in General Relativity, the Levi-Civita connection [4, 20]. With this connection we demand the following things:

- The connection coefficients are torsion free, that is $\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\nu\mu}$.

- Metric compatibility, that is $\nabla_{\mu}g_{\nu\sigma} = 0$.

The Levi-Civita connection is the only connection which will have both such properties. This is known as the fundamental theorem of Riemannian geometry. From here we can derive an expression for the connection coefficients for the Levi-Civita connection in terms of the metric. If we consider the covariant derivative of the metric $\nabla_{\rho}g_{\mu\nu}$ and permute the three indices [4, 5]:

$$
\begin{aligned}
\nabla_{\rho}g_{\mu\nu} &= \partial_{\rho}g_{\mu\nu} - \Gamma^{\lambda}_{\rho\mu}g_{\lambda\nu} - \Gamma^{\lambda}_{\rho\nu}g_{\mu\nu} = 0, \\
\nabla_{\mu}g_{\nu\rho} &= \partial_{\mu}g_{\nu\rho} - \Gamma^{\lambda}_{\mu\nu}g_{\lambda\rho} - \Gamma^{\lambda}_{\mu\rho}g_{\nu\lambda} = 0, \\
\nabla_{\nu}g_{\rho\mu} &= \partial_{\nu}g_{\rho\mu} - \Gamma^{\lambda}_{\nu\rho}g_{\lambda\mu} - \Gamma^{\lambda}_{\nu\mu}g_{\rho\lambda} = 0.
\end{aligned}
\tag{2.28}
$$

Now if we take the first equation and subtract the second and third from it, then multiply by the inverse metric we will obtain the desired expression for the coefficients [4]

$$
\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}g^{\lambda\sigma}\left(\partial_{\mu}g_{\nu\sigma} + \partial_{\nu}g_{\sigma\mu} - \partial_{\sigma}g_{\mu\nu}\right).
\tag{2.29}
$$

This is the expression we will use as the definition of the connection coefficients.

2.5 Parallel Transport

In order to understand the curvature on the manifold we first need a new concept called parallel transport. Parallel transport is the operation of moving a tensor along a path on the manifold in such a way that the covariant derivative of the tensor along the path is zero. That is any changes to the tensor as a result of moving along the curve are due to the structure of space, not due to the tensor itself changing.

If we consider a curve γ on the manifold M given by a parameterization $x^\mu(\lambda)$ with λ an affine parameter [5] and with the tangent vector to the curve of $T^\mu = \frac{dx^\mu}{d\lambda}$ we can define the directional covariant derivative to be [4]

$$\frac{D}{d\lambda} = \frac{dx^\mu}{d\lambda}\nabla_\mu.$$
(2.30)

We now say that a vector V^μ is parallel transported along γ when

$$\frac{DV^\mu}{d\lambda} = 0.$$
(2.31)

This can be rephrased as [4]

$$\frac{d}{d\lambda}V^\mu + \Gamma^\mu_{\sigma\rho}\frac{dx^\sigma}{d\lambda}V^\rho = 0.$$
(2.32)

We are now ready to define one of the key ideas in general relativity: a geodesic path. A geodesic is the shortest path between two points, and brings the idea of a straight line in flat space into the curved spaces we will use in relativity. In keeping with this idea, we will define a geodesic to be a path for which the tangent vector T^μ is parallel transported along the path [4]. That is

$$\frac{D}{d\lambda}T^\mu = \frac{D}{d\lambda}\frac{dx^\mu}{d\lambda} = 0.$$
(2.33)

Inserting this into (2.32) gives the standard geodesic equation:

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0.$$
(2.34)

These results can also be obtained using variational methods [4], however we will not reproduce such proofs here.

2.6 World Lines

A world line is a path γ as measured on the spacetime manifold between events. We can then define the type of the world line by considering what physical phenomena can occur on that path through spacetime. Since the speed of light is the highest attainable speed we will find that spacetime intervals can be divided into cases where normal objects can traverse the distance in spacetime, only light can traverse the distance in spacetime, or no physical phenomena can travserse the distance [5, 12, 4].

We have, for the tangent vector T^μ to a curve γ with parameter λ the following three possibilities:

---

[5]An affine parameter is a parameter which is related to the proper time τ by real coefficients a and b in the form $\lambda = a\tau + b$. We will introduce the proper time soon.

- If $g_{\mu\nu}T^\mu T^\nu = 0$ then the curve is light-like or null. Paths such as these are the paths which light travels along.

- If $g_{\mu\nu}T^\mu T^\nu < 0$ then the curve is time-like. These paths are the paths which objects with mass travel along, at less than the speed of light.

- If $g_{\mu\nu}T^\mu T^\nu > 0$ then the curve is space-like. These paths are impossible for any object to travel along, and thus are completely physically disparate.

If we choose the parameterization of the curve γ such that $g_{\mu\nu}T^\mu T^\nu = -1$ then the parameter is the proper time for an object, τ. This is the time measured by an observer moving on the path γ through spacetime [5].

For a timelike path we can now rewrite the geodesic equation (2.34) in terms of the four-velocity [4] $U^\mu = dx^\mu/d\tau$

$$U^\nu \nabla_\nu U^\mu = 0. \tag{2.35}$$

2.7 Riemann Curvature Tensor

In order to measure the curvature of a manifold we need a new tensor which measures the change of a vector moving on the manifold around a closed loop. We choose a closed loop in order to see how the manifold varies from flat space: if the vector is unchanged by its journey around the loop then we should be in flat space [5, 4].

For an arbitrary vector V^μ we have that the Riemann curvature tensor is defined as [4]

$$(\nabla_\varphi \nabla_\sigma - \nabla_\sigma \nabla_\varphi)V^\mu = R^\mu{}_{\nu\varphi\sigma} V^\nu. \tag{2.36}$$

This definition of the Riemann curvature tensor is only true for connections which have torsion free coefficients [15].

While we will not prove this definition, we can give some motivation for how it works. The covariant derivative effectively measures how much a vector moves away from a path on which it is parallel transported. On flat space, this deviation would be zero, and thus the curvature tensor would be zero. The terms involving the covariant derivative measure the change of a vector compared to its parallel transport in that direction. Thus by considering two different directions we can obtain the loop mentioned above. This also gives the condition that in flat space the tensor should be zero, as $\nabla_\varphi \nabla_\sigma - \nabla_\sigma \nabla_\varphi = \partial_\varphi \partial_\sigma - \partial_\sigma \partial_\varphi = 0$ in flat space [5, 4, 23].

If we now perform the computation in (2.36) with an arbitrary vector V^λ we can obtain an expression for the Riemann curvature tensor in terms of the connection coefficients [4, 23]. First consider $\nabla_\mu \nabla_\nu V^\lambda$

$$
\begin{aligned}
\nabla_\mu \nabla_\nu V^\lambda &= \nabla_\mu(\partial_\nu V^\lambda + \Gamma^\lambda_{\nu\sigma} V^\sigma) \\
&= \partial_\mu \partial_\nu V^\lambda + (\partial_\mu \Gamma^\lambda_{\nu\sigma}) + \Gamma^\lambda_{\nu\sigma} \partial_\mu V^\sigma + \Gamma^\lambda_{\mu\varphi}(\partial_\nu V^\varphi + \Gamma^\varphi_{\nu\sigma} V^\sigma) - \Gamma^\sigma_{\mu\nu}(\partial_\sigma V^\lambda + \Gamma^\lambda_{\varphi\sigma})
\end{aligned} \tag{2.37}
$$

now by exchanging μ and ν and using the properties of the connection we can find

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu)V^\lambda = (\partial_\mu \Gamma^\lambda_{\nu\sigma} - \partial_\nu \Gamma^\lambda_{\mu\sigma} + \Gamma^\lambda_{\mu\lambda}\Gamma^\varphi_{\nu\sigma} - \Gamma^\lambda_{\nu\lambda}\Gamma^\varphi_{\mu\sigma})V^\sigma. \tag{2.38}$$

Thus we can give an expression for the Riemann curvature tensor in terms of the connection coefficients. This is the expression most commonly used to calculate the tensor

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_{\mu}\Gamma^{\rho}{}_{\nu\sigma} - \partial_{\nu}\Gamma^{\rho}{}_{\mu\sigma} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\mu\sigma}. \tag{2.39}$$

2.7.1 Properties of the Riemann Curvature Tensor

We will list some of the important properties of the Riemann curvature tensor. These are easier to see in the version of the tensor with 4 lower indices

$$R_{\rho\sigma\mu\nu} = g_{\rho\lambda}R^{\lambda}{}_{\sigma\mu\nu}. \tag{2.40}$$

- Antisymmetry in the first two indices $R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$ and in the last two indices $R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}$.

- Symmetric when exchanging the first pair of indices with the second $R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$.

- The sum of the cyclic permutations of the last three indices is zero $R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0$

From these properties it is possible to show that in n dimensional space the Riemann curvature tensor must have N independent components [5], where N is

$$N = \frac{1}{12}n^2(n^2 - 1). \tag{2.41}$$

Remark 2.7.1. *When $n = 1$ we have that $N = 0$ and thus there can be no curvature for a line. For a surface however, with $n = 2$, we find that $N = 1$. This single independent component is related to the Gaussian curvature [5, 13] of a surface from classical differential geometry. We have that*

$$K = -\frac{R_{1212}}{|g_{\mu\nu}|} \tag{2.42}$$

with $|g_{\mu\nu}|$ being the determinant of the metric tensor and K the Gaussian curvature. With $n = 4$ we have 20 independent components of the Riemann curvature tensor.

Finally another important property is the Bianchi identity

$$\nabla_{\omega}R_{\rho\sigma\mu\nu} + \nabla_{\rho}R_{\sigma\omega\mu\nu} + \nabla_{\sigma}R_{\omega\rho\mu\nu} = 0. \tag{2.43}$$

We can also write this by using the antisymmetry properties of $R_{\rho\sigma\mu\nu}$ as

$$\nabla_{[\omega}R_{\rho\sigma]\mu\nu} = 0. \tag{2.44}$$

2.7.2 Ricci Tensor

If we now contract the first and third indices of the Riemann curvature tensor we can form the Ricci tensor [4]

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}. \tag{2.45}$$

Due to our choice of the connection this is the only contraction of the Riemann curvature tensor which yields interesting results. All other contractions are either zero or related to the Ricci tensor [4]. The Ricci tensor is also symmetric in its indices.

2.7.3 Ricci Scalar

If we now raise one of the indices of the Ricci tensor and then contract over the indices we will obtain the Ricci scalar

$$R = R^{\mu}{}_{\mu} = g^{\mu\nu}R_{\mu\nu}. \tag{2.46}$$

2.7.4 Einstein Tensor

We now can derive the Einstein tensor, which is the tensor we will use to describe the curvature of space in the Einstein equation. If we take the Bianchi identity (2.43) and contract twice we will obtain a similar identity in terms of the Ricci tensor and scalar

$$g^{\nu\sigma}g^{\mu\lambda}(\nabla_{\lambda}R_{\rho\sigma\mu\nu} + \nabla_{\rho}R_{\sigma\lambda\mu\nu} + \nabla_{\sigma}R_{\lambda\rho\mu\nu}) = 0,$$
$$\nabla^{\mu}R_{\rho\mu} - \nabla_{\rho}R + \nabla^{\nu}R_{\rho\nu} = 0. \tag{2.47}$$

Now if we define the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \tag{2.48}$$

we can see that the contracted Bianchi identity must imply that [5, 4]

$$\nabla^{\mu}G_{\mu\nu} = 0. \tag{2.49}$$

2.8 The Energy Momentum Tensor

The fundamental idea of General Relativity is that the curvature of spacetime is caused by the nearby mass and energy at that point [12, 4]. We now have the mathematical description of how the curvature of space is measured. Thus, we now need to develop an object to measure the amount of mass and energy near a point. This is the energy momentum tensor.

The energy momentum tensor is a type $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ symmetric tensor which measures the flux of four momentum p^{μ} across a surface keeping one coordinate constant [4]. In 3 dimensional flat space this would represent the forces pointing in three orthogonal directions along three orthogonal planes at a point [5].

In 4 dimensional general relativity we have to take into account energy as well, as due to the classic equation from special relativity for the rest energy of matter

$$E = mc^2 \tag{2.50}$$

we can view energy as having a mass. Thus in the 4 dimensional case we take the energy momentum tensor to be in matrix form

$$T^{\mu\nu} = \begin{pmatrix} T^{00} & T^{01} & T^{02} & T^{03} \\ T^{10} & T^{11} & T^{12} & T^{13} \\ T^{20} & T^{21} & T^{22} & T^{23} \\ T^{30} & T^{31} & T^{32} & T^{33} \end{pmatrix}. \tag{2.51}$$

We can now identify the components of this equation [12, 4]

- T^{00} is the energy density at the point, which we will call ρ.

- $T^{0i} = T^{i0}$ for the three spatial indices is the momentum density at the point.

- Terms on the diagonal T^{ii} give the spatial components of the pressure p_i.

- All other off diagonal terms represent the sheering terms, that is forces on the planes not in the direction of the plane. These are created due to things like viscosities.

An example of this, and the one used most in modeling physical situations in general relativity is of the perfect fluid, which we will take to have an energy momentum tensor of [4, 12]

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{2.52}$$

in the rest frame of the fluid. We now give a tensorial version of this equation which is valid in any frame [4]

$$T^{\mu\nu} = (\rho + p)U^{\mu}U^{\nu} + p g^{\mu\nu} \tag{2.53}$$

where we define U^{μ} to be the constant four velocity of the particles within the fluid.

Another important property of the energy momentum tensor is that it must be conserved, that is

$$\nabla_{\mu} T^{\mu\nu} = 0. \tag{2.54}$$

This is related to the conservation of energy and momentum in an inertial frame [4, 12].

2.9   The Einstein Equation

At its core, general relativity is about relating the curvature of spacetime to the matter distribution in space. This is illustrated by the Einstein equation [4, 12]

$$G_{\mu\nu} = k T_{\mu\nu} \tag{2.55}$$

for some constant k which we will soon find. The Einstein equation is the field equation for general relativity, describing exactly how the curvature of spacetime happens due to the presence of mass.

Our choice of the Einstein tensor in this equation is due to requiring a tensor based on the curvature but always having a zero covariant derivative. We see the latter from (2.49), and the former is due to derivation of the Einstein tensor in terms of the Riemann curvature tensor and eventually the metric.

In order to find the constant k of proportionality between the two tensors we recall the field equation for Newtonian gravity

$$\nabla^2 \Phi = 4\pi G \rho. \tag{2.56}$$

In the case of a weak gravitational field, we must have that Einstein's gravitational theory reduces to the Newtonian theory.

If we contract (2.55) on both sides we will obtain [4, 5]

$$R = -kT \tag{2.57}$$

which we can use to rewrite (2.55) as

$$R_{\mu\nu} = k \left(T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right). \tag{2.58}$$

We now need to see how to recover Newton's theory of gravitation for weak gravitational fields and slowly moving objects. In a weak field case, we will be able to assume that spacetime is mostly flat like 4 dimensional Minkowski space and the metric will be $\eta_{\mu\nu}$ other than a small perturbation, which we will call $h_{\mu\nu}$ [4, 23]. We can now decompose a weak field metric into components

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{2.59}$$

with $|h_{\mu\nu}| << 1$. Due to taking the determinant of $h_{\mu\nu}$ to be very small, we will ignore products of $h_{\mu\nu}$ with itself. Thus to first order we can find the inverse metric

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \tag{2.60}$$

where we now use the Minkowski metric to raise and lower indices.

If we also assume that any objects with mass are moving at much less than the speed of light, then we can say for the three spacial coordinates x^i and the time coordinate t that in terms of the proper time τ [4]

$$\frac{dx^i}{d\tau} << \frac{dt}{d\tau}. \tag{2.61}$$

As such, we can simplify the geodesic equation by assuming that terms involving the slowly changing spatial coordinates do not make a large difference to the path, and write

$$\frac{dx^\mu}{d\tau^2} + \Gamma^\mu_{00} \left(\frac{dt}{d\tau} \right)^2 = 0. \tag{2.62}$$

In Newtonian gravity we would also have that the field is time independent, thus, we can take that $\partial_0 g_{\mu\nu} = 0$ and therefore compute the connection coefficients for this weak field. Due to (2.62) we only need calculate the connection coefficent corresponding to the time coordinate

$$\begin{aligned} \Gamma^\mu_{00} &= \frac{1}{2} g^{\mu\lambda} (\partial_0 g_{\lambda 0} + \partial_0 g_{0\lambda} - \partial_\lambda g_{00}) \\ &= -\frac{1}{2} g^{\mu\lambda} \partial_\lambda g_{00}. \end{aligned} \tag{2.63}$$

In the case of (2.59) we have, to first order in $h_{\mu\nu}$ [4]

$$\Gamma^{\mu}_{00} = -\frac{1}{2}\eta^{\mu\lambda}\partial_{\lambda}h_{00} \tag{2.64}$$

and thus we can rewrite the geodesic equation (2.62)

$$\frac{d^2 x^{\mu}}{d\tau^2} = \frac{1}{2}\eta^{\mu\lambda}\partial_{\lambda}h_{00}\left(\frac{dt}{d\tau}\right)^2. \tag{2.65}$$

If we now recall the special relativistic result that $\frac{dt}{d\tau} = \gamma$ then in the slow moving frame $\gamma \approx 1$, and thus we can use the time coordinate t, as t and τ will coincide. Therefore

$$\frac{d^2 x^{\mu}}{dt^2} = \frac{1}{2}\eta^{\mu\lambda}\partial_{\lambda}h_{00}. \tag{2.66}$$

If we now consider just the spatial coordinates, then $\eta^{\mu\nu} = \delta^{\mu}_{\nu}$ for 3 dimensional flat space. Thus, we can write, with i from 1 to 3

$$\frac{d^2 x^i}{dt^2} = \frac{1}{2}\partial_i h_{00}. \tag{2.67}$$

Now if we recall that the acceleration due to a gravitational potential Φ in Newtonian gravity is $a = -\nabla\Phi$ then we see that (2.67) is in the same form, if we choose $h_{00} = -2\Phi$. This is then equivalent to choosing the first coefficient of the metric as [4]

$$g_{00} = -(1 + 2\Phi). \tag{2.68}$$

Finally, let us consider a collection of dust, that is a collection of particles for which the energy momentum tensor is $T_{\mu\nu} = \rho U_{\mu}U_{\nu}$ (this is the case of (2.53) without pressure) and we consider the inertial frame where $U^{\mu} = (1, 0, 0, 0)$. Then we will have that [4, 5]

$$T_{00} = \rho \tag{2.69}$$

and all other coefficients will be zero. The trace of this energy momentum will be $T = g^{00}T_{00} = -\rho$ and thus we can insert these into (2.58) to gain an approximation for k.

$$\begin{aligned} R_{00} &= k\left(T_{00} - \frac{1}{2}Tg_{00}\right) \\ &= \frac{1}{2}k\rho \end{aligned} \tag{2.70}$$

since we are taking that $g_{00} \approx \eta_{00} = -1$ in the Newtonian limit. If we now consider the definition of the Ricci tensor in terms of the Riemann curvature tensor we can obtain an approximate expression in terms of the metric and its perturbation as in (2.59)

$$R_{00} = R^i_{\ 0j0} = \partial_j\Gamma^i_{00} - \partial_0\Gamma^i_{j0} + \Gamma^i_{j\lambda}\Gamma^{\lambda}_{00} - \Gamma^i_{0\lambda}\Gamma^{\lambda}_{j0} \tag{2.71}$$

with i and j running from 1 to 3 [4]. If we now note that products of Γ are products of the metric perturbation and are therefore very small we can ignore these in the Newtonian limit. Also, the

time derivative ∂_0 is automatically zero due to the static field. Thus we must have that the Riemann curvature tensor in the form $R^i{}_{0j0} = \partial_j \Gamma^i_{00}$ and if we now use (2.64) we can calculate

$$
\begin{aligned}
R_{00} &= -\frac{1}{2} \delta^{ij} \partial_i \partial_j h_{00} \\
&= -\frac{1}{2} \nabla^2 h_{00}.
\end{aligned}
\tag{2.72}
$$

Thus we can write

$$
\nabla^2 h_{00} = -k\rho.
\tag{2.73}
$$

However since we set $h_{00} = -2\Phi$ as in (2.68) we recover Poisson's equation (2.56) when we set $k = 8\pi G$. Thus, we can give the correct expression for the Einstein equation

$$
R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}.
\tag{2.74}
$$

In the vacuum case where $T_{\mu\nu} = 0$ we have using (2.58)

$$
R_{\mu\nu} = 0
\tag{2.75}
$$

which is known as the vacuum Einstein equation.

The Einstein equation presents us with 10 partial differential equations for the functions of the metric and their derivatives given a distribution of matter in a region. There is a slight simplifying constraint in the Bianchi identity, $\nabla^\mu G_{\mu\nu} = 0$, which reduces the number of independent equations to 6. The reduction of the constraints by 4 is due to the metric coefficients needing to be invariant under coordinate transformations for the 4 coordinates in spacetime [4].

These equations are generally incredibly difficult to solve, and any analytical solutions will inevitably make use of symmetry properties of the physical situation at hand. We will now look at one such very useful situation with a large degree of symmetry, that is a non rotating spherical uniform mass. While this situation is very unlikely ever to exist in nature the results it obtains are perfect for classical tests of general relativity and we will use it for this purpose. The metric for this situation is called the Schwarzschild metric.

2.10 Schwarzschild Metric

As stated above, the Schwarzschild metric is for spherically symmetric situations, and as such we will choose the most logical coordinate system for such a situation, polar coordinates $\{t, r, \theta, \varphi\}$. In polar coordinates the metric is as follows [4]

$$
ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2
\tag{2.76}
$$

with $d\Omega^2 = d\theta^2 + \sin^2\theta \, d\varphi^2$, M the total mass of the object causing the gravitational field and G Newton's gravitational constant.

Here we must be careful to note that the radial coordinate r is not the distance from the origin of the coordinate system to a point, instead it is related to the area of a 2-sphere on the manifold around the gravitating body. Since the curvature of spacetime will not be flat for the mass distribution modeled

by the Schwarzschild metric, the distance that a radial coordinate would measure between spheres would not be the same as the distance that a physical observer would measure. Instead, if we use that [12]

$$r = \left(\frac{A}{4\pi} \right)^{\frac{1}{2}} \tag{2.77}$$

with A the area of a sphere then an observer will measure (for constant t, θ, and φ) the infinitesimal distance between two spheres to be

$$ds = \frac{dr}{\sqrt{1 - \frac{2GM}{rc^2}}} \tag{2.78}$$

where we introduce the factor of c^2 to give the units in meters. Here we see that as we get further from the source of gravity the measured radii and coordinate radii become closer to being equal.

The factor of $2GM$ is defined to be the Schwarzschild radius. Bodies for whom all their mass is contained within this radius will collapse to form a black hole, although we will not look at such things in this project [5, 4].

This metric is the unique solution of Einstein's equations in a vacuum, for a spherically symmetric gravitational field. This result is called Birkhoff's theorem, a proof of which can be found in [4].

3 | Gravitational Waves

In Newton's gravitational theory the gravitational field is instantly established by the presence of mass. In Einstein's theory however, this is no longer the case; gravitation can only move at at most the speed of light, like every other physical phenomena or force. This then begs the question of how do changes in the gravitational force propagate to effect objects in spacetime. The answer is that changes in a gravitational field are transmitted through gravitational waves.

In the previous chapter we used the weak field approximation to derive the constant in the Einstein equation. If we relax the condition that the fields must be time independent this formalism has another use as the method of visualizing gravitational wave emission. In this chapter we will show how such emission works, following closely the discussion in Cheng's book [5].

3.1 Linearized Gravitation

The idea of a gravitational wave is strongly influenced by electromagnetic waves within classical electrodynamics.

Recall from chapter two the equation for the metric in flat space with a perturbation (2.59)

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \tag{3.1}$$

We can view this equation as being a tensor field propagating on flat Minkowski space [5] at the speed of light. Similar to before we can define the Riemann curvature tensor to be (dropping higher order terms of $h_{\mu\nu}$) [5]

$$R_{\alpha\mu\beta\nu} = \frac{1}{2}(\partial_\alpha\partial_\nu h_{\mu\beta} + \partial_\mu\partial_\beta h_{\alpha\nu} - \partial_\alpha\partial_\beta h_{\mu\nu} - \partial_\mu\partial_\nu h_{\alpha\beta}) \tag{3.2}$$

the Ricci tensor to be

$$R_{\mu\nu} = \eta^{\alpha\beta}R_{\alpha\mu\beta\nu} = \frac{1}{2}(\partial_\alpha\partial_\nu h^\alpha{}_\mu + \partial_\mu\partial_\alpha h^\alpha{}_\nu - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h) \tag{3.3}$$

and the Ricci scalar to be

$$R = \partial_\mu\partial_\nu h^{\mu\nu} - \Box h \tag{3.4}$$

where we take $\Box = \partial_\mu\partial^\mu$ to be the d'Alembert operator and h the trace of the metric perturbation. Finally we can write the Einstein tensor [5, 4]

$$\begin{aligned}
G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}R \\
&= \frac{1}{2}(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\nu - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial_\rho\partial_\lambda h^{\rho\lambda} + \eta_{\mu\nu}\Box h).
\end{aligned} \tag{3.5}$$

We note that the Einstein tensor is linear in $h_{\mu\nu}$ as we would expect.

3.2 Gravitational Waves

Now, much as in classical electrodynamics we will choose the gauge transformation for the fields to be the Lorentz gauge. This gauge is guarantied to be invariant under changes of frames, due to the Lorentz invariance of the gauge condition [10].

If we recall in electrodynamics using the potential formulation, a potential given by $\vec{A}, V$ was unchanged under the following transformations with a scalar function λ [10]

$$\vec{A}' = \vec{A} + \nabla \lambda$$
$$V' = V - \frac{\partial \lambda}{\partial t}, \tag{3.6}$$

that is $\vec{A}', V'$ still represent the same physical situation. Such transformations are called gauge transformations. We can obtain a similar result for $h_{\mu\nu}$ by realizing that the coordinate invariant nature of relativity we desire means that we should be able to make small changes to the coordinates. These changes should keep $\eta_{\mu\nu}$ the same, but not $h_{\mu\nu}$. Mathematically, if we consider that the coordinates are shifted by a small amount $x^{\mu'} = x^\mu + \chi^\mu(x)$ where $\chi^\mu(x)$ are four arbitrary functions similar in magnitude to the metric perturbation [12], then we will obtain this condition. This is much like the situation with λ in (3.6). Now by considering the transformation of the metric as per (2.20) we can derive the gauge transformation for the $h_{\mu\nu}$ field [5]

$$h_{\mu'\nu'} = h_{\mu\nu} - \partial_\mu \chi_\nu - \partial_\nu \chi_\mu. \tag{3.7}$$

By choosing the gauge to be the Lorentz gauge, in same way as in electrodynamics, we can simplify the resultant expressions for the Ricci tensor and scalar to obtain a wave equation for the perturbation of the metric. The Lorentz gauge is

$$\partial^\mu \bar{h}_{\mu\nu} = 0 \tag{3.8}$$

with $\bar{h} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$ the trace reversed perturbation. This then simplifies the Ricci tensor to

$$R_{\mu\nu} = -\frac{1}{2}\Box h_{\mu\nu} \tag{3.9}$$

and the Ricci scalar to

$$R = -\frac{1}{2}\Box h. \tag{3.10}$$

Finally, in this way we can write the Einstein equation (2.74) as a wave equation in $\bar{h}_{\mu\nu}$

$$\Box \bar{h}_{\mu\nu} = 16\pi G T_{\mu\nu}. \tag{3.11}$$

The solution in terms of a retarded field is similar to electrodynamics [5]

$$\bar{h}_{\mu\nu}(\vec{x}, t) = 4G \int d^3\vec{x}' \frac{T_{\mu\nu}(\vec{x}', t - |\vec{x} - \vec{x}'|/c)}{|\vec{x} - \vec{x}'|}. \tag{3.12}$$

In order to simplify the discussion of the propagation of the waves we will take the energy momentum tensor to be zero for the rest of this chapter. This leaves us with the linearized Einstein equation in the vacuum case

$$\Box \bar{h}_{\mu\nu} = 0. \tag{3.13}$$

In this case we can see that, because $\Box \eta_{\mu\nu} = 0$,

$$\Box h_{\mu\nu} = 0. \tag{3.14}$$

Since we have that the metric perturbation field satisfies the wave equation in the Lorentz gauge, we can view the perturbations as being waves spreading out from a source of strong gravity. Thus we can consider $h_{\mu\nu}$ to be a plane wave solution [5, 12]

$$h_{\mu\nu}(x) = \epsilon_{\mu\nu} e^{ik_\alpha x^\alpha} \tag{3.15}$$

with $\epsilon_{\mu\nu}$ being a 4 by 4 symmetric matrix with constant entries which describes the amplitude of the components of the wave and $k^\alpha = (\omega, \vec{k})$ a 4-wavevector. Now by substituting into (3.14) we find that $k^2 \epsilon_{\mu\nu} e^{ik_\alpha x^\alpha} = 0$ which then implies [4, 12]

$$k^2 = k_\alpha k^\alpha = -\omega^2 + \vec{k}^2 = 0 \tag{3.16}$$

meaning that, in order for plane waves to solve (3.14) their wave vector must be a null vector, traveling at c.

Finally, due to the conditions given above we also must have that [5]

$$k^\mu \epsilon_{\mu\nu} = 0 \tag{3.17}$$

which is the condition that the waves are transverse.

3.3   Transverse Traceless Gauge

Even after applying the gauge condition of (3.8) we still have some freedom to choose coordinates. If we take that the gauge functions χ_μ are also constrained by [5]

$$\Box \chi_\mu = 0 \tag{3.18}$$

then we can simplify the polarization tensor such that it is traceless [5]

$$\epsilon^\mu{}_\mu = 0 \tag{3.19}$$

and such that

$$\epsilon_{0\mu} = \epsilon_{\mu 0} = 0. \tag{3.20}$$

This condition on the gauge functions is called the transverse traceless gauge. In this gauge the gravitational wave will have 2 polarization states, which we will show now.

Before we imposed the transverse traceless gauge condition $\epsilon_{\mu\nu}$ had 10 independent components, however (3.17), (3.19) and (3.20) fix 8 of these components, leaving 2 possible choices of polarization states. If we consider a wave propagating along the z direction $k^\mu = (\omega, 0, 0, \omega)$ then using the conditions above we can see that in general the metric perturbation will have the form [5, 12]

$$h_{\mu\nu}(z,t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{i\omega(z-t)}. \tag{3.21}$$

And thus the two polarization states will be

$$\epsilon^{(+)}_{\mu\nu} = h_+ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.22}$$

and

$$\epsilon^{(\times)}_{\mu\nu} = h_\times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.23}$$

with h_+ and $h_\times$ being the "plus" and "cross" amplitudes, respectively.

3.4   Effect of Gravitational Waves on Particles

In order to observe the effects of gravitational waves we need to understand how they will affect a distribution of particles. Due to the equivalence principle, we must have that single particles cannot detect the effects of a gravitational wave on their own, as in an inertial frame we should be able to transform the effects of gravity away. Thus, we need to have at least two particles and need to consider their separation to see any gravitational wave effects.

Consider two particles, one at the origin and another a small distance from it on the x axis $(0, \xi, 0, 0)$. If a "plus" polarized gravitational wave passes between them they will observe a distance change of [5]

$$\begin{aligned} ds &= \sqrt{g_{\mu\nu} dx^\mu dx^\nu} = \sqrt{g_{11}}\xi \\ &\approx \left[\eta_{11} + \frac{1}{2}h_{11} \right]\xi = \left[1 + \frac{1}{2}h_+ \sin\omega(t-z) \right]\xi. \end{aligned} \tag{3.24}$$

If we consider the same setup but with the other particle on the y axis $(0, 0, \xi, 0)$ then we will find

$$ds = \left[1 - \frac{1}{2}h_+ \sin\omega(t-z) \right]\xi \tag{3.25}$$

and thus we see that there is an elongation along the x axis and a shortening along the y axis. The effects of a "cross" polarized wave are the same, except at 45° to the "plus" polarized case.

4 | Classical Tests of General Relativity

Einstein's theories of relativity make bold claims about the nature of reality, and after they were published physicists were immediately interested in verifying the claims made by the theory experimentally. In this chapter we will present the three "classical" tests of general relativity, which were the first tests to be done after Einstein published his theories. These observations are made on bodies within our solar system, and mostly use the Schwarzschild metric obtained at the end of the first chapter.

4.1   The Precession of Mercury's Perihelion

Celestial mechanics based on Newtonian gravitational theory had been incredibly successful in describing the motion of the planet's within the solar system. There was one exception to this however: the precession of Mercury's perihelion. The perihelion is defined to be the point of a planets orbit within an ellipse which is closest to the sun [5]. Over time, the axis of the ellipse of the orbit moves. This is mostly due to the gravitational effects of the other planets in the solar system. However, when such effects are taken into account, there is still a precession of 43 arcseconds per century.

   This effect had been observed around about 1850 by astronomers, and many had tried and failed to explain this additional precession of Mercury's orbit; for example by postulating the existence of a smaller planet between Mercury and the sun. It was only after general relativity had been formulated that we can explain this additional contribution to the precession [5].

   In order to derive this effect we require the orbital equations derived from the Schwarzschild metric (2.76) using Lagrangian mechanics. We choose the Lagrangian for a particle in a gravitational field modeled by the Schwarzschild metric to be [5]

$$\mathscr{L} = \left(\frac{ds}{d\tau}\right)^2 = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \tag{4.1}$$

where we choose τ to be the proper time and $\dot{x}$ is the derivative with respect to the proper time. We recall from classical mechanics that a constant of motion will be the case when [11]

$$\frac{\partial \mathscr{L}}{\partial x^i} = 0 \tag{4.2}$$

for a particular coordinate x^i. In the case of the Schwarzschild metric we apply these to t and to φ [12]

$$\frac{d}{d\tau}\left(\left(1 - \frac{2GM}{r}\right)\frac{dt}{d\tau}\right) = 0$$
$$\frac{d}{d\tau}\left(r^2\frac{d\varphi}{d\tau}\right) = 0 \tag{4.3}$$

and find that those coordinates are constant in time. For the time coordinate, this corresponds to conservation of energy and we will give this constant the symbol E [12, 4]. For the φ coordinate this corresponds to conservation of angular momentum [12] which we will call L.

By recalling from special relativity that $ds^2 = -d\tau^2$ we can write the Lagrangian in the Schwarzschild case for massive particles as [5]

$$\mathscr{L} = -\left(1 - \frac{2GM}{r}\right)\dot{t}^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\varphi}^2 = -1 \tag{4.4}$$

where we set $\theta = \frac{\pi}{2}$ as the orbits will always stay in that plane.

Now, by putting E and L into the Lagrangian (4.4) we can obtain

$$-\frac{E^2}{1 - 2GM/r} + \frac{\dot{r}^2}{1 - 2GM/r} + \frac{L^2}{m^2 r^2} = -1 \tag{4.5}$$

with m the mass of the particle. Then, if we multiply by $\frac{1}{2}m(1 - 2GM/r)$ and rearrange we find

$$\frac{1}{2}m\dot{r}^2 + \left(1 - \frac{2GM}{r}\right)\frac{L^2}{2mr^2} - \frac{GmM}{r} = \frac{mE^2}{2} - \frac{1}{2}m \tag{4.6}$$

Now, if we define a new constant $\mathscr{E}$ to be

$$\mathscr{E} = m\frac{E^2 - 1}{2} \tag{4.7}$$

we can rewrite (4.6) to be in the form of a Newtonian energy balance equation [5]

$$\frac{1}{2}m\dot{r}^2 + \left(1 - \frac{2GM}{r}\right)\frac{L^2}{2mr^2} - \frac{GmM}{r} = \mathscr{E}. \tag{4.8}$$

At this point it is possible to formulate an orbital equation for Mercury in terms of (4.8) however we will not do so, and simply state the result. If we define e to be the eccentricity of the orbit, $\alpha = L^2/GMm^2 = (1 + e)r_{\text{min}}$ with r_{min} the radial distance at the perihelion and $\epsilon = 6GM/2\alpha$ then the radius of the planets orbit will be given by

$$r = \frac{\alpha}{1 + e\cos((1 - \epsilon)\varphi)}. \tag{4.9}$$

We can see that this will not return to the same point after 1 orbit (when $\varphi = 2\pi$), but after slightly greater than one orbit (when $\varphi = 2\pi/(1 - \epsilon)$). The entire orbit will advance per revolution by

$$\delta\varphi = \frac{6\pi GM}{(1 - \epsilon)r_{\text{min}}}. \tag{4.10}$$

We now input the values we know for the quantities: the Schwarzschild radius of the sun is $2GM = 2.95\text{km}$, Mercury's eccentricity is $e = 0.206$ and the perihelion is at $r_{\text{min}} = 4.6 \times 10^7\text{km}$ [5] then we can get the numerical value of the advance per revolution as

$$\delta\varphi = 5 \times 10^{-7} \text{ radians per revolution} \tag{4.11}$$

which we then can convert to arcseconds per century to get 43 arcseconds per century, in perfect agreement with observations.

4.2 Deflection of Light by the Sun

In 1919, Arthur Eddington observed a star on the edge of the sun as a solar eclipse was occurring. This star was actually behind the sun, and the light from the star was being bent by the gravitational field of the sun. This provided the second experimental test of general relativity, the effect of gravity on light rays.

Recall from chapter 2 that light travels on a light like world line, with $ds^2 = 0$. If we choose this direction to be along the coordinates such that $\varphi = \theta = 0$ then the metric will take the form [5]

$$ds^2 = g_{00}c^2\mathrm{d}t^2 + g_{rr}\mathrm{d}r^2 = 0 \tag{4.12}$$

where we include the speed of light term for clarity in the following steps.

From this metric we can find the effective speed of light that an observer sees [5]

$$c(r) = \frac{dr}{dt} = c\sqrt{-\frac{g_{00}(r)}{g_{rr}(r)}}. \tag{4.13}$$

If we now note from the Schwarzschild metric (2.76) that $g_{rr} = -(g_{00})^{-1}$ then we can define an index of refraction for the observer by

$$n(r) = \frac{c}{c(r)} = \sqrt{-\frac{g_{rr}(r)}{g_{00}(r)}} = \frac{1}{-g_{00}(r)} = \left(1 - \frac{2GM}{rc^2}\right)^{-1}. \tag{4.14}$$

If we now consider a light wave propagating along the x axis with the gravitational field along the y axis then a section of the wave front which is closer to the source will travel more slowly than a section further from the source. The amount of this bending, given in terms of a deflection angle from the straight path the light would take without gravity, is given by [5]

$$\delta\varphi = \frac{4GM}{c^2 r_{\min}} \tag{4.15}$$

where in this case M is the solar mass and $r_{\min}$ is the closest that the light comes to the large object bending it.

For the test that Eddington performed this resulted in a deflection of 1.74 arcseconds, which is exactly what was observed.

A more modern test of this phenomena is looking for gravitational lensing of very distant objects by massive objects in between. A lens equation for light bending is

$$\beta = \theta - \frac{D_{ls}}{D_s D_l}\frac{4GM}{\theta c^2} \tag{4.16}$$

with β the angle from the real object to the observer, θ the angle from the image to the observer, D_{sl} the distance from the source to the mass causing deflection, D_s the distance from source to observer and D_l the distance from the mass to the observer.

Such events produce Einstein rings in the special case when $\beta = 0$, that is the source, the massive object acting as a lens and the observer are all aligned. The angular radius in this case will be

$$r_A = \sqrt{\frac{D_{ls}}{D_s D_l}\frac{4GM}{c^2}}. \tag{4.17}$$

Thus we see how astronomers can use gravitational lensing observations to "weigh" objects which act as the lens, such as a blackhole which lenses an object at a known distance.
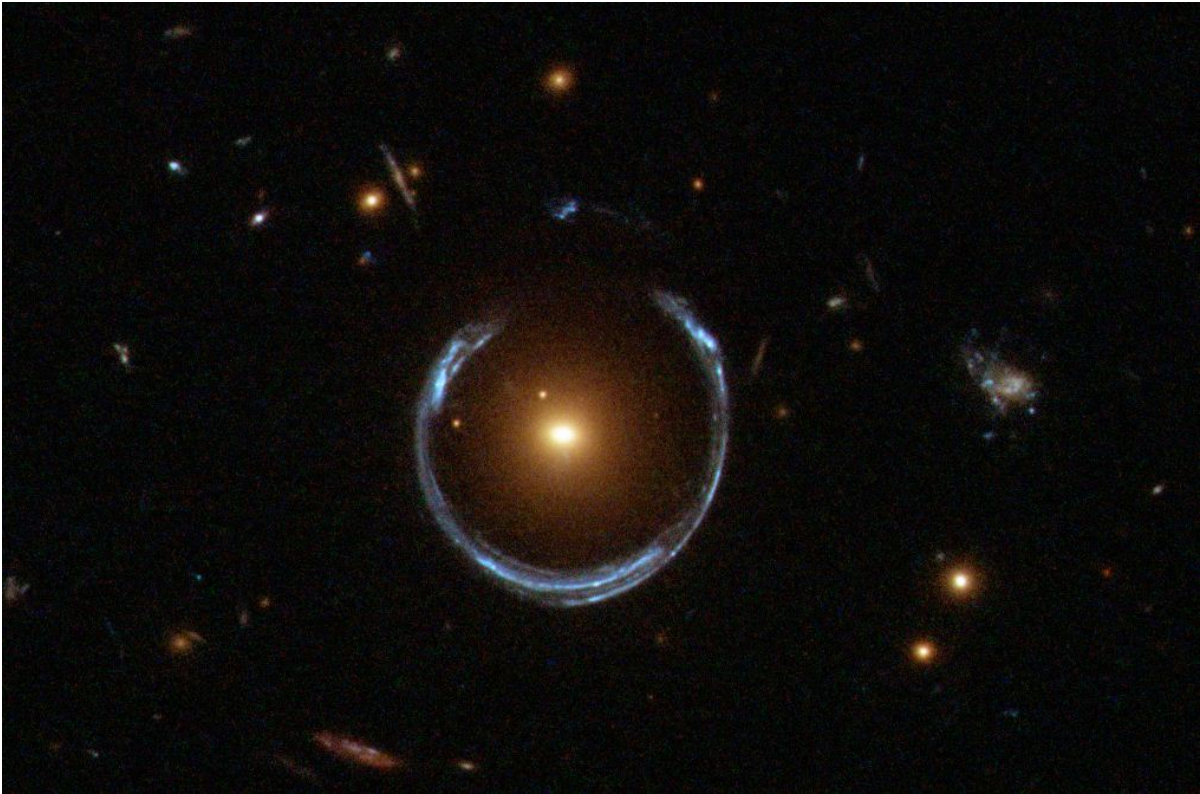
Figure 4.1: A picture of an Einstein ring taken by the Hubble space telescope. Light from the blue galaxy is bent by the red galaxy in the foreground.

4.3 Gravitational Redshift

Gravitational redshift is an effect where light emitted closer to a large object will appear redshifted when viewed further from that object. It is caused by gravitational time dilation, as frequency is a time based phenomena.

In order to derive the effect, consider an observer who is stationary in a Schwarzschild spacetime. For this observer $U^i = 0$ that is, the spatial part of the four velocity is zero. Thus, because for objects with mass we must have $U_\mu U^\mu = -1$ we can get U^0 [4]

$$U^0 = \left(1 - \frac{2GM}{r}\right)^{-\frac{1}{2}}. \tag{4.18}$$

This observer will measure the frequency of the photon along a null geodesic as

$$\omega = -g_{\mu\nu} U^\mu \frac{dx^\nu}{d\lambda}. \tag{4.19}$$

Now in terms of U^0 this relation is [4]

$$\omega = \left(1 - \frac{2GM}{r}\right)^{\frac{1}{2}} \frac{dt}{d\lambda}. \tag{4.20}$$

If we now substitute for the definition of E from (4.3) then we will find

$$\omega = \left(1 - \frac{2GM}{r}\right)^{-\frac{1}{2}} E \tag{4.21}$$

where we take E to be defined in terms of the affine parameter λ for light. From this we can see that ω will take different values depending only on the radius, as E is conserved throughout the path. If we recall that the radial coordinate is measured by spheres from the source of gravitation we see that the effect is dependent on the photon's height in the gravitational potential, as we expected [4].

For a photon emitted at r_1 and observed at r_2 we can take the ratio of values of (4.21) to obtain the frequency shift

$$\frac{\omega_2}{\omega_1} = \left(\frac{1 - 2GM/r_1}{1 - 2GM/r_2}\right)^{\frac{1}{2}}. \tag{4.22}$$

Now, if we consider the case where $r >> 2GM$ we can give a useful approximation

$$\frac{\omega_2}{\omega_1} = 1 - \frac{GM}{r_1} + \frac{GM}{r_2} = 1 + \Phi_1 - \Phi_2 \tag{4.23}$$

where $\Phi = -GM/r$ is the Newtonian potential.

The experiment to test these predictions was the last of the classical tests of general relativity, and is called the Pound-Rebka experiment. In 1959, Robert Pound and his PhD student Glen Rebka suggested that it would be possible to measure the gravitational redshift of light on earth [18]. By accounting for the special relativistic frequency shift of

$$f = \sqrt{\frac{1 - v}{1 + v}} f_0 \tag{4.24}$$

which is contrary to the effects of gravitational frequency shifts, it would be possible to measure the effects of light falling down the elevator shaft in Harvard university's physics building [18]. By moving the apparatus upwards away from the receiver at the bottom of the elevator shaft the effects of (4.24) and (4.22) could be balanced, allowing Pound and Rebka simply to observe a signal at exactly the same frequency as the emitted signal.

In order to do this Pound and Rebka had to rely on very high frequency gamma waves, as otherwise it would have been impossible to detect the very small frequency shift due to the change in gravitational potential. Additionally, they had to use a specific crystal structure for the gamma source, in order to reduce the recoil of the atom within the source during the emission of the gamma rays which would cause additional Doppler effects which would be difficult to measure consistently.

The Pound Rebka experiment was highly successful test of general relativity's predictions, and later experiments using similar techniques would bring the error bounds to less than 1% of the expected values [22].

5 | Gravitational Wave Tests

Increasingly, gravitational waves are being seen as an exciting method of both testing the predictions of general relativity [22] and observing effects in astronomy that conventional light based observations would be completely incapable of observing [9]. We will present two tests of general relativity using gravitational waves.

5.1 Gravitational Wave Interferometry

When we derived the effects of gravitational waves on particles in chapter 3, we saw that gravitational waves had the effect of stretching or compressing an object depending on their polarization and direction compared to the object. A simple test of gravitational waves is therefore to build a device which is sensitive to such changes in its geometry. Such a device is an interferometer.

An interferometer works by producing light of a single wave length from a laser, splitting the light beam into two halves using a beam splitter, sending the beams down two long paths of equal length at 90° to each other and then recombining them. If the beams have been unaffected by their journey they will recombine perfectly. However, if they have not they will form an interference pattern on arrival at the detector.
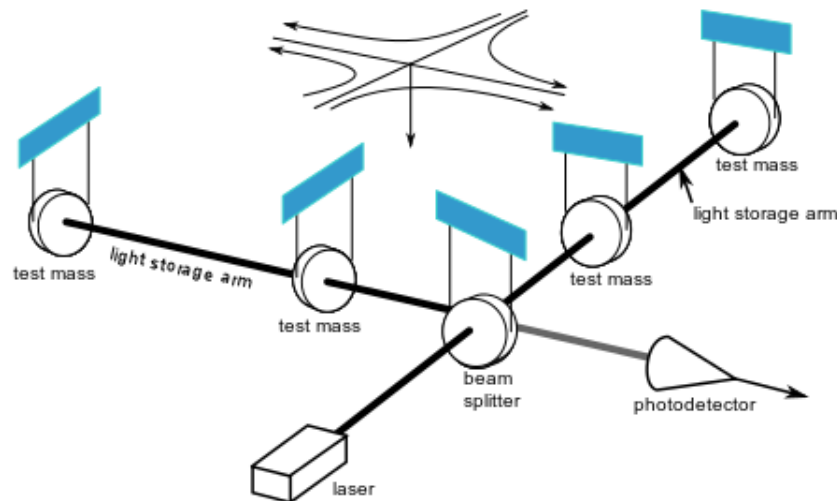

Figure 5.1: A diagram of a laser interferometer, adapted from `http://www.aei.mpg.de/~yanbei/page1/page1.html`

Currently, such interferometers, such as the Laser Interferometer Gravitational Observatory (LIGO) have not yet seen evidence of gravitational waves, however as the sensitivity increases it is likely

they will do so [22].

One of the most important future uses of gravitational wave interferometry will be gravitational wave astronomy [19], where sources of very strong gravity which cannot be seen using conventional EM radiation astronomy may be observed by gravitational waves.

Unfortunately, as of yet there is no direct evidence for gravitational waves. However, there is indirect evidence for gravitational waves in a strong gravitational situation.

5.2   Binary Pulsars

During a radio survey of pulsars by the Arecibo Radio Telescope in the 1970s a binary system containing two pulsars was discovered. Due to the fast orbit of the two pulsars around each other, gravitational radiation from the orbit slowly carries energy away from the system, decreasing the orbital period.

It is possible to derive an expression for the energy loss due to gravitational radiation by considering the mass moments of the system [5]. We will simply state that the rate of energy loss is

$$\frac{dE}{dt} = \frac{128G}{5}\omega^6 M^2 R^4 \tag{5.1}$$

with M the mass of the pulsars which are approximately the same, and ω the angular frequency. From this equation we can calculate the orbital decay of the system.

If the total energy of the system is [5]

$$E = MV^2 = \frac{GM^2}{2R} \tag{5.2}$$

then by considering that $MV^2/R = GM^2/(2R)^2$ we can obtain an equation for V^2

$$V^2 = \frac{GM}{4R} \tag{5.3}$$

and thus the total energy is

$$E = -\frac{GM^2}{4R}. \tag{5.4}$$

If we now rearrange (5.3) for R and note that for an orbit $V^2 = 2\pi R/P_b$ where P_b is the orbital period then

$$R^3 = \frac{GM}{16\pi^2}P_b^2 \tag{5.5}$$

and thus we can write

$$E = -M\left(\frac{\pi MG}{2}\right)^{\frac{2}{3}} P_b^{-\frac{2}{3}}. \tag{5.6}$$

Now if we rearrange (5.6) for P_b and differentiate with respect to time we can obtain the rate of orbital period decrease

$$\dot{P}_b = -\frac{3P_b}{2E}\left(\frac{dE}{dt}\right) \tag{5.7}$$

which after substitution of (5.6) and (5.1) becomes

$$\dot{P}_b = -\frac{48\pi}{5} \left(\frac{4\pi GM}{P_b} \right)^{\frac{5}{3}}.$$ (5.8)

By measuring P_b over time, and examining the results the pulsar's orbital period was seen to be decreasing. The observation of the energy decay over several decades was found to be within 0.3% accuracy of the result obtained by considering gravitational waves. This is a strong sign that gravitational waves are a phenomenon which do exist.

6 | Relativity and the Global Positioning System

Relativity plays a significant role in the operation of the Global Positioning System, from both the fundamental postulates behind how the system works to understanding how errors are introduced to timing and signal propagation. The fact that the GPS is as accurate as it is is a testament to the accuracy of Einstein's theories. A full description of how the GPS works is very much beyond the scope of this text (a very detailed description is available in the book containing [3]), and thus we will focus on the applications of relativity in GPS theory.

6.1   Fundamental operation of the GPS

GPS fundamentally relies on the consistency of the speed of light in any inertial frame. The gravitational field around the earth is taken to be weak enough that spacetime is flat and as such light will move along straight lines.

The global positioning satellites send out coordinated pulses of information at the same time. Four satellites are required in total to allow the receiver to determine its position in spacetime. From the information about the satellites' positions, and by assuming the consistency of the speed of light, a receiver can work out its position on the surface of the earth by solving the system of vector equations obtained from the known positions $\vec{r}_i$ of the satellites and the times t_i that the signal was in transit:

$$c^2(t - t_i)^2 = |\vec{r} - \vec{r}_i|^2. \tag{6.1}$$

While the relativistic effects on the system are fairly small, the clocks on the satellites only need to be out of synchronisation by 3ns for there to be an error of 1m to the calculation of position. Thus we need to correct for such errors for the system to be usable.

6.2   Relativistic Effects on the GPS

There are four main effects that relativity has on the GPS [3].

- Time dilation due to the motion of the satellites

- Gravitational time dilation

- The Sagnac effect

- Doppler shifts in radio frequencies

Most of these effects can be derived from the Schwarzschild metric (2.76) or from a simpler rotating flat space metric. We will discuss these individually.

6.2.1 Time Dilation

There are two sources of time dilation for GPS satellites: the special relativistic effect of the high velocity of the satellites and the general relativistic effect of the lowering of the gravitational potential.

Special Relativistic Time Dilation

If we consider a stationary observer on (but not moving with) the earth's surface measuring the velocity of a GPS satellite they find that the velocity of a satellite, v_s is approximately 3900m/s [5]. Thus by using the special relativistic time dilation formula

$$\gamma_s = \frac{1}{\sqrt{1 - \frac{v_s^2}{c^2}}} \tag{6.2}$$

we find that the fractional change in time compared to a motionless clock is 0.85×10^{-10}. Here we can safely ignore the motion of a clock due to the rate of the earth, as it has a corresponding value of γ which is over 100 times smaller than for the satellites [5].

Gravitational Time Dilation

Gravitational time dilation states that a clock in a higher gravitational potential will run faster than a clock in a lower potential. As such there will be an effect contrary to the special relativistic effect slowing the satellite clocks down, due to the satellites being at a higher potential than the surface of the earth.

This effect is derived in the same way as (4.20) in the Schwarzchild metric. If we imagine ω as defining ticks of a clock then we can obtain, for the fractional change in time of a clock in the gravitational field

$$\frac{\Phi_E - \Phi_s}{c^2} \tag{6.3}$$

with Φ_E the potential at the earth's surface and Φ_s the potential at the satellite. By taking $\Phi = -\frac{GM}{r}$ (the standard Newtonian potential) we find that this effect corresponds to a fractional change in time of -5.2×10^{-10} [12]. This is far larger than the special relativistic result. Thus we see that the GPS depends not only on special relativistic physics but also on general relativity.

6.2.2 Relativistic Doppler Effect

Due to the high speed of the orbiting satellites, relativistic Doppler effects have to be considered when detecting signals from the satellites. The relativistic frequency shift is given by [12]

$$f = f_0 \frac{\sqrt{1 - v_s^2}}{1 - v_s \cos \alpha} \tag{6.4}$$

where v_s is the velocity of the satellite, α is the angle of the light ray from the satellite's motion. For the main $L_1 = 1575.42$MHz carrier frequency which is used for most of the transmission of GPS data this gives potential offsets of up to ± 10KHz which the receivers must account for.

6.2.3 The Sagnac Effect

In our discussion so far we have ignored the rotation of the earth. While compared with the speed of light the rotation of the earth is small, it still produces noticeable effects on the time of propagation of the light rays carrying data from the satellites to the receivers. If a satellite is ahead of the rotation of the earth in the sky the time for the ray to travel will be less than for a satellite behind the rotation of the earth [3, 2].

If we recall the Minkowski metric for flat space from special relativity in cylindrical coordinates:

$$-ds^2 = -(cdt)^2 + dr^2 + r^2 d\varphi^2 + dz^2 \tag{6.5}$$

We now transform (6.5) into a new metric where the azimuthal coordinate rotates with constant speed ω_e by transforming the variables as follows.

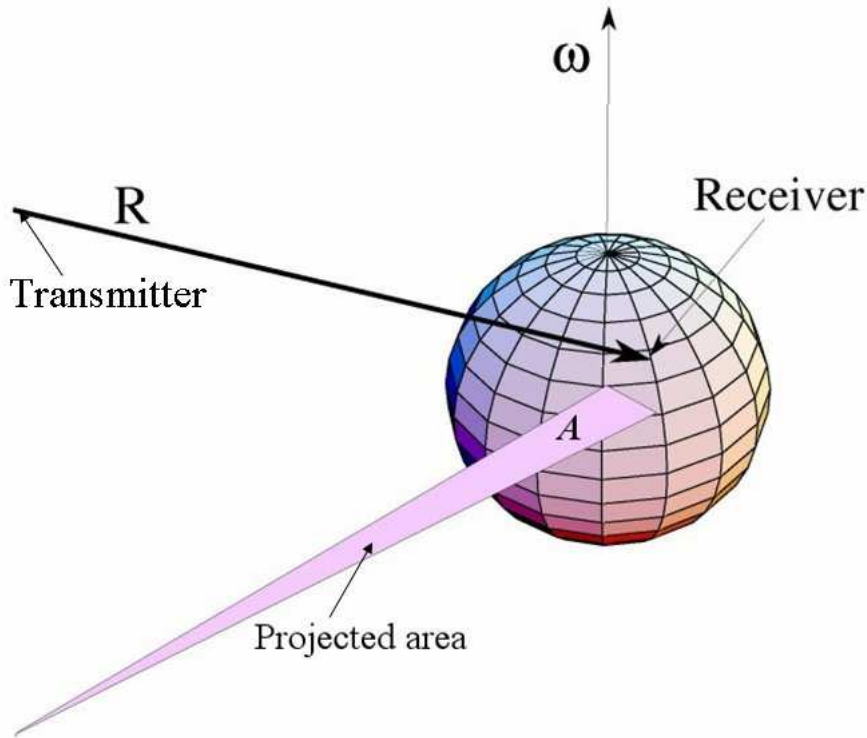$$t = t', \quad r = r', \quad \varphi = \varphi' + \omega_e t', \quad z = z' \tag{6.6}$$

Figure 6.1: A visualisation of the projection of A_z taken from Ashby's article [1]

This transformation produces the metric used in the Sagnac effect [2]

$$-ds^2 = -\left(1 - \frac{\omega_e^2 r'^2}{c^2}\right)(cdt')^2 + 2\omega_e r'^2 d\varphi' dt' + (d\sigma')^2 \tag{6.7}$$

with $(d\sigma')^2 = (dr')^2 + (r'd\varphi')^2 + (dz')^2$. As we recall from chapter two, light travels such that $ds^2 = 0$ and thus we can form an equation for dt' which we can solve to give the time of travel for a light ray. If

we ignore terms $\frac{\omega_e^2 r'^2}{c^2} \ll 1$ we can solve for dt'

$$(cdt')^2 - \frac{2\omega_e r'^2 d\varphi'(cdt')}{c} - (d\sigma)^2 = 0 \tag{6.8}$$

and thus by solving the quadratic for dt'

$$cdt' = d\sigma + \frac{\omega_e r'^2 d\varphi'}{c}. \tag{6.9}$$

Now, if we integrate dt' we can obtain the time of travel for light in this spacetime

$$\int_{\text{path}} dt' = \int_{\text{path}} \frac{d\sigma'}{c} + \frac{2\omega_e}{c^2} \int_{\text{path}} dA'_z \tag{6.10}$$

where dA_z is the projection of the path taken by an observer on the surface of the earth projected onto a triangular infinitesimal area on a plane through the equator. Thus we observe that the Sagnac effect is dependent on how close you are to the equator: for observers at the pole no such effect would be measured, whereas for observers at the equator this could amount to a discrepancy in measured time from the satellite of up to ±207ns.

6.3   Other Relativistic Effects

There are a few other effects which relativity has on the global positioning system, which do not play as much of a role as compared to the ones above.

These are as follows [2, 3]

- Tidal effects on satellite clocks

- The non-spherical nature of the earth produces a field which is not quite like the Schwarzschild metric. This produces quadrupole effects on the satellite clocks.

- Shapiro delay of the light rays

- Relativistic frame dragging due to the rotation of the earth

7 | Conclusions

General relativity continues to stand up to great number of experimental tests, many more than we have had time to survey in this project [22]. The direct observational evidence and direct uses of general relativity in technology we use every day certainly point to the validity of the theory in most gravitational circumstances.

However, this is not to say that general relativity is a completely proven theorem. Increasingly, observations of large scale astronomical phenomena are demanding physics which may not be contained within general relativity. An example of this is the observation of dark matter and dark energy, which may require reformulations of the Einstein equation in order to explain the gravitational fields which seem to be observed.

Additionally, there are other theories of gravitation which ever more sensitive observations will be able to test the validity of. As an example of this, we can consider $f(R)$ gravity, where the Einstein-Hilbert action [1]

$$S = \int \sqrt{-|g_{\mu\nu}|} \frac{R}{16\pi G} d^4 x \tag{7.1}$$

is altered to include a new scalar function $f(R)$ [14, 8]

$$S = \int \sqrt{-|g_{\mu\nu}|} \frac{R + f(R)}{16\pi G} d^4 x \tag{7.2}$$

which would make particles cluster together quicker, producing the effects of dark matter gravitationally. There are many theories such as these, and putting a bound on the possibilities of alternate theories of gravitation is an important effort for experimental physicists [22].

On the whole though, the predictions of general relativity fit with almost all observational data found. The fundamental ideas of general relativity, that mass curves spacetime which then creates gravitation is not in doubt at all. Also, the coordinate independent physics using manifolds and tensors that general relativity introduced are very deeply embedded into how fundamental physics works, and will certainly not be discarded any time soon. Thus, I feel it fair to say that general relativity has been, and will remain, an excellent description of gravitation in our universe.

---

[1]Using this action it is possible to derive the Einstein equation (2.55) using variational principles

Bibliography

[1]   N. Ashby. "Relativistic Effects in the Global Positioning System". In: (2006). URL: `http://aapt.org/doorway/TGRU/articles/Ashbyarticle.pdf`.

[2]   N. Ashby. "Relativity in the global positioning system". In: *Living Rev. Relativity* 6.1 (2003). URL: `http://www.livingreviews.org/lrr-2003-1`.

[3]   N. Ashby and JJ Spilker. "Introduction to relativistic effects on the Global Positioning System". In: *Global Positioning System: Theory and applications.* 1 (1996), pp. 623–697.

[4]   S.M. Carroll. *Spacetime and geometry: An introduction to general relativity.* Addison Wesley, 2004.

[5]   T.P. Cheng. *Relativity, gravitation, and cosmology: a basic introduction.* Oxford University Press, 2010.

[6]   Y. Choquet-Bruhat. *General relativity and the Einstein equations.* Oxford University Press, 2008.

[7]   Mihalis Dafermos. "General Relativity and the Einstein Equations". In: *The Princeton companion to mathematics* (2008), pp. 483–493.

[8]   Antonio De Felice and Shinji Tsujikawa. "f(R) Theories". In: *Living Reviews in Relativity* 13.3 (2010). DOI: `10.12942/lrr-2010-3`. URL: `http://www.livingreviews.org/lrr-2010-3`.

[9]   Chris L. Fryer and Kimberly C.B. New. "Gravitational Waves from Gravitational Collapse". In: *Living Reviews in Relativity* 14.1 (2011). URL: `http://www.livingreviews.org/lrr-2011-1`.

[10] David Griffiths. *Introduction to electrodynamics.* Prentice Hall, 1999.

[11] Louis N Hand and Janet D Finch. *Analytical mechanics.* Cambridge University Press, 1999. ISBN: 9780521573276.

[12] J Hartle. *Gravity : an introduction to Einstein's general relativity.* Addison-Wesley, 2003. ISBN: 9780805386622.

[13] L.D. Landau and E.M. Lifshitz. *The classical theory of fields.* Butterworth-Heinemann, 1975.

[14] Baojui Li. "Testing gravity on cosmological scales". In: *Astronomy & Geophysics* 53.4 (2012), pp. 4.37–4.41. ISSN: 1468-4004. DOI: `10.1111/j.1468-4004.2012.53437.x`. URL: `http://dx.doi.org/10.1111/j.1468-4004.2012.53437.x`.

[15] David Lovelock and Hanno Rund. *Tensors, differential forms, and variational principles.* Courier Dover Publications, 1975.

[16] Stephen M. Merkowitz. "Tests of Gravity Using Lunar Laser Ranging". In: *Living Reviews in Relativity* 13.7 (2010). URL: `http://www.livingreviews.org/lrr-2010-7`.

[17] J J O'Connor and E F Robertson. *History of General Relativity.* URL: `http://www-history.mcs.st-and.ac.uk/HistTopics/General_relativity.html`.

[18] R. V. Pound and G. A. Rebka. "Gravitational Red-Shift in Nuclear Resonance". In: *Phys. Rev. Lett.* 3 (9 Nov. 1959), pp. 439–441. DOI: `10.1103/PhysRevLett.3.439`. URL: `http://link.aps.org/doi/10.1103/PhysRevLett.3.439`.

[19] B.S. Sathyaprakash and Bernard F. Schutz. "Physics, Astrophysics and Cosmology with Gravitational Waves". In: *Living Reviews in Relativity* 12.2 (2009). DOI: `10.12942/lrr-2009-2`. URL: `http://www.livingreviews.org/lrr-2009-2`.

[20] P. Szekeres. *A course in modern mathematical physics: Groups, Hilbert Space and Differential Geometry*. Cambridge University Press, 2004.

[21] J. Weber. *General Relativity and gravitational waves*. Dover Publications, 2004.

[22] Clifford M. Will. "The Confrontation between General Relativity and Experiment". In: *Living Reviews in Relativity* 9.3 (2006). URL: `http://www.livingreviews.org/lrr-2006-3`.

[23] N.M.J. Woodhouse. *General Relativity*. Springer, 2006.