

Formulation of Einstein Field Equation Through Curved Newtonian Space-Time

Austen Berlet

Lord Dorchester Secondary School
Dorchester, Ontario, Canada

Abstract

This paper discusses a possible derivation of Einstein's field equations of general relativity through Newtonian mechanics. It shows that taking the proper perspective on Newton's equations will start to lead to a curved space time which is basis of the general theory of relativity. It is important to note that this approach is dependent upon a knowledge of general relativity, with out that, the vital assumptions would not be realized.

Note: A number inside of a double square bracket, for example $[[1]]$, denotes an endnote found on the last page.

1. Introduction

The purpose of this paper is to show a way to rediscover Einstein's General Relativity. It is done through analyzing Newton's equations and making the conclusion that space-time must not only be realized, but also that it must have curvature in the presence of matter and energy.

2. Principal of Least Action

We want to show here the Lagrangian action of limiting motion of Newton's second law ($F=ma$). We start with a function q mapping \mathbb{R} to \mathbb{R}^n space of n dimensions and we equip it with a standard inner product.

$$q : \mathbb{R} \rightarrow (\mathbb{R}^n, (\cdot, \cdot)) \quad (1)$$

We take a function (q) between q_0 and q_1 and look at the ds of a section of the curve. We then look at some properties of this function (q). We see that the classical action of the functional (L) of q is equal to $\int ds$, L denotes the systems Lagrangian.

$$L[q] = \int ds \quad (2)$$

We expand out the $\int ds$ in (2) and find

$$L[q] = \int ds = \int \sqrt{(dq, dq)} = \int dt \sqrt{\left(\frac{dq}{dt}, \frac{dq}{dt}\right)} \quad (3)$$

Which is the inner product of two velocity vectors.

$$\int dt \sqrt{(\dot{q}, \dot{q})} \quad (4)$$

Now we set $t = t(s)$, where t is our old parameter and $t(s)$ is our new parameter, to show that L is parameter invariant. To do this we look at the inner product of \dot{q} . Which leads us to

$$(q, q) = 2\pi R \quad (5)$$

Then we find:

$$q(t) = R \begin{pmatrix} \cos 2\pi t \\ \sin 2\pi t \end{pmatrix} \quad (6)$$

We add the time dimension of $q(t)$ because we want to work with the idea of space-time so we combine space and time to get an idea of what this will be. Now we must search for the geometric parameter. We do this by taking \dot{q} and expand it out.

$$\frac{dq}{dt} = \frac{dq}{ds} \frac{ds}{dt} \quad (7a)$$

$$= \int \frac{dt}{ds} ds \sqrt{\left(\frac{dq}{ds} \frac{ds}{dt}, \frac{dq}{ds} \frac{ds}{dt} \right)} \quad (7b)$$

$$= \int \frac{dt}{ds} \frac{ds}{dt} \sqrt{\left(\frac{dq}{ds}, \frac{dq}{ds} \right)} \quad (7c)$$

$$= \int ds \sqrt{\left(\frac{dq}{ds}, \frac{dq}{ds} \right)} \quad (7d)$$

This is our normalized geometric parameter

$$\left(\frac{dq}{ds}, \frac{dq}{ds} \right) = 1 \quad (8)$$

This is true for every s . Now it is important to solve the variation problem (I say it is important to solve the variation problem because I have prior knowledge that this leads us in the right direction), which simply put is finding curves with the smallest slopes to tangents (smallest variation in the function). To do this we take a function $f(x)$ and say (we use x for the general case but we will switch it to q later)

$$f(x) = f(x_0 + \delta x) \quad (9)$$

From this we add

$$+\delta x - f'(x) \quad (10a)$$

And

$$+\sigma(\delta x^2) \quad (10b)$$

We see this and we have an idea, the Lagrangian of q and the change in q (δq) gives us a solution to the variation problem. So we claim that with a Taylor expansion.

$$L[q + \delta q] = L[q] + \sigma(\delta q^2) \quad (11a)$$

If we calculate this out we find

$$L[q] = \int dt \ell(q, \dot{q}) \quad (11b)$$

$$= \int dt L(q + \delta q, \dot{q} + \delta \dot{q}) \quad (11c)$$

$$= \int dt \left[\ell(q + \dot{q}) + \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} + \sigma(\delta q, \delta \dot{q}) \right] \quad (11d)$$

It is apparent that this works iff:

$$\int dt \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right) \delta q = 0 \quad (11e)$$

Thus

$$0 = \int dt \delta q \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right) \Bigg|_{t=0}^{t=1} \quad (11f)$$

We then find the Euler-Lagrange equation.

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0 \quad [[1]] \quad (11g)$$

Because

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{d}{dt} m \dot{q} = 0 \quad (12)$$

The Euler-Lagrange equation of minimized action just decreases the motion of Newton's equation thus $\ddot{q} = 0$. Now we claim that if $t=1$ then

$$\int dt f(t)g(t) = 0 \quad (13)$$

This is true for every $f(t)$

For example

$$\ell = \sqrt{(\dot{q}, \dot{q})} \quad (14)$$

If we say that $t=0$, then $g(t)=0$ and we once again see that

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0 \quad (15)$$

Now we use geometric paratization

$$\left(\frac{dq}{dt}, \frac{dq}{dt} \right) = 1 \quad (16a)$$

Or

$$(\dot{q}, \dot{q}) = 1 \quad (16b)$$

Which is true for every (t). We use our paratization to find:

$$\frac{\partial L}{\partial \dot{q}} = \frac{\partial \dot{q}}{\partial \sqrt{(\dot{q}, \dot{q})}} = \dot{q} \quad (17)$$

With properties of $q(t) = at + b \mid a, b \in \mathbb{R}^n$ we see that

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \ddot{q} = 0 \quad (18)$$

Which shows us that the first time derivative is a horizontal time (this is of course the smallest tangent slope possible).

3. Stationary Curves

When we looked at the Larangian we found that the limiting action was $\ddot{q} = 0$ this means that the 1st time-derivative of the function q is a straight line. We can consider a line to be a stationary (or flat) curve, so we can use the properties we found above.

Take the inner product of (u,v)

$$(u, v) = \sum_{i=1}^3 u_i v_i \quad (19)$$

Where the matrix of g is

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

We write $g(u,v)$ for the general product

$$g(u,v) \equiv u^i g_{ij} v^j \quad (20)$$

Such that

$$g^i_j = g_{ji} \quad (21)$$

We get a horizontal u matrix and a vertical v matrix

$$(-u-)(g) \begin{pmatrix} | \\ v \\ | \end{pmatrix}. \text{ We now show this as a product of our function } (q), \text{ because curved space}$$

(g) depends on a point on q .

$$g(q)(u,v) \quad (22)$$

Which can also be written as

$$\sum_{ab} g(q)_{ab} u^a v^b \quad (23)$$

And in stationary (flat) curves.

$$g(q)_{ab} = \delta_{ab} \quad [[6]] \quad (24)$$

If we look again at the functional (L) of our function (q) when compared to what we found in (24) we come up with

$$L[q] = \int dt \sqrt{g(q)_{ab} \dot{q}^a \dot{q}^b} \quad (25a)$$

We can remove the root to find the minimal.

$$L[q] = \int dt [g(q)_{ab} \dot{q}^a \dot{q}^b] \geq 0 \quad (25b)$$

We now want to add our functional (L) to the change in $L, (\delta L)$. We do this in order to find the correction components for when the curve of our manifold is not stationary.

$$L + \delta L = \int \left[\frac{\partial g(q)_{ab}}{\partial m} \delta q^m \dot{q}^a \dot{q}^b + 2g(q) \frac{d}{dt} (\delta q)^a \dot{q}^b \right] \quad (26)$$

When we calculate this out we find that. (comma denotes a derivative)

$$L + \delta L = \int dt \delta q^m \left[g_{ab, m} \dot{q}^a \dot{q}^b - \frac{d}{dt} (2g_{mb} \dot{q}^b) \right] \quad (27)$$

Now it starts to get interesting when we see that.

$$L + \delta L = \int dt \delta q^m \left[g_{ab, m} \dot{q}^a \dot{q}^b - 2g_{mb, n} \dot{q}^n \dot{q}^b - 2g_{mb} \ddot{q}^b \right] \quad (28)$$

I say this is interesting because our term in the square brackets equals out to zero.

$$0 = g_{ab, m} \dot{q}^a \dot{q}^b - 2(g_{mb, n} \dot{q}^n) \dot{q}^b - 2g_{mb} \ddot{q}^b \quad (29)$$

It is important to now write this out as

$$g_{mb} \ddot{q}^b + \frac{1}{2} (2g_{mb, a} - g_{ab, m}) \dot{q}^a \dot{q}^b = 0 \quad (30a)$$

Such that $2g_{mb, a} = g_{mb, a} + g_{ma, b}$, we know that this is zero because we are dealing with stationary curves, which is

$$g_{mb} \ddot{q}^b + \frac{1}{2} (g_{mb, a} + g_{ma, b} - g_{ab, m}) \dot{q}^a \dot{q}^b = 0 \quad (30b)$$

In such a way that $BB^{-1} = 1$ (B representing a matrix) and $g_{ab}g^{bc} = \delta^{ca}$ for the inverses.

[[6]]

We then take (30b) and multiply by g^{nm} we get.

$$\ddot{q}^n + \frac{1}{2} g^{nm} (g_{mb, a} + g_{ma, b} - g_{mb, n}) \dot{q}^a \dot{q}^b = 0 \quad (30d)$$

We now use this to define our correction components, which are Christoffal symbols of the second kind. The corrections are meant for tangents spaces to manifolds. If the manifold is curved then we need corrections because the tangent spaces exist in 2 times as many dimensions as the manifold. This is to ensure that the tangent spaces cannot intersect.

$$\Gamma_{ab}^n \equiv \frac{1}{2} g^{nm} (g_{mb, a} + g_{ma, b} - g_{mb, n}) \quad (31)$$

In our correction components we see some important properties $g_{ab} = \delta_{ab}$ as well as $g^{ab} = \delta^{ab}$ also $g_{ab,m} = 0$ and $\ddot{q}'' = 0$. We will show an example in polar coordinate system. We take a $g \in \mathbb{R}^n$; where $g(r, \varphi) = (q^1, q^2)$

$$g(r, \varphi) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \quad (32)$$

We now use this with our correction components at different points on our function (q).

$$\ddot{q}^1 + \Gamma_{ab}^1 \dot{q}^a \dot{q}^b \quad (33a)$$

$$q^2 + \Gamma_{ab}^2 \dot{q}^a \dot{q}^b \quad (33b)$$

It would be useful at this time to analyze the (r, φ) coordinate system with our functional (L). We want to do this because we want to compare stationary curves in Cartesian and Polar coordinate systems in order to extract properties of the coordinate systems.

$$\int dt(\dot{r}, \dot{\varphi}) \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \dot{r} \\ \dot{\varphi} \end{pmatrix}$$

This works out to be:

$$\int dt(r^2 + r^2\dot{\varphi}^2) \quad (34)$$

And because

$$r^2\dot{\varphi} = L \quad (35)$$

It follows that

$$\frac{dL}{dr} = 2r\dot{\varphi}^2 \quad (36)$$

And

$$\ddot{r} - r\dot{\varphi}^2 = 0 \quad (37)$$

We look at these and we see that there are some properties to notice. They are important because they help us understand motion in polar coordinates.

$$\frac{dL}{d\dot{r}} = 2\dot{r} \quad (38)$$

$$\frac{dL}{d\phi} = 0 \quad (39)$$

$$\frac{dL}{d\dot{\phi}} = 2r^2\dot{\phi} \quad (40)$$

We then see forming

$$\ddot{r} - r \frac{L^2}{r^4} = 0 \quad (41a)$$

Where

$$\ddot{r} = \frac{L}{r^3} \quad (41b)$$

Also

$$2 \cdot 2r\dot{r} + \dot{\phi} + 2r^2\ddot{\phi} = 0 \quad (42)$$

We then notice that

$$\ddot{r} - r\dot{\phi} \cdot \dot{\phi} = 0 \quad (43)$$

$$\ddot{\phi} - 2\frac{1}{r}\dot{r}\dot{\phi} = 0 \quad (44)$$

Which describes motion with constant velocity, which for this coordinate system is the equivalent of a stationary curve. From these we find that $\dot{\phi} = 2\pi$, $r = 0$, $\phi(t) = 2\pi t$ and $r(t) = R$. Also we should take

$$\Gamma_{22}^1 = r \text{ and } \Gamma_2^2 = \frac{1}{r}$$

where all unrelated r vanish. Now we will continue by calculating some straight lines on a sphere, which are called geodesics (thus we are looking for the geodesic equations). It is of interest to do this because we see that on a sphere parallel lines always intersect but geodesics do not, so we need a way to make sense of the coordinate changes. We start with

$$q(t) = R \begin{pmatrix} \cos 2\pi t \\ \sin 2\pi t \end{pmatrix} \quad (45)$$

and work with this a bit and we see that

$$\int \sqrt{\dot{q}, \dot{q}} dt = \int \sqrt{(2\pi R) \begin{pmatrix} \cos 2\pi t \\ \sin 2\pi t \end{pmatrix}, i} dt \quad (46)$$

It then follows that this equals

$$\int 2\pi R \Gamma_1 \quad (47)$$

if $dt = 2\pi R$ and $t = 0$. We shall now write this in terms of our polar coordinate system (this is the natural coordinate system for a sphere) and we find that

$$\int dt \sqrt{\dot{r}^2 + r^2 + \dot{\varphi}^2} = \int dt \sqrt{R^2 (2\pi)^2} = 2\pi \quad (48)$$

where again $t = 0$. If we look at a sphere we see the coordinate conversion is

$$x = R \cdot \sin(\theta) \cos(\varphi)$$

$$y = R \cdot \sin(\theta) \sin(\varphi)$$

$$z = R \cdot \cos(\theta)$$

Then

$$x^2 + y^2 + z^2 = \dots = R^2$$

With this we see the matrix

$$g_{ab} = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{pmatrix}$$

where the unit sphere has a matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}$$

If we say that $q^1 = \theta$ and $q^2 = \varphi$, and look once again at the functional (L) of our function (q), then,

$$L[q] = \int dt (\dot{\theta}, \dot{\varphi}) \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix} \begin{pmatrix} \dot{\theta} \\ \dot{\varphi} \end{pmatrix} \quad (49)$$

Because of the matrix multiplication I showed earlier, we get.

$$L[q] = \int dt(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \quad (50)$$

The term $(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2)$ we will let equal to $\ell(\theta, \psi, \dot{\theta}, \dot{\phi})$. Now when we look at ℓ we find out that:

$$\frac{\partial \ell}{\partial \theta} = 2 \cos(\theta) \sin(\theta) \dot{\phi}^2 \quad (51a)$$

The first time derivative of this is

$$\frac{d}{dt} \frac{\partial \ell}{\partial \theta} = 2\ddot{\theta} \quad (51b)$$

We now find two important properties. They are important because their consequences show us interesting things in Cartesian coordinates when they are converted.

$$\ddot{\theta} - \cos(\theta) \sin(\theta) \dot{\phi}^2 = 0 \quad (52)$$

and

$$\dot{\phi} \sin \theta = 0 \quad (53)$$

which of course shows us that $\ddot{\theta} = 0$ and leads us to believe that

$$\frac{\partial L}{\partial q} = 0 \quad (54)$$

and

$$\frac{\partial L}{\partial \dot{\phi}} = 2 \sin \theta \dot{\phi} \quad (55)$$

This is true when ϕ is constant, which gives us the matrix

$$g_{ab} = \begin{pmatrix} g_{zt}, g_{z\phi} \\ g_{\phi t}, g_{z\phi} \end{pmatrix}$$

4. Differential Geometry

We want to start working with differential geometry because General Relativity is a theory that has space-time represented as a manifold. Differential geometry is of course the natural mathematics for a manifold, especially in high dimensions. Now we

incorporate the nabla, which is a differential operator. We will take the covariant derivative of this to be:

$$\nabla_{\dot{x}} \dot{x} = \partial_{\dot{x}} \dot{x}^n + \Gamma_{ab}^n \dot{x}^a \dot{x}^b = 0 \quad (56)$$

These differential operators are very useful because as you can see they incorporate the correction components that we defined earlier. Now, we have invented this term,

$$(\partial_{\dot{x}} \dot{x}^n + \Gamma_{ab}^n \dot{x}^a \dot{x}^b = 0)$$

but we need to know what this means in order to understand the covariant derivative of our differential operator. We see that

$$\dot{x}^n = \frac{d}{dt}(\dot{x}^n) = \dot{x}^m \frac{\partial}{\partial m} \dot{x}^n = \dot{x}^m \frac{d}{dx^m} \dot{x}^n \quad (57)$$

but we are missing a connection to our earlier work and it is this

$$\dot{x}^m \frac{d}{dx^m} \dot{x}^n = \partial_{\dot{x}} \dot{x}^n \quad (58)$$

We then say that in a flat space $\Gamma \equiv 0$ and $\partial_{\dot{x}} \dot{x} = 0$. This is understandable, seeing that the correction components were created to help with the problem of tangent spaces on a curved manifold. Our next step will be to look at the parallel transport of x along itself. To start we will say that

$$\nabla_{\dot{x}} x = 0 \quad (59)$$

where

$$\dot{x} = x \quad (60)$$

If we take our equation (56) we find that

$$x^m \partial x^a + \Gamma_{ab}^n x^a x^b = 0 \quad (61)$$

Now we define the Riemann Curvature tensor to be

$$R(x, y)z \equiv \left[\nabla_x \nabla_y \right] z - \nabla_{[x, y]} x \quad (62a)$$

The $\left[\nabla_x \nabla_y \right]$ is the usual commutator of operators (which in this case is $R(x, y)z$.)

$$[\nabla_x \nabla_y] = \nabla_x \nabla_y - \nabla_y \nabla_x \quad (62b)$$

The Riemann curvature tensor is a tensor field defined by $R(x,y)z = [\nabla_x \nabla_y]z - \nabla_{[x,y]}x$ at every point. We also define our vector fields $[x,y]$ to be.

$$[x,y]^a \equiv x^m \partial_m Y^a - Y^m \partial_m X^a \quad (63)$$

This is a Schrodiner space-time structure. Now we define the manifold (M) to be the set of all points that are locally homomorphic to \mathbb{R}^n space of n dimension such that

$u \rightarrow f \rightarrow v \subseteq \mathbb{R}^n$ where v is a chart and \mathbb{R}^n is an open subset or equal to the chart (v). If

we take a curve $\gamma : \mathbb{R} \rightarrow M$ which is like saying that $\mathbb{R} \rightarrow \mathbb{R}^n$. Now if we assume

$\gamma(0) = \rho$ then we can say that $\mathbb{R} \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$ the first mapping is a γ action and the

second is a f action we call this $f \circ \gamma$. This is important when we say

$$\dot{\gamma} \rho : C^\infty(\mathbb{R}^n) \rightarrow \mathbb{R} \quad (64)$$

which will work out to

$$\dot{\gamma} \rho f := (f \circ \gamma) \quad (65)$$

which with a limit of t=0 work out to

$$\dot{\gamma} \rho f := \frac{d}{dt} (f \circ \gamma) |_{t=0} \quad (66)$$

We now want to put this into nicer looking terms

$$\frac{df}{dx^u} \cdot \frac{d\gamma}{dt} \quad (67)$$

which we will rewrite as

$$\frac{d\gamma}{dt} \cdot \frac{\partial}{\partial x^u} f \quad (68)$$

using the chain rule.

We now recognize the term $\left(\frac{\partial}{\partial x^u} \right)$ which using compact notation is ∂_u . This is because

the tangent space of the Manifold has a basis of:

$$\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n} \right)$$

$$\partial_u \equiv \frac{\partial}{\partial x^u} \quad (69)$$

Now we take a vector x at a point p

$$x = x^u \partial_u \quad (70)$$

If we then take $f \in C^\infty(M)$, belonging to the smooth manifold, we get

$$xf = x^u \partial_u f = \partial_x f \quad (71)$$

I now define a tensor to be a multi linear map such that $T : T_p M \times T_p M \times \dots \times T_p M \rightarrow \mathbb{R}$

where $T(x,y,\dots,z)$ is linear and $C^\infty(M)$ smooth manifolds, are linear maps. Now we take the vector fields x,y and define

$$[x, y] \equiv xy - yx \quad (72a)$$

(where square brackets denote the comutator of operators) to be a vector, if we take (72a) and use a Lie bracket we find

$$[x, y]f = x(yf) - y(xf) \quad (72b)$$

5. Linear Connections

We use the covariant derivative of the differential operator (∇) (which is a vector, we do this to see the properties of such a vector) where $\nabla_x f = xf$

We see three properties

$$\nabla_x^i(fy) = (\nabla_x f)y + f\nabla_x y \quad (73)$$

$$\nabla_{fx} y = f\nabla_x y + yZ = \nabla_x z + \nabla_y z \quad (74)$$

$$\nabla_x(y + w) = \nabla_x y + \nabla_x w \quad (75)$$

where $\nabla_x y$ is again a vector. Now we want to find the components of our tensor so we take the vector fields x,y and say

$$T(x, y) = T(x^a \partial_a y^b \partial_b) \quad (76a)$$

We can then move x^a and y^b so

$$X^a Y^b T(\partial_a, \partial_b) \quad (76b)$$

We then define $T_{ab} \equiv T(\partial_a, \partial_b)$, so that ab are our tensor components (we define this just to simplify notation). This of course means that

$$T(x, y) = X^a Y^b T_{ab} \quad (76c)$$

We then take our vector $\nabla_x y$ and with the knowledge we just gained we say that it is equal to

$$\nabla_x y = \nabla x^a \partial_a (y^b \partial_b) \quad (77a)$$

Which we could write as

$$X^a (\nabla_{\partial_a} (Y^b \partial_b)) \quad (77b)$$

We want to ignore ∂_b for a moment and look at the term

$$x^a (\nabla_{\partial_a} y^b) \quad (77c)$$

which we then expand out to be

$$x^a ((\nabla_{\partial_a} y^b) \partial_b + y^b (\nabla_{\partial_a} \partial_b)) \quad (77d)$$

Now we put our correction components in

$$x^a (\partial_a y^b) \partial_b + x^a y^b \Gamma_{ba}^c \partial_c \quad (77e)$$

As we stated earlier $\nabla_x x = 0$ so we now use that with what we found in (77e). This gives us

$$x^a \partial_a x^c + x^a x^b \Gamma_{ba}^c = 0 \quad (78)$$

Which work iff

$$\ddot{x}^c + \Gamma_{ba}^c \dot{x}^a \dot{x}^b = 0 \quad (79)$$

6. Curvature Tensor

The curvature of space-time is found in General Relativity through tensor analysis so naturally a curvature tensor is necessary. So we define our curvature tensor to be

$$R(x, y; z) = [\nabla_x \nabla_y] z - \nabla_{[x, y]} z \quad (80)$$

The manifold mapping assigns the operator $R(x, y; z) = -R(y, x; z)$ [[2]], so we only have to check for linearity in z . We do this by multiplying z by a constant, which will be the function (f) . We do this by

$$R(x, y, f; z) = \nabla_x \nabla_y (fz) - \nabla_y \nabla_x (fz) - \nabla_{[x, y]}(fz) \quad (81)$$

With a few lines [[3]] we can prove that it works out to $fR(x, y; z)$. This proves linearity and shows us that R is indeed a tensor. Now this allows us to move onto torsion. We say then

$$T(x, y) = \nabla_x y - \nabla_y x - [x, y] \quad (82)$$

which for Einstein's theory we define torsion to be zero $T \equiv 0$ which means

$$T_{(x, y)} = 0 \quad (83)$$

This also means that $\nabla_x z = \nabla_z x$ which is important when we write our curvature tensor as

$$R(x, z; x) = \nabla_x \nabla_z x - \nabla_z \nabla_x x - \nabla_{[x, z]} x \quad (84)$$

Now because $\nabla_x x = 0$ the terms $(\nabla_z \nabla_x x)$ and $\nabla_{[x, z]} x$ cancel out, and because $\nabla_z x = \nabla_x z$ we write this as

$$R(x, z; x) = \nabla_x \nabla_x z \quad (85)$$

7. 3D Oscillator

We want to start to work towards General Relativity so we need to make connections from Newtonian to Einstein Gravity. We will now look at the 3-D oscillator

$$\ddot{z}^a = -(\partial^a \partial_b \gamma) z^b \quad (86)$$

It makes sense to look at this because if we look at the two particles falling in a gravitational field we see tidal acceleration. This looks like an oscillator. Now if we

assume $x^\mu = (x_0, x_1, x_2, x_3)$ where $x_0 \equiv t$ [[4]] and we look at our curve in classical (Newtonian) mechanics we would say

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^3 \quad (87a)$$

$$\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \quad (87b)$$

But now if we introduce the concept of space-time we see our curve γ to be

$$\tilde{\gamma} : \mathbb{R} \rightarrow \mathbb{R}^4 \quad (88a)$$

$$\tilde{\gamma}(t) = \begin{pmatrix} t \\ x(t) \\ y(t) \\ z(t) \end{pmatrix} \quad (88b)$$

Now we will reintroduce (85) and write it as

$$(\nabla_x \nabla_x z)^a = (R(x, z; x))^a \quad (89)$$

If we then say that $x = \partial_o$ we see

$$(R(\partial_o, \partial_b, \partial_o))^a z^b \quad (90)$$

such that

$$(\nabla_x \nabla_x z)^a = R^a_{obo} z^b \quad (91)$$

If we look again at the 3-D oscillator

$$\ddot{z}^a = -(\partial^a \partial_b v) z^b \quad (92)$$

We see that (92) and (91) look very similar and we get an idea, what if gravity is an effect of the curvature of space-time.

$$R^a_{obo} \stackrel{!}{=} -\partial^a \partial_b v \quad (93)$$

If we say that (Δv) is matter density is

$$\Delta v = \mathfrak{S} \quad (94)$$

iff

$$R^a{}_{obo} = \mathfrak{S} \quad (95)$$

For this to work we sum over the indices (Einstein summation convention) to say that

$$\sum_{ab} -\partial^a \partial_b V = \mathfrak{S} \quad (96)$$

and iff

$$\sum_{ac} R^a{}_{bcd} = R_{bd} \quad (97)$$

This means that

$$R_{oo} = \mathfrak{S} \quad (98)$$

which is the Ricci tensor, and here represents Newtonian gravity in geometric form. This tells us that the curvature of space-time is geometrically, what Newtonian gravity looks like.

8. Einstein Tensor

We will look at our Ricci curvature tensor and we say that this is nice, but what I would really like to see is the Einstein tensor. We will start by saying that R_{oo} (98) can also be written as R_{ab} . And then we must find the metric tensor. To do this we look again at a ds but this time of a tensor, which is also known as a Riemann metric (once again I will use a hand waving technique and say that we do this because we already know it will work).

We use what is called a generalized Pythagorean theorem. Our metric in a Newtonian \mathbb{R}^3

space has a matrix $\begin{pmatrix} g_{11} & g_{21} & g_{31} \\ g_{12} & g_{22} & g_{32} \\ g_{13} & g_{23} & g_{33} \end{pmatrix}$ which will work out to be

$$ds^2 = g_{11}dx_1^2 + g_{12}dx_1dx_2 + g_{13}dx_1dx_3 + g_{21}dx_2dx_1 + g_{22}dx_2^2 + g_{23}dx_2dx_3 + g_{31}dx_3dx_1 + g_{32}dx_3dx_2 + g_{33}dx_3^2 \quad (99)$$

If we are in Euclidean space where

$$g_{ab} = \delta_{ab} \quad (100)$$

then we use the usual form of Pythagorean theorem

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 \quad (101)$$

which means that our metric matrix is diagonal with values at (g_{11}, g_{22}, g_{33}) and all other elements are zero. Now we need to have the component known as scalar curvature so we define it as

$$R \equiv g^{ab} R_{ab} \quad (102)$$

We can now create the Einstein tensor G_{ab} which is

$$G_{ab} = R_{ab} - \frac{1}{2} R g_{ab} \quad (103)$$

9. Field Equation

Now that we have found the Einstein tensor we are getting very close to unlocking all of his general theory of relativity. To start we will say

$$G_{ab} = \kappa T_{ab} \quad (104)$$

Where T_{ab} is the stress-energy tensor this can be written as

$$G_{ab} = R_{ab} - \frac{1}{2} R g_{ab} = \kappa T_{ab} \quad (105)$$

We work with this a bit and come up with

$$R_{ab} = \frac{1}{2} g_{ab} R + \kappa T_{ab} = \frac{1}{2} (2T_{ab} - g_{ab} T_{ab}) \quad (106)$$

which we will write as

$$\frac{1}{2} \kappa [2T_{ab} + (T^a_b + T^i_j)] \quad (107)$$

which simplifies to

$$\frac{1}{2} \kappa (T_{ab} + T^i_j) \quad (108)$$

Finally we say that

$$R_{ab} = \frac{1}{2} \kappa \rho \quad (109)$$

where ρ is matter density. If we then say that $\kappa=8\pi$ (this is a large hand wave but a necessary one) we get an idea

$$R_{ab} = 4\pi \rho \quad (110)$$

This will give us numerical values, and because we said that $G_{ab} = \kappa T_{ab}$ (101) we can say

$$G_{ab} = 8\pi T_{ab} \quad (111)$$

This is the Einstein field equation, it encompasses 16 partial differential equations, which describe gravitational effects. But because G_{ab} and T_{ab} are symmetric it reduces to 10 differential equations and adds 4 differential identities (this fact is of no consequence to this paper but I think that it is interesting to understand Einstein's work and how he discovered it). Next we add a Lorenz transformation Λg_{ab} . Einstein set this equal to zero when he created the cosmological constant.

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab} \quad (112)$$

10. Conclusion

We find that using an accelerated reference frame with Newton's equations we can find the curvilinear coordinates. These are indeed very much like the curvature equations of general relativity. Using the principle of equivalence we can work in this matter to arrive at the field equation of general relativity.

11. Acknowledgments

This paper is from a series of mentoring sessions with Frederic Schuller at Perimeter Institute. I used the same approach, and much of the same mathematical techniques to get to the Ricci Curvature Tensor (which is where he left off) and continued from there to the Field Equation of General Relativity. These mentoring sessions took place during the two-week period of the International Summer School for Young Physicists (ISSYP).

[[1]]. We find this to be true by looking at the principal of least action

$$\begin{aligned}
 L &= T - V = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots \\
 &= \int_0^1 dt \sqrt{\frac{dq^i}{dt}(t), \frac{dq}{dt}(t)} = \int_0^1 dt \ell(q, \dot{q}) \\
 L + \delta L &= \int dt \ell(q + \delta q, (\dot{q} + \delta \dot{q})) \\
 &= \int dt \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{d}{dt}(\delta q) \right) + \int dt \ell(q, \dot{q}) \\
 &= \int dt \left(\frac{\partial L}{\partial q} \delta q - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) \right) - \left[\frac{\partial L}{\partial \dot{q}} \delta q \right] \\
 &= \int dt \delta q \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \\
 \Rightarrow \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} &= 0
 \end{aligned}$$

[[2]].

$$\begin{aligned}
 R(x, y; z) &= \left[\nabla_x \nabla_y \right] z - \nabla_{[x, y]} z \\
 &= -\left[\nabla_y \nabla_x \right] z + \nabla_{[y, x]} z \\
 &= -R(y, x; z)
 \end{aligned}$$

[[3]].

$$\begin{aligned}
R(x,y,f;z) &= \nabla_y \nabla_x (fz) - \nabla_x \nabla_y (fz) - \nabla_{[x,y]}(fz) \\
&= \nabla_y (f \nabla_x z + (xf)z) - \nabla_x (f \nabla_y z + (yf)z) + f \nabla_{[x,y]} z + ([x,y]f)z \\
&= fR(x,y)z + (y(xf) - x(yf) + [x,y]f)z \\
&= fR(x,y;z)
\end{aligned}$$

This idea is indeed strange because our curvature tensor is made up of covariant derivatives that are not linear operators over the ring of smooth functions.

[[4]]

We say that x^u is the parameters of the manifold. In this case it is a 4 manifold of (x_0, x_1, x_2, x_3) where $x_0 = t$

[[5]]

This is true if the parameter of our manifold is $x^u = (x_1, x_2, x_3, \dots, x_n)$

[[6]] δ_{ab} is the Krocknecker delta which is used for orthonormality 1 for $a=b$, 0 for $a \neq b$.