University of Cape Town
phy3004w Project 2020

Exploring the Shores of the General Relativistic Ocean*

Joshua Passmore

19 October 2020

Abstract

This report shows that the Einstein field equations of general relativity may be heuristically obtained, via the extension of Newtonian gravity, modelling spacetime as a four-dimensional pseudo-Riemannian manifold. Initially providing the physical motives for general relativity and explaining the Einstein equivalence principle, we connect the physics of gravity to differential geometry. The mathematics of differential geometry is necessary to describe the theory, therefore we have attempted to provide the rudiments of differential geometry so that final-year physics majors are able to follow the lines of reasoning presented.

1 Introduction

Einstein's general theory of relativity (GR) is a wonderful geometric theory of gravitation and spacetime (see [9] pg. 1). The theory came into being approximately ten years after the formulation of the special theory of relativity (SR), and after the various attempts of trying to fit gravity into the "new-age" relativistic picture. Before the advent of relativity, it was the Newtonian formulation of gravity that reigned supreme. Perhaps one of the greatest triumphs for Newtonian theory was, in an attempt to account for Uranus' unusual path, the correct prediction of the existence and position of the planet Neptune [10].

In 1907, two years after the birth of SR, Planck conjectured, using Einstein's mass-energy equivalence relation, that all energy must gravitate (create its own gravitational field) [10]. Newtonian gravity could not account for this proposition, Newtonian gravity was a pre-relativistic theory and the mass-energy equivalence relation was a direct consequence of SR. This was alarming, Newtonian gravity had been a greatly successful theory, yet it could not be united in a natural manner with SR [5].

The key to resolve this issue lay in Einstein's equivalence principle. SR is a theory of how one correctly transforms, by preserving the universality of the speed of light and the principle of relativity, between inertial reference frames. While working on the mathematical formalism of how to correctly transform between frames of reference that are allowed to accelerate, i.e. the general case of SR, Einstein was led naturally to a theory of gravitation since his equivalence principle forms a strong relationship between gravity and accelerated reference frames. In pursuing his general theory of relativity, Einstein combined the mathematics into a single tensor equation, known as the Einstein field equation, which relates the spacetime curvature to the mass-energy distribution within the region of interest [7].

In this report we will be addressing the proposition that the Einstein field equation can be heuristically obtained, via the extension of Newtonian gravity, modelling spacetime as a four-dimensional pseudo-Riemannian manifold with the techniques of differential geometry. Furthermore, we aim to make the material that is presented accessible to final-year physics majors.

2 Ideas that Led Einstein to the General Theory of Relativity

To provide historical background and motivation for the general theory of relativity, in this section we cover some of the fundamental concepts which emerged in the late 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries and eventually culminated in the general theory of relativity.

*The title is a play-on-words from the first chapter of *Cosmos*, by Carl Sagan, entitled *The Shores of the Cosmic Ocean.*

2.1 Ideas from Special Relativity

Before we begin to consider the theory of general relativity, we will briefly mention the ideas of the special theory of relativity. In this report we assume that the reader has some familiarity with these ideas of SR, so we do not attempt to give an in-depth exposition of SR; rather we aim to remind the reader of these ideas to set the tone for the rest of the material presented, we also assume the reader is familiar with the Einstein summation convention. SR is built upon two postulates: (i) the special principle of relativity and (ii) the constancy of the velocity of light in vacuum for all inertial observers [4]. As a consequence of these postulates, we view space and time as one entity that comes equipped with a special function called the *Minkowski metric*. The Lorentz transformations are the correct linear transformations between inertial frames such that if one knows the spacetime coordinates x^α for an event E in one inertial frame, then one may obtain the coordinates $x^{\alpha\prime}$ for the *same* event E in a different inertial frame such that $x^{\alpha\prime} = \Lambda x^\alpha$ and Λ is a Lorentz transformation matrix [11].

Consider two vectors x^α, y^α in Minkowski space $\mathbb{R}^4$. Mathematically, the Minkowski metric is defined as a *bilinear* and *symmetric* map $\eta : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ such that [8]

$$\eta(x^\alpha, y^\beta) = -x^0y^0 + x^1y^1 + x^2y^2 + x^3y^3.$$

The special property about $\eta(x^\alpha, y^\beta)$ is that it has the *same* value in *all* inertial frames; $\eta(x^\alpha, y^\beta)$ is invariant under Lorentz transformations. η also satisfies the requirements of an inner product on $\mathbb{R}^4$, with one caveat; it is easy to see it is not true that $\eta(x^\alpha, y^\beta) \geqslant 0$ for all $x^\alpha, y^\alpha \in \mathbb{R}^4$ — this means that η is not positive-definite. Furthermore, η can be associated with a 4×4 symmetric matrix

$$\eta_{\alpha\beta} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

such that $\eta(x^\alpha, y^\beta) = \eta_{\alpha\beta}x^\alpha y^\beta$, we often refer to $\eta_{\alpha\beta}$ as the Minkowski metric and say it has signature $(-+++)$. The invariance of $\eta(x^\alpha, y^\beta)$ means that, for all $x^\alpha, y^\alpha \in \mathbb{R}^4$, $\eta_{\alpha\beta}x^\alpha y^\beta = \eta_{\alpha\beta}x^{\alpha\prime}y^{\beta\prime}$ where $x^{\alpha\prime} = \Lambda x^\alpha$ and $y^{\beta\prime} = \Lambda y^\beta$ for all Lorentz transformation matrices Λ.

2.2 The Galilean and Einstein Equivalence Principles

An important concept for GR is the equivalence of gravitational and inertial mass, which is known as the Galilean equivalence principle [10]. The idea that an object's gravitational and inertial mass are equivalent goes back to Galileo, and has the implication that all massive objects (regardless of their mass) accelerate equally under the influence of a gravitational field. The Eötvös experiment was a major result in the 19^{th} century which confirmed, to a great degree of accuracy, the equivalence of gravitational and inertial mass [1] [4].

In 1911, Einstein published a paper in which he produced his *hypothesis as to the physical nature of the gravitational field* ([2] § 1.). To paraphrase from that paper, we consider a homogeneous gravitational field where the acceleration due to gravity has magnitude γ and let there be a stationary coordinate system K, orientated such that the gravitational field is directed in the negative direction of the z-axis. In a space free from gravitational fields consider another coordinate system K' moving with uniform acceleration (with magnitude γ) in the positive direction of its z-axis. Relative to both coordinate systems, objects which are not subjected to the action of other objects, have the following equation of motion:

$$\frac{d^2z}{dt^2} = -\gamma.$$

This is clear for the accelerated system K', but for the system K, which is stationary in a homogeneous gravitational field, this follows from the Galilean equivalence principle: all bodies in a gravitational field are equally accelerated.

Einstein realised that these two coordinate systems are physically *exactly equivalent*; we can regard system K as being, in a region free from gravitational fields, uniformly accelerated, and we can regard system K' as being stationary in a region where there is a homogeneous gravitational field. The figure below illustrates this idea:

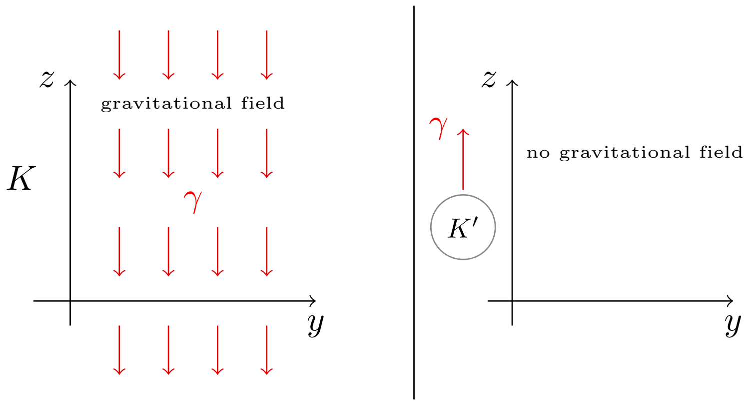

Figure 1: Here we illustrate the Einstein equivalence principle. The coordinate system K is stationary in a region where there is a homogeneous gravitational field of magnitude γ pointing in the negative direction of the z-axis. The coordinate system K' is in a region without gravitational fields, but it is uniformly accelerating with magnitude γ in the positive direction of its z-axis. The equivalence principle asserts that these two coordinate systems are physically exactly equivalent.

Einstein took this one step further and stated that for two such coordinate systems K and K', we assume that the laws of nature, as observed from these two systems, are in entire agreement with each other. Hence by understanding the physical processes that take place relative to a system of uniform acceleration, we are able to understand the same physical processes which take place in a homogeneous gravitational field. This is what we understand as the *Einstein equivalence principle*.

2.3 A Generalisation of the Principle of Relativity

The special principle of relativity states that the laws of physics hold good, in their simplest form, in all inertial reference frames. Ernst Mach was the first to point out, via epistemological reasoning, that this idea needed to be extended to *all* frames of reference ([4] § 2.). In [4] Einstein points out that this generalisation of the principle of relativity is well-grounded on the following physical fact.

We consider a coordinate system K which is inertial; an object which is unaffected by the action of others, as observed from K, moves in a straight line with constant speed. We consider a second coordinate system K' which is moving relative to K with constant acceleration in a straight line. Relative to K' an object that is unaffected by the action of others would be accelerating such that its state of acceleration is completely independent of the material composition of the mass. An observer situated at rest relative to K' can then not know whether they are actually in an accelerated frame of reference or just in a homogeneous gravitational field, this we know from the Einstein equivalence principle. Therefore we may regard the system K' as being *stationary* in a homogeneous gravitational field and in this case the laws of physics hold good in their simplest form with respect to K'. Both systems of reference K and K' then have the property that physical laws hold good in their simplest form with respect to these two systems. Einstein stated the general principle of relativity as: *The laws of physics must be of such a nature that they apply to systems of reference in any kind of motion* ([4] § 2.).

We will see that as a consequence of this, the Einstein equivalence principle, combined with the generalisation of the principle of relativity, manifests itself in differential geometry as the *local-flatness theorem* and hence it is the Einstein equivalence principle which provides the link between gravity and differential geometry. It is due to this connection between the Einstein equivalence principle and the local-flatness theorem, that we propose the field equations of GR can be heuristically obtained by modelling spacetime as a four-dimensional structure using the techniques of differential geometry. However, before discussing the similarities between gravity and differential geometry, we look at some of the mathematics that is required.

3 Intuitive Understanding of the Mathematics of General Relativity

Here we will briefly outline some of the mathematics that is required to understand the rudiments of GR and attempt to do so in an intuitive manner so as to make the material accessible to undergraduate physics students. GR is a geometric theory of spacetime ([9] pg. 1), the Einstein equations are the fundamental equations of the theory which govern the structure and dynamics of spacetime. Mathematically, the Einstein equations are described using a tensorial relationship in a four-dimensional manifold. Since one of the aims of this report is to make accessible these ideas of GR to undergraduates, we need working definitions of structures such as manifolds, and objects such as tensors.

We can think of an n-dimensional manifold as a continuous (not necessarily Euclidean) space that comes equipped with a set of coordinates. An example of a two-dimensional manifold is the surface of the unit sphere $\mathbb{S}^2$, this space is embedded in Euclidean 3-space and comes equipped with its own "curved" coordinate system. $\mathbb{S}^2$ is itself not a vector space since it is not "flat" (more precisely, the space is not closed under addition nor scalar multiplication), however at any point $p \in \mathbb{S}^2$ the neighbourhood around it is *approximately* Euclidean. This means that the tangent plane (which is a vector space) to $p \in \mathbb{S}^2$ is *locally* a good approximation of the space $\mathbb{S}^2$.

One of the more difficult concepts for physics students to grasp is that of a tensor, one of the reasons is that it is quite abstract and students are often left puzzled as to what the difference between a tensor and a square matrix is. In the undergraduate physics modules, students only ever approach tensors of second rank, and the tensors are almost always represented as square matrices. This is very misleading as students begin to think that every square matrix is then a tensor, which is false. The definition that we use of a tensor is the one which defines how the components of the tensor transform under a coordinate transformation. There are indeed different, yet equivalent, definitions of tensors; such as defining a tensor as a multilinear form, which maps elements from a vector space and its dual space to a set of scalars (see [8] for an in-depth introduction to tensors). This definition is more rigorous but unnecessarily abstract for understanding the basics of GR. We will make use of the definition as given by [7]:

Consider a k-dimensional manifold $\mathcal{M}$ with a coordinate system $x^\alpha = \{x^1, x^2, \ldots, x^k\}$, we define a tensor of type (n, m) as an object (with $r = n+m$ indices) $T^{\alpha\ldots\beta}{}_{\gamma\ldots\delta}$ which, under a coordinate transformation $x^\alpha \to x^{\alpha\prime}$, transforms as

$$T^{\alpha'\ldots\beta'}{}_{\gamma'\ldots\delta'} = \frac{\partial x^{\alpha'}}{\partial x^\alpha} \cdots \frac{\partial x^{\beta'}}{\partial x^\beta} \frac{\partial x^\gamma}{\partial x^{\gamma'}} \cdots \frac{\partial x^\delta}{\partial x^{\delta'}} \, T^{\alpha\ldots\beta}{}_{\gamma\ldots\delta}.$$

The number of superscripts on T is equal to n and the number of subscripts equal to m, we also refer to superscripts as contravariant indices and subscripts as covariant indices. The number r is called the rank of the tensor.

Next we outline some special kinds of tensors. Contravariant vectors are tensors of type (1, 0) and covariant vectors (dual vectors) are tensors of type (0, 1). The metric tensor (often just abbreviated as "the metric") is a second rank covariant tensor $g_{\alpha\beta}$ that is used to define the inner product between two vectors a, b:

$$(a|b) := g_{\alpha\beta} a^\alpha b^\beta.$$

We've seen that in SR, the inner product between two vectors in Minkowski space is not necessarily positive-definite. If GR is the generalisation of SR, then we must assume that $(a|b)$ is not necessarily positive-definite and hence we say that the metric $g_{\alpha\beta}$ is not positive-definite. A manifold where the metric $g_{\alpha\beta}$ is positive-definite is called a Riemannian manifold, a manifold where the metric is not positive-definite is called pseudo-Riemannian, hence in GR we always work with a pseudo-Riemannian manifold. Another special property of the metric is that it's symmetric, $g_{\alpha\beta} = g_{\beta\alpha}$, and can be used to raise and lower vector indices: $a_\mu = g_{\mu\nu} a^\nu$ and $b^\mu = g^{\mu\nu} b_\nu$ (where $g^{\mu\nu}$ is the metric inverse) (see [7] pg. 3). One slight subtlety when working with vectors (and tensors in general) on manifolds is that they're not actually defined in the manifold itself, but rather vectors (and tensors) are defined in the tangent planes of the manifold. This is because, in general, the manifold is not a linear space, whereas the tangent space is. The figure below illustrates this idea:

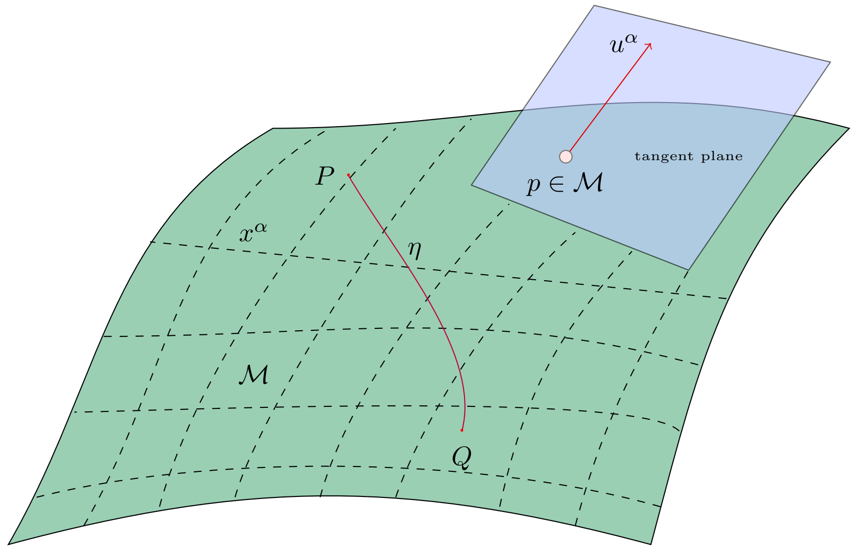

Figure 2: Here we show a schematic representation of a two-dimensional manifold $\mathcal{M}$ with a coordinate system x^α (shown with the dashed lines). We consider a point $p \in \mathcal{M}$. A vector u^α at p is defined in the tangent plane at p and not in $\mathcal{M}$. The red curve η represents the geodesic between the points Q and P.

We now look at how one differentiates tensors, this is known as *covariant differentiation*. Tensors in separate tangent planes cannot be combined in a meaningful manner, hence one needs to transport a tensor from one tangent plane to another so that the tensors concerned are in the same tangent plane and can be combined. Covariant differentiation allows for the differentiation of a tensor field (we may think of this as a generalisation of a vector field) by adding another structure to the manifold called the *connection* $\Gamma^\alpha_{\mu\beta}$. For example, consider a manifold $\mathcal{M}$ with coordinates x^α and a vector field v^α defined in the tangent space of the manifold. The absolute differential of v^α is defined by (see [7] pg. 5)

$$\begin{aligned} Dv^\alpha &:= \left(\frac{\partial v^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\mu\beta} v^\mu\right) dx^\beta \\ &= \left(v^\alpha_{,\beta} + \Gamma^\alpha_{\mu\beta} v^\mu\right) dx^\beta \\ &= v^\alpha_{;\beta}\, dx^\beta \end{aligned}$$

where we've introduced the notation $\partial v^\alpha / \partial x^\beta := v^\alpha_{,\beta}$ and $v^\alpha_{;\beta} := (v^\alpha_{,\beta} + \Gamma^\alpha_{\mu\beta} v^\mu)$. The object $v^\alpha_{;\beta}$ is a tensor and is called the covariant derivative of v^α.

As a practical example, we consider the figure above. We have two points Q and P in the manifold $\mathcal{M}$. Suppose η is a curve in the manifold that connects these two points. The distance s between Q and P is given by

$$s = \int_Q^P ds$$

where ds is the infinitesimal line-element in the manifold; $ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$. We want the curve η to be a geodesic, this means we want the distance s to be minimised. Suppose η is parametrised in the coordinate system x^α by the affine parameter λ. This just means that $|\dot{x}^\alpha| := |dx^\alpha / d\lambda| = 1$ for all λ. Then

$$ds = \sqrt{g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta} d\lambda \implies s = \int_{\lambda(Q)}^{\lambda(P)} \sqrt{g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta} d\lambda.$$

The (affine) parametrisation $x^\alpha(\lambda)$ for the geodesic η is then obtained by solving the following *geodesic equation*

$$\ddot{x}^\alpha = -\Gamma^\alpha_{\beta\gamma} \dot{x}^\beta \dot{x}^\gamma. \tag{1}$$

The last bit of mathematics that one needs to understand the rudiments of GR is the *geodesic deviation equation*. We consider a manifold and introduce a family of geodesics η_s with $s \in [0, 1]$ as is depicted below. Each curve is parametrised by λ in the coordinates x^α. We collectively describe this family using the relations $x^\alpha(s, \lambda)$. The vector field $\partial x^\alpha / \partial \lambda$ is tangent to the geodesics.

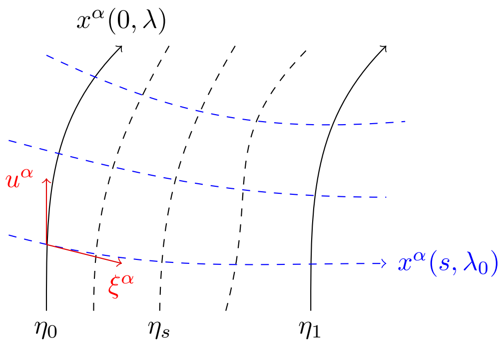

Figure 3: A family of geodesics η_s is shown with parametrisation $x^\alpha(s,\lambda)$. The blue curves represent the s family of curves which are not necessarily geodesics. The vector $u^\alpha(\lambda)$ is tangent to the geodesic η_0 and $\xi^\alpha(\lambda)$ is the deviation vector between η_0 and η_1. This figure is inspired by [7] pg. 17.

For each λ, we can keep λ constant and let s vary to obtain the family of curves highlighted in blue above. The vector field $\partial x^\alpha/\partial s$ is tangent to these curves. We define

$$u^\alpha(\lambda) = \frac{\partial x^\alpha}{\partial \lambda}|_{s=0} \text{ and } \xi^\alpha(\lambda) = \frac{\partial x^\alpha}{\partial s}|_{s=0}.$$

We call ξ^α the deviation vector between the geodesics η_0 and η_1. The geodesic deviation equation provides an expression for the acceleration of this deviation vector:

$$\frac{D^2\xi^\alpha}{d\lambda^2} = -R^\alpha_{\beta\gamma\delta}u^\beta u^\delta \xi^\gamma \tag{2}$$

where $R^\alpha_{\beta\gamma\delta}$ is the Riemann curvature tensor. Qualitatively, this equation states that curvature in the manifold (as represented by the Riemann tensor) produces a relative acceleration between two neighbouring geodesics (see [7] pg. 18). The Riemann tensor is also related to the connection in the following way:

$$R^\alpha{}_{\beta\gamma\delta} = \Gamma^\alpha{}_{\beta\delta,\gamma} - \Gamma^\alpha{}_{\beta\gamma,\delta} + \Gamma^\alpha{}_{\mu\gamma}\Gamma^\mu{}_{\beta\delta} - \Gamma^\alpha{}_{\mu\delta}\Gamma^\mu{}_{\beta\gamma}. \tag{3}$$

Now that we have some of the basic mathematics required, we can begin to explore gravitational physics.

4 Exploring Newtonian Gravity

Newton's famous inverse square law states that, given a particle A (centred at the origin) with gravitational mass M_g and a particle B with gravitational mass m_g, then the attractive gravitational force of A on B is given by

$$\boldsymbol{F}(\boldsymbol{r}) = -G\frac{M_g m_g}{r^2}\hat{\boldsymbol{r}}$$

where $\boldsymbol{r}$ is the position vector of particle B. Provided that the only force acting on particle B is this gravitational force, then we may write

$$\boldsymbol{F}(\boldsymbol{r}) = m_i\ddot{\boldsymbol{r}} = -G\frac{M_g m_g}{r^2}\hat{\boldsymbol{r}}$$
$$\implies \ddot{\boldsymbol{r}} = -\frac{GM_g}{r^2}\hat{\boldsymbol{r}}\left(\frac{m_g}{m_i}\right),$$

where m_i is the inertial mass of particle B. By the Galilean equivalence principle we write

$$\ddot{\boldsymbol{r}} = -\frac{GM}{r^2}\hat{\boldsymbol{r}} \quad \text{(where we've dropped the subscript } g \text{ from } M).$$

This is the equation for the gravitational acceleration of particle B in the gravitational field of particle A. As was stated previously (and is now clear by the equation above) a consequence of the Galilean equivalence principle is that the acceleration of particle B is independent of its mass. Since the choice of particle B is arbitrary (provided it has mass), the quantity

$$-\frac{GM}{r^2}\hat{\boldsymbol{r}}$$

actually represents the gravitational field of particle A, thus we write

$$\boldsymbol{g}(\boldsymbol{r}) = -\frac{GM}{r^2}\hat{\boldsymbol{r}} \quad \text{(gravitational field of particle } A\text{)}.$$

In analogy to Coulomb's law in electrostatics, the gravitational force on a point mass m in the gravitational field $\boldsymbol{g}$ can then be expressed as $\boldsymbol{F}(\boldsymbol{r}) = m\boldsymbol{g}(\boldsymbol{r})$. More generally, if we consider a finite (and fixed) volume of matter that is continuous, then the gravitational field of this mass distribution is given by the integral

$$\boldsymbol{g}(\boldsymbol{r}) = -G\int_{\mathcal{V}} \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}(\boldsymbol{r}-\boldsymbol{r}')dV', \tag{4}$$

where the integral is evaluated through the volume source $\mathcal{V}$ and $\rho(\boldsymbol{r}')$ is the mass-density function of the source.

However, upon closer inspection of this integral, we see that once a mass distribution is brought into empty space, it sets up its gravitational field everywhere instantaneously (we call this instant-action at a distance). This implies that the information of the gravitational field would travel faster than the speed of light. This clearly is in contradiction to the special theory of relativity. It was Poincarѐ who correctly conjectured in 1905 that gravity should propagate through space at finite velocity, and more precisely at the speed of light. In analogy to electromagnetism, he also conjectured that this propagation would be in the form of gravitational waves. (See [10]).

The gravitational field (4) is conservative, which means that there exists a scalar potential field $\phi(\boldsymbol{r})$ such that

$$\boldsymbol{g}(\boldsymbol{r}) = -\nabla\phi(\boldsymbol{r}). \tag{5}$$

Moreover from Eq.(4) one may obtain the differential form of Gauss's law for gravity (see [11] for a formal derivation for the analogous Gauss's law in electrostatics)

$$\nabla \cdot \boldsymbol{g} = -4\pi G\rho. \tag{6}$$

Substituting Eq.(5) into the equation above (6), we arrive at the Poisson equation:

$$\nabla^2\phi = 4\pi G\rho. \tag{7}$$

5 A Heuristic Approach to a Geometric Theory of Gravity

As was claimed in section 2.3, the Einstein equivalence principle, combined with the generalisation of the principle of relativity, manifests itself in differential geometry as the local-flatness theorem. Since we now have the mathematics that is required, we state the theorem: [7]

Theorem 1 (Local-flatness)**.** For a given point (event) P in spacetime, it is always possible to find a coordinate system x^α such that the metric tensor evaluated at P yields the Minkowski metric, and the connection evaluated at P yields nullity. Symbolically, it is always possible to find a coordinate system such that

$$g_{\alpha\beta}(P) = \eta_{\alpha\beta} \quad \text{and} \quad \Gamma^{\alpha}_{\beta\gamma}(P) = 0.$$

This theorem implies that the Einstein equivalence principle fixes local spacetime to be Minkowski, and gravity describes how various localities of spacetime are connected [10]. From this point forward we will model spacetime as a four-dimensional pseudo-Riemannian manifold.

5.1 The Relation Between the Metric and the Gravitational Potential

Consider a region of spacetime where there is a continuous mass distribution $\rho(\boldsymbol{r}')$, according to Newtonian gravity, this mass distribution emits a gravitational field $\boldsymbol{g}(\boldsymbol{r})$ which is associated with a potential $\phi(\boldsymbol{r})$ such that Eq.(7) is satisfied. Sticking with Newtonian gravity, suppose we have a particle in this gravitational field with a spatial position $\boldsymbol{x}$. By the Galilean equivalence principle we write

$$\ddot{\boldsymbol{x}} = \boldsymbol{g}. \tag{8}$$

If we assume that the mathematical properties of spacetime satisfy that of a pseudo-Riemannian manifold, then we need to form a relationship between Eq.(8) and the geodesic equation (1) (see [9] pg. 82). In this case we want to identify the gravitational field $\boldsymbol{g}$ as providing the geodesic for a particle in the field. In 1913, Einstein and Grossman worked on a tensor theory of gravitation and realised that gravity can be mathematically described using differential geometry (see [3]), this means that there needs to be a relationship between $\boldsymbol{g}$ and $\Gamma^{\alpha}_{\beta\gamma}\dot{x}^{\beta}\dot{x}^{\gamma}$. Making use of dimensional analysis, we consider the physical dimensions of Eqs.(8) & (1): (we use the symbol L for the dimensions of length and t for the dimensions of time)

$$[\boldsymbol{g}] = \frac{L}{t^2} \quad \text{and} \quad [\Gamma^{\alpha}_{\beta\gamma}\dot{x}^{\beta}\dot{x}^{\gamma}] = \frac{L}{t^2},$$

a few more lines show that

$$[\Gamma^{\alpha}_{\beta\gamma}] = \frac{1}{L}.$$

Our assumption is that, in some sense, $\boldsymbol{g}$ and $\Gamma^{\alpha}_{\beta\gamma}\dot{x}^{\beta}\dot{x}^{\gamma}$ are equivalent (or at least proportional), thus we expect their ratio to be constant. In particular

$$\frac{[\boldsymbol{g}]}{[\Gamma^{\alpha}_{\beta\gamma}]} = \left(\frac{L}{t}\right)^2,$$

which we expect to be constant. The only fundamental constant in nature with these dimensions is the speed of light squared. This leads us to the following relation (see [9] pg. 83)

$$\frac{1}{c^2}\boldsymbol{g} \sim \Gamma^{\alpha}_{\beta\gamma}. \tag{9}$$

(Throughout the report we will use the symbol "$\sim$" to mean that two quantities are related to each other in some way. This is indeed vague, but note that the connection is a more complex structure on the manifold than the gravitational field; so we can't necessarily equate the two.). This is a natural way of adjoining special relativity, using the constant c^2, with gravity (as described by $\boldsymbol{g}$) and geometry (as described by $\Gamma^{\alpha}{}_{\beta\gamma}$). On a pseudo-Riemannian manifold, where the metric is symmetric, the connection Γ is related to the metric tensor in the following way [7]

$$\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2}g^{\alpha\mu}(g_{\mu\beta,\gamma} + g_{\mu\gamma,\beta} - g_{\beta\gamma,\mu}), \tag{10}$$

in this case the connection is fully determined by the metric and its partial derivatives. We refer to the $\Gamma^{\alpha}_{\beta\gamma}$ as the Christoffel symbols. Eq.(5) and relation (9) imply

$$\Gamma(g) \sim -\frac{1}{c^2}\nabla\phi \quad \text{(noting that the connection is a function of the metric } g\text{)} \tag{11}$$

and so the components of the metric tensor $g_{\alpha\beta}$ must contain the gravitational potential ϕ. From Eq.(10) we see that

$$\Gamma^{\alpha}_{\beta\gamma} \propto \frac{1}{2}\partial_{\mu}g_{\alpha\beta}. \tag{12}$$

Combining relations (11) & (12) we have

$$\begin{aligned} \frac{1}{2}\partial_{\mu}g_{\alpha\beta} &\sim -\frac{1}{c^2}\partial_{\mu}\phi \\ \implies \frac{1}{2}g_{\alpha\beta} &\sim -\frac{1}{c^2}\phi. \end{aligned} \tag{13}$$

Under a static and spherically symmetric gravitational field, emitted by some spherical mass M, the Newtonian potential is given by (see [9] pg. 69)

$$\phi = -G\frac{M}{r}.$$

Substituting this into relation (13) yields

$$g_{\alpha\beta} \sim \frac{2GM}{c^2 r}.$$

This seems like a reasonable relation for the metric tensor of spacetime. However in the limit $M \to 0$, we would expect the spacetime to be Minkowski space and thus the metric $g_{\alpha\beta}$ should be the Minkowski metric $\eta_{\alpha\beta}$. But this is not the case since $g_{\alpha\beta} \to 0$ as $M \to 0$ which is not Minkowski. Therefore to ensure that we have Minkowski space when $M \to 0$ we can do the following:

$$g_{\alpha\beta} \sim \begin{cases} -1+\frac{2GM}{c^2 r} & \text{if } \alpha = \beta = 0 \\ 1+\frac{2GM}{c^2 r} & \text{if } \alpha = \beta = 1,2,3 \\ 0 & \text{otherwise.} \end{cases}$$

This satisfies the symmetric requirement of the metric and the signature $(-+++)$. Notice that this is equivalent to

$$g_{\alpha\beta} \sim \begin{cases} -1-\frac{2}{c^2}\phi & \text{if } \alpha=\beta=0 \\ 1-\frac{2}{c^2}\phi & \text{if } \alpha=\beta=1,2,3 \\ 0 & \text{otherwise.} \end{cases}$$

We should point out that in the above relation, the gravitational potential ϕ need not correspond to a spherically symmetric gravitational field; the choice of gravitational field could be arbitrary. However we do require that, physically, $\phi \to 0$ as $|\boldsymbol{r}| \to \infty$. Let us define $\varphi := \frac{\phi}{c^2}$ so that

$$g_{\alpha\beta} \sim \begin{cases} -1-2\varphi & \text{if } \alpha=\beta=0 \\ 1-2\varphi & \text{if } \alpha=\beta=1,2,3 \\ 0 & \text{otherwise.} \end{cases}$$

This is our ad hoc relationship for the metric tensor of spacetime where φ now represents the relativistic gravitational potential. In order to see if this relation between the metric and the gravitational potential is consistent with our knowledge of physics, let us use this metric to determine the trajectory of a particle in a weak gravitational potential – that is $|\varphi| << 1$. Consider a spacetime with coordinates $\{x^0, x^1, x^2, x^3\} = \{ct, x^1, x^2, x^3\}$. Our goal is to determine the trajectory of a particle moving from point Q to point P. Suppose the particle follows a path η that is parametrised in the coordinate system x^α by the affine parameter λ. Since the only interaction this particle is subject to is the gravitational one, gravity provides the geodesic η for this particle; the parametrisation $x^\alpha(\lambda)$ is obtained by solving the geodesic equation (1). Using the metric $g_{\alpha\beta}$ above and the conditions for non-relativistic motion in a weak gravitational field, let us investigate the role that gravity plays in the geodesic eqution.

First we change variables and use $u^\alpha := \dot{x}^\alpha$ so that the geodesic equation becomes

$$\dot{u}^\alpha + \Gamma^\alpha_{\beta\gamma} u^\beta u^\gamma = 0.$$

For non-relativistic motion, $u^0 >> |\boldsymbol{u}|$ and $|\varphi| << 1$ ($\boldsymbol{u}$ is the 3-velocity of the particle in λ), we have (see [9] pg. 84)

$$\dot{u}^\alpha + \Gamma^\alpha_{\beta\gamma} u^\beta u^\gamma \simeq \dot{u}^\alpha + \Gamma^\alpha_{00} u^0 u^0 = 0 \quad \text{and} \quad \Gamma^0_{00} = 0.$$

These two equations above imply $\dot{u}^0 = 0$ and $u^0 = D$ where D is a constant. But $u^0 := \frac{d}{d\lambda}(x^0) = c\frac{dt}{d\lambda} = D$. Let us focus on the spatial coordinates, where we use Latin indices for indexing over the set $\{1, 2, 3\}$, the geodesic equation reads

$$\frac{du^i}{d\lambda} + \Gamma^i_{00} u^0 u^0 = 0.$$

Noting that $d\lambda = \frac{c}{D} dt$, a little rearrangement shows that

$$\frac{d^2 x^i}{dt^2} = -c^2 \Gamma^i_{00}.$$

We now focus on the Christoffel symbols

$$\begin{aligned} \Gamma^i_{00} &= \frac{1}{2} g^{ij} (\overbrace{g_{j0,0} + g_{j0,0}}^{=0} - g_{00,j}) \\ &= -\frac{1}{2} g^{ij} (g_{00,j}). \end{aligned}$$

Recall that $g_{00} = -1 - 2\varphi$, hence $g_{00,j} = \partial_j g_{00} = -2\partial_j \varphi$, then

$$-\frac{1}{2} g^{ij} (-2\partial_j \varphi) = \partial^i \varphi \quad \text{(the metric raises the index on the derivative operator).}$$

This allows us to write

$$\begin{aligned} \frac{d^2 x^i}{dt^2} &= -c^2 \partial^i \varphi \\ &= -\partial^i (c^2 \varphi) \\ &= -\partial^i \phi \\ \iff \ddot{\boldsymbol{x}} &= -\nabla \phi = \boldsymbol{g}. \end{aligned}$$

Thus the relationship between our current metric and the gravitational potential yields the correct form for the particle's equation of motion, in the non-relativistic and weak-gravitational regime, when we consider the geodesic equation for the particle. Another important equation for pseudo-Riemannian manifolds, as we have seen, is the geodesic deviation equation. This is important for us, under the assumption that the mathematical properties of spacetime satisfy the requirements of a pseudo-Riemannian manifold, then gravity (which is very closely related to acceleration) should be related to the curvature of the manifold, this follows directly from the geodesic deviation equation (2).

5.2 Gravity and the Geodesic Deviation Equation

Consider a four-dimensional spacetime manifold where there is a gravitational potential $\varphi = \varphi(x^\alpha(s,t))$ and t represents time. In a similar manner to the geodesic deviation equation, we consider two families of curves that are described collectively by $x^\alpha(s,t)$. For the moment we will only concentrate on the spatial coordinates $x^i = \{x^1, x^2, x^3\}$. We define

$$u^i := \frac{\partial x^i}{\partial t} \quad \text{and} \quad \xi^i := \frac{\partial x^i}{\partial s}.$$

Our goal is to derive an expression for $\frac{\partial^2 \xi}{\partial t^2}$ and compare this to the geodesic deviation equation (2) (see [9] pgs. 85 and 86). By Clairaut's theorem we can write

$$\frac{\partial \xi^i}{\partial t} = \frac{\partial u^i}{\partial s},$$
$$\text{hence} \quad \frac{\partial^2 \xi^i}{\partial t^2} = \frac{\partial}{\partial s}\frac{\partial}{\partial t} u^i = \frac{\partial}{\partial s}\frac{\partial^2 x^i}{\partial t^2}.$$

We also note that the gravitational acceleration of an object is proportional to the gradient of the potential:

$$\frac{\partial^2 x^i}{\partial t^2} = -\frac{\partial \varphi}{\partial x^i}$$

which implies

$$\begin{aligned}\frac{\partial^2}{\partial t^2}\xi^i &= -\frac{\partial}{\partial s}\left(\frac{\partial \varphi}{\partial x^i}\right)\\ &= -\frac{\partial^2 \varphi}{\partial x^j \partial x^i} \cdot \frac{\partial x^i}{\partial s} \quad \text{(using chain rule)}\\ &= -\frac{\partial^2 \varphi}{\partial x^j \partial x^i}\xi^i. \end{aligned} \tag{14}$$

We may re-write the last line (14) as

$$\frac{\partial^2}{\partial t^2}\xi^i = -(\partial_j \partial_i \varphi)\xi^i \tag{15}$$

and compare this to

$$\frac{D^2}{d\lambda^2}\xi^\alpha = -\left(R^\alpha_{\beta\gamma\delta} u^\beta u^\delta\right)\xi^\gamma. \tag{16}$$

We can think of Eq.(15) as the analogous geodesic deviation equation in a three-dimensional space with gravitational potential φ, whereas Eq.(16) is the geodesic deviation equation for an arbitrary manifold. If we want to describe gravitation using differential geometry, then we need to investigate the similarities between Eqs.(15) & (16). First we can change the Latin indices to Greek indices $i \to \gamma$ and $j \to \alpha$ and by comparison to the geodesic deviation equation (16):

$$\partial_j \partial_i \varphi \to \partial^\alpha \partial_\gamma \varphi$$

where α is a superscript in order to match the superscript on the Riemann tensor. In this case we write

$$R^\alpha_{\beta\gamma\delta} u^\beta u^\delta \sim \partial^\alpha \partial_\gamma \varphi.$$

Tensorial contraction allows us to express this relationship with the Ricci tensor,

$$R_{\beta\delta} := R^\gamma_{\beta\gamma\delta} \quad \text{(Ricci curvature tensor)},$$

which brings us to

$$R_{\beta\delta} u^\beta u^\delta \sim \partial^\alpha \partial_\alpha \varphi. \tag{17}$$

Consider a three-dimensional space for the moment. The object $\partial^\alpha \partial_\alpha \varphi$ is equivalent to $\nabla^2\varphi$, and we know from Eq.(7) that this is proportional to the mass distribution ρ in the region. We can infer from this that there must be a *relationship between the Ricci curvature tensor $R_{\beta\delta}$ and the mass distribution within the region*:

$$R_{\beta\delta} \sim \rho.$$

However the density distribution of matter ρ in the spacetime region under consideration needs to be generalised into a second rank covariant tensor in order for there to be a formal tensorial relationship between the Ricci tensor and the matter distribution. This is what we investigate next.

5.3 Touch Down at the Field Equations

In [4] (§ 4.) Einstein argues that the metric components $g_{\alpha\beta}$ describe the gravitational field. Continuing with this proposition, if we want to obtain a geometric theory of gravity, we would need a set of differential equations in which we solve for the metric. If the relationship (17) described above is true, then in a region where there is no mass or energy we'd have

$$\partial^\alpha \partial_\alpha \varphi = 0 \implies R_{\beta\gamma} = 0.$$

Thus, mathematically speaking, when there is no matter or energy in the manifold, the structure of the manifold is purely governed by the equation

$$R_{\beta\gamma} = 0. \tag{18}$$

We compare these equations (18) to the Laplace equation $\nabla^2\phi = 0$ in Newtonian gravity. The implications of this are actually quite deep. The Ricci tensor is a function of the Riemann tensor and the Riemann tensor is a function of the connection (see Eq.(3)). We know from Eq.(10) that the Christoffel connection is a function of the metric $g_{\mu\nu}$ and its partial derivatives $\partial_\gamma g_{\mu\nu}$, hence Eq.(18) represents a system of partial differential equations in which the unknown functions are the components of the metric. By solving Eq.(18) for a given set of boundary conditions, we obtain all the information we need about the structure of the manifold when there is no matter or energy present. We need to generalise these equations (18) to account for space which is not vacuum — a space where there is energy and matter present.

In analogy to equations (18) being compared to the Laplace equation, we want to generalise this concept so that we have an equation involving the Ricci tensor $R_{\beta\gamma}$ on one side of the equation and another tensorial expression on the other side, so that we can compare this to the Poisson equation (7). In [4] (§ 16.) Einstein notes that "inert mass is nothing more or less than energy, which finds its complete mathematical expression in a symmetrical tensor of second rank, the energy tensor." To generalise the matter distribution ρ, we would require a second rank covariant symmetric tensor (since the Ricci tensor is symmetric), for the moment let us refer to this "new" tensor as $P_{\alpha\beta}$ so that we can be explicit in stating that we want

$$R_{\alpha\beta} = P_{\alpha\beta}\ , \tag{19}$$

such that when there is no matter or energy present, $P_{\alpha\beta}$ vanishes and we are left with Eq.(18). The stress-energy tensor $T_{\alpha\beta}$ is a result from SR, such that for any collection of particles and any electromagnetic fields, the contravariant components $T^{\alpha\beta}$ associated with this system are interpreted as (see [6] pg. 16)

$$T^{\alpha\beta} = \begin{pmatrix} \begin{matrix}\text{pressure}\\ \text{tensor}\end{matrix} & \begin{matrix}\text{momentum}\\ \text{density}\end{matrix} \\ & \\ \begin{matrix}\text{energy-flux}\\ \text{density}\end{matrix} & \begin{matrix}\text{energy}\\ \text{density}\end{matrix} \end{pmatrix}.$$

This interpretation holds good in any inertial frame, we may think of the stress-energy tensor as generalising the stress tensor from classical mechanics. $T_{\alpha\beta}$ may also be understood as relating the energy density ρ_E of a physical system to an observer with 4-velocity $u^\alpha := dx^\alpha/d\tau$ where τ is the proper time (the time as measured in the observers inertial frame) such that (see [6] pg. 16)

$$\rho_E = T_{\alpha\beta} u^\alpha u^\beta.$$

Moreover $T_{\alpha\beta}$ *is* symmetric and there is a covariant expression for energy and momentum conservation:

$$T^{\alpha\beta}{}_{;\beta} = 0.$$

It can be shown that, for the Ricci tensor, $R^{\alpha\beta}{}_{;\beta} \neq 0$ (see [10]). Therefore the Ricci tensor alone cannot be proportional to the stress-energy tensor. Associated with the Ricci tensor is the Ricci scalar R, which is defined by taking the trace of the Ricci tensor:

$$R := R^{\alpha}{}_{\alpha}.$$

Einstein found that the object $R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R$ is proportional to $T_{\alpha\beta}$ [10]:

$$R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R \propto T_{\alpha\beta}\ ,$$

the main impetus for this is that

$$\left(R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R\right)_{;\beta} = 0 \quad \text{(see [9])}.$$

We let λ be the constant of proportionality so that

$$R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = \lambda T_{\alpha\beta}. \tag{20}$$

In analogy to Eq.(18) being compared to the Laplace equation, we would like to have only $R_{\alpha\beta}$ on the left hand side of Eq.(20) so that we may compare this to the Poisson equation (7). We write

$$R_{\alpha\beta} = \lambda T_{\alpha\beta} + \frac{1}{2}g_{\alpha\beta}R. \tag{21}$$

When there is no matter or energy present we would like the right hand side of Eq.(21) to vanish so that we are left with Eq.(18), however this is not clear from the equation above. We apply the metric inverse $g^{\beta\alpha}$ to both sides:

$$\begin{aligned} g^{\beta\alpha}R_{\alpha\beta} &= \lambda g^{\beta\alpha}T_{\alpha\beta} + \frac{R}{2}g^{\beta\alpha}g_{\alpha\beta} \\ \iff R &= \lambda T + \frac{R}{2}\delta^{\beta}{}_{\beta}\ , \end{aligned}$$

where $T := T^{\beta}{}_{\beta} = g^{\beta\alpha}T_{\alpha\beta}$ is Laue's scalar [4]. We note that, by the summation convention, $\delta^{\beta}{}_{\beta} = 4$, therefore

$$R = \lambda T + 2R \iff R = -\lambda T.$$

Going back to Eq.(21):

$$\begin{aligned} R_{\alpha\beta} &= \lambda T_{\alpha\beta} - \frac{\lambda}{2}g_{\alpha\beta}T \\ \iff R_{\alpha\beta} &= \lambda\left(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T\right). \end{aligned} \tag{22}$$

We note that when there is no energy and matter present, $T_{\alpha\beta} = 0$ (hence $T = 0$) and thus Eq.(22) reduces to Eq.(18). This is in complete alignment to the Newtonian case where if there's no matter present, $\rho = 0$, then the Poisson equation reduces to the Laplace equation. This means that Eq.(22) is the generalisation we have been searching for. To be explicit we show the comparison:

$$R_{\alpha\beta} = \lambda\left(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T\right) \quad\longleftrightarrow\quad \nabla^2\phi = 4\pi G\rho.$$

The correct constant of proportionality in S.I. units is $\lambda = 8\pi G/c^4$, this brings us to the field equations of GR:

$$R_{\alpha\beta} = \frac{8\pi G}{c^4}\left(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T\right).$$

In reference to Eq.(19) we have $P_{\alpha\beta} := \frac{8\pi G}{c^4}\left(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T\right)$. Going back to Eq.(20) we may equally write

$$G_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta}$$

where $G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R$ is the Einstein tensor. The field equations are most commonly expressed in this form, this single tensor equation consists of ten nonlinear partial differential equations in which the unknown functions are the components of the metric $g_{\alpha\beta}$. By solving these equations one uncovers the structure of the region of spacetime under consideration and yields information about the gravitational field, since it is the metric $g_{\alpha\beta}$ which represents the gravitational field in GR.

6 Conclusions

By starting with the fundamental mathematics of differential geometry, and linking the physics of gravity to this mathematics via Einstein's equivalence principle and the local-flatness theorem, we were able to heuristically build our way from Newtonian gravity to the Einstein equation. A crucial step came in comparing the equation $R_{\alpha\beta} = 0$, in our geometric formalism, to the Laplace equation $\nabla^2\phi = 0$, in Newtonian theory. Knowing that Newtonian gravity can be completely described by the Poisson equation $\nabla^2\phi = 4\pi G\rho$, which is a generalisation of the Laplace equation when there is matter present, we realised that if we can generalise the equation $R_{\alpha\beta} = 0$ to account for space which is not vacuum, then we would arrive at a complete description of the manifold when there is energy and matter present.

Our approach was not, in no doubt, mathematically rigorous — the formal mathematics of the theory is highly sophisticated and the author hopes that one day he will be able to sincerely appreciate the mathematical beauty of GR. From a heuristic perspective, however, our approach was appropriate as it favoured an intuitive understanding for how one may connect differential geometry to gravity.

References

[1] Eötvös, L. (1890). On the Gravitation Produced by Earth on Different Substances. *The Abraham Zelmanov Journal.* Vol. 1, 2008. Originally published as "Über die Anziehung der Erde auf verschiedene Substanzen". Mathematische und Naturwissenschaftliche Berichte aus Ungarn, 1890, Bd. 8, S. 65 – 68.

[2] Einstein, A. (1911). On the Influence of Gravitation on the Propagation of Light. *The Principle of Relativity.* New York: Dover Publications. Originally published as "Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichte", *Annalen der Physik*, 35, 1911.

[3] Einstein, A. Grossman, M. (1913). *Outline of a Generalized Theory of Relativity and of a Theory of Gravitation.* Leipzig: Teubner.

[4] Einstein, A. (1916). The Foundation of the General Theory of Relativity. *The Principle of Relativity*, New York: Dover Publications. Originally published as "Die Grundlage der allgemeinen Relativitätstheori", *Annalen der Physik*, 49, 1916.

[5] Einstein, A. (1921). A Brief Outline of the Development of the Theory of Relativity. *Nature*: Nature Publishing Group.

[6] Israel, W. (1973). *Relativity, Astrophysics and Cosmology.* Dordrecht, Holland: D. Reidel Publishing Company.

[7] Poisson, E. (2004). *A Relativists Toolkit, The Mathematics of Black-Hole Mechanics.* Cambridge: Cambridge University Press.

[8] Jeevanjee, N. (2015). *An Introduction to Tensors and Group Theory for Physicists.* 2nd ed. Switzerland: Springer International Publishing.

[9] Böhmer, C.G. (2017). *Introduction to General Relativity and Cosmology.* Singapore: World Scientific.

[10] Ni, W.T. (2017). Chapter 2 Genesis of general relativity — A concise exposition. *One Hundred Years of General Relativity, From Genesis and Empirical Foundations to Gravitational Waves, Cosmology and Quantum Gravity.* Singapore: World Scientific.

[11] Griffiths, D.J. (2017). *Introduction to Electrodynamics.* 4th ed. Cambridge: Cambridge University Press.