GENERAL RELATIVITY AND THE NEWTONIAN LIMIT

ALEXANDER TOLISH

Abstract. In this paper, we shall briefly explore general relativity, the branch of physics concerned with spacetime and gravity. The theory of general relativity states that gravity is the manifestation of the curvature of spacetime, so before we study the physics of relativity, we will need to discuss some necessary geometry. We will conclude by seeing how relativistic gravity reduces to Newtonian gravity when considering slow-moving particles in weak, unchanging gravitational fields.

Contents

1. What is General Relativity?

Until the begining of the twentieth century, physicists had unquestioningly accepted the notion that the universe was described by the Euclidean geometry developed two thousand years ago. In other words, they believed that space and time were both flat and conceptually distinct, allowing for precise definitions of length and simultaneity. However, Maxwell's theory of electromagnetism and experiments regarding the speed of light raised great doubts about these assumptions. It soon became clear that quantities classically regarded as invariant under an observer's change between inertial coordinate systems are not invariant at all. In 1905, Albert Einstein submitted his theory of special relativity (see [1]) to answer these failures; Euclidean space and the traditional Pythagorean distance formula were replaced by more abstract structures. The theory's primary shortcoming is its willful silence on gravity and accelerating frames of reference, which Einstein rectified with general relativity.

Gravitational mass m_g (which describes how gravity produces a force on a body) is, as far as physicists can tell, identical to inertial mass m_i (which describes how the body reacts to force). Therefore we simply call both mass. If a particle of mass $m = m_g = m_i$ is separated from a particle of mass M by a displacement $\vec{r}$ and subject to no forces but gravity, then according to Newton's laws of gravity and motion, it experiences a force $F_g = \frac{GMm_g\hat{r}}{r^2}$ (where $\hat{r}$ is the normalized vector in the direction of $\vec{r}$) which causes an acceleration $\frac{d^2\vec{r}}{dt^2} = \frac{F_g}{m_i}$. Combining the two, we find

$$\frac{d^2\vec{r}}{dt^2} = \frac{1}{m_i}\frac{GMm_g\hat{r}}{r^2} = \frac{GM\hat{r}}{r^2}\frac{m_g}{m_i} = \frac{GM\hat{r}}{r^2}. \tag{1.1}$$

Therefore mass has no visible effect on the body's responce to gravity; all particles behave the same way in the same gravitational field. This is remarkably different from gravity's nearest classical relative, electrostatics. Coulomb's law states that a particle of charge q separated from a particle of charge Q by displacement $\vec{r}$ will feel a force $F_e = \frac{KQq\hat{r}}{r^2}$, but it will still undergo acceleration $\frac{d^2\vec{r}}{dt^2} = \frac{F_g}{m_i}$; combining these, we now find

$$\frac{d^2\vec{r}}{dt^2} = \frac{1}{m_i}\frac{KQq\hat{r}}{r^2} = \frac{KQq\hat{r}}{m_i r^2} \tag{1.2}$$

A body's response to an electrostatic field depends on both charge and mass. Therefore, two different particles will act in very different ways. The other two kinds of force, the strong nuclear and weak nuclear forces, yield even more convoluted equations of motion. Gravity is somehow unique amongst the physical forces. This distinction prompted Einstein to make the hypothesis that lies at the heart of general relativity: gravity is not simply a phenomenon that takes place in spacetime; rather, gravity *is* spacetime, warped by the presence of mass and energy.

Einstein quantified this revelation with the Einstein field equation. We can interpret Newton's law of gravity as an equation describing the strength of a scalar gravitational field as a function of a mass distribution; similarly, we can say that Einstein's equation tells us how space and time curve as a function of a more abstract mass-energy distribution. If we assume the principle of least action, which states, roughly speaking, that particles in motion do not travel by needlessly complex paths, we can use the curvature to find the equation of motion for a set of initial conditions. While these findings taken alone make relativity an interesting exercise, they do not necessarily make it a substitue for Newtonian gravity: centuries' worth of experiments defend Newton's laws (at least with regard to slow-moving, light bodies). However, the relativistic equations of motion actually contain the classical equations; when we make approximations corresponding to weak, static gravitational fields and light, slow-moving particles, we recover Newton's classical laws of gravity and motion, making general relativity very attracitve as a genuine physical theory.

We will begin by establishing the mathematical structure that best models the physical universe–the differentiable manifold, along with its associated vector and dual vector spaces. We will then proceed to develop the calculus of tensors, a class of functions on manifolds which are invariant under coordinate transformations. Imposing a special metric tensor on a manifold gives it a great deal of useful structure. To begin, a given metric tensor lends itself to a unique derivative operator. This allows us to see what happens as we slide a tangent vector around the manifold, which in turn offers us a way of measuring curvature. We can also use the derivative to measure the straightness of a curve on the manifold. The straightest possible curves are called geodesics and are of great practical importance.

At this point we will introduce physics in the form of the Einstein equation. Unfortunately, it must be taken axiomatically: Einstein did not derive it from abstract mathematics, but instead chose it because it effectively describes the physical world. However, we will attempt to provide some motivation for and justification of the equation by studying how general relativity encompasses the special limiting case of Newtonian gravity. By placing a few restrictions on our relativistic mass-energy distribution and making a few approximations in our calculations, we can demonstrate that Newton's laws of gravity and motion can be considered as special cases of general relativity.

2. Mathematical Preliminaries: Manifolds

Since general relativity is the study of spacetime itself, we want to start with as few assumptions about spacetime as possible. Therefore, rather than immediately associating spacetime with $\mathbb{R}^n$, we wish to find a more general structure.

Definition 2.1. A *smooth n-dimension manifold* M is a set with a finite family of subsets $\{U_\alpha\}$ and bijections $\{\phi_\alpha : U_\alpha \to \mathbb{R}^n\}$ that satisfy the following three conditions.

- $\{U_\alpha\}$ cover M.
- $\phi_\alpha(U_\alpha)$ is open for all α.
- If $U_\alpha \cap U_\beta \neq \emptyset$, then $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is a diffeomorphism.

The pair (U_α, ϕ_α) is called a chart or a local coordinate system. The set of all charts is called an atlas.

Essentially, a smooth manifold is a structure that looks like $\mathbb{R}^n$ locally, but not globally. Every neighborhood of a point on the manifold can be described by some local coordinate system (U_α, ϕ_α), and if a point is contained in multiple coordinate systems, then it will have a neighborhood whose images in the different coordinate systems are diffeomorphic, assuring that a change of coordinates will not be too traumatic.

Definition 2.2. A *scalar field* f on M is a function which assigns a real number to every point in the manifold.

We have not yet defined the concept of distance on M, so we cannot immediately speak of the continuity or differentiablity of scalar fields on M–but we can for fields on $\mathbb{R}^n$. Let p be a point in M and (U_α, ϕ_α) be a coordinate system containing p. Then we say that the scalar field f is continuous at p if the function $(f \circ \phi_\alpha^{-1}) : \mathbb{R}^n \to \mathbb{R}$ is continuous at $\phi_\alpha(p)$. It is natural, of course, to worry that f might be continuous with respect to one coordinate system but not another, but, because we can diffeomorph between coordinate systems, this problem will not arise. If f is continuous at every p in M, we say that f is continuous on M. Similarly, we can define what it means for a field to be differentiable or smooth on M. We will denote the set of smooth scalar fields on a manifold F^∞.

Definition 2.3. A *curve* γ on M is a continuous function $\gamma : \mathbb{R} \to M$. We define continuity through the curve's image in coordinate systems, as we did with scalar fields. A curve on a manifold which we interpret as spacetime is called a *worldline.*

We are now in a position to define vectors. We are most interested in the spaces of vectors tangent to the manifold. If an n-dimensional manifold is embedded in a higher-dimensional space like $\mathbb{R}^k$, it is not difficult to express tangent vectors as vectors in the superspace. However, we are not assuming our manifolds to be embedded, so we would like to come up with a new definition which depends only on the intrinsic nature of the manifold. A vector tangent to M at p expresses the idea of motion along M through p, as does the directional derivative. We therefore associate the two.

Definition 2.4. Let M be a manifold containing point p. From the definition of a manifold, we know that p has a neighborhood contained within one coordinate system where $\mathbb{R}^n$ is described by coordinates $\{x^\mu\}$. The space spanned by the partial derivatives in this coordinate system $\{\frac{\partial}{\partial x^\mu}\}$ (or, more often, $\{\partial_\mu\}$) is called the *tangent space* V_p of p. Naturally, a single element of V_p is called a *tangent vector*.

It is not hard to see that the tangent space is indeed a vector space. Furthermore, tangent vectors as so defined can be used to describe directional derivatives. We must simply diffeomorph a curve from a manifold into a coordinate system to see that the curve's derivative can be expressed as a vector in this space; similarly, given a ventor in the tangent space at a point, we can easily construct a curve passing through that point so that its directional derivative equals the vector.

Vectors are geometrical objects; we use coordinates to describe them, but they are ultimately coordinate-independent. If $v = \sum v^\mu \partial_\mu$ in $\{x^\mu\}$, then by the chain rule, in $\{x'^\nu\}$, $\partial_\mu = \sum_\nu \frac{\partial x'^\nu}{\partial x^\mu} \partial'_\nu$; we thus require the *vector transformation law*

$$v'^\nu = \sum_\mu v^\mu \frac{\partial x'^\nu}{\partial x^\mu} \tag{2.5}$$

so that $\sum v^\mu \partial_\mu = \sum v'^\nu \partial'_\nu$, as we would expect.

A *vector field* on M is a function that assigns a tangent vector to every point on M. A vector field is *smooth* if its components are smooth functions in every coordinate system. So far, our definitions and equations have been given in terms of an individual vector, but it is easy to pointwise-generalize them to smooth vector fields as well.

Definition 2.6. The *dual (cotangent) space* V_p^* is the set of linear functions (or *dual vectors*) $\omega : V_p \to \mathbb{R}$.

If ω_1 and ω_2 are both dual vectors, then an arbitrary linear combination $a\omega_1 + b\omega_2$ is clearly also a linear function from V_p to $\mathbb{R}$, so the dual space is itself a vector space. We can say more than that, though. Since V_p^* is a vector space, $(a\omega_1 + b\omega_2)(v) = a\omega_1(v) + b\omega_2(v)$. We can consider v as a function acting on ω rather than vice versa, defining $v(a\omega_1 + b\omega_2) = av(\omega_1) + bv(\omega_2)$. It follows that a vector v is a linear function $v : V_p^* \to \mathbb{R}$; the tangent space forms the dual space to its own dual space, or $V_p = V_p^{**}$. We can repeat the process indefinitely to identify V_p with V_p^{**} and V_p^{****} and so on. Similarly, we can identify V_p^* with V_p^{***} and other odd dual spaces.

Since we have defined the dual space in terms of the tangent space, we can use the tangent space to find a convenient basis of the dual space. We define the *cobasis* $\{dx^\mu\}$ so that $dx^\mu(\partial_\nu) = \partial_\nu(dx^\mu) = \delta^\mu_\nu$ (where δ^μ_ν is the Kronecker delta).

It is easy to use the linearity of both vector spaces to show that $\{dx^\mu\}$ spans the space of dual vectors. By analogy to the vector transformation law, we find the *dual transformation law*: if $\{x^\mu\}$ and $\{x'^\nu\}$ are two coordinate systems and $\omega = \sum w_\mu dx^\mu$, then

$$w'_\nu = \sum_\mu w_\mu \frac{\partial x^\mu}{\partial x'^\nu}. \tag{2.7}$$

A *dual field* on M assigns a dual vector to every point on M. A dual field is *smooth* if the action of the field on any smooth vector field is a smooth scalar field. Like with vector fields, it is easy to show that everything we've established for dual vectors holds true for smooth dual fields.

3. Mathematical Preliminaries: Tensors and Metrics

We have defined manifolds and myrriad vector spaces upon them. Now we wish to operate upon our constructions. We can do so with tensors: just as vectors and dual vectors can be thought of as functions of one another, tensors can be thought of as groups of vectors and dual vectors that act upon other vectors and dual vectors.

Definition 3.1. Let M be a smooth manifold containing point p. We define a *tensor* T of type (n, m) on M at p as a multilinear function

$$T : \overbrace{V_p^* \times \ldots \times V_p^*}^{n} \times \overbrace{V_p \times \ldots \times V_p}^{m} \to \mathbb{R}.$$

We find that vectors are $(1, 0)$ tensors, duals are $(0, 1)$ tensors, and scalars are $(0, 0)$ tensors. It is easy to see that linear combinations of tensors will also be tensors, so the set of all tensors of type (n, m) form a vector space. Because tensors are multilinear, we can use the bases of vector and dual spaces to determine a basis for the tensor space. But first, we must introduce one of the most important tensor operators, the outer product.

Definition 3.2. Let T be a (n, m) tensor and S be a (k, l) tensor. Then their outer product $T \otimes S$ is the $(k + n, l + m)$ tensor

$$\begin{aligned}(T\otimes S)&(\omega_1, \ldots, \omega_{k+n}; v_1, \ldots, v_{l+m})\\ &= T(\omega_1, \ldots, \omega_n; v_1, \ldots, v_m)S(\omega_{n+1}, \ldots, \omega_{n+k}; v_{m+1}, \ldots, v_{m+l})\end{aligned} \tag{3.3}$$

Choose a coordinate system at p and let $\{\partial_\mu\}$ be the canonical basis of V_p and $\{dx^\nu\}$ be the cobasis of V_p^*. Then a type (n, m) basis tensor is a tensor of the form

$$\partial_{i_1} \otimes \ldots \otimes \partial_{i_n} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_m}.$$

It is not hard to show that basis tensors really do span the space of tensors. We will use the Einstein summation convention as introduced in [2]: Greek characters in a tensor's indices refers to the tensor's components and are to be summed over all possible values when they appear in both upper and lower indices, while Latin characters refer to a full tensor. For the sake of brevity, we will often refrain from explicitly writing basis tensors when expanding in components. Using this convention, we can write a (n, m) tensor T as

$$T = (T^{\mu_1,\ldots,\mu_n}_{\nu_1,\ldots,\nu_m})(\partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_n} \otimes dx^{\nu_1} \otimes \ldots \otimes dx^{\nu_m}) = T^{i_1,\ldots,i_n}_{j_1,\ldots,j_m}$$

Although we have introduced tensors as maps into $\mathbb{R}$, their multilinearity tells us that tensors can also be regarded as maps amongst other tensor spaces. For example, if we take a (n, m) tensor and insert j duals into j dual slots and k vectors into k vector slots, then the result will be a multilinear function from $n-j$ copies of V_p^* and $m-k$ copies of V_p into $\mathbb{R}$; that is, it will be a $(n-j, m-k)$ tensor.

If we take tensors to be maps from V_p and V_p^* into $\mathbb{R}$, then it follows that, because these spaces are independent of coordinate systems, tensors should also be coordinate independent. The chain rule tells us how the basis vectors and duals (and thus the basis tensors) transform, so we must compensate by adjusting the tensor components. Following the process that gave us the vector transformation law, we find the *tensor transformation law*:

$$T^{\mu'_1,\ldots,\mu'_n}_{\nu'_1,\ldots,\nu'_m} = \frac{\partial x'^{\mu_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x'^{\mu_n}}{\partial x^{\mu_n}} \frac{\partial x^{\nu_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\nu_m}}{\partial x'^{\nu_m}} T^{\mu_1,\ldots,\mu_n}_{\nu_1,\ldots,\nu_m}. \tag{3.4}$$

Definition 3.5. Let T be an (p, q) tensor on an n-dimensional smooth manifold. The *contraction* of T with respect to the αth and βth indices is the $(p-1, q-1)$ tensor

$$T^{i_1,\ldots,\sigma,\ldots,i_n}_{j_m,\ldots,\sigma,\ldots,j_m} = \sum_{\sigma=1}^{n} T(\ldots, dx^\sigma, \ldots; \ldots, \partial_\sigma, \ldots)$$

where $\{\partial_\sigma\}$ and $\{dx^\sigma\}$ are a basis of V and the associated basis of V^*, inserted into the βth vector slot and αth dual slot. It is clear from the definition of $\{dx^\sigma\}$ and the linearity of tensors that the contraction does not depend on which basis is used.

We can define a *tensor field* on M just as we defined vector and dual fields. A tensor field is *smooth* if its action on any collection of smooth vector and dual fields is a smooth scalar field. Again, everything that applies to a lone tensor also applies to tensor fields.

Definition 3.6. A *metric* g is a smooth $(0, 2)$ tensor field that is both *symmetric* and *nondegenerate*. That is, $g(v_1, v_2) = g(v_2, v_1)$ and $g(v_1, v) = 0$ for all v only when $v_1 = 0$.

Although many important metrics are positive definite, we shall not require that this be generally true. Expanding the metric with the Einstein convention, we find $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$; assuming the basis tensors, we often simply identify the metric tensor with its component matrix, $g_{\mu\nu}$.

The metric is one of the most powerful tensors we can introduce to a manifold. By interpreting it as a measure of distance between infinitesimally close points, we can generalize important concepts like length and magnitude to manifolds.

Definition 3.7. Once we have endowed a manifold with a metric, we can define an associated *inner product* $\langle v_1, v_2 \rangle$ on the tangent space. If g is a metric on M and v_1 and v_2 are both elements of V_p, then $\langle v_1, v_2 \rangle = g(v_1, v_2)$. In coordinates, $g_{ij} v_1^i v_2^j$.

Note that the two vectors must lie in the same tangent space: we shall see in the next section that it takes more preparation to analyze vectors tangent at distinct points.

With an inner product, we can find a natural isomorphism between tangent space and dual space. Associate a dual vector ω with a vector v if, for every vector u,

$\omega(u) = \langle v, u \rangle$. Let's find the form of the map in a coordinate system

$$\omega_\mu u^\mu = \omega(u) = g_{\mu\nu} v^\mu u^{nu} \tag{3.8}$$

Since u is arbitrary, we can say that the dual vector associated with v has components $\omega_\nu = g_{\mu\nu} v^\mu$. This process is called *lowering the index* of v. We can define an *inverse metric* $g^{\mu\nu}$ as the inverse of the matrix $g_{\mu\nu}$. It follows that $v^i = g^{ij}\omega_j$: this is *raising the index* of ω. We can use the metric and inverse metric to raise and lower the indices of arbitrary tensors; for example, $T^{\alpha\ \gamma\delta\epsilon}_{\ \beta} = g_{\beta\mu} g^{\delta\nu} g^{\epsilon\rho} T^{\alpha\mu\gamma}_{\ \ \ \ \nu\rho}$.

4. Curvature

If gravity is the manifestation of curved spacetime, then we will need a mathematically rigorous, quantitative notion of curvature. We will find it in the Riemann curvature tensor. To measure curvature, we need to know how tangent vectors behave when slid around the manifold; this information is contained within a special map, the covariant derivative.

Definition 4.1. A *covariant derivative* or *derivative operator* on a manifold M is a map ∇ from smooth (n, m) tensor fields $T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ to smooth $(n, m+1)$ tensor fields $\nabla_{j_{m+1}} T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ that satisfies the following conditions.

It is linear:
$\nabla_{j_{m+1}}(\alpha A^{i_1,\dots,i_n}_{j_1,\dots,j_m} + \beta B^{i_1,\dots,i_n}_{j_1,\dots,j_m}) = \alpha\nabla_{j_{m+1}}(A^{i_1,\dots,i_n}_{j_1,\dots,j_m}) + \beta\nabla_{j_{m+1}}(B^{i_1,\dots,i_n}_{j_1,\dots,j_m})$

It obeys the Leibnitz rule for derivatives:

$$\nabla_{j_{m+1}}(A^{i_1,\dots,i_n}_{j_1,\dots,j_m} \otimes B^{i_1,\dots,i_n}_{j_1,\dots,j_m}) = (\nabla_{j_{m+1}} A^{i_1,\dots,i_n}_{j_1,\dots,j_m}) \otimes (B^{i_1,\dots,i_n}_{j_1,\dots,j_m}) + (A^{i_1,\dots,i_n}_{j_1,\dots,j_m}) \otimes (\nabla_{j_{m+1}} B^{i_1,\dots,i_n}_{j_1,\dots,j_m})$$

It commutes with contraction:
$\nabla_{j_{m+1}}(A^{i_1,\dots,c,\dots,i_n}_{j_1,\dots,c,\dots,j_m}) = (\nabla_{j_{m+1}} A)^{i_1,\dots,c,\dots,i_n}_{j_1,\dots,c,\dots,j_m}$

It acknowledges vectors as directional derivatives on scalar fields:
For all v in V_p and f in F^∞, $v^i \nabla_i f = v(f)$

It is torsion free (it commutes with itself on smooth scalar fields):
For all f in F^∞, $\nabla_i \nabla_j f = \nabla_j \nabla_i f$

Since this definition gives us conditions on the covarient derivative rather than instructions how to calculate it, it follows that there are many different maps that can be called derivative operators. We are left wondering whether there are any useful features all derivatives share.

Lemma 4.2. *Let ∇ and $\widetilde{\nabla}$ be two derivative operators, f be a smooth scalar field, and $T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ be a smooth tensor field, all defined at a point p in M. Then the following conditions are true.*

(1) $(\nabla - \widetilde{\nabla})$ *is also a derivative.*
(2) $(\nabla - \widetilde{\nabla})_i(f) = 0$.
(3) $(\nabla - \widetilde{\nabla})_{j_{m+1}}(f T^{i_1,\dots,i_n}_{j_1,\dots,j_m}) = f(\nabla_{j_{m+1}} T^{i_1,\dots,1_n}_{j_1,\dots,j_m} - \widetilde{\nabla}_{j_{m+1}} T^{i_1,\dots,i_n}_{j_1,\dots,j_m})$.
(4) $(\nabla - \widetilde{\nabla})_{j_{m+1}}(T^{i_1,\dots,i_n}_{j_1,\dots,j_m})$ *depends only on* $T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ *'s value at p.*

Proof. The first three statements are easily derived from the definition of covariant derivatives. The proof of the fourth statement is more complicated and uses a technique we will need again later, so it is worth showing.

Let $H^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ be a smooth tensor field equal to $T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ at p but otherwise arbitrary. Then $T^{i_1,\dots,i_n}_{j_1,\dots,j_m} - H^{i_1,\dots,i_n}_{j_1,\dots,j_m} = Q^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ is a new smooth tensor field equal to zero at p. If $\{\theta^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m}\}$ is the set of basis tensors, then $Q^{i_1,\dots,i_n}_{j_1,\dots,j_m} = f^{\nu_1,\dots,\nu_m}_{\mu_1,\dots,\mu_n}\theta^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m}$ ($f^{\nu_1,\dots,\nu_m}_{\mu_1,\dots,\mu_n}$ is not a tensor, but a double-indexed set of scalar fields that give the component coefficients of the difference tensor, so $f^{\nu_1,\dots,\nu_m}_{\mu_1,\dots,\mu_n}|_p = 0$ for every set of indices). Now let us differentiate this tensor.

$$\begin{aligned}
&(\nabla_{j_{m+1}} - \widetilde{\nabla}_{j_{m+1}})(Q^{i_1,\dots,i_n}_{j_1,\dots,j_m})|_p \\
&\quad = \nabla_{\nu_{m+1}}(f^{\nu_1,\dots,\nu_m}_{\mu_1,\dots,\mu_n}\theta^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m})|_p - \widetilde{\nabla}_{j_{m+1}}(f^{\nu_1,\dots,\nu_m}_{\mu_1,\dots,\mu_n}\theta^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m})|_p \\
&\quad = f^{\nu_1,\dots,\nu_m}_{\mu_1,\dots,\mu_n}|_p(\nabla_{\nu_{m+1}}\theta^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m}|_p - \widetilde{\nabla}_{\nu_{m+1}}\theta^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m}|_p) \\
&\quad = 0
\end{aligned}$$

The last equality tells us that $(\nabla_{j_{m+1}} - \widetilde{\nabla}_{j_{m+1}})T^{i_1,\dots,i_n}_{j_1,\dots,j_m} = (\nabla_{j_{m+1}} - \widetilde{\nabla}_{j_{m+1}})H^{i_1,\dots,i_n}_{j_1,\dots,j_m}$. However, we only know one thing about $H^{i_1,\dots,i_n}_{j_1,\dots,j_m}$: it has the same value as $T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ at p. We can therefore say that the difference of derivatives depends only on the value of a tensor at a single point. □

The derivative difference $(\nabla_i - \widetilde{\nabla}_i)$ linearly maps dual vectors (not necessarily dual fields: we saw above that this derivative depends only on a tensor's value at a single point) to $(0,2)$ tensors; specifying two vectors takes us linearly to $\mathbb{R}$, so the derivative difference can be thought of as a multilinear map $(\nabla_i - \widetilde{\nabla}_i) : V_p^* \times V_p \times V_p \to \mathbb{R}$. We have not been rigorous about coordinate independence, so we cannot call this a true $(1,2)$ tensor, but it is close enough that we are justified in using a similar notation, $C^c_{a,b}$. It follows that

$$\nabla_a \omega_b = \widetilde{\nabla}_a \omega_b + C^c_{ab}\omega_c.$$

The gradient of a scalar field $\nabla_b f$ is a dual vector, we can substitute it into the above equation to find

$$\nabla_a \nabla_b f = \widetilde{\nabla}_a \widetilde{\nabla}_b f + C^c_{ab}\nabla_c f. \tag{4.3}$$

Since derivatives commute on scalar fields, the derivatives in equation (4.3) are symmetric, so C^c_{ab} is symmetric in its lower indices as well: $C^c_{ab} = C^c_{ba}$. If v is a vector, then $v^b\omega_b$ is a scalar field for any dual vector ω; letting $(\nabla_a - \widetilde{\nabla}_a)$ act on $v^b\omega_b$, we can combine statement 2 of the lemma with the Leibnitz rule to find

$$\nabla_a v^b = \widetilde{\nabla}_a v^b - C^b_{ac}v^c, \tag{4.4}$$

where C^c_{ab} is the same as above. This pattern can be easily generalized to arbitrary tensors:

$$\nabla_a T^{b_1,\dots,b_n}_{c_1,\dots,c_m} = \widetilde{\nabla}_a T^{b_1,\dots,b_n}_{c_1,\dots,c_m} + \sum_j C^d_{ac_j} T^{b_1\dots,b_n}_{c_1,\dots,d,\dots,c_m} - \sum_i C^{b_i}_{ad} T^{b_1,\dots,d,\dots,b_n}_{c_1,\dots,c_m} \tag{4.5}$$

Definition 4.6. Let (U_α, ϕ_α) be a local coordinate system on manifold M with bases $\{\partial^\mu\}$ and $\{dx_\mu\}$ and $T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ be a tensor with components $T^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m}$. Then the *ordinary derivative* $\partial_{j_{m+1}}$ is the derivative operator associated with (U_α, ϕ_α) such that $\partial_{j_{m+1}} T^{i_1,\dots,i_n}_{j_1,\dots,j_m}$ is the tensor with components $\partial(T^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_m})/\partial x^{\nu_{m+1}}$.

Because the ordinary derivative is a set of partial derivatives, it easily obeys the covariant derivative conditions we specified above. However, it obviously depends on the coordinate system we are working in, so the ordinary derivative is not as powerful as we would like. All the same, we can use it to find a more useful derivative by substituting it into equation (4.4):

$$\nabla_a v^b = \partial_a v^b + \Gamma^b_{ac} v^c, \tag{4.7}$$

The *Christoffel symbol* Γ^b_{ac} is, in essence, no more than a special case of the derivative difference C^b_{ac} defined above, but it happens to be an extreemly important special case. If we define ∇ as a coordinate-independent covariant derivative, then ∇'s Christoffel symbols in a certain coordinate system tell us how to obtain ∇ from the ordinary (partial) derivatives of that coordinate system. Since the ordinary derivatives are so easy to calculate, we can effectively identify a covarient derivative with its Christoffel symbols in a certain coordinate system.

Now, so far we have not specified any specific coordinate-independent derivative as particularly useful. We are finally in a position to use the metric to do so.

Theorem 4.8. *Let g_{bc} be a metric on M. There exists a unique covariant derivative ∇, the Levi-Civita covariant derivative, such that $\nabla_a g_{bc} = 0$.*

Proof. Assume there exists some ∇ such that $\nabla_a g_{bc} = 0$. We will show that, given a coordinate system, there exists a unique, well-defined set of Christoffel symbols describing ∇_a. To begin, expand our assumption with (4.5):

$$0 = \nabla_a g_{bc} = \partial_a g_{bc} + \Gamma^d_{ab} g_{dc} + \Gamma^d_{ac} g_{bd}.$$

Although the Christoffel symbols are not true tensors because they are coordinate dependent, they do behave like tensors within one coordinate system. Thus it is not too egregrius an abuse of notation to interpret these products as the lowering of the indices of the Christoffel symbols:

$$\Gamma_{cab} + \Gamma_{bac} = \partial_a g_{bc}. \tag{4.9}$$

We can permute the indices to find

$$\Gamma_{cba} + \Gamma_{abc} = \partial_b g_{ac}, \tag{4.10}$$

$$\Gamma_{bca} + \Gamma_{acb} = \partial_c g_{ab}. \tag{4.11}$$

By adding (4.9) and (4.10) and subtracting (4.11) and employing the Christoffel symbols' symmetry in their second and third indeces, we can solve for one symbol. We can then raise the first index to give an equation for the true symbol:

$$\Gamma^c_{ab} = \frac{1}{2} g^{cd}(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}). \tag{4.12}$$

As we saw above, the ordinary derivatives and Christoffel symbols determine the full derivative; a given coordinate system carries with it the ordinary derivatives, and these combine with the metric to give unique Christoffel symbols. It follows that any set of numbers satisfying equation (4.12) will serve as the Christoffel symbols of an appropriate derivative, justifying our assumption that the derivative exists. □

From now on, unless we state otherwise, the symbol ∇ will refer to the Levi-Civita covariant derivative.

An explicit calculation of the derivative's symbols themselves depends on our choice of coordinate system. However, the only real assumption we made in the proof was that $\nabla_a g_{bc} = 0$, which does not depend on coordinates; we can conclude that the full derivative ∇_a is coordinate independent and that the Christoffel symbols depend on coordinates in such a way that they cancel the ordinary derivatives' dependence on coordinates.

Now that we have a metric and a metric-specific derivative, there are multiple ways to derive a quantitative way of measuring the curvature of the metric. Unfortunately, all of these methods still bear the scars of their evolution from geometric intuition to mathematical rigour: they retain little of the pictoral simplicity that inspired them in the first place. One aproach is the path dependence or independence of vector translation across the manifold. In flat Euclidean space, a vector is unchanged after sliding around a closed curve. In a curved space, such as the surface of a sphere, this is not so: imagine a vector on the equator of the Earth, pointing east. Slide it (without rotating it) along a line of latitude to the North pole, then down the opposite line of latitude back onto the equator. Slide it along the equator to its starting point: it will now be pointing west. Such a manipulation cannot be done on $\mathbb{R}^n$.

Since the Levi-Civita derivative is a well-behaved tensor that measures motion in a certain direction on a manifold, our definition of curvature should be related to the commutator of derivatives. Using the Leibniz rule, we can easily show that for any smooth scalar field f, dual field ω_c, and basis vector fields a and b,

$$[\nabla_a, \nabla_b](f\omega_c) = (\nabla_a\nabla_b - \nabla_b\nabla_a)(f\omega_c) = f(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c,$$

Following the logic of Lemma 4.2, we find that $[\nabla_a, \nabla_b]\omega_a$ depends only on ω_a at a single point, not in a neighborhood. Therefore, it is a linear map from dual vectors to $(0, 3)$ tensors; specifying three vectors takes us to $\mathbb{R}$, so we can say that $[\nabla_a, \nabla_b] : V_p^* \times V_p \times V_p \times V_p \to \mathbb{R}$ is a $(1, 3)$ tenosr. Since the Levi-Civita derivative is coordinate-independent, this commutator is indeed a true tensor.

Definition 4.13. The *Riemann tensor* on M is the $(1, 3)$ tensor R^d_{abc} such that for any cobasis dual field ω_c on M $R^d_{abc}\omega_d = \nabla_a\nabla_b\omega_c - \nabla_b\nabla_a\omega_c$ (where the indices simply refer to coordinates of ω). The Riemann tensor is obviously antisymmetric in its first two lower indices. If the Riemann tensor is zero at every point on M, then M is said to be a *flat manifold.*

By substituting the Christoffel expansion (4.7) into the definition of the Riemann tensor, we find an explicit equaiton for the Riemann tensor:

$$R^\sigma_{\mu\nu\lambda} = \partial_\nu\Gamma^\sigma_{\mu\lambda} - \partial_\mu\Gamma^\sigma_{\nu\lambda} + \Gamma^\xi_{\mu\lambda}\Gamma^\sigma_{\xi\nu} - \Gamma^\xi_{\nu\lambda}\Gamma^\sigma_{\xi\mu}. \tag{4.14}$$

Example 4.15. Let M be $\mathbb{R}^n$ covered by one chart with the standard coordinates and the traditional Euclidean metric $g_{\mu\nu} = \delta^\nu_\mu$. Because the metric is the same throughout all M, (**??**) tells us that all the Christoffel symbols are zero. By (4.7), then, the Levi-Civita derivative ∇_μ is simply the partial derivative $\partial_m u$; partial derivatives commute, so $[\nabla_a, \nabla_b]\omega = 0$ for all vector fields and dual fields. Therefore $R^d_{abc} = 0$ as well. This space is flat.

Example 4.16. Let M be $\mathbb{R}^4$ covered by one chart with the standard coordinates. Furthermore, let us endow M with the metric

$$g_{ij} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This is the spacetime of special relativity. It is called *Minkowski space* and the metric, often denoted $\eta_{\mu\nu}$, is called the *Minkowski metric.* We can use the same method as above to show that Minkowski space is flat.

For the sake of convenience, we introduce a few new fields closely associated with the Riemann tensor.

Definition 4.17. The *Ricci tensor* R_{ac} is the contraction of the Riemann tensor with respect to the second lower and the upper index: $R_{\alpha\gamma} = R^{\beta}_{\alpha\beta\gamma}$. The *scalar curvature* R is the contraction of the Ricci tensor: $R = R^{\mu}_{\mu} = g^{\mu\nu}R_{\mu\nu}$. The *Einstein tensor* G_{ab} is the tensor with the components $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$.

We are almost in a position to use these measures of curvature to do physics. First, however, we will demonstrate a few useful relations. Recall the Jacobi identity:

$$[[\nabla_a, \nabla_b], \nabla_c] + [[\nabla_b, \nabla_c], \nabla_a] + [[\nabla_c, \nabla_a], \nabla_b] = 0,$$

where $[\cdot, \cdot]$ is the commutator. It follows from the covariant derivative's definition that it satisfies the Jacobi identity. We can manipulate this equation to find the *(second) Bianchi identity*:

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

Multiply both sides of the Bianchi identity by $g^{\nu\sigma}g^{\mu\lambda}$, rearange the second Riemann tensor using antisymmetry, and contract twice to find $\nabla^\mu R_{\rho\mu} - \nabla_\rho R + \nabla^\nu R_{\rho\nu} = 0$, or, renaming dummy indices,

$$\nabla^\mu R_{\rho\mu} - \frac{1}{2}\nabla_\rho R = 0. \tag{4.18}$$

If we apply this equation to the Einstein tensor, we find the important identity

$$\nabla^a G_{ab} = 0, \tag{4.19}$$

Where $\nabla^a G_{ab}$ is simply the tensor $\nabla_a G_{ab}$ with its first index raised.

Now that we have a way to describe a manifold's global structure, we are entitled to ask if we can generalize Euclidean concepts such as straight lines and length to curved spaces. Intuitively, a straight line shouldn't bend with respect to itself; we can formalizing this by requiring that the tangent vectors of a geodesic–the generalization of a straight line to curved spaces–have a zero derivative with respect to themselves..

Definition 4.20. Let M be a manifold with a metric. A *geodesic* on M is a curve $\gamma(\lambda)$ whose tangent vectors $v^a(\lambda)$ satisfy the equation $v^a\nabla_a v^b = 0$ for all λ.

If we describe γ with local coordinates $\{x^\mu\}$, straightness manifests itself as linearity: $\frac{d^2x^\mu}{d\lambda^2} = 0$. We can expand this equation with the multidimensional chain rule to see that

$$\frac{d^2x^\mu}{d\lambda^2} = \frac{dx^\nu}{d\lambda}\partial_\nu\frac{dx^\mu}{d\lambda} = 0.$$

We wish to make this equation coordinate-independent. We can do this by replacing the partial derivative ∂_ν with the covariant derivative ∇_ν. If we then expand ∇_ν with (4.7), we find the *geodesic equation*:

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0. \tag{4.21}$$

Because a general metric is not positive definite, an attempt to measure the length of a curve in the traditional way may result in a negative number. Because the thought of a negative distance is so disturbing, we rename this number the curve's proper time.

Definition 4.22. Let M be a manifold with a metric g_{ab}. The *proper time* τ of a curve $\gamma(\lambda)$ with tangent vectors $v^a(\lambda)$ is given by the following integral:

$$\tau = \int_\gamma \sqrt{g_{\mu\nu}v^\mu(\lambda)v^\nu(\lambda)}d\lambda. \tag{4.23}$$

The proper time along a curve is essential in defining relativistic quantities like velocity and momentum, as we shall soon see.

5. Einstein's Equation

We are finally in a position to use our knowledge of differential geometry and tensor calculus to describe the physical world. But before we begin to explore the bridge between the mathematics and the physics, Einstein's field equation, we must introduce a few important physical quantities, culminating in the stress-energy tensor. But before we begin, it is worth introducing some notation often used in relativity. If a point in a spacetime manifold is described by a local coordinate system $\{\partial_\mu\}$, it is standard to use the first coordiante x^0 to describe time and the second, third, and fourth coordinates $x^{\mu'}$ to describe space.

Definition 5.1. A body's *four-velocity* U^a with respect to a coordinate system $\{x^\mu\}$ is the vector with the components $U^\mu = \frac{dx^\mu}{d\tau}$, where τ is the proper time of the body's motion.

Definition 5.2. A body's *four-momentum* p^μ with respect to a coordinate system $\{x^\mu\}$ is the vector with the components $p^\mu = mU^\mu$, where m is the body's mass and U^a is the body's four-velocity.

Definition 5.3. The *stress-energy tensor* T^{ab} is the $(2,0)$ tensor whose (μ,ν) component $T^{\mu\nu}$ is given by the flux of μ^{th} component of the four-momentum vector across a surface of constant x^ν coordinate. Thus T^{00} is the flow of relativistic mass through time, the spatial mass density; $T^{\mu 0}$ is the flow of the μ^{th} component of spatial momentum through time, the spacial μ-momentum density; $T^{0\nu}$ is the flow of relativistic mass through a surface of constant coordinate x^ν, the ν-flux of mass; the rest of the components are the mechanical stresses.

We can differentiate T^{ab} to see how it changes in certain direction. The quantity $\nabla_\nu T^{\mu\nu}$ accounts for changes in the μ^{th} component of momentum in every direction and, within one coordiante system, the components of momentum must be conserved. Therefore conservation of momentum is enshrined in the tensor equation

$$\nabla_\nu T^{\mu\nu} = 0. \tag{5.4}$$

The stress-energy tensor serves a role in general relativity similar to that of mass distribution in Newtonian physics; it tells space how to deform, creating what we observe as gravity. Similarly, the Riemann tensor and its contractions are conceptually related to the classical gravitational field. Let's work on extending this analogy between Newtonian and relativistic gravitation. If Φ is the gravitational field and ρ is the mass density, then classical field equation of gravity, derived from Newton's law by advanced calculus, is

$$\nabla^2\Phi = 4\pi G\rho, \tag{5.5}$$

where G is the constant from Newton's law of gravitation. We might hope that the equations governing relativity are not too far removed from the equation of classical gravity. Let us try–admittedly naievely–simply substituting relativistic quantities in for their classical analogues. Doing so will not yeild the true Einstein equation, but it will point us in the right direction. Relating the mass distribution to the stress-energy tensor and the gravitational field to the Ricci tensor (we choose the Ricci tensor and not the Riemann tensor so that both sides of the resulting tensor equation would be of the same type), we find

$$R_{\mu\nu} = \alpha T_{\mu\nu} = \alpha g_{\mu\mu} g_{\nu\nu} T^{\mu\nu}. \tag{5.6}$$

Rather than require that the constant of proportionality be the same in both theories of gravity, we have replaced it with the arbitrary constant α. In fact, Einstein himself briefly considered using this equation, but its faults soon became apparent. If we try to enforce the conservation of momentum by requiring that $\nabla_b T^{ab}$, it follows that $\nabla^b R_{ab} = 0$; by (4.18), $\nabla_a R = 0$. We can contract (5.6) to show that this implies $\nabla_a T = 0$ (where T is the contraction of the stress-energy tensor), or T is constant throughout space. This is an unrealistic requirement to place on the universe, so we deduce that (5.6) is not true: the Ricci tensor is not proportional to the stress-energy tensor.

However, we know another $(0, 2)$ tensor intricately related to the Riemann tensor which does have a zero derivative, honoring the conservation of momentum: the Einstein tensor G_{ab}. If we try using it as the modern description of the gravitational field, we find the *Einstein equation*:

$$G_{\mu\nu} = \alpha T_{\mu\nu}.$$

The Einstein equation is usually given in one of the following equivalent forms:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \alpha T_{\mu\nu} \tag{5.7}$$

or

$$R_{\mu\nu} = \alpha(T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu}). \tag{5.8}$$

Although the Einstein equation is slightly more complicated than (5.6), it satisfies the conservation of momentum and any other such tests applied to it. We therefore take it to be our leading candidate for the relativistic gravity equation. However, the most important test still remains: since relativity is a physical theory, it can never be proved with absolute certainty. It can only be checked against physical observations. Four hundred years of science stand behind Newton's law of gravity and its field equation (5.5) when studying classical phenomena. Therefore, if (5.7) is supposed to supersede (5.5) as the general law of gravitation, it must include (5.5) as its limit when describing slow-moving, relatively light bodies.

6. The Newtonian Limit

Newtonian gravity consists of two equations: one tells us how matter responds to gravity, and the other tells us how matter produces gravity. The first equation, derived from Newton's second law of motion, says

$$\vec{a} = -\nabla\Phi, \tag{6.1}$$

where Φ is the gravitational potential, ∇ is the Euclidean gradient operator, and $\vec{a}$ is acceleration through space. The second equation, derived from Newton's law of gravity, says

$$\nabla^2\Phi = 4\pi G\rho, \tag{6.2}$$

where ∇^2 is the Euclidean Laplacian operator, G is a physical constant, and ρ is the mass density distribution. We wish to show that, when we are considering classical systems (that is, particles are neither too fast moving nor too massive, and the field is weak and static), general relativity (as embodied in the Einstein equation (5.7) and the geodesic equation (4.21)) reduces to these two statements.

We are considering relatively simple, stable matter-energy distributions that don't carry a net flow of force, momentum, or energy through space. Therefore, refering to (5.3), we see that the only nonzero component of the stess-energy tensor is T^{00}, the flux of the 0^{th} component of momentum (mass-energy) in the x^0 (time) direction. The flow of mass-energy through time is simply the spatial mass-energy density, ρ. Therefore the stress-energy tensor is given (or at least very closely approximated) by

$$T_{ij} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.3}$$

By Einstein's equation, the only nonzero component of the Einstein tensor will be G_{00}. Therefore, in this analysis we will be concerned primarily with the $(0, 0)$ components of tensors.

If the gravitational field is weak enough, then spacetime will be only slightly deformed from the gravity-free Minkowski space of special relativity (see 4.16), and we can consider the spacetime metric as a small perturbation from the Minkowski metric:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| << 1. \tag{6.4}$$

Since the components of the correction tensor h_{ij} are so small, we can use η_{ij} and its inverse to raise and lower indices without including any h term. In particular, we can raise the indices of h_{ij} to h^{ij}; we find that $(\eta^{\mu\sigma} - h^{\mu\sigma})g_{\sigma\nu} = \delta^{\nu}_{\mu}$ to within first order of h_{ij}, so we set $\eta^{\mu\nu} - h^{\mu\nu} = g^{\mu\nu}$, defining an inverse metric.

Let's begin by showing that the geodesic equaiton (4.21) reduces to Newton's second law (6.1). The geodesic equation describes the worldline of a particle acted upon only by gravity. When we require that a particle move slowly, as we do in the Newtonian limit, we mean that the time-component of the particle's four-velocity dominates all the other components. That is, if the 0^{th} component of a certain coordinate system is the time component and $\{x^{\mu'}\}$ are the space components, then

$$\frac{dx^{\mu'}}{d\tau} << \frac{dx^0}{d\tau}.$$

Recall the geodesic equation, using proper time as the parameter of the worldline:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0.$$

The second term hides a sum in ρ and σ over all indices. Because the particle in question is moving slowly, every term containing one or two spatial four-velocity components will be dwarfed by the term containing two time components. We can therefore take the approximation

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{dx^0}{d\tau}\right)^2 = 0. \tag{6.5}$$

Equation (4.12) tells us how to calculate the Levi-Civita covariant derivative's Christoffel symbols with respect to a given coordinate system. In particular,

$$\begin{aligned}\Gamma^\mu_{00} &= \frac{1}{2}\eta^{\mu\nu}\left(\partial_0 g_{\nu 0} + \partial_0 g_{0\nu} - \partial_\lambda g_{00}\right)\\ &= -\frac{1}{2}\eta^{\mu\nu}\partial_\nu g_{00},\end{aligned}$$

because the field is static and the time derivative $\partial_0 g_{\mu\nu}$ is zero. We know that $g_{00} = \eta_{00} + h_{00} = 1 + h_{00}$, so

$$\Gamma^\mu_{00} = -\frac{1}{2}\eta^{\mu\nu}\partial_\nu h_{00}. \tag{6.6}$$

Substituting (6.6) back into (6.5), we find

$$\frac{d^2x^\mu}{d\tau^2} = \frac{1}{2}\eta^{\mu\nu}\partial_\nu h_{00}\left(\frac{dx^0}{d\tau}\right)^2. \tag{6.7}$$

$\partial_0 g_{00} = 0$, so $\partial_0 h_{00} = 0$ as well. Therefore $\frac{d^2x^0}{d\tau^2} = 0$ and $\frac{dt}{d\tau}$ is constant: a particle in a classical gravitational system feels time progress onwards at a steady rate, as was unquestioningly assumed until Einstein introduced the theory of special relativity. Turning our attention to the spatial ($\mu \neq 0$) equations, we find

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

Multiply both sides of the equation by $\left(\frac{d^2\tau}{dt^2}\right)^2$. We know from the derivative rule for inverses that this cancels the $\left(\frac{d^2t}{d\tau^2}\right)^2$ on the right side; the left side looks like $\frac{d^2x^{\mu'}}{dt^2}$ expanded by the chain rule. Therefore, we find

$$\frac{d^2x^{\mu'}}{dt^2} = \frac{1}{2}\partial_{\mu'}h_{00}.$$

If we require $h_{00} = -2\Phi$ (or, equivalently, $g_{00} = -(1 + 2\Phi)$), then this equation becomes

$$\frac{d^2x^{\mu'}}{dt^2} = -\partial_{\mu'}\Phi,$$

which is simply (6.1) in component form.

We have seen how, in classical conditions, a particle's relativistic worldline through a gravitational field looks like the trajectory plotted by Newton's second law of motion, or that (4.21) contains (6.1). We still have to show that a classical mass-energy distribution will curve spacetime in a way that produces the gravitational field predicted by Newton's law of gravity so that (5.8) contains (6.2). Recall that we require our stress-energy tensor to be of the form (6.3). We find the trace of T_{ij} by contraction:

$$T = T^{\mu}_{\mu} = \eta^{\mu\mu} T_{\mu\mu} = \eta^{00} T_{00} = -\rho.$$

Substituting T_{ij}, T, and g_{ij} into (5.8), we find

$$R_{00} = \frac{1}{2}\alpha\rho. \tag{6.8}$$

But R_{00} is just the contraction $R^{\mu}_{0\mu 0}$; recalling the expansion of the Riemann tensor (4.14), we can write

$$R_{00} = R^{\mu}_{0\mu 0} = \partial_{\mu}\Gamma^{\mu}_{00} - \partial_0 \Gamma^{\mu}_{\mu 0} + \Gamma^{\mu}_{\mu\xi}\Gamma^{\xi}_{00} - \Gamma^{\mu}_{o\xi}\Gamma^{\xi}_{\mu 0}.$$

The second term of the expansion is a time derivative; the field is static, so this term is zero. The third and fourth terms are of the form $\Gamma^{c}_{ab}\Gamma^{f}_{de}$. We are interested in expressions linear in the metric, but (4.12) tells us that Γ^{c}_{ab} is already first-order in g_{ij}, so $\Gamma^{c}_{ab}\Gamma^{f}_{de}$ will be second-order in g_{ij} and thus negligible in our calculations. Therefore, to first-order in g_{ij},

$$R_{00} = \partial_{\mu}\Gamma^{\mu}_{00}. \tag{6.9}$$

We have already calculated this Christoffel symbol above. Combining (6.6), (6.8), and (6.9), we find

$$\eta^{\mu\nu}\partial_{\mu}\partial_{\nu} h_{00} = -\alpha\rho. \tag{6.10}$$

All the off-diagonal entries of the Minkowski metric matrix are zero, so we need not concern ourselves with the corresponding terms in our Einstein sum. Since the gravitational field is static, its time derivative $\partial_0 h_{00}$ is also zero. Therefore, the only nonzero terms are those with a double spatial index $\mu'\mu'$ and $\eta^{\mu'\mu'} = 1$, so we can explicitly expand (6.10):

$$\partial_1\partial_1 h_{00} + \partial_2\partial_2 h_{00} + \partial_3\partial_3 h_{00} = -\alpha\rho. \tag{6.11}$$

The expression $\partial_1\partial_1 h_{00} + \partial_2\partial_2 h_{00} + \partial_3\partial_3 h_{00}$ is simply the Euclidean Laplacian operator $\nabla^2 h_{00}$. Furthermore, we saw above that we can relate the relativistic geodesic equation to Newton's second law if we require $h_{00} = -2\Phi$. Now we can simplify (6.11) into a very suggestive form:

$$2\nabla^2\Phi = \alpha\rho. \tag{6.12}$$

Equation (6.12) has the same mathematic form as the field variation of Newton's law of gravity, (6.2). All we have left to do is set the constant $\alpha = 8\pi G$ to find

$$\nabla^2\Phi = 4\pi G\rho. \tag{6.13}$$

In a classical system, the equations of relativity–the geodesic and Einstein equations–reduce to the equations of Newtonian gravity–Newton's second laws of motion and gravity. General relativity thus passes its most important test: in agreeing with Newtonian gravity, it agrees with centuries' worth of experimental data. However, we have done more than check that relativity contains Newton's laws; we have found a single satisfactory value for the constant α, allowing us to finally write Einstein's equation in its fullest form:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{6.14}$$

or, equivalently,

$$R_{\mu\nu} = 8\pi G\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right). \tag{6.15}$$

One might argue that we were wrong to assume that the proportion between the stress-energy tensor and Einstein's tenosr, α, is indeed a constant. There is no mathematical way to refute such an argument–we cannot derive physical equations, we must find and constantly refine them. However, we have good reason to believe that it is a fixed constant. General relativity makes many interesting physical predictions, such as black holes and gravitational waves; although such topics require mathematical preparation beyond the scope of this paper, they are based on Einstein's equation as formulated above. Scientific observations are in tremendous agreement with Einstein's equation as we found it.

Acknowledgements. I owe a great debt of gratitude to my mentor, Artem Pulemotov, for the countless hours he spent with me offering much-needed guidance, reviewing material, and revising my work. It takes a great deal of time and energy to introduce the subtleties of mathematical writing into the cluttered mind of a physics student, but he always found time to work with me. Without Artem's constant help and advice, this paper would certainly not have been possible.

References

[1] Albert Einstein. *The Principle of Relativity.* Trans. W. Perrett and G. B. Jeffery. Reprinted in *On the Shoulders of Giants.* Ed. Stephen Hawking. Running Press. 2002.

[2] Robert M. Wald. *General Relativity.* University of Chicago Press. 1984.