An Introduction to the General Theory of Relativity

Zachary Hall

July 11, 2017

Abstract

We present a paper that is primarily pedagogical in spirit, with the first author having written the preliminary draft as he worked through conceptual difficulties he had with the general theory of relativity when first learning it, that aims to be a standalone introduction to the basic statements of the general theory of relativity. The paper is conclusion of a series of three projects the authors have undertaken in an attempt to improve the pegagogy of the General Theory, and draws heavily upon the author's previous work (Hall, Stachel) discussing Newton-Cartan gravitational theory and extending it to predict frame-dragging effects in the low-velocity, near field limit. As an intended standalone pedagogical reference, the authors re-present various discussions of other authors' and sythnesize these discussions into a larger discussion of the theory, augmenting the work of others where we feel appropriate. The original contribution to the pegagogical content is a radar-techinque method of building the so-called Rindler wedge, inspired by Hermann Bondi's k-calculus. Additionally, we build upon the user-interactive introductory lesson to the kinematics of special relativity that can be found on the Center for Einstein Studies website, introducing concepts and language dependent on the basic kinematics of SR and important for understanding the basics of general relativity and further analyzing its consequences. We include a brief discussion of General Relativity's background independent nature and constraints it places on constructions of theories of quantum gravity.

1 Introduction

The general theory of relativity has a certain mystique surrounding it, that makes it seem an innaccessible theory for all but those most gifted of minds. We do not deny that there are abstract difficulties in grasping the basic statements of the general theory. However, we maintain that many of the difficulties learners encounter can be ameloriated by a careful analysis of the evidence and motivating insights informing the general theory. By examining the historical development of the theory, and analyzing the two foundational insights of the theory – the kinematic structure of special relativity and the gravitational equivalence principle – we maintain that the pedagogy of the general theory can be improved, and the foundational understanding of the theory for both teachers and learners alike deepened as a result. In this paper, we seek to give a discussion that, combined with an understanding of the basics of special relativity and a thorough reading of our auxillary work of (Hall, Stachel), serves as such a pedogogical introduction to the basic ideas underlying the theory.

In Section 1, we present a brief review of the argument presented in (Hall, Stachel) that the gravitational equivalence principle allows us to geometrize gravity, and subsquently predict the existence of magnetic-type gravitational forces owing to frame-dragging of the spacetime, independently of the the postulates of special relativity. In Section 2, we present an overview of the basics of the kinematics and and dynamics of special relativity. We recommend that any reader not comfortable with the basics of special relativity work through a user-interactive, visual introduction to the special theory on the Center for Einstein Studies Website. In Section 3, we present a discussion of the reference frame of an observer with uniform acceleration in special relativity, in order to build intuition as to what to expect from the general theory in light of the gravitational equivalence principle. In particular, we augment the discussion of Rindler coordinates using ideas from the k-calculus of Hermann Bondi's "Relativity and Common Sense," first using such ideas to construct the class of Rindler observer's for a given proper acceleration. In Section 4, we first present a motivation for the particular form of the Einstein Field Equations in vacuum, and then present the Schwarzchild solution in relation to the theory of gravitation reviewed in Section 1 and in (Hall, Stachel), as first done by Havas (1964). We then proceed in Section 5 to present the field equations in the presence of matter and energy while explicating the conservation law of the general theory and also include an example of the conservation law. In Section 6, we discuss briefly the conceptual of feature of GR of background independence.

2 Non-flat Newtonian spacetime - a brief review

Briefly reviewing (Hall, Stachel), it can be shown that the gravitational equivalence principle, which asserts the essential equivalence of gravitation and inertia, can be used to reformulate Newtonian gravitational theory chronogeometrically and also extend this chronogeometric treatment to predict the frame-dragging effects of the full general theory, this treatment being valid in the near-field, low velocity regime.

In traditional Newtonian theory, for a given gravitational field $\vec{g}$ and test particle,

$$m_I \vec{a} = m_G \vec{g}$$

, where $\vec{a}$ is the 3-acceleration of a test particle w.r.t. the center of mass of the matter giving rise to the gravitational field $\vec{g}$, m_I is a determiner of the 3-acceleration of test particle in response to a given external force, i.e. the test particle's inertial mass, and m_G is the gravitational mass or, in analogy with electromagnetism, gravitational "charge" of the test particle that determines the extent of the gravitational force acting on the test particle. Experimentally it has been shown to a high degree of precision that $m_I = m_G$ and thus $\vec{a} = \vec{g}$ for every test particle.

This means that, locally, the effects of a gravitational field and of being in an accelerated frame of reference are indistinguishable. This forces us to reconceptualize the motional tendencies of matter as influenced by the force of gravity, in Newtonian theory, as merely the inertial motional tendencies of matter with respect to so-called inertial frames of reference in Newtonian theory. To be in a frame in which, locally, there is a so-called gravitational force must be interpreted as being in a frame which has a particular non-zero four-acceleration.

The reformulation of accelerertion due to gravity really being the natural, inertial motional tendencies of matter can be stated mathematically with the definition of a non-metric affine space and of a non-metric affine connection, $\Gamma^{\kappa}_{\nu\mu}$, and the closely associated operation of the covariant derivative, defined by $D_\nu B^\mu = \partial_\nu B^\mu + \Gamma^{\mu}_{\kappa\nu} B^\kappa$, where $_{,\gamma} := \partial_\gamma = \frac{\partial}{\partial x^\gamma}$ and B^μ is an arbitrary vector field defined on the manifold. This is done in (Hall, Stachel). After such definitions are made, the four-force-free equations of motion become the autoparallel equation applied to the four-velocity of a world line in an affine spacetime:

$$W^\nu D_\nu W^\kappa = 0 \rightarrow$$
$$\frac{d^2 X^\kappa}{d\lambda^2} + \Gamma^{\kappa}_{\nu\mu} \frac{dX^\mu}{d\lambda} \frac{dX^\nu}{d\lambda} = 0 \tag{1}$$

where the preferred parameter λ of the curve can be identified with the absolute Newtonian time t via $\lambda = T = ct$ for convenience. More generally,

$$A^\kappa = W^\nu D_\nu W^\kappa = \frac{d^2 X^\kappa}{d\lambda^2} + \Gamma^{\kappa}_{\nu\mu} \frac{dX^\mu}{d\lambda} \frac{dX^\nu}{d\lambda}$$

is the four-acceleration of the curve, with the four-force (non-gravitational force) being defined as $F^\kappa = mA^\kappa$.

It can then be seen that Newton's first Law as he originally concieved it can be thought of as the law of autoparallel transport in an affine flat, or globally Euclidian, spacetime, wherein all $\Gamma^{\kappa}_{\nu\mu} = 0$ and thus $\frac{d^2 X^\kappa}{d\lambda^2} = 0$. Once it is recognized that $\vec{a} = \vec{g}$ for every test particle, it is conceptually simpler to acribe such three-acceleration as due to non-vanishing $\Gamma^{\kappa}_{\nu\mu}$ in the frame of reference with zero three-acceleration w.r.t. the center of mass of the source matter, i.e. convential Newtonian inertial frames. This already allows the reconceptualization of Newtonian gravitation, where the only non-zero affine connection components, Γ^{i}_{00}, are solved in terms of the affine Ricci tensor, defined exactly in analogy with the metric Ricci tensor, only with a non-metric affine connection:

$$R_{00} = \frac{4\pi G\rho}{c^2}$$

This, as shown in (Hall, Stachel) and originally by Cartan (1923), reproduces all predictions of Newtonian dynamics including gravitation, but with the gravitational force reconceptualized chronogeometrically.

We can interpret the condition R_{00} as the divergence of the affine Reimmann tensor $A^{\lambda}_{\mu\kappa\nu}$, as can be checked from the definition of the affne Ricci tensor

$$R_{\mu\nu} := A^{\lambda}_{\mu\lambda\nu}$$

the definition of the affine Reimann tensor

$$A^{\lambda}_{\mu\kappa\nu} := \Gamma^{\mu}_{\kappa\mu,\nu} - \Gamma^{\lambda}_{\nu\mu,\kappa} + \Gamma^{\lambda}_{\nu\epsilon}\Gamma^{\epsilon}_{\mu\kappa} - \Gamma^{\mu}_{\delta\kappa}\Gamma^{\delta}_{\nu\mu}$$

and that the only non-zero connection components are the Γ^{i}_{00}. This is analogous to traditional Newtonian gravitational theory, which can be written as

$$\sum_i \varphi_{,i,i} = \vec{\nabla}^2\varphi = 4\pi G\rho$$

where $,i := \frac{\partial}{\partial x^i}$ and φ is the gravitational potential, from which the gravitational force can be derived via

$$\vec{g} = -\vec{\nabla}\varphi$$

We may go beyond the predictions of conventional Newtonian gravitational theory with the aide of the tetrad formalism, wherein all quantities are projected upon a set of four basis vector fields defined throughout the manifold – three spatial, $e^{\nu}_{(i)}$, and one temporal, $e^{\nu}_{(0)}$, taking the quantities defined from spatial hypersurface to spatial hypersurface. The tetrad field is more rigorously defined in (Hall, Stachel), for the purposes of review we state here that we also define a dual basis vector field $e^{(\alpha)}_{\nu}$ by the conditions $e^{(\alpha)}_{\nu}e^{\nu}_{(\beta)} = \delta^{(\alpha)}_{(\beta)}$ and $e^{(\alpha)}_{\nu}e^{\kappa}_{(\alpha)} = \delta^{\kappa}_{\nu}$. Projecting W^{κ} onto this tetrad, $W^{\kappa} = W^{(\alpha)}e^{\kappa}_{(\alpha)}$, with $W^{(0)} = 1$ and $W^{(i)} = \frac{w^{(i)}}{c}$, where w^i is the three-velocity in the i-th direction of the given tetrad, defining the covariant derivative by $D_\nu B^\mu = \partial_\nu B^\mu + \Gamma^{\mu}_{\kappa\nu}B^{\kappa}$, the tetrad components of the connection (t.c.c. for short) by $\Gamma^{(\gamma)}_{(\alpha)(\beta)} := e^{(\gamma)}_{\kappa}e^{\nu}_{(\alpha)}D_\nu e^{\kappa}_{(\beta)}$, and defining $\frac{D}{d\lambda} := W^\nu D_\nu$ equation (1) becomes

$$\left(\frac{DW}{d\lambda}\right)^{(\gamma)} + \Gamma^{(\gamma)}_{(\alpha)(\beta)}W^{(\alpha)}W^{(\beta)} = 0$$

Because Euclidian three-geometry and absolute time are essential to the framework of Newtonian physics, we can impose conditions on the t.c.c. ensuring these aspects of the theory are present – i.e. we can impose compatibility conditions between the affine structure of the manifold and the chronogeometry of the theory. This is done in (Hall, Stachel) and results in the neccessary vanishing of all t.c.c. except for the $\Gamma^{(m)}_{(0)(0)}$, $\Gamma^{(m)}_{(0)(n)}$, and $\Gamma^{(m)}_{(n)(0)}$, where (m) and (n) indicate purely spatial indices. The apriori specificition of absolute time also allows us to identify $T = ct$ with the preferred parameter λ, as it is postulated in advance to progress the same for all world lines. The $\Gamma^{(m)}_{(0)(n)}$ and $\Gamma^{(m)}_{(n)(0)}$ are related to the anholonomic object of the spacetime,

$$\Omega^{(\gamma)}_{(\alpha)(\beta)} = \frac{1}{2} e^{\nu}_{(\alpha)} e^{\mu}_{(\beta)} (e^{(\gamma)}_{\mu,\nu} - e^{(\gamma)}_{\nu,\mu})$$

in the case of a torsionless connection (Stachel-Papapetrou , 1978), by

$$\Gamma^{(\gamma)}_{[(\alpha)(\beta)]} = -\Omega^{(\gamma)}_{(\alpha)(\beta)}$$

It thus follows that, in a holonomic spacetime wherein all $\Omega^{(\gamma)}_{(\alpha)(\beta)} = 0$, that $\Gamma^{(m)}_{(0)(n)} = \Gamma^{(m)}_{(n)(0)}$. As elaborated upon in (Hall, Stachel) this gives a natural geometric description of the Coriolis and centrifugal forces. Even in spacetimes that are not holonomic, if we know the anholonomic object, we need only know the $\Gamma^{(m)}_{(0)(n)}$ to know the $\Gamma^{(m)}_{(n)(0)}$.

It can be shown, from the tetrad components of the affine Ricci tensor given in Stachel-Papapetrou (1978), that the $\Gamma^{(m)}_{(0)(0)}$ are directly related to the $R_{(0)(0)}$ and the $\Gamma^{(m)}_{(0)(n)}$ to the $R_{(0)(n)}$.

If we then posit that, in the case of active dynamics, the $\Gamma^{(m)}_{(0)(0)}$ and the $\Gamma^{(m)}_{(0)(n)}$ are derivable from a scalar and vector potential, respectively,

$$\Gamma^{(m)}_{(0)(0)} = \delta^{(m)(j)} \partial_{(j)} \varphi$$

and

$$\Gamma^{(m)}_{(0)(n)} = \delta^{(m)(j)} [\partial_{(j)} A_{(n)} - \partial_{(n)} A_{(j)}]$$

,

and assume that, analogous to the way in which the electric-type Newtonian gravitational force can be geometrically redescribed by $R_{(0)(0)} = \frac{4\pi G}{c^2}\rho$, the $\Gamma^{(m)}_{(0)(n)}$ are derivable by $R_{(0)(n)} = \frac{4\pi G}{c^2}\rho\frac{v^{(n)}}{c}$, for the case of an holonomic spacetime, we then arrive at

$$R_{(0)(0)} = \delta^{mj} \partial_{mj} \phi = \nabla^2 \varphi = \frac{4\pi G \rho}{c^2}$$

and

$$R_{(0)(n)} = \delta^{mj} \partial_{mj} A_{(n)} = \nabla^2 A_{(n)} = \frac{4\pi G \rho v^{(n)}}{c^3}$$

with the condition $\partial_j(\delta^{mj} A_m) = \nabla \cdot \vec{A} = 0$. Solving the above field equations for a sphere of uniform density and constant angular rotation, as done in (Hall, Stachel), then gives a magnetic-type gravitational field, most propertly intepreted as the dragging of the inertial frames of spacetime, in mathematical agreement with the near-field, low velocity approximation scheme of Rindler (1964). While the chronogeometry of Newtonian spacetime can be said to be characterized by affine structure (affine flat, or Euclidian, in the original conceptual framework of Newtonian mechanics, affine non-flat when informed by the equivalence principle), spatial intervals are postulated to be Euclidian, and independent of temporal intervals – hence we cannot charactarize such a spacetime by a four-dimensional metric, as in the special and general theories of relativity. More succintly, it is said that the chronogeometry of Newtonian spacetime has affine structure, but no four-metric structure. This will be elaborated upon in Section 4.

3 Special relativity - a brief overview

Here we offer an overview of the basic ideas of the special theory of relativity. The authors have constructed a lesson for building up a more intuitive understanding of the special theory, using the aid of three user-interactive Java applications, that we recommend readers at all conceptually uncomfortable with the special theory work through first. The lesson and the applications were inspired by Hermann Bondi's book "Relativity and Common Sense," (Rindler , 1964) and can be found on the Center for Einstein Studies website at the time of publication.

The basic insight of Eintein's theory we now call special relativity is that the appropriate interpretation of the Michelson-Morely experiment is that the kinematic structure of spacetime is not the Newtonian-Galiliean structure, with absolute time independent of the Euclidian three-geometry of space, presupposed by all of physics prior to special relativity, but rather a kinematic structure wherein the possible relations between the assigning of spaces and times of events by inertial observers is characterized by an invariant spacetime interval, $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$. This encodes that the speed of light, c, is independent of the velocity of its source and thereby the same for all observers. In particular, all electromagnetic phenomena are mediated along *null cones* defined by $ds^2 = 0$, with the null cone structure of the spacetime being invariant, i.e. agreed upon by all observers in the spacetime. The transformations that leave this interval, ds^2, between distinct events in the spacetime invariant are called Lorentz transformations and are defined as all changes of coordinates (if passive transformations) $x^\nu \to x^{'\nu}$ or permutation of spacetime points under the same spacetime coordinate system (if active transformations) $p \to p'$ such that $\eta_{\mu\nu}dx^{'\mu}dx^{'\nu} = \eta_{\mu\nu}dx^\mu dx^\nu$ where $\eta_{\mu\nu} = diag\{1, -1, -1, -1\}$ is the four-metric characterizing the chronogeometric nature of the spacetime and dx^μ is the four dimensional infinitesimal vector defining the spatio-temporal separation of two events in a given frame of reference. It is then the case that the Lorentz transformations include: (1) spatial translations of fixed (non-time-varying) value, as this trivially leaves the infinitesimal *separations* of spacetime points invariant (2) spatial rotations of a fixed (non-time varying) angle, as this merely permutes the spatial points rotationally, and it is well known that this leaves inifintesimal separations of spatial points invariant and (3) *Lorentz boosts*, which relate the assignment of spaces and times of events between observers moving relative to each other with constant velocity and are of the form

$$\begin{pmatrix} x^{'0} \\ x^{'1} \\ x^{'2} \\ x^{'3} \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

, where the 0th coordinate is always the temporal coordinatization of events times the speed of light c, v is the magnitude of the relative velocity of the two observers, $\beta := \frac{v}{c}$, and $\gamma := \frac{1}{\sqrt{1-\beta^2}}$. The above is a Lorentz boost along the x^1 direction – boosts along the other two directions take the same form but with cross-terms in the temporal and boosted spatial direction in each case considered. Just as in conventional Newtonian theory, in special relativity, there is a preferred class of observers – the force-free, inertial observers, for whom the chronogeometric structure of spacetime is characterized by the form of the invariant $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$ infinitesimal interval between events. Lorentz boosts relate the spacetime coordinate systems of the members of this class of observers.

The expression for a boost not along one of the spatial coordinate axes is more complicated, but the basic geometric structure of the spacetime of special relativity can be understood by a selection of coordinate axes such that any relative motion considered is along one of the spatial axes specified. It can be shown that if we consider that some inertial observer observes an object moving at v relative to them in some direction – let's call this direction the x^1 direction for simplicity – and we boost the various positions and times this inertial observer ascribes to this moving object (defined by $\frac{dx^1}{dt} = v$) into a frame of reference moving at $-u$ relative to this observer along the x^1 direction, then the positions and times this relatively moving observer will ascribe the moving object will be such that $v' = \frac{dx'^1}{dt} = \frac{u+v}{1+\frac{uv}{c^2}}$. Thus, if observer B moves relative to observer A at $+u$ along some direction and determines an object to be moving at $+v$ along that same direction considered, observer A will measure a velocity of that object of $u \oplus v = \frac{u+v}{1+\frac{uv}{c^2}}$. This tells us that the general formula for the special-relativistic collinear addition of velocities is

$$v_1 \oplus v_2 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}$$

From this, it can be easily shown that if we define the *rapidity* ϕ of a moving object in some inertial frame of reference by $\frac{v}{c} = \beta = \tanh\phi$ in some inertial frame of reference, since

$$\tanh(\phi_1 + \phi_2) = \frac{\tanh\phi_1 + \tanh\phi_2}{1 + \tanh\phi_1 \tanh\phi_2}$$

, the change in rapidity for a collinear boost is something all inertial observers agree upon: ϕ + dϕ = ϕ' + dϕ', where the primes indicate the same quantities in a new frame of reference after a particular transformation of the Lorentz group. With this definition of rapidity, it also follows that $\gamma = \cosh\phi$ and thus the Lorentz boosts take on the form

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

Lorentz boosts can then be interpreted as hyperbolic rotations of the positions of events and times of events in the direction of relative velocity. Successive collinear boosts become continual hyperbolic rotations, and the interval $v \in [-c, c]$ maps to $\phi \in [-\infty, \infty]$.

In terms of the invariant spacetime interval ds, if $ds^2 > 0$ the infinitesimal *proper time* between these two events is $d\tau = \frac{ds}{c}$, as it is the purely temporal interval that would be measured by an inertial observer for whom the two events happened at the same place. If $ds^2 < 0$, the events cannot occur at the same place for *any* inertial observer. However, there does exist an inertial observer who determines these events to occur at the same time, in this case $d\sigma = \sqrt{-ds^2}$ is the infinitesimal *proper distance* between the events, as it is the magnitude of the purely spatial separation measured by an observer for whom these events are simultaneous. For a given world line of some arbitrary (not neccessarily inertial) observer in some other observer's inertial frame of reference, $\Delta\tau = \int d\tau$ is the amount by which a clock (a clock is just a controlled sequencing of events) will have advanced. The basic chronogeometric structure of the spacetime of special relativity is such that the straighest, i.e. inertial, paths through the spacetime, are also the *longest* paths as measured by the proper time, as this expression for the amount by which a clock will have advanced is maximized for a world line that is at rest at the origin of an inertial frame of reference in which $\Delta\tau = \int d\tau = \int \sqrt{dt^2 - \left(\frac{dx}{c}\right)^2} = \int \frac{dt}{\gamma}$ is valid. This is the essence of the twin paradox in special relativity.

The kinematics and dynamics of special relativity are expressed naturally through the four-vector formalism. We may define the four-velocity of a world line − which is a collection of positions and times at which an object is– as $U^\mu = \frac{dX^\mu}{d\tau} = \gamma(1, \vec{u})$ where $\vec{u}$ is the three-velocity of the world line in the given frame of reference, $\vec{u} = \frac{d\vec{x}}{dt}$. We can then define the four-momentum as $P^\mu = m_0 U^\mu$, where m_0 is the mass of the object considered. Because the four-velocity is defined in terms of proper time, which all observers agree upon, and four-momentum is proportional to four-velocity, it follows that the norm squared of this four-vector, $\eta_{\mu\nu} P^\mu P^\nu$, should be frame-invariant. In the rest frame of an object of mass m_0, we have $\eta_{\mu\nu} P^\mu P^\nu = m_0^2 c^2$, and more generally $\eta_{\mu\nu} P^\mu P^\nu = \gamma^2 m_0^2 c^2 - (\gamma m_0 \vec{u}) \cdot (\gamma m_0 \vec{u})$. Since it should be that the norm squared is a Lorentz invariant, it should be that $m_0^2 c^2 = \gamma^2 m_0^2 c^2 - (\gamma m_0 \vec{u}) \cdot (\gamma m_0 \vec{u})$.

We can now assign a meaning to the temporal component of an object's four-momentum in that object's rest frame, $|P^\mu| = m_0 c$. If we multiply this by c, it has units of energy. We can then postulate that this energy is the energy an object has simply by virtue of its passage through time, $E_{rest} = m_0 c^2$. This postulate has been vindicated by the phenomena of nuclear fusion and nuclear fission. In addition, further considerations make such a definiton of rest energy attractive for a self-consistent framework for describing momentum, energy, and force in special relativity. This postulate is a general equivalance of matter and energy, and additionally predicts that the thermal energy a body has contributes to its effective mass. For example, during an inelastic collision wherein non-conservative forces are applied, kinetic energy is converted by those non-conservative forces into thermal energy internal to the system, and this, according to the special theory of relativity, increases the effective mass of the system.

The quantity $\vec{p} = \gamma m_0 \vec{u}$ can be defined as the *three-momentum* of an object, as mathematical analysis can show that the total three-momentum of a system is conserved after collisions in all frames of reference. If we then define the three-force acting on an object as

$$\vec{f} = \frac{d\vec{p}}{dt} \tag{2}$$

, and adopt the classical definiton of work,

$$dW = \vec{f} \cdot \vec{dx} = \vec{f} \cdot \vec{u} dt$$

, it can be shown that $dW = m_0 c^2 d\gamma$ (Rindler , 1964). Hence if we define the kinetic energy an object has, i.e. the energy it has by virtue of its motion, as $K = (\gamma - 1)m_0 c^2$, it follows that

$$\Delta K = \int dW = \int \vec{f} \cdot \vec{u} dt \tag{3}$$

, and the *total* energy an object has is given by $E = E_{rest} + K = \gamma m_0 c^2$. This, with our definiton of three-momentum, gives an energy-momentum relationship of $E^2 = (pc)^2 + (m_0 c^2)^2$. Mathematically, $K = (\gamma - 1)m_0 c^2$ reduces to the classical $K = \frac{p^2}{2m}$ when $p << m_0 c \leftrightarrow v << c$. We note that, with the definition (2) of three-force, the three-force and the classical three-acceleration, $\vec{a} = \frac{d\vec{u}}{dt}$ are proportional in a frame comoving with the object, which is the only frame in which the physical interaction giving rise to the non-zero three-force can be directly measured.

Just as the selection of a preferred class of observers (inertial) in conventional Newtonian theory can be interpreted as the law of autoparallel transport of a four-velocity vector in a globally (affinely) Euclidian spacetime, i.e. an affine flat spacetime, the selection of a preferred class of inertial observers in special relativity means that the spacetime of special relativity is affine flat, as well. The essential difference is that the relations of spacetime points in special relativity is characterized by a hyperbolic chronogeometry rather than the absolute time and Euclidian three-geometry postulated of Galiliean-Newtonian spacetime. This spacetime, the spacetime of special relativity, is called Minkowski spacetime, after German mathematician Hermann Minkowski, who first realized special relativity could be intepreted (chrono)geometrically as a globally flat, hyperbolic spacetime. The task of formulating the field equations of general relativity then becomes how to impose non-flat affine structure on a spacetime with hyperbolic chronogeometric structure. This we will do in Section 4. Before that, though, let us build some intuition as to what to expect from the gravitational equivalence principle, – that the effects of gravitation and being in an accelerated reference frame are locally indistinguishable – and considerations of acceleration in affine flat, Minkowski spacetime.

4 Acceleration in Minkowski spacetime

Let us consider an observer accelerating at a constant rate in Minkowski spacetime. We must first define what this means. Because changes in rapidity are agreed upon by all observers, and changes in proper time are agreed upon by all observers, a world line with a constant acceleration is one where the rapidity of that world line is constantly increasing as a function of the proper time of that world line:

$$\frac{d\phi}{d\tau} = l$$

where l is, for now, an undetermined constant, although we see that, since rapidity is unitless, l must have units of inverse time.

We may as well consider the initial time of the observer's acceleration to be at $\tau = 0$, i.e. the zero time by the clock of the observer with this world line, thus giving

$$\phi = l\tau$$

The world line is characterized by

$$\beta(\tau) = \tanh(l\tau) \to v = c\tanh(l\tau)$$

in some inertial frame of reference. To find the position x of the world line as a function of the proper time of the world line, we note that $\frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau} = v\gamma$.

It is also well known that $\gamma = \cosh\phi = \cosh(l\tau)$, giving

$$\frac{dx}{d\tau} = c\sinh(l\tau)$$

We may now interpret the constant l by noting that

$$\frac{d^2x}{d\tau^2} = cl\cosh(l\tau)$$

and thus

$$\frac{d^2x}{d\tau^2}|_{\tau=0} = cl$$

What is the quantitity $\frac{d^2x}{d\tau^2}$ at $\tau = 0$? At $\tau = 0$, $v = 0$ and thus the frame with spacetime coordinates (t,x) is *comoving* with the accelerated world line at $\tau = 0$. Thus, in the infinitesimal temporal interval for the accelerated observer around $\tau = 0$, $\frac{d^2x}{d\tau^2}$ is the acceleration *felt* by an observer in the accelerated reference frame, i.e. our intuitive notion of what it means to accelerate. We thus give $\frac{d^2x}{d\tau^2}|_{\tau=0}$ the name g, noting that it represents the three-acceleration of the world line in a comoving frame. Thus we see that $l = \frac{g}{c}$, which does indeed have units of inverse time. We may now find coordinate x of the world line as a function of the world line's proper time by

$$\int_{x_0}^{x} dx' = c\int_{\tau_0}^{\tau} \sinh\frac{g\tau'}{c}d\tau'$$

If we say that $\tau_0 = 0$ and $x_0 = 0$, i.e. construct a situation in which, at $\tau_0 = 0$, the frame with spacetime coordinates (t,x) is both comoving and spatially coincident with the constantly-accelerating world line, we have

$$x(\tau) = \frac{c^2}{g}\left[\cosh\frac{g\tau}{c} - 1\right] \tag{4}$$

We also use $dt = \gamma d\tau$ to arrive at

$$\int_{t_0}^{t} dt' = \int_{\tau_0}^{\tau} \cosh \frac{g\tau'}{c}$$

If we set $t_0 = 0$, i.e. say that the frame with spacetime coordinates (t,x) is momentarily coincident and comoving with the constantly accelerating world line at $t = 0$, we have

$$t(\tau) = \frac{c}{g} \sinh \frac{g\tau}{c} \tag{5}$$

The points $\big(\mathrm{t}(\tau),\mathrm{x}(\tau)\big)$ parameterized by the proper time of the world line define the world line as determined by the inertial observer using spacetime coordinate system (t,x), and is a hyperbola including the point (0,0) asymptoting to lines with slope $\frac{dx}{dt} = \pm c$ as $\tau \to \pm\infty$.

We now seek to characterize the reference frame of an observer following the accelerated world line. Following the ideas of Hermann Bondi's k-calculus as presented in Bondi (1964) and used extensively in the lesson presently on the Center for Einstein Studies website inspired by this book – we can say that, in a special-relativistic conceptual framework informed by the fact that the speed of light is independent of its source, the only meaningful account an observer can give of when and where an event happens is by measuring that event with radar. That is, by sending a radar pulse at some change in the universe, noting the time by their watch at which the radar pulse was sent, and recieving that radar pulse back, and noting the time by their watch at which this happened. Let us consider this process for the accelerating observer measuring events that the inertial observer says happen at t = 0. In this case, the accelerating observer would have to send out a pulse at some proper time $\tau = -\tau_0$ and recieve it back at $\tau = \tau_0$. The observer would then deduce that the event was $\pm\tau_0$ light-seconds away, depending on the direction he or she initially sent the radar pulse.

Let us then mark the events $\big(\mathrm{t}(-\tau_0),\mathrm{x}(-\tau_0)\big)$ and $\big(\mathrm{t}(\tau_0),\mathrm{x}(\tau_0)\big)$ in the inertial frame and consider that they are connected by a time-forward null cone and a time-backward null cone, respectively, to some event along the inertial observer's hyperplane of simultaneity t=0. We know that, in inertial frames of reference in special relativity, any entity propagating along null cones propogates *globally* with speed ±c along some spatial direction of the null cone. Let us, then, connect the events $\big(\mathrm{t}(-\tau_0),\mathrm{x}(-\tau_0)\big)$ and $\big(\mathrm{t}(\tau_0),\mathrm{x}(\tau_0)\big)$ with null cones *in the inertial frame*, and see at what value of x they intersect with each other *in the inertial frame*. This will allow us to see how the position (x,0) in the inertial frame is measured by the accelerating observer – particularly, if the accelerating observer uses radar, he or she will say this event was ± cτ_0 away.

Light pulses being sent along the -x direction will have slope $\frac{dx}{dt} = -c$ in the inertial frame, and will thus be of the form

$$ct - ct_1 = -(x - x_1) \tag{6}$$

,while light pulses being sent along the +x direction will have slope $\frac{dx}{dt} = c$ in the inertial frame, and will be of the form

$$ct - ct_1 = x - x_1 \tag{7}$$

where t_1, x_1 are the spacetime coordinates of a particular event on the light pulse. Equations (6) and (7) are of the point-slope form of a null geodesic in the inertial reference frame. We now consider that the accelerating observer is looking "backward", i.e. in the direction away from which they are accelerating. In this case, the light pulse sent at $\tau = -\tau_0$ is of the form of equation (6), and the event $\big(\mathrm{t}(-\tau_0),\mathrm{x}(-\tau_0)\big)$ is certainly on this light pulse, as this is where the observer was when the light pulse was sent. In the same way, the pulse recieved at $\tau = \tau_0$ will be of the form (7) and include the point $\big(\mathrm{t}(\tau_0),\mathrm{x}(\tau_0)\big)$.

We now use equations (4) and (5) to show that the light pulses connecting a point (x,0) to the accelerating world line are related to the proper time interval τ_0 by

$$x = \frac{c^2}{g}\big[\cosh\frac{g\tau_0}{c} - \sinh\frac{g\tau_0}{c} - 1\big]$$

renaming $x \to x_b$, reminding us that this relation holds only for events "behind" the accelerating observer, we have

$$x_b = \frac{c^2}{g}\big[e^{-\frac{g\tau_0}{c}} - 1\big] \tag{8}$$

Performing the same analysis for events with $x > 0$ by using equation (6) with the event $\big(\mathrm{t}(\tau_0),\mathrm{x}(\tau_0)\big)$ and equation (7) with the event $\big(\mathrm{t}(-\tau_0),\mathrm{x}(-\tau_0)\big)$, and calling this x position measured by radar x_f, we have

$$x_f = \frac{c^2}{g}\big[e^{\frac{g\tau_0}{c}} - 1\big] \tag{9}$$

As $\tau_0 \to \infty$, $x_b \to -\frac{c^2}{g}$, i.e. the event (x,t) $= (-\frac{c^2}{g},0)$ is *infinitely* far away according to the accelerating observer's radar technique – they would have had to have sent the pulse at $\tau = -\infty$ and would recieve it back at $\tau = \infty$. It is easy to see, then, that the partial null cone of this event extended in the spatial direction the world line is accelerating (+x) defines the lines to which the constantly accelerating observer asymptotes as $\tau \to \pm\infty$ according to the inertial observer – see Figure 1 in the Appendix. On the other hand, the accelerating observer has no trouble measuring any of the events towards which he or she is accelerating.

We can see with Figure 1 that events lying on the line $ct = x + \frac{c^2}{g}$ with $x > -\frac{c^2}{g}$ also take infinite time to interact with the accelerating observer, as any interaction mediated along the future null cones of these events in the direction of the accelerating observer is coincident with the null spacetime world line $ct = x + \frac{c^2}{g}$, while the accelerating observer could have easily interacted with these events along the time-backward segments of their null cones. See Figure 2. In the same way, we can see that it would have taken the accelerating observer sending a causal signal along his or her -x oriented, time-forward null cone at $\tau = -\infty$ to interact with any events along the null spacetime world line $ct = -x - \frac{c^2}{g}$ with $x > -\frac{c^2}{g}$, while these events could easily interact with the accelerating observer along their time-forward null cones. See Figure 3. Any event inside of the wedge defined by the two null spacetime world lines of $ct = \pm x \pm \frac{c^2}{g}$ and the condition $x > 0$ can be causally connected with the accelerating observer along both their time-forward and time-backward null cones. What about the relation of the events outside of this wedge to the accelerating observer? We can actually break these down into three classes.

If the events have spacetime coordinates according to the inertial observer in the spacetime region defined by $ct > x + \frac{c^2}{g}$ and $ct > -x - \frac{c^2}{g}$ (now defined as Region I), these events can *never* causally interact with the accelerating observer, but the accelerating observer can causally interact with the events. See Figure 4. If the events according to the inertial observer in the spacetime region defined by $ct < x + \frac{c^2}{g}$ and $ct < -x - \frac{c^2}{g}$ (let us define this as Region II), then the observer can *never* causally interact with the events, but the events can causally effect the observer. See Figure 5. The remaining case is when the events lie in the spacetime region defined by the inertial observer as $ct > x + \frac{c^2}{g}$ and $ct < -x - \frac{c^2}{g}$ (let us define this as Region III) – in this case neither the events nor the observer can ever causally interact with each other. See Figure 6. Thus, we see that the partial null cone of the event (-$\frac{c^2}{g}$,0) in the +x direction defines the limit of when and where there is two-way causal interaction possible along null cones between events and the accelerating observer considered here. This is called the *Rindler Horizon.* Since a two-way causal interaction is necessary for determining when and where an event happens by the radar technique of (Bondi , 1964), the events lying beyond the Rindler horizon aren't even meaningfully *in* the frame of reference of the accelerating observer, as this observer has no way of assigning spacetime coordinates to these events. Thus, whatever frame of reference we can define for an accelerating observer can neccessarily only map the wedge of spacetime in which two-way causal interactions are possible, often called the Rindler wedge. The observer moving with constant acceleration is often called a Rindler observer, here we call this observer the fiducial Rindler observer, as it is the observer whose perspective we are considering.

Now we can see result which is, at first, quite surprising. We might intuitively say that an observer some distance separated from the fiducial Rindler observer, this observer must accelerate at *the same* rate as the fiducial Rindler observer in order to remain at that same distance from the fiducial observer. However, if we consider an observer accelerating at g in a comoving frame but at an initial position x_0 such that $-\frac{c^2}{g} < x_0 < 0$, then this observer is only within the Rindler horizon of the fiducial observer for some interal of their proper time η centered about $\eta = 0$ – see Figure 7 in the Appendix. But we have already said that any events beyond the Rindler horizon are not even *in* the reference frame of the fiducial observer – certainly it can't be that the observer remains at a constant distance with respect to the fiducial Rindler observer if this observer has left the reference frame of the fiducial observer. The solution to this difficulty is to say that all observers in the fiducial Rindler observer's reference frame, hereafter referred to as *Rindler observers*, must share the same Rindler horizon. By inspection of Figure 7, it is clear that this means that different observers at different positions must have different accelerations. Thus, we can repeat our analysis for a given Rindler observer starting at position x_0 and allow g to vary as a function of this initial position $g \to g(x_0)$ and require that $g(0) = g_0$ is the same as the acceleration g previously considered. Here we are simply renaming $g \to g_0$ in equations (4), (5), (8), and (9), and allowing the acceleration g more generally to be a function of initial position in the inertial reference frame.

The analogues of equations (4) and (5) then become

$$x(\tau) = x_0 + \frac{c^2}{g_(x_0)}\big[\cosh\frac{g(x_0)\tau}{c} - 1\big] \tag{10}$$

and

$$t(\tau) = \frac{c}{g(x_0)} \sinh \frac{g(x_0)\tau}{c} \tag{11}$$

Performing the same analysis that lead us to the Rindler horizon in the first place with these new expressions gives an analogue of equation (8) of

$$x_b = x_0 + \frac{c^2}{g(x_0)} \left[e^{-\frac{g(x_0)\tau_0}{c}} - 1 \right] \tag{12}$$

We now require the limit of the above expression as $\tau_0 \to \infty$ is $-\frac{c^2}{g_0}$, as before. This gives

$$g(x_0) = \frac{g_0}{\frac{x_0 g_0}{c^2} + 1} \tag{13}$$

which degenerates as $x_0 \to -\frac{c^2}{g_0}$, i.e. a Rindler observer *at* the Rindler horizon must have infinite acceleration to stay in the Rindler wedge, and goes to 0 as $x_0 \to \infty$.

This can also be understood from the perspective of the inertial observer. Consider the case shown in Figure 7, of two separated world lines with the same constant acceleration. In this case, from the perspective of the inertial observer, as t increases, both observers are moving more and more quickly. Thus, by length contraction, the separation of two accelerating observers, which is constant as seen by the inertial observer, at some later time, is shorter than the proper separation of the two observers, i.e. the separation as measured by an observer who is comoving with the two accelerating observers at that later time. Thus the proper separation of the two accelerating observers increases as a function of time – if a material rod were connecting the two observers, it would soon break. The only way for the proper separation of accelerating observers to remain constant is for the accelerating observers to be converging *from the perspective of the inertial observer.* This is only possible if the observers further along the direction of acceleration accelerate less rapidly, with the amount of acceleration as a function of position given by equation (13).

We can also come to some startling conclusions about the nature of time in an accelerated reference frame. If an observer further along the direction of acceleration attempts to measure, via radar, the time at which events separated by a constant proper time interval along the fiducial Rindler observer's occur, this observer will find that these times are separated by a larger interval of their own proper time. To prove this mathematically, we need to find the times of intersection of the world line of the measuring (non-fiducial) Rindler observer with the time-forward and time-backward null cones of a particular event of a particular proper time on the fiducial Rindler observer's – the time of this event as determined by the non-fiducial, measuring Rindler observer is then half way between these times of intersection. This process can then be repeated for a succession of events on the fiducial Rindler observer's world line that are separated by an interval of constant proper time, and this interval can be compared to the resulting separation of the times of the events as determined by the measuring observer. This is all made much clearer by Figure 8. It turns out that finding the intersection times algebraically gives rise to a transcendental equation, i.e. one that cannot be solved analytically. Instead, we solve this problem numerically and plot the results in Figure 8. Looking at Figure 8, we do indeed see that the observer further along the direction of acceleration measures times of events such that these times have a larger interval of separation than the interval of proper time separation of the events measured along the fiducial Rindler observer's world line.

We can repeat this analysis, but for the case of the fiducial Rindler observer measuring the times of events along the world line of some Rindler observer closer to the Rindler horizon, such that the separation of these observers is the same is in the case considered in Figure 8. This we show in Figure 9, which shows us that the nature of the effect is the same, but that its magnitude is larger than in the case considered in Figure 8. Referencing Bondi (1964), we see that such a radar technique is sufficient for determining the route-dependence of time in special relativity. Thus, in the same way, this radar technique seems to tell us that *an interval of a Rindler observer's proper time corresponds to a larger interval of the fiducial Rindler observer's coordinate time closer to the Rindler horizon.* Thus, if an observer were to be in the Rindler observer's frame of reference and move first closer to the Rindler horizon, and then return to the fiducial observer's point in the reference frame, the clock of the observer who spent some time near the Rindler horizon would be behind, in offset, the clock of the fiducial observer. By behind in offset, we mean that the clocks, upon returning to the same point, would continue to run at the same rate, but that one would be a constant interval behind the other, i.e. it would have aged less owing to the different histories of the two world lines considered. We can also predict from Figures 8 and 9, that radiation emitted by Rindler observers will be redshifted as it advances in the direction of acceleration, as the intervals of time between the emission of maxima of the electromagnic waves comprising the radiation are longer intervals compared to intervals of proper time further along the Rindler wedge. On the basis of the gravitational equivalance principle, we can then predict that, if two world lines start from the same spacetime point and come back together at some other point in the spacetime, with one of the world lines spending more of its proper time nearer to a gravitating mass, the clock of that world line will be behind that of the other in offset. We may also predict, on the basis of the gravitational equivalence principle, that radiation climbing out of a gravitational well will be redshifted.

To make more quantitative statements about the fiducial Rindler observer's reference frame, it is neccessary to map, in some way, the spacetime coordinates (t,x) in the Rindler wedge of the inertial observer to coordinates of constant coordinate time of the fiducial Rindler observer, T, and constant X coordinate of the Rindler observer. These coordinates can then be used to construct a metric of the coordinates (T,X) from their relation to (t,x) and the requirement that the spacetime interval is $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$. This can be done by comparing the spatiotemporal coordinate assignments of the events (t,x) for a momentarily coincident and comoving inertial frame at all points of the fiducial Rindler observer's world line. This allows us to characterize the relative distances and times of events along the world lines of the various Rindler observers as determined by such a class of comoving, coincident inertial observers. This is done in Semay (2006), and results in a relation of coordinates of

$$ct = (X + \frac{c^2}{g}) \sinh \frac{gT}{c}, \tag{14}$$

$$x = (X + \frac{c^2}{g}) \cosh \frac{gT}{c} - \frac{c^2}{g} \tag{15}$$

and a resulting expression for the spacetime interval of

$$ds^2 = \left(1 + \frac{g_0 X}{c^2}\right)^2 c^2 dT^2 - dX^2 - dY^2 - dZ^2 \tag{16}$$

Figure 10 shows lines of constant X, which are world lines of Rindler observer's, and of constant T, in the inertial observer's coordinate system (t,x). We can see that, in the limit of the position X going to the Rindler horizon, an infinite amount of the Rindler observer's coordinate time T is 0 amount of time t, and as $X \to \infty$, a finite amount of the Rindler observer's coordinate time T is an infinite amount of time t. If an observer in the Rindler observer's frame of reference quasistatically (i.e. such that it is the case that this observer is always approximately comoving and coaccelerating with a Rindler observer at their position) moves from the origin X = 0, to near the Rindler horizon, and returns to the origin X = 0, this observer's clock will lag that of the fiducial Rindler observer in offset, according to the inertial observer with coordinate system (t,x). Because the comparison of clocks is a physical measurement, it should be that this result does not depend on the inertial observer's coordinate system, and this result is what the two observers (the fiducial Rindler observer and the observer who moved near the Rindler horizon) actually report. Thus, Figure 10 allows us to visualize the positional dependence of the relative passage of proper time of various Rindler observers. Indeed, the metric (16) gives us this prediction – in the fiducial Rindler observer's reference frame (T,X), one second of proper time for a stationary observer is equal to one second of proper time at X = 0, i.e. for the fiducial Rindler observer, with the more general relation given by $d\tau = (1 + \frac{g_0 X}{c^2})dT$. We note that, throughout the process of comparing these observer's clocks, four-forces must have been applied, as the observers are accelerating. Further, since the observers nearer the Rindler horizon have a greater proper acceleration, the observer who came near the Rindler horizon and returned to X = 0 had to have more four-force applied throughout this process, and this is the observer whose clock lags, in offset, the clock of the fiducial Rindler observer.

The Rindler observer's coordinate position X agrees with the inertial observer's coordinate distance x at t = T = 0, whereas we saw in introducing the Rindler horizion, the radar distance of Rindler horizon according to the Rindler observer is *infinite*. Thus the expressions (14), (15), and (16) are not defined in terms of radar distance. Indeed, these expressions were, from the outset, constructed in terms of momentarily coincident and comoving inertial observers, of whom the observer with coordinate system (t,x) is one – at t = x = 0.

With this method of defining distances, it follows that the Rindler horizon always appears to be at $X = -\frac{c^2}{g_0}$ for all values of T. In the full four dimensional treatment, the Rindler horizon is at $X = -\frac{c^2}{g_0}$ regardless of the values of Y and Z – thus, the fiducial Rindler observer sees a "plane" form $\frac{c^2}{g}$ in the direction away from which they are accelerating, at which it appears that time "comes to a stop" and from which all radiation is infinitely redshifted. The expressions (14), (15), and (16) are actually modified forms of what are, aptly, called *Rindler coordinates*, after their originator. Rindler coordinates are related to (14), (15) by moving the origin of the inertial observer with coordinate system (t,x) to the Rindler horizon. In this case, the spacetime interval is of the form

$$ds^2 = \frac{g_0^2 X^2}{c^4} c^2 dT^2 - dX^2 - dY^2 - dZ^2 \tag{17}$$

and the Rindler coordinate mapping is shown in Figure 11. In this mapping, the proper acceleration of a Rindler observer , originally given by (13), is simply given by $g = \frac{c^2}{X}$. The observers who are force-free will accelerate, in this frame of reference, opposite the direction of acceleration of the frame of reference, and the proper time as given by the metric (17) or (16) will be maximized. That is to say, it is still the case that the straightest, inertial path in special relativity is the longest such path as given by the proper time of the observer following that path. It is simply that this chronogeometric structure of special relativity has been redescribed from the perspective of a non-inertial observer. This is shown mathematically in Semay (2006).

5 The General Theory

We now present the the field equations of the general theory of relativity and a conceptual motivation for their particular form. The basic task for formulating a spacetime theory of gravitation is to combine the chronogeometric structure of Minkowski spacetime demanded by the independence of the propogation velocity of electromagnetic radiation of the source velocity with the gravitational equivalence principle that demands we impose non-flat affine structure on the spacetime. How is the chronogeometric structure of the Minkowski spacetime specified? The chronogeometry is conveniently represented by the *4-metric* $\eta_{\mu\nu} = diag\{1, -\frac{1}{c^2}, -\frac{1}{c^2}, -\frac{1}{c^2}\}$, if the four coordinates are the set of the three ordinary spatial coordinates and ordinary time t, in some inertial frame of reference which is the same in all frames of reference related to this frame by Lorentz transformations. The invariant spacetime interval is then $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ in all such frames. We can, with equal validity, define the 4-metric as $\eta_{\mu\nu} = diag\{c^2, -1, -1, -1\}$, as all four coordinates are, as before, of the same nature. In this latter case, we have made them spatial; in the previous case we made them temporal. It is merely a matter of personal taste whether to take all spatio-temporal dimensions and resulting spacetime interval as spatial or temporal in nature – all that matters is that, once a convention is chosen, it is followed consistently in a given treatment of the theory. We may also define the 0th coordinate to be the speed of light times the ordinary time t, and the four-metric becomes $\eta_{\mu\nu} = diag\{1, -1, -1, -1\}$. This makes all spatio-temporal coordinates spatial before imposing metric structure on their interrelations. We may also, prior to imposing a four-metric of $\eta_{\mu\nu} = diag\{1, -1, -1, -1\}$, define the three spatial coordinates to be the ordinary spatial coordinates divided by the speed of light c, making all coordinates temporal in nature before imposing metric structure on their interrelations.

In special relativity, gravitation is not considered a part of the chronogeometric structure of the spacetime, but rather treated as a traditional three-force, as in conventional Newtonian theory. We have already seen that, even without considering the postulates of special relativity, we should regard gravitation as a modification of the structure of the spacetime itself. The question is now: how do we relate the chronogeometric structure of the spacetime – a structure conveniently characterized by a 4-metric – to the dynamical, inertio-gravitational structure of the spacetime imposed by the affine connection?

Mathematically, there exists a well-defined notion of an affine connection on a manifold that is *fully compatible* with a metric of that manifold, i.e. such that the covariant derivative defined by the affine connection of the metric itself vanishes:

$$D_\lambda g_{\mu\nu} = 0$$

where $g_{\mu\nu}$ is the metric of the manifold, such that $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ gives the infinitesimal length ds between points x_0^ν and $x_0^\nu + dx^\nu$.

It is useful to define the inverse metric $g^{\mu\nu}$: $g^{\mu\nu} g_{\nu\kappa} = \delta^\mu_\kappa$. The metric and its inverse can be used to convert contravariant tensors to covariant tensors and vice versa if such tensors are defined in metric manifolds.

It can be shown that a mathematical definition of metric-compatible connection components relates to the metric and its inverse via

$$\Gamma^{\mu}_{\kappa\nu} = \frac{1}{2} g^{\mu\gamma}(g_{\kappa\gamma,\nu} + g_{\gamma\nu,\kappa} - g_{\kappa\nu,\gamma})$$

These values are called the *Christoffel symbols of the second kind.*

If we take the affine connection components of a manifold to be derivable from the metric of a manifold in this way, it can be shown that the autoparellel paths given by equation (1) are also the *extremal* paths through the manifold, i.e. the paths between fixed points x_1 and x_2 in the manifold for which the line element $\int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{g_{\mu\nu} dx^{\mu} dx^{\nu}}$ are *stationary.* The terms stationary and extremal are defined formally and mathematically, but conceptually, for our purposes, it amounts to saying that the integrated length of the path between two points x_1 and x_2 is either the *longest* or *shortest* such paths, as defined by the metric, between the two fixed points.

Let us consider the foundation of geometry, Euclidian geometry. Euclidian geometry can, in our language, be characterized as an affine-flat manifold of n-dimensions, as the affine connection was defined in terms of comparing neighboring tangent spaces wherein Euclid's postulates are valid. Thus, if Euclid's postulates are globally valid, no affine connection is needed – in other words, all connection components must be zero. If we say that Euclidian distances are characterized by the intuitive $d\sigma^2 = dx^2 + dy^2 + dz^2$, in three-space, or, more generally $d\sigma^2 = dx^i dx^i$, we may characterize the Euclidian space by a metric $g_{\mu\nu} = \delta_{\mu\nu}$ and $g^{\mu\nu} = \delta^{\mu\nu}$. For this metric, all components of the Christoffel symbols clearly vanish, and it is the case that the straightest paths in a Euclidian space (autoparallels) are also the shortest paths as measured by its metric (geodesics).

We may also mathematically describe the autoparallel / geodesics of the geometry of the surface of a sphere by emedding the sphere in a Euclidian three-space and relating the infinitesimal distances in the Euclidian three-geometry to infinitesimal changes in latitude and longitude along that surface, $d\varphi$ and $d\theta$, respectively. Such an analysis gives $g_{\mu\nu} = diag\{R^2, R^2 \sin^2\theta\}$ with inverse metric $g^{\mu\nu} = diag\{\frac{1}{R^2}, \frac{1}{R^2 \sin^2\theta}\}$, where R is the radius of the sphere considered. Then, objects constrained to move along the surface of such a sphere will follow the geodesic / autoparallel paths as defined by the Christoffel symbols and equation (1), which is also the geodesic equation when the metric and the connection are made fully compatible.

In the case of general relativity, it must be that, locally, special relativity applies. In particular, for some observer with no four-force it must be that a frame of reference not accelerating and momentarily comoving with the four-force-free observer, commonly called a local inertial frame (LIF), the 4-metric must be of the form of Minkowski spacetime, i.e. of signature {1,-1,-1,-1 } (or {-1,1,1,1}). However, the inability to make the affine connection components vanish globally in any frame of reference in Newtonian gravitational theory tells us that if the connection components are derivable from the metric, the global metric – i.e. the metric tensor *field* across the manifold – must vary from point to point in the manifold. Further, we have seen that the curvature of the spacetime as defined by the affine Riemmann curvature tensor in general will vary from point to point on the manifold. Indeed, Newtonian gravitation is reinterpreted as the divergence of this tensor. Thus, if we require that the affine connection and the metric be fully compatible, we cannot specify any conditions on the four-metric apriori, other than that it must be locally transformable into the form of the Minkowski metric. Such a metric is an example of a *pseudo-Riemmannian* metric.

We now offer the motivation for making the metric and the connection compatible. In special relativity, we have seen that the straightest possible path between two events is the longest in terms of the proper time of an observer. Mathematically this is captured by the signature of the Minkowski metric. In special relativity, the straightest path is also the *force-free, inertial path.* In our reconceptualization of Newtonian gravitational theory, the straightest, force-free, inertial paths are described by the affine connection and the autoparallel equation (1). Since it is the case that, if the metric and connection are fully compatible, autoparallel paths leave the integrated intervals characterized by the metric extremal (usually either maximized or minimized), and since we have shown that for the affine-flat Minkowski spacetime, the autoparallel (inertial) paths maximize proper time, and we require that the metric be locally transformable into the form of the Minkowski metric, it follows that, *if we take the metric and the connection to be fully compatible, the straight, four-force-free paths through the spacetime between two events maximize the proper time elapsed between those two events.*

Thus, if we are to synthesize the two insights informing the general theory of relativity – the gravitational equivalence principle demanding we impose non-flat affine structure on the spacetime, and the extension of the Galliliean principle of equivalence to electromagnetic phenomena leading to a chronogeometry conveniently described by a four-metric – it appears that this is the approach we must take. The proper time of a timelike world line ($g_{\mu\nu}dx^\mu dx^\nu > 0$ for all points on the world line) is the time as measured by the watch of the observer following that world line, defined, as in special relativity, by $d\tau = \frac{ds}{c}$. The proper distance of spacelike world lines, ($g_{\mu\nu}dx^\mu dx^\nu < 0$ for all points on the world line) is then defined, as in special relativity, by $d\sigma = \sqrt{-ds^2}$. We parameterize timelike worldlines by the proper time of those world lines, so that the autoparallel / geodesic equation (1) is parameterized by the proper time of the geodesic considered.

The field equations of general relativity in vacuum become, then, as in Newtonian gravitational theory informed by the gravitational equivalence principle, the vanishing of the Ricci tensor:

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

The difference between the vacuum field equations of general relativity and the vacuum field equations of non-flat Newtonian spacetime is simply that the affine connection components used to compute the Ricci tensor are derivable from the 4-metric of the manifold as the Christoffel symbols of the second kind, and the first task of solving and interpreting the field equations becomes solving for the 4-metric which satisfies the vanishing of the Ricci tensor formed from the Christoffel symbols. Because the 4-metric characterizes the chronogeometry of the manifold, and we can specify no conditions on the 4-metric apriori, we cannot, in the general theory of relativity, specify anything of the nature of space and time before solving the field equations, which in vacuum are simply the vanishing of the Ricci tensor.

5.1 The Schwarzchild Solution

We now present the Schwarzchild solution to the Einstein vacuum field equations, $R_{\mu\nu} = 0$, and the key steps in its derivation, as done by (Rindler , 1964). The Schwarzchild solution is the solution for the metric and connection components for a spherically symmetric mass distribution, and is one of few exact solutions to the Einstein field equations – nonetheless it is conceptually elucidating.

Apriori we can say nothing about the metric $g_{\mu\nu}$, other than that is is pseudo-Riemmannian. However, by spherical symmetry we can say that we expect the three-space to deviate, if at all, from Euclidian geometry radially with respect to the spherically symmetric distribution. A Euclidian line-element in spherical-polar coordinates has the metric $d\sigma^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2$, giving a metric of $g_{ij} = diag\{1, r^2, r^2 \sin^2\theta\}$. This, as the reader can verify, already gives rise to non-zero Christoffel symbols, as the local description of what it means to be "in the θ direction" or "in the φ direction" varies from point to point in the Euclidan 3-manifold in such a coordinatization. The reader can also verify that the Riemmann tensor vanishes for this metric. Since we only expect, if any distortion of the 3-space from Euclidian, it only to be in the radial direction, we make the Ansatz that the spatial part of the 4-metric is of the form $d\sigma^2 = \beta(r)dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2$.

From our considerations of constant acceleration in special relativity and the gravitational equivalence principle, we expect that the temporal coordinate of a frame of reference in which there is a uniform gravitational field vary in its relative passage in different positions in such a frame of reference. Because the spherically symmetric matter / energy distribution will induce a spherically symmetric, i.e. radial, gravitational field (actually non-zero affine connection components) for a frame of reference defined globally with zero three-acceleration with respect to the matter n itself, we expect that the passage of time be relatively different for observers at different radial distances from the center of mass of the source mass. We thus allow the proper time interval for a purely temporal coordinate interval to be related via $d\tau = \alpha(r)dt$.

We thus expect the spacetime interval in Schwarzchild spacetime to be of the form $ds^2 = \alpha(r)c^2dt^2 - \beta(r)dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\varphi^2$, or for the non-zero components of the Schwarzchild metric to be

$$\begin{aligned} g_{00} &= \alpha(r)c^2, & g_{11} &= -\beta(r) \\ g_{22} &= -r^2, & g_{33} &= -r^2\sin^2\theta \end{aligned} \tag{18}$$

The first two represent the effect of the spherically symmetric matter / energy distribution on the chronogeometric structure of the spacetime while the second two are relics of how we coordinatize three-geometries in a spherically symmetric way. It can then be shown, by solving $R_{\mu\nu} = 0$ in terms of the above metric components, that $\alpha(r) = \frac{1}{\beta(r)} = (1 - \frac{r_s}{r})$, where $r_s := \frac{2GM}{c^2}$ is the *Schwarzchild radius* and M is the total mass of the distribution considered.

Let us see what effect this has on the relative passage through time of two world lines that maintain a constant position with respect to the center of mass of the source mass, but at different radial positions, by looking at the proper time interval for purely $ds^2 = (1 - \frac{r_s}{r})c^2dt^2 \rightarrow d\tau = \sqrt{1 - \frac{r_s}{r}}dt$. From this, it is obvious that the metric is defined is such a way that, for an observer infinitely far away from the source mass and at rest with respect to its center of mass, an interval of coordinate time (locally measured time) is equal to an interval of proper time, while intervals of coordinate time closer to the center of mass of the source mass are *shorter* intervals of proper time, just as we would expect from our considerations of acceleration in special relativity and the gravational equivalence principle. As $r \rightarrow r_s$, an infinite interval of coordinate time becomes a zero interval of proper time. Semay (2006) shows that, locally, the chronogeometric effects of being in an accelerated reference frame in Minkowski spacetime and of being static in a gravitational field are indistinguishable for $r >> r_s$.

The Schwarzchild radius is, under familiar circumstances to terrestrial observers, usually much smaller than the radius of the spherically symmetric gravitating body. For example, the Schwarzchild radius of Earth is about 9 millimeters. Because we have been considering solutions to the *vacuum* field equations, this metric does not apply in the regions where there is matter and energy. It is, however, possible to have the Schwarzchild radius of a matter / energy distribution "showing" – this happens whenever all the matter and energy is contained within the Schwarzchild radius of that distribution of matter and energy, which is physically possible. This gives rise to a black hole.

It appears that at the surface of the black hole ($r = r_s$) "time comes to a stop," with respect to the time at $r = \infty$. Are we to take this interpretation at face value, and attempt to interpret what the metric tells us for $r < r_s$ as well? It can be shown that in fact it is merely the manner in which we have coordinatized the chronogeometric structure of such a spacetime - coordinatized for a global frame of reference with zero three-acceleration w.r.t. the source matter – that gives rise to such a strange apparent result. It turns out that regions of the spacetime with $r < r_s$ in this coordinate system are analogous to Regions I, II, and III in Figure 11 – events still reside in these regions, so to speak, but there lacks a two-way causal connection between these events and those in regions with $r > r_s$, and thus, just as Regions I, II, and III are not even *in* the coordinate system / frame of reference of the Rindler observer, the regions of the spacetime with $r < r_s$ are not even in the frame of reference of any observer with $r > r_s$. Thus, the surface defined by $r = r_s$ is called the *event horizon* of a black hole, in analogy with the Rindler horizon we encountered in Zection 3. Time may "come to a stop" for an observer on the event horizon compared to an observer at $r = \infty$, but, just as the observer at the Rindler horizon needs an infinite acceleration to remain in the Rindler wedge, such an observer would need infinite four-acceleration, and thus infinite four-force, to remain at the event horizon. Since this is impossible, it cannot be that time ever meaningfully comes to a stop in this case. Rather, in this case, the observer at the event horizon will inevitably cross the event horizon and enter a region of the spacetime not even meaningfully described by the metric (18).

We can, however, consider that two observers start at some radius r_0, one of the observers remains at r_0, and the other moves quasistatically nearer to the Schwarzchild radius and then returns to r_0, then the observer who came near the event horizon will have their clock lag the one who remained at r_0 in offset. To actually make this measurement, four-forces must have been applied throughout the process, with the observer coming near the event horizon requiring greater four-forces throughout their trip, as in Section 3. We thus pose the question: might it be possible, then, to determine the effect on the relative passage of times of clocks purely in terms of the nongravitational dynamics those clocks have undergone throughout their histories? We can answer this right away by considering the twin paradox in special relativity. In this case, the difference in clock readings depends on at what speed and for how long a twin left another twin and returned to that twin. Acceleration and thus four-force is required for this twin to return, however the magnitude of the acceleration and four-force required to be applied is independent of the amount of time for which the twin travels away from the other, and hence independent of the final discrepency in the watches of the two twins. Because the spacetime of general relativity is hyperbolic as well, and asymptotes to Minkowski spacetime sufficiently far from any source matter, we can trivially conclude that in the general theory, as well, it is a consequence of the chronogeometric structure on which non-gravitational dynamics unfold and not a consequence of the non-gravitational dynamics themselves that there is a path dependence of the reading of a clock.

Figures 12 and 13 show how, for static observers w.r.t. the source matter, the same interval of coordinate time compares with the proper time $\tau = \frac{s}{c}$, i.e. coordinate time as $r \to \infty$, at different positions r. Figure 12 shows this up to $10r_s$ and Figure 13 up to $30r_s$. We can see that closer to the Schwarzchild radius, the relative passage of coordinate time changes more rapidly as a function of position, representing the increasing magnitude of the measured gravitational field for static observers. It should be remembered that Figures 12 and 13 are defined for static observers only, and these observers neccessarily have a net four-force and four-acceleration. Thus, Figures 12 and 13 cannot be used to compare the potential comparison of clocks between orbiting observers at different distances form the source matter, for these observers are four-force-free. Although the metric (18) cannot be used to characterize the entirety of the spacetime, it is possible to recoordinatize the chronogeometry to describe the entirety of this solution, and the result is very strange indeed. This is done in Rindler (1964).

We can also see that the matter and energy does force the spatial geometry to be non-Euclidian in the surrounding vacuum, as described by the modification of the radial coordinate in the spatial metric. It turns out that the spatial geometry is of the same form as that for the surface of what is called *Flamm's paraboloid*, which can be greatly helpful in visualizing the spatial curvature of Schwarzchild spacetime (UCLA Physics Demoweb , 1998).

The non-zero Christoffel symbols, and hence also the affine connection components, are:

$$\begin{aligned} \Gamma^t_{rt} &= \frac{r_s}{2r^2(1-\frac{r_s}{r})}, & \Gamma^r_{tt} &= \frac{GM}{r^2}(1-\frac{r_s}{r}) \\ \Gamma^r_{rr} &= -\frac{r_s}{2r^2(1-\frac{r_s}{r})}, & \Gamma^r_{\theta\theta} &= -r + r_s \\ \Gamma^r_{\varphi\varphi} &= \Gamma^r_{\theta\theta}\sin^2\theta, & \Gamma^\theta_{r\theta} &= \Gamma^\varphi_{r\varphi} = \frac{1}{r} \\ \Gamma^\theta_{\varphi\varphi} &= -\sin\theta\cos\theta, & \Gamma^\varphi_{\theta\varphi} &= \cot\theta \end{aligned} \tag{19}$$

(Havas , 1964). The first three arise strictly because of the distorting influence of the matter and energy of the spacetime on the chronogeometric structure of the spacetime, the last three arise strictly because of the manner in which we have coordinatized the three-geometry, and the fourth and fifth are a combination of the spatial curvature and the manner in which we have coordinatized the three-geometry. Looking at the autoparallel (and in this case, also geodesic) equation (1), we see that the Γ^r_{rr} results in freely falling particles accelerating more rapidly radially inward towards the mass whenever there is already non-zero radial velocity owing to the radial spatial curvature, with the magnitude of this effect given by $\Gamma^r_{rr}v_r^2$. We also see that, as a particle moves radially outward from the source mass, the passage through time in this coordinatization of the spacetime increases, as $\Gamma^t_{rt} > 0$. This agrees with our examination of how clocks should compare in their readings at different radial distances from the source mass. As shown below, the $\Gamma^r_{\theta\theta}$ and thereby also $\Gamma^r_{\varphi\varphi}$ is less negative than in the case without spatial curvature, and this, along with the Γ^r_{rr}, will modify the paths traced out by orbiting test particles. In particular, these connection components lead to the advance of the periapsis and apoapsis, i.e. the semi-major axis, of a bound elliptical orbit – the ability of general relativity to account for the perihelion advance of Mercury, one of the theory's early triumphs, comes from this spatial curvature. It also leads to the instability, in general, of circular orbits. In the general theory, the autoparallel / geodesic equation is now parameterized by the *proper time* of the world line being described, as is the definition of four acceleration, i.e.

$$A^\mu := V^\nu D_\nu V^\mu$$

is still the definition of the four-acceleration, as in curved Newtonian spacetime, but the definition of four-velocity is carried over from special relativity, $V^\mu := \frac{dX^\mu}{d\tau}$.

In terms of the metric, electromagnetic radiation propagates, as in special relativity, along null-geodesics, such that $ds^2 = 0$. It can be shown (Rindler , 1964) that such paths are not Euclidian straight lines in a frame of reference with zero three-acceleration w.r.t. the center of mass of the source matter, as is assumed in Newtonian theory. This gives rise to the phenomenon of *gravitational lensing.*

The non-zero connection components of a spherically symmetric, static mass distribution in the non-metric, non-flat affine reformulation of Newtonian gravity, wherein Euclidian spatial geometry and absolute time were assumed apriori, are:

$$\begin{aligned} \Gamma^r_{tt} &= \frac{GM}{r^2}, & \Gamma^r_{\theta\theta} &= -r \\ \Gamma^r_{\varphi\varphi} &= \Gamma^r_{\theta\theta}\sin^2\theta, & \Gamma^\theta_{r\theta} &= \Gamma^\varphi_{r\varphi} = \frac{1}{r} \\ \Gamma^\theta_{\varphi\varphi} &= -\sin\theta\cos\theta, & \Gamma^\varphi_{\theta\varphi} &= \cot\theta \end{aligned} \tag{20}$$

The first term is the only term resulting from the distorting influence of the matter and energy in the spacetime, the rest are relics of our coordinatazation of the three-space. We now take a low-velocity limit of the Schwarzchild solution in equations (19) by sending $c \to \infty$, as this is equivalent to considering that all motions have a three-velocity, in any given frame of reference, v such that $v << c$. Doing so gives the connection components (20), which are the non-magnetic-type connection components given in (Hall, Stachel), and sends $r_s \to 0$. Thus, in regions of spacetime surrounding static, spherically symmetric matter / energy distributions where $r >> r_s$ and all motions are such that $v << c$, the Newtonian limit is obtained mathematically. Thus, the low-velocity limit is neccessarily that of (Hall, Stachel) when the radius R of spherical mass considered is much larger than that body's Schwarzchild radius, $R >> r_s$.

As Havas (1964) shows, when taking the low-velocity limit by sending $c \to \infty$, the Schwarzchild metric *degenerates* to two independent pseudo-4-metrics (psuedo because they have a determinant of 0) – one with only a temporal component, from the g_{00} of the Schwarzchild, and one with three diagonal spatial components, from the g_{ii} of the Schwarzchild metric. This can be said to happen because the $g_{00} = c^2 - \frac{2GM}{r} \to \infty$ as $c \to \infty$ while the g_{ii} remain finite in this limit. Thus the g_{00} can only be meaningfully compared to itself, and the g_{ii} only meaningfully compared to themselves. These pseudo-4-metrics can then be used to formulate the Newtonian limit of the Schwarzchild solution in language even closer to that of the metric characterization of the full general theory of relativity.

We add the additional stipulation that the formulation of gravitation, including gravitomagnetic effects, of (Hall, Stachel) is a *near field* limit because, in dragging the spacetime, the moving matter / energy content cannot instantaneously drag the spacetime. It must be that, everywhere in the spacetime, the frame-dragging effects of gravitomagnetism are observed to be propogating away from the moving source matter at the speed of light. Thus, we must additionally constrain the region considered to be such that these effects are propagated effectively instantaneously, thus limiting the region considered to being sufficiently near to the moving source matter. For the case of a rotating sphere of uniform density detailed in (Hall, Stachel), the radial extent of the near-field region may be defined as being such that the distance r from the rotating sphere, divided by the speed of light c, is less than some fraction of the period of rotation, and of this order of magnitude. I.e., the near-field region is on the order of $r < cT$, where T is the period of rotation.

5.2 The Field Equations in the Presence of Matter and Energy

We now present the field equations more generally. To do so, we need to relate the chronogeometric structure of the spacetime to the matter / energy content of the spacetime. To do this, we need to relate the Ricci tensor to a tensor representing the matter and energy content of the spacetime. Let us first explain what this tensor, which is called the stress-energy tensor (SET), is. We have seen from (Hall,Stachel) that we should expect that both the matter / energy content as well as the momentum content of the spacetime modify the chronogeometric structure of the spacetime. Thus, a a natural choice for a simple matter distribution of density ρ_0 in that distribution's rest frame is

$$T^{\mu\nu} = \rho_0 U^\mu U^\nu$$

The units of this expression are $\frac{mass}{length^3}\frac{length}{time}\frac{length}{time} = \frac{mass\frac{length}{time^2}}{length^2}$. Written in this way, it is clear each component has units of pressure. We can transform the units also to $\frac{mass\frac{length}{time}}{length^2 time}$ to see that this is equivalent to momentum flux (change in momentum per unit time per unit area). Perhaps the simplest form is $\frac{mass(\frac{length}{time})^2}{length^3}$, showing each component of the SET has units of energy density, and that energy density, pressure, and momentum flux are all equivalent in physical units.

The 00 component is $T^{00} = \gamma^2 \rho c^2$ and represents the energy density of the mass – one factor of γ gives the relativistic increase in energy as a function of relative velocity, and the other the increase in effective density from length contraction. The 0i component is $T^{0i} = \gamma^2 \rho v^i c$, which is the momentum-density multiplied by the speed of light to give commensurate units (interpreted as the flux of three-momentum-density through a surface of constant x^0, i.e. temporal direction with spatial units).

The (μ, ν) component of the SET more generally can be interpreted as the flux of the μth component of the four-momentum density of an infinitesimal mass/energy element across a surface of constant ν coordinate value. Thus the 00 component represents the "flux" of the mass/energy component of four-momentum density through time (which is energy density), the i0 component is the "flux" of the ith component of the three-momentum density through time, the 0i component the flux of mass/energy density component of four-momentum density through a surface of constant coordinate x^i – this is the same as the 0i component–, the ii component represents the flux of momentum density in the i-th direction across a surface of constant coordinate x^i – since the flux of momentum is the same thing as a force, we see that this represents a pressure, with similar interpretations for the ij, i.e. shearing, components of the SET (Tolish , 2010). The components of the Stress-Energy Tensor with purely spatial indices, then, represent the stress tensor σ^{ij} of conventional Newtonian physics. The pressure terms,

$$T^{ii} = \gamma^2 \rho_0 v^i v^i$$

are positive or greater than zero for a single mass regardless of the sign of v^i.

In formulating the field equations of how this stress-energy-momentum content is related to the spacetime structure, it is desirable that both the conservation of momentum and energy, two long-held principles of physics, which is to say the conservation of *four-momentum*, holds locally. Since $T^{\mu\nu}$ represents the flux of all components of four-momentum through all directions of the manifold, we should require that the the spacetime structure is related to the stress-energy-momentum content contained therein such that the four-diverenge of four-momentum flux vanish – i.e. if, in one direction, it is that there is more three-momentum flowing in a certain direction, it must be that that the flow of four-momentum is less through some other surface in some other direction. The formal mathematical statement of this requirement is that

$$D_\nu T^{\mu\nu} = T^{\mu\nu}_{;\nu} = 0 \tag{21}$$

In flat spacetime, the $\mu = 0$ component of this says

$$\partial_0 T^{00} + \partial_i T^{0i} = 0$$

i.e. if the temporal flux of matter, i.e. mass-energy density at a point of a spacetime increases in the temporal direction relative to a spacetime point, it can only do so by there being a negative momentum gradient in the spatial directions at that spacetime point, and a decrease in the amount of mass-energy density at a spacetime point in the temporal direction is neccessarily accompanied by a positive momentum gradient in the spatial directions at that spacetime point.

This is the ordinary Newtonian equation of continuity, and expresses conservation of mass. It can easily be shown to be equivalent to the ordinary Newtonian expression:

$$\partial_0 T^{00} + \partial_i T^{0i} = 0 \rightarrow \frac{\partial}{\partial(ct)}\gamma^2\rho_0 c^2 + \vec{\nabla}\cdot\gamma^2\rho_0\vec{v}c = 0 \rightarrow \frac{\partial}{\partial t}\rho_0 + \vec{\nabla}\cdot\rho_0\vec{v} = 0$$

(if $\frac{\partial\gamma^2}{\partial t} \approx 0$). In flat spacetime, the $\mu = j$ component of equation (21) says

$$\partial_0 T^{j0} + \partial_i T^{ji} = 0$$

Remembering that the ji components represent stresses and pressures, this says that if the momentum-density in the j-th direction becomes greater in the temporal (0th) direction relative to the spacetime point considered, it can only occur by flux of three-momentum in the jth direction through all the spatial surfaces, this momentum flux being the same thing as a net force per unit area. Thus the $\mu = j$ component of the conservation law in flat spacetime is the same thing as Newton's Second Law. Thus, in flat spacetime, the conservation law $T^{\mu\nu}_{,\nu} = 0$ expresses conservation of mass and Newton's Second Law. Newton's First Law is included implicitly by the spacetime being flat. Newton's Third Law stands as an independent postulate, and has the consequence that the center of mass of a closed system follows autoparallels of the spacetime – whether flat or curved. See Figures 14 and 15. However, in the case of gravitation, a component of Newton's Third Law is reconceptualized: the reciprical nature of gravitational interactions then becomes simply that everything modifies the spacetime structure in exactly the same fashion, rather than two gravitating particles mutually pulling each other from autoparallel transport in a flat spacetime. This holds for affine spacetime theories of gravitation, with the chronogeometry being tied to the matter/energy content by $R_{(0)(0)} = \frac{4\pi G}{c^2}\rho$ or this plus $R_{(0)(n)} = \frac{4\pi G}{c^2}\rho\frac{v^{(n)}}{c}$, and the resulting magnitude of the affine connection for a frame of reference globally defined to have zero three-acceleration w.r.t. the center of mass of the source matter being linear in both ρ and $\rho\frac{v^{(n)}}{c}$ (Hall, Stachel).

Collectively, what equation (21) says is that the energy content of the spacetime, the "stuff" of the universe, is not spontaneously created at a point of the spacetime. Thus, in formulating the field equations of general relativity, if we wish the spacetime and the energy content contained therein to be related in such a way that energy and momentum are still locally conserved, we should require that whatever relations are imposed between the chronogeometric structure, characterizing the local relations of spacetime points, and the SET satisfy equation (21).

Since we saw that the vacuum field equations were $R_{\mu\nu} = 0$, we might try $R^{\mu\nu} = \kappa T^{\mu\nu}$, where $R^{\mu\nu} := g^{\mu\rho}g^{\nu\sigma}R_{\rho\sigma}$, and solve for κ by requiring that the solutions approximate to the Newtonian limit in the appropiate circumstances.

If we are to require that $T^{\mu\nu}_{;\nu} = 0$, it would then be that $R^{\mu\nu}_{;\nu} = 0$. However, it is a known identity from Riemannian geometry that $(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R)_{;\nu} = 0$, where $R := g_{\mu\nu}R^{\mu\nu}$, which would imply, if $R^{\mu\nu}_{;\nu} = 0$, that $R_{;\nu} = R_{,\nu} = 0$. Defining $T := g_{\mu\nu}T^{\mu\nu}$, it would then follow that $T_{,\nu}$ equals zero, an unreasonable requirement to put on the stress-energy-momentum content of the universe. Since $(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R)_{;\nu} = 0$ is identically satisfied for any Riemannian geometry, a good candidate that automatically satisifies $T^{\mu\nu}_{;\nu} = 0$ is

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = \kappa T^{\mu\nu}$$

It can be shown by linear approximation procedures in solving for the metric $g_{\mu\nu}$ and resulting connection components, that if we require that solutions to these field equations match with the vacuum field equations at the vacuum / non-vacuum interface and agree with the Newtonian limit, we must take $\kappa = -\frac{8\pi G}{c^4}$ (Tolish , 2010). Whereas in Tolish (2010), natural units with c = 1 and a metric convention of {-,+,+,+} was used, we here use arbitrary units with metric convention of {+,-,-,-}, hence the sign difference and the factor of $\frac{1}{c^4}$ in the field equations with these conventions. The geometric factor 8π is twice that of (Hall, Stachel) reviewed in Section 1 owing to the factor of $\frac{1}{2}$ in definition of the Christoffel symbols. The spacetime unit factor is $\frac{1}{c^4}$ rather than $\frac{1}{c^2}$ as in (Hall, Stachel) because of the way in which we have defined the SET with the four-velocity twice– the 00 component of the SET is ρc^2, and thus this additional factor of c^2 must be cancelled in order to give agreement with (Hall, Stachel).

The field equations then become

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = -\frac{8\pi G}{c^4}T^{\mu\nu} \tag{22}$$

Einstein searched for field equations of this form early in his quest to reformulate gravity chronogeometrically as informed by the gravitational equivalence principle, but initially abandoned this search, as he thought they lead to physically distinct solutions for different choices of coordinates. As detailed in Stachel (2014), he only returned to his search for the above, generally covariant field equations after reconsidering a foundational assumption in the argument that originally dissuaded him from considering them: that the coordinates themselves have any physical, rather than merely mathematical, significance. Indeed, it has been argued convincingly by Stachel (2014) that, in light of the hole argument, for the field equations (22) to be valid, it must be that the spacetime is a *geometry* rather than an *algebra* (geometry and algebra being differentiated here in terms of the individuation or lack of individuation of the spacetime points, the spacetime is still most properly a *chronogeometry*). The spacetime is a geometry in the sense that the spacetime points have no inherent individuality, are of the same nature, and it is the relations between these points that the field equations describe. Different coordinatizations of the spacetime are then redescriptions of the same relations between spacetime points, i.e. the same solutions to the Einstein field equations, despite having different coordinates for the same points.

It is a well known feature of Eintein's equations that they are non-linear: the solution for the sum of two matter-energy-momentum distributions is not the sum of the solutions for those matter-energy distributions considered independently. Thus, Newton's Third Law for gravitation fails in general relativity. If we consider two masses, one having double the amount of matter-energy content as the other, it is no longer true that this more massive mass gives the other twice the three-acceleration – via modifying the spacetime structure – as the other gives it.

5.3 The Conservation Law In Curved Spacetime - An Example

Let us work through an example of applying the conservation of stress-energy as demanded by equation (21) that recovers a phenomenon described by conventional Newtonian dynamics and gravitation. Consider a static sphere of uniform density ρ and radius R. How must the non-gravitational force acting on the top of an object resting on the sphere compare to the non-gravational force acting on the bottom of that object? We expect from conventional Newtonian theory that the force directed radially away from the sphere at the top of the object differ from that at the bottom by the amount $\Delta\vec{F} = m\vec{g}$, where m is the total mass of the object considered and $\vec{g}$ is the local gravitational field, given by $\vec{g} = -\frac{GM}{r^2}\hat{r}$, where M is the total mass of the sphere considered.

We can recover this from considering the conservation law of general relativity, which states

$$T^{\mu\nu}_{;\nu} = T^{\mu\nu}_{,\nu} + \Gamma^{\mu}_{\lambda\nu}T^{\lambda\nu} + \Gamma^{\nu}_{\lambda\nu}T^{\mu\lambda} = 0$$

In the Newtonian limit, we have seen that we expect the only non-vanishing $\Gamma^{\mu}_{\nu\lambda}$ possible for any frame of reference be the Γ^{r}_{00}, causing us to look at the $\mu = r$ terms to give gravitationally non-trivial (i.e. not 0=0 or simple four-momentum conservation in flat spacetime) results:

$$T^{r\nu}_{;\nu} = T^{ri}_{,i} + T^{r0}_{,0} + \Gamma^{r}_{00}T^{00} = 0 \rightarrow$$

$$T^{ri}_{,i} + \frac{d}{d(ct)}(\gamma^2 v^r c) = -\rho\Gamma^{r}_{tt}$$

If the object considered is at rest, then the second term vanishes (the momentum in the object's rest-frame remains zero), and it must be that

$$T^{ri}_{,i} = -\rho\Gamma^{r}_{tt}$$

Remembering that the T^{rr} is the pressure in the radial direction and the T^{ij} shearing stresses, if we assume there is no shearing in the neighborhood of the object, we have

$$\frac{\partial}{\partial r}T^{rr} = -\rho\Gamma^{r}_{tt}$$

From the definition of the stress tensor, the quantity $T^{rr} = \frac{F^r}{\partial A^r}$, where ∂A^r is an infinitesimal area normal to the radial direction and F^r is the radial, non-gravitational force in a distribution. Assuming the length scales of the object considered are small enough that neither the T^{rr} nor the $\Gamma^{\mu}_{\nu\lambda}$ change appreciably, this gives

$$\Delta\frac{F^r}{\partial A^r} = \frac{\partial}{\partial r}T^{rr}\Delta r = -\rho\Gamma^{r}_{tt} \rightarrow$$

$$\Delta F^r = -\rho\Gamma^{r}_{tt}\Delta r\Delta A$$

$\rho\Delta r\Delta A$ is the total mass m considered in the small volume $\Delta r\Delta A$. In the Newtonian limit (with the additional stipulation $r >> r_s$) this becomes

$$\Delta F^r = -\frac{GM}{r^2}m$$

which is exactly the amount predicted by Newtonian mechanics.

Interpreted in the notation of the conservation law,

$$T^{rr}_{,r} + \Gamma^r_{00} T^{00} = 0$$

we say that the change of the momentum flux vertically is such that the infinitesimal piece of matter considered gives these pieces a continual four-acceleration corresponding to the Γ^r_{00}

This (non-gravitational) force gradient gives the object a four-acceleration of

$$\frac{F^\kappa}{m} = A^\kappa = V^\nu D_\nu V^\kappa = \Gamma^r_{00} V^0 V^0 \delta^\kappa_r = \Gamma^r_{00} c^2 \delta^\kappa_r = \frac{GM}{r^2} \delta^\kappa_r$$

, which is also the net non-gravitational force per unit mass acting on such an object. Indeed, in order to have such a four-acceleration, such a non-gravitational force gradient must exist throughout the object. We see that, if the object is in a frame of reference wherein the Γ^r_{00} vanish – a freely falling object – there is no such force gradient across the object. We can also see that the spatial curvature of the Schwarzchild solution predicts a small decrease in the magnitude of this force gradient. Under the same assumptions (object at rest, no shearing) accounting for spatial curvature:

$$\Delta \frac{F^r}{\partial A^r} = \frac{\partial}{\partial r}(T^{rr}) \Delta r = -\Gamma^r_{00} T^{00} \Delta r - 2\Gamma^r_{rr} T^{rr} \rightarrow$$

$$\delta \Delta F^r = -2\Gamma^r_{rr} T^{rr} \Delta r \Delta A = 2 \frac{r_s}{2r^2(1 - \frac{r_s}{r})} T^{rr} \Delta V \rightarrow$$

$$\delta \Delta F^r = T^{rr} \Delta V \frac{r_s}{r^2} \frac{1}{1 - \frac{r_s}{r}} \leftrightarrow \frac{\delta \Delta F^r}{\Delta A} = T^{rr} \Delta r \frac{r_s}{r^2} \frac{1}{1 - \frac{r_s}{r}}$$

This decrease in pressure differential is proportional to the spatial curvature, the height of the object, and the pressure in the radial direction at the location of the object, and results conceptually from the fact that the relative stretching of the relations of points in the radial direction (w.r.t. Euclidian space) means that relative changes in momentum flux in the radial direction across surfaces of constant radial coordinate are less drastic than they would be without this spatial curvature.

In the same way, looking carefully at the expression

$$T^{r\nu}_{;\nu} = T^{r\nu}_{,\nu} + \Gamma^r_{\lambda\nu} T^{\lambda\nu} + \Gamma^\nu_{\lambda\nu} T^{r\lambda} = 0 \rightarrow T^{rj}_{,j} + T^{r0}_{,0} + \Gamma^r_{00} T^{00} + 2\Gamma^r_{rr} T^{rr} + \Gamma^0_{r0} T^{rr} = 0$$

we see that there is a differential pressure change of

$$\frac{\delta \Delta F^r}{\Delta A} = -T^{rr} \Gamma^0_{r0}$$

compared to if $\Gamma^0_{r0} = 0$. Referencing equation (19) Γ^0_{r0} is positive in a Schwarzchild geometry, and so this decreases the pressure gradient observed in the radial direction for the case of a static matter/energy distribution, and thus increases the *magnitude* of the pressure gradient. This can be understood conceptually by the fact that the pressure in the radial direction is the same thing as the momentum flux in the radial direction, and momentum flux involves a derivative w.r.t. time, $\frac{\partial}{\partial t}$, and the relative passage of the coordinate time is faster further along the radial direction of the Schwarzchild geometry, thus resulting in an effective increase in the magnitude of the pressure further along the radial direction.

Over the long term, matter and energy in many astrophysical systems distributes itself according to the principle of hydrostatic equilibrium in the following way: In the $\mu = i$ terms of the conservation law,

$$T^{i\nu}_{;\nu} = T^{i0}_{,0} + T^{ij}_{,j} + \Gamma^{i}_{\lambda\nu}T^{\lambda\nu} + \Gamma^{\nu}_{\lambda\nu}T^{i\lambda} = 0$$

the distribution of four-momentum content will be such that the $T^{ij}_{,j}$, i.e. variations in pressures and stresses in infinitesimal neighborhoods of spacetime points, and the spacetime structure terms $\Gamma^{i}_{\lambda\nu}T^{\lambda\nu} + \Gamma^{\nu}_{\lambda\nu}T^{i\lambda}$ are related in such a way that the $T^{i0}_{,0}$ vanish. This is because vanishing $T^{i0}_{,0}$ means that the three-momenta of the constituents of the distribution are not changing and hence remain zero in their rest frames, and thus the distribution is stable when the four-momentum distribution and the space-time structure set up by that four-momentum distribution are related such that $T^{i0}_{,0} = 0$ is true everywhere in the distribution, and once this distribution is realized, it will remain so. If the spacetime structure and the pressure/stress gradients in any region of the spacetime are set up such that $T^{ij}_{,j} + \Gamma^{i}_{\lambda\nu}T^{\lambda\nu} + \Gamma^{\nu}_{\lambda\nu}T^{i\lambda} \neq 0$, $T^{i0}_{,0}$ will be non-zero by exactly the amount required by the conservation law, the four-momentum distribution and spacetime structure will evolve together.

In summary, the conservation law given by equation (21) tells us, given the chronogeometric and affine structure of the spacetime, what the local differential changes in stresses, pressures, forces, and momenta must be throughout the source matter, while the field equations (22) tell us what the chronogeometric and affine structure must be given such a distribution of stresses, pressures, momenta, and mass/energy. The reader may ask: which do we solve first? The answer is we can only require that both are satisfied throughout the spacetime, and solve self-consistently. Appropriate solutions to the field equations are the ones in which $T^{\mu\nu}_{;\nu} = 0$ is true at every point of the spacetime given the distribution of four-momentum density and the interrelations of spacetime points.

6 Background Independence, Quantum Gravity

We have seen that, in determining solutions to the field equations as presented in Section 4.2, nothing of the nature of space and time can be specified apriori. Instead, the very notion of what "when and where" means is determined by "when and where" processes are! All we can do is solve *self-consistently* as demanded by the field equations for the relations between the defining notions of the spatiotemporal structure of the universe and the particular spatiotemporal situations of matter and energy. This is the sense in which general relativity is *background independent* – it assumes no fixed, background spacetime structure that serves as a stage on which all physics unfolds, rather, the stage itself, i.e. the spacetime, becomes a dynamical entity, evolving with the matter and energy content contained therein. This is a major conceptual break from the manner in which we have reasoned about the physical world around us previously, and leads to deep conceptual issues for how to unify quantum-mechanical and general relativistic descriptions of nature. In all formulations of quantum mechanics so far, even the special relativistic and immensely successful quantum field theory, a fixed background spacetime structure is presupposed.

We note that all evidence for the quantum-mechanical nature of the universe comes from *non-gravitational phenomena*. Our best theory of gravity – general relativity – is informed by the evidence of Lorentz invariance (the Michelson-Morely experiment) and the gravitational equivalence principle (equality of inertial and gravitational masses), and the appropriate interpretation of the theory founded thereupon and subsequently confirmed by further experimentation, is that the spacetime itself is a dynamical entity, about which nothing can be specified apriori. The task of quantum gravity, then, should be to formulate a description of space and time wherein quantum mechanical entities themselves modify the structure of the spacetime, in such a way that "when and where" a quantum mechanical entity exists and what "when and where" means quantum mechanically are solved self-consistently, and in which the predictions of both quantum field theory and general relativity are reproduced in appropriate regimes. This has been discussed in greater detail by several authors, including Stachel (2006b).

References

É Cartan (1923) *Sur les variétés à connexion affine et la théorie de la relativité généralisée* Annales scientifiques de l'É.N.S. 3^e série, tome 40 (1923), p. 325-412

A. Papapetrou and J. Stachel (1978) *"A New Lagrangian for the Vacuum Einstein Equations and Its Tetrad Form" General Relativity and Gravitation* 9, 1075 (1978).

J. Stachel (2006a) *Einstein's Intuition and the Post-Newtonian Approximation* World Scientific (2006), pp. 12

W. Rindler (1964) *Essential Relativity* New York: Springer-Verlag (1977).

H. Bondi (1964)*Relativity and Common Sense: A New Approach to Einstein* New York: Dover Publications (1964)

UCLA Physics Demoweb - Curved Spacetime. July 15, 1998.

P. Havas (1964) *Four-Dimensional Formulations of Newtonian Mechanics and Their Relation to the Special and the General Theory of Relativity* Reviews of Modern Physics - October 1964.

J. Stachel (2014) *The Hole Argument and Some Physical and Philosophical Implications* Living Rev. Relativity, 7, (2014), 1

C. Semay (2006) *Observer with a constant proper acceleration* Physics Education

J. Stachel (2006b) *Prolegomena to any future Quantum Gravity* in D. Oriti (Ed.) *Approaches to quantum gravity*

A. Tolish (2010) *General Relativity and the Newtonian Limit* University of Chicago VIGRE REU 2010

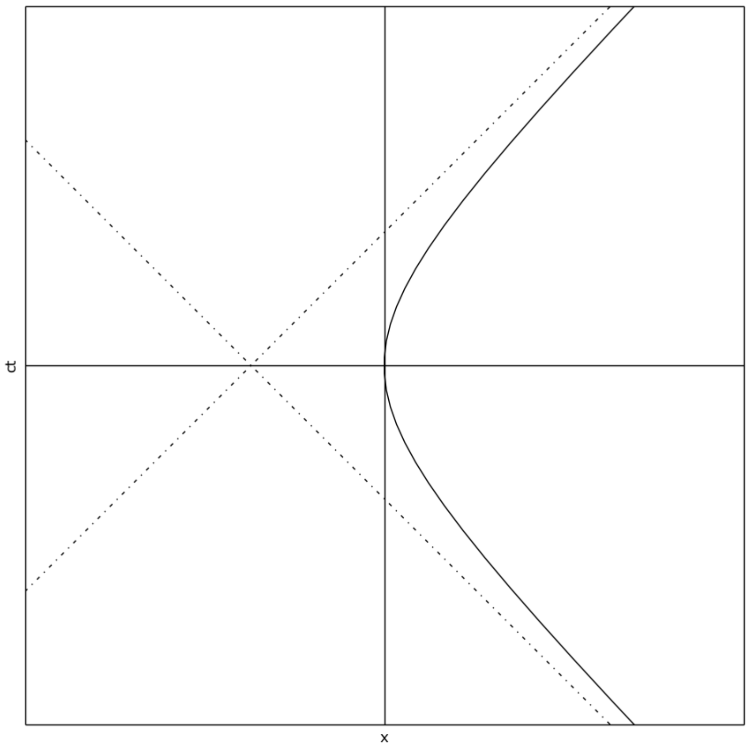

Figure 1: An observer accelerating with respect to a momentarily coincident and comoving inertial observer. The two dotted lines are emanating from (t,x) = $(0, \frac{-c^2}{g})$ – this event the accelerating observer says is infinitely far away according to radar measurement.

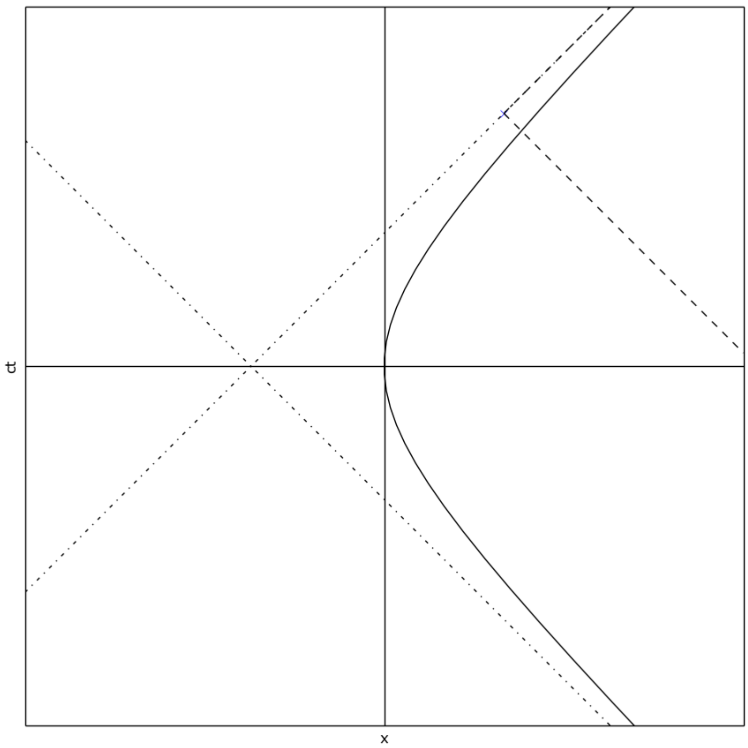

Figure 2: Events along line $ct = x + \frac{c^2}{g}$ with $x > -\frac{c^2}{g}$ take infinite time to interact with accelerating observer, but could easily have been affected by the accelerating observer.

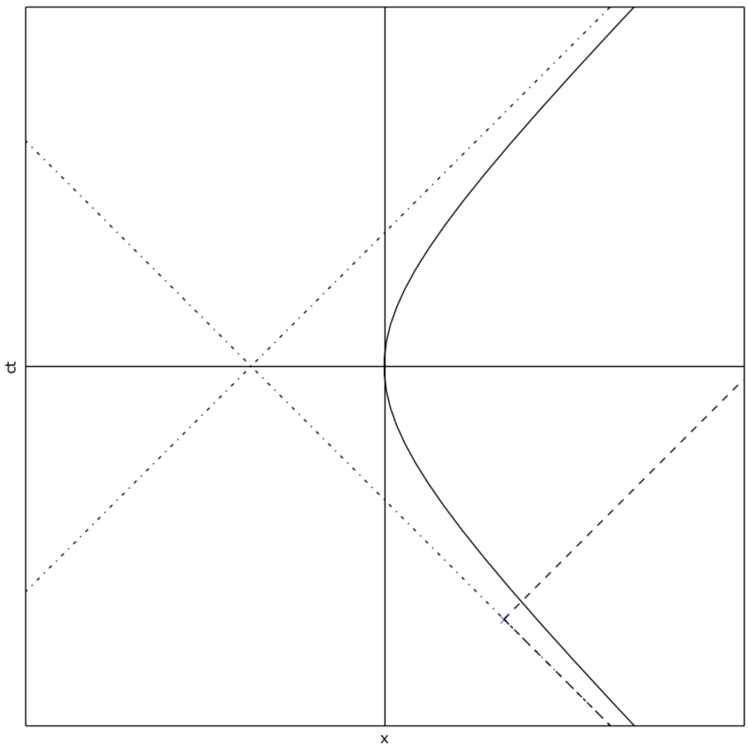

Figure 3: Events along line $ct = -x - \frac{c^2}{g}$ with $x > -\frac{c^2}{g}$ can easily affect the accelerating observer, but the accelerating observer would have had to have sent out a causal signal at $\tau = -\infty$ to affect the events.

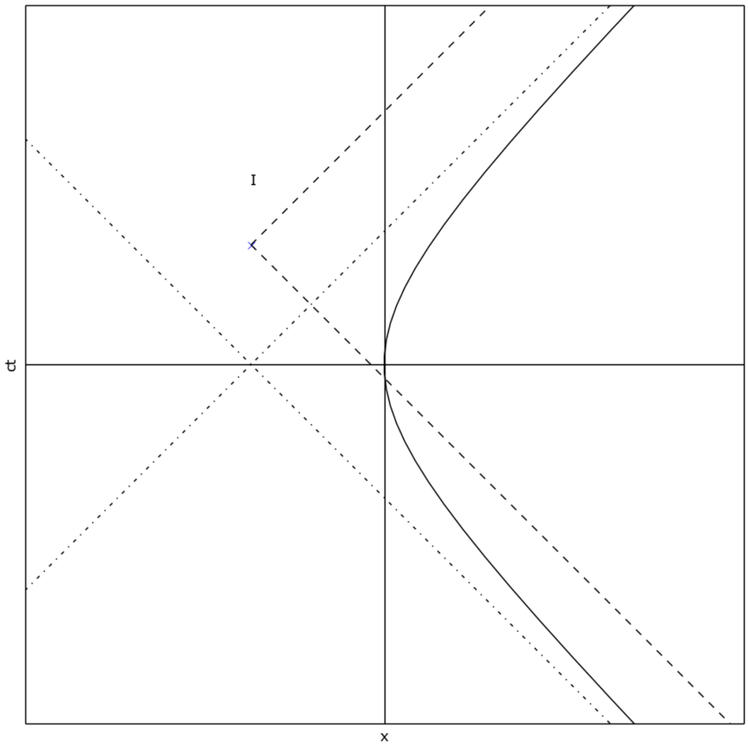

Figure 4: Events in the spacetime region defined by $ct > x + \frac{c^2}{g}$ and $ct > -x - \frac{c^2}{g}$ (Region I) can never affect the accelerating observer, but can easily be affected by the accelerating observer

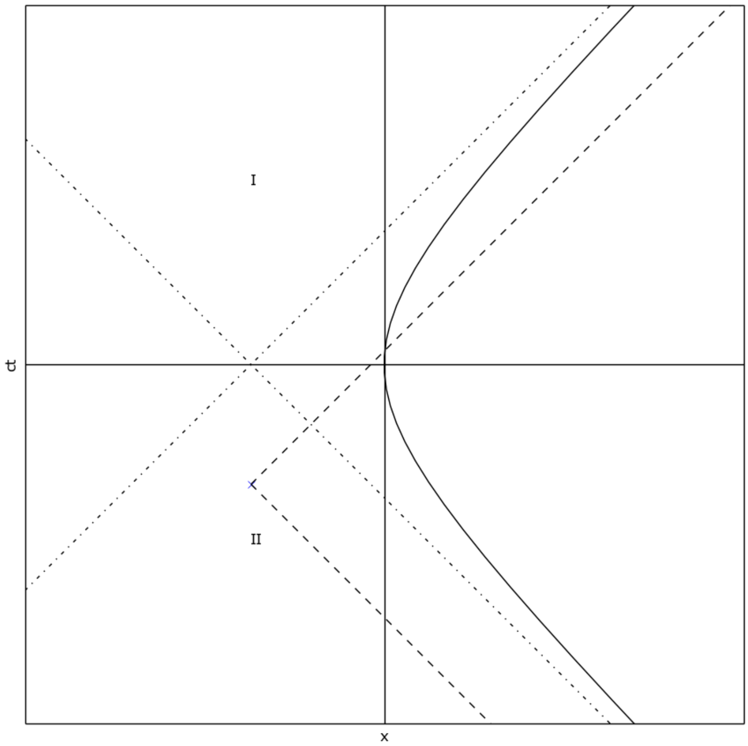

Figure 5: Events in the spacetime region defined by $ct < x + \frac{c^2}{g}$ and $ct < -x - \frac{c^2}{g}$ (Region II) can easily affect the accelerating observer, but can never be affected by the accelerating observer

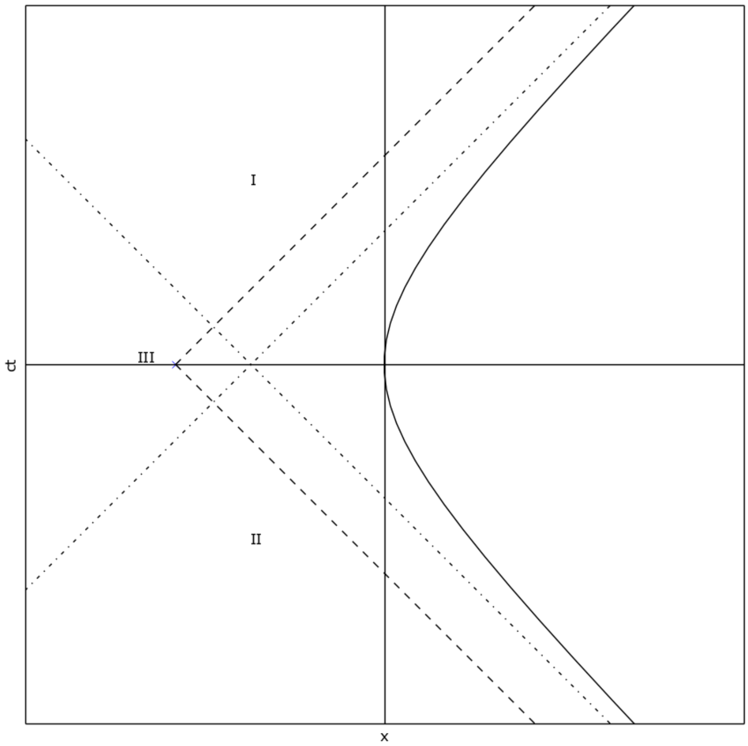

Figure 6: Events in the spacetime region defined by $ct > x + \frac{c^2}{g}$ and $ct < -x - \frac{c^2}{g}$ (Region III) cannot affect or be affected by the accelerating observer.

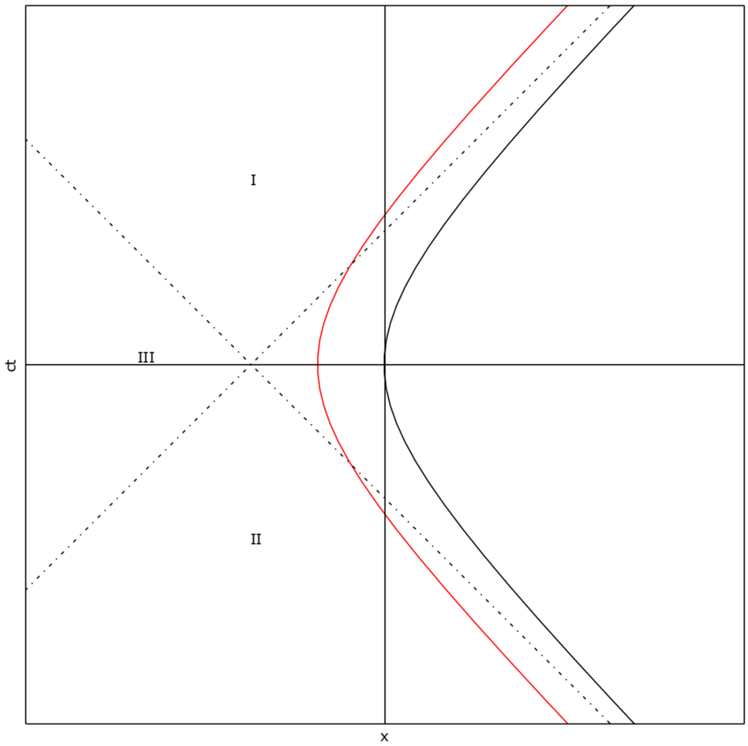

Figure 7: An observer accelerating at the same rate g as the observer initially at $x = 0$ with $v = 0$ but initially at $-\frac{c^2}{g} < x < 0$ with $v = 0$ leaves the reference frame of the observer originally considered.

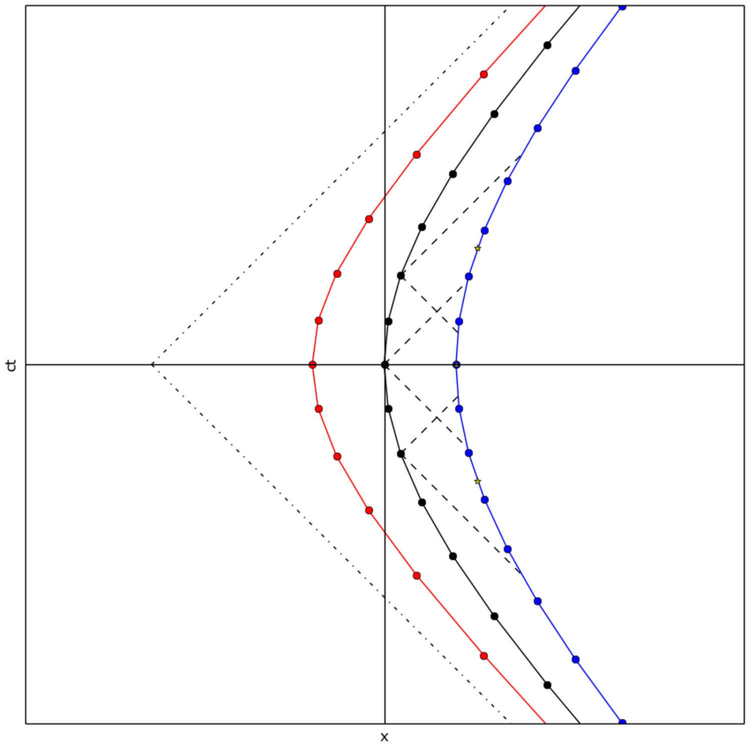

Figure 8: If an observer further along the direction of acceleration determines the times of the events of constant proper time interval of the fiducial Rindler observer by radar measurement, the proper time intervals of the times determined by the measuring observer (yellow stars) are longer than the proper time intervals of the events on the fiducial Rindler observer's world line (black dots from which the partial null cones emanate).

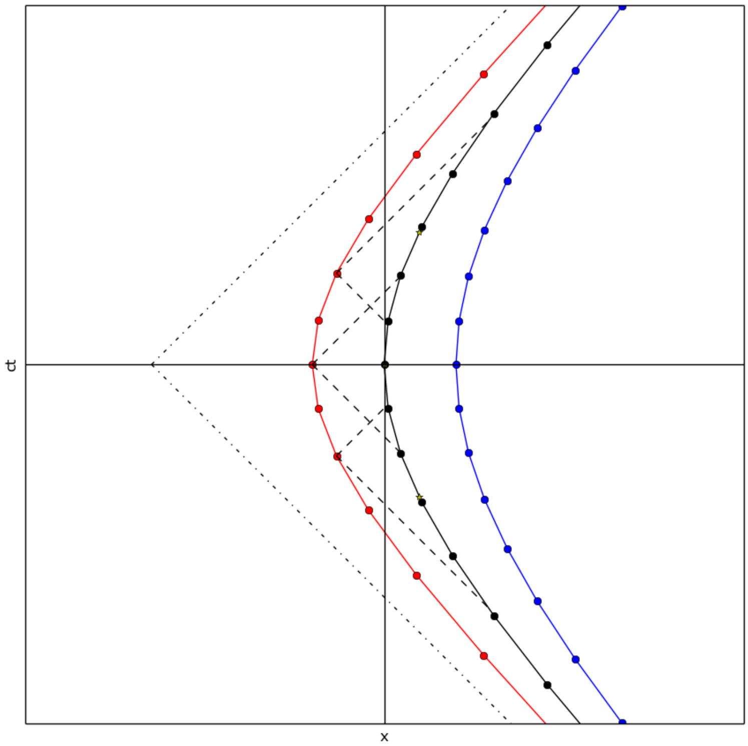

Figure 9: Just as in Figure 8, if the fiducial Rindler observer determines the times of the events of constant proper time interval of an observer closer to the Rindler horizon, the proper time intervals of the times determined by the measuring (in this case fiducial) Rindler observer (yellow stars) are longer than the proper time intervals of the events on the world line of the observer being measured (red dots from which the partial null cones emanate). In this case, the effect is greater than in the case shown in Figure 8.

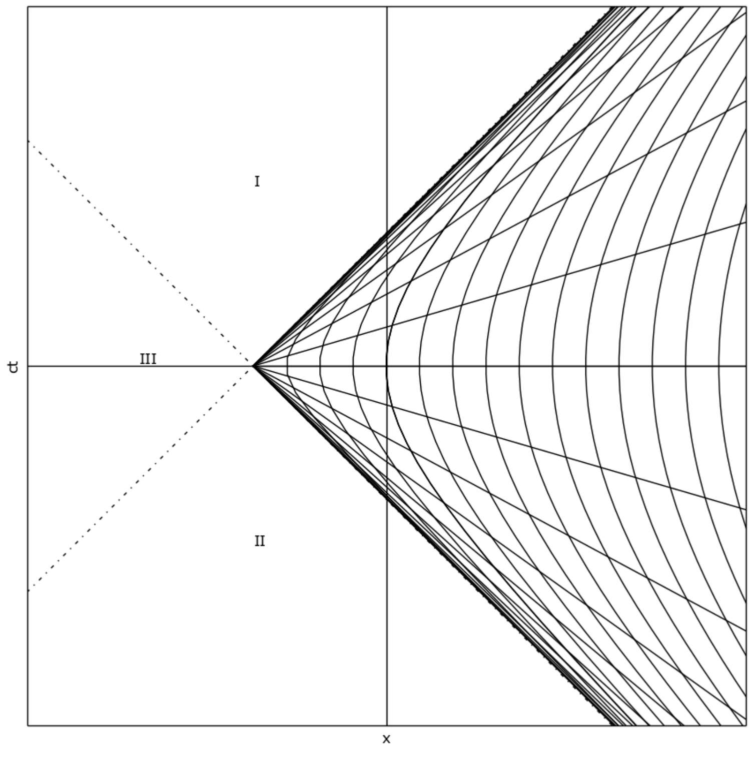

Figure 10: The mapping of lines of constant T, and of constant X as given by equations (14) and (15), to the inertial observer's coordinate system (t,x). A time interval ΔT is a greater interval of time Δt further away from the Rindler horizon.

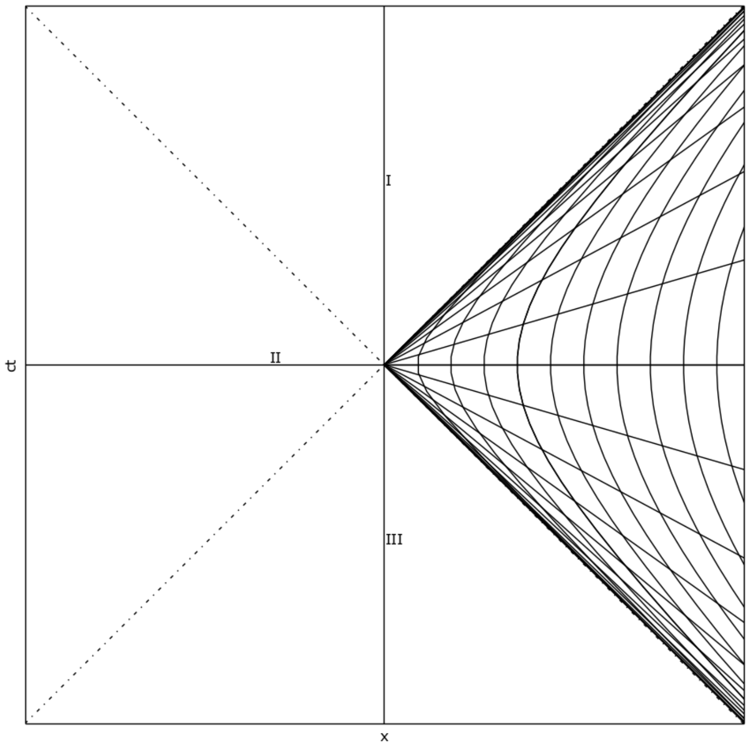

Figure 11: Rindler coordinates mapped into a local inertial frame with spacetime coordinate system (t,x). The Rindler horizon is at X = x = 0, and the proper acceleration of the Rindler observers is given by $g = \frac{c^2}{X} = \frac{c^2}{x}$.

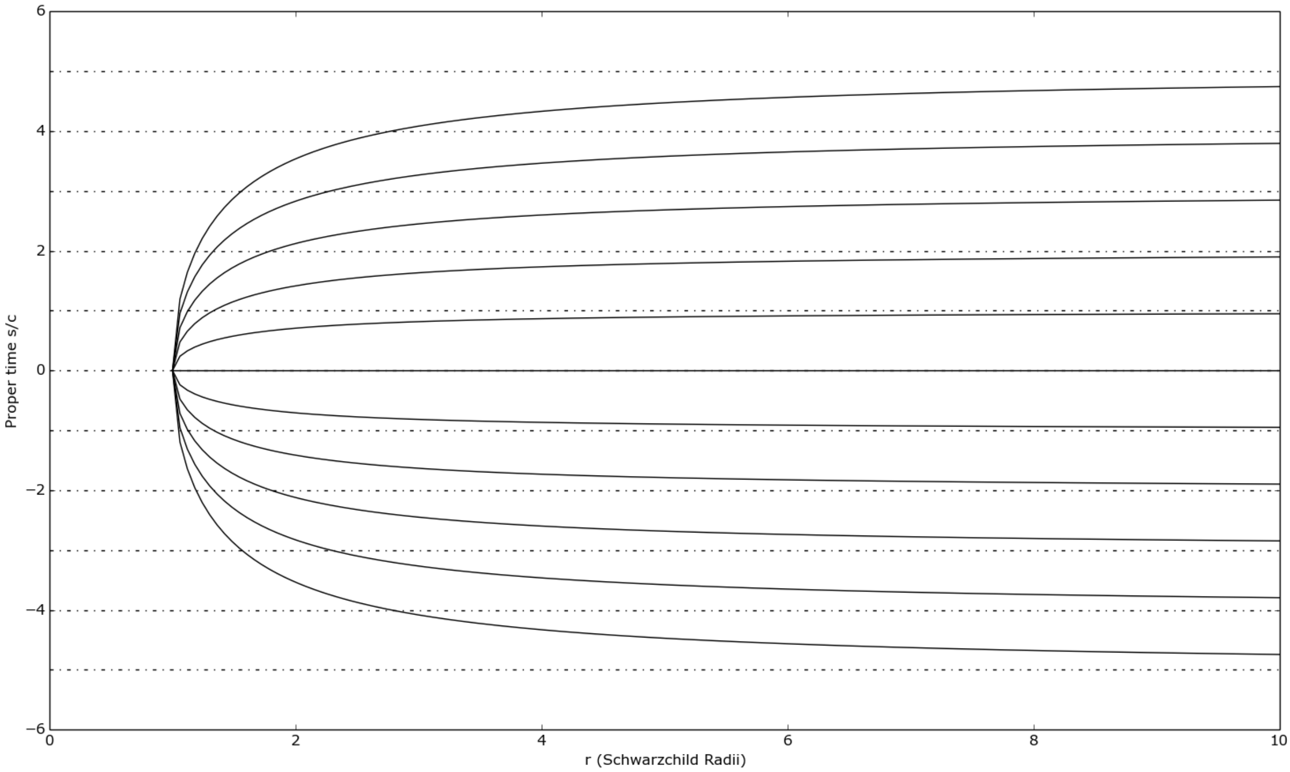

Figure 12: Lines of constant coordinate time t for static observers mapped into proper time $\frac{s}{c}$. The dotted lines are lines of constant proper time. Closer to the Schwarzchild radius, the relative passage of coordinate time changes more rapidly as a function of position, representing the increasing magnitude of the measured gravitational field for static observers as given by equations (19).

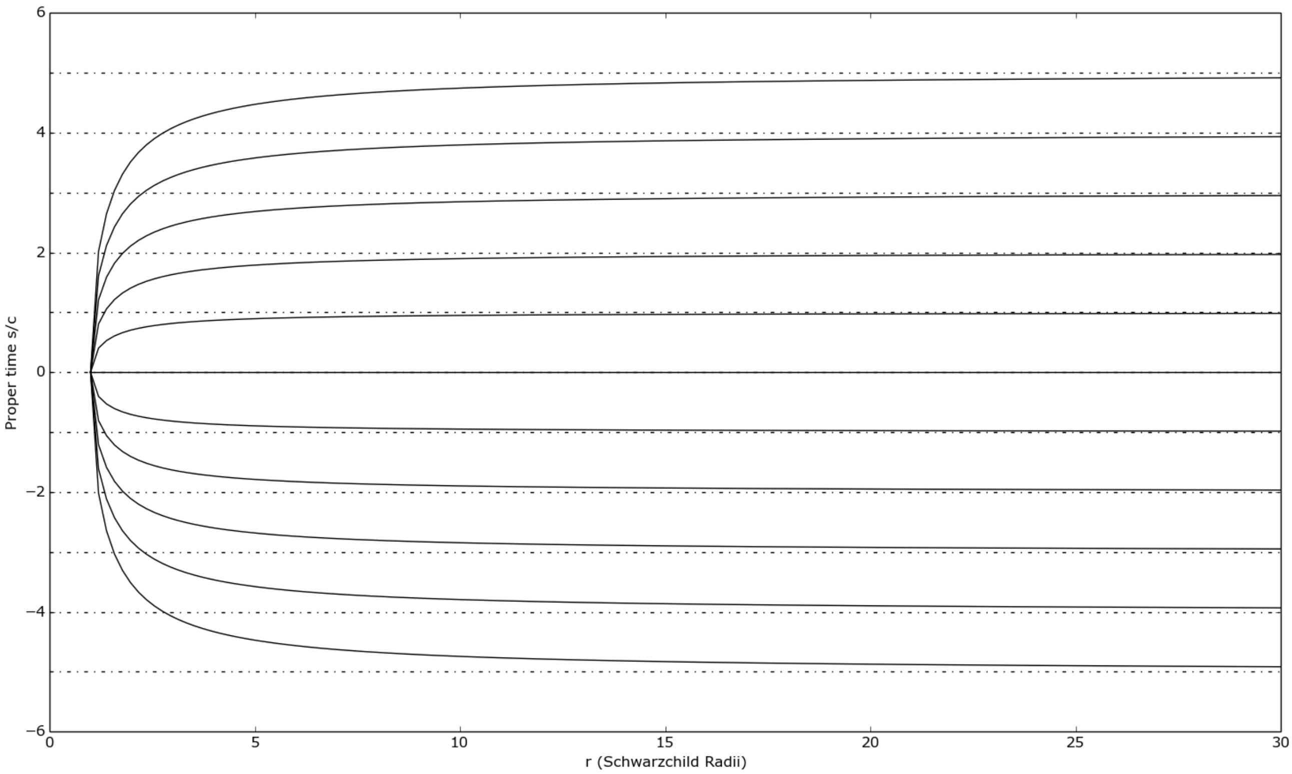

Figure 13: The same as shown in Figure 12, but shown out to $r = 30r_s$.

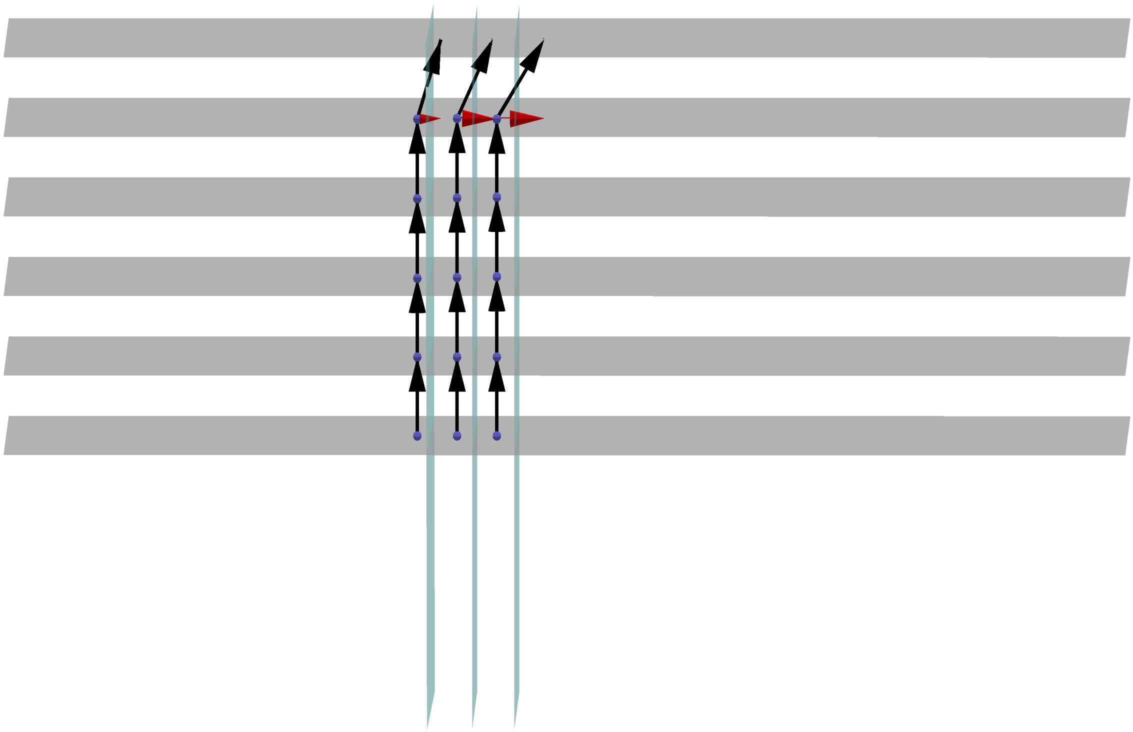

Figure 14: The above shows consistency with the conservation law $T^{\mu\nu}_{,\nu}$ in flat spacetime – there is a change in momentum density, and this change comes from a flux of momentum through spatial hypersurfaces, i.e. $T^{j\nu}_{,\nu} = 0$, equivalent to Newton's Second Law, is satisfied, and there is a positive momentum gradient at a region of the spacetime, and a corresponding decrease in the mass-energy density in this region, i.e. $T^{0\nu}_{,\nu} = 0$, equivalent to mass conservation, is satisfied. However, Newton's Third Law does not hold if the spacetime region shown constitutes a closed system. Newton's Third Law postulates that, for the above to be seen in region of flat spacetime, the increase in three-momentum shown must result in a decrease in three-momentum in another region of the spacetime, i.e. the accelerating particles must "push off" something or be "pulling towards" something.

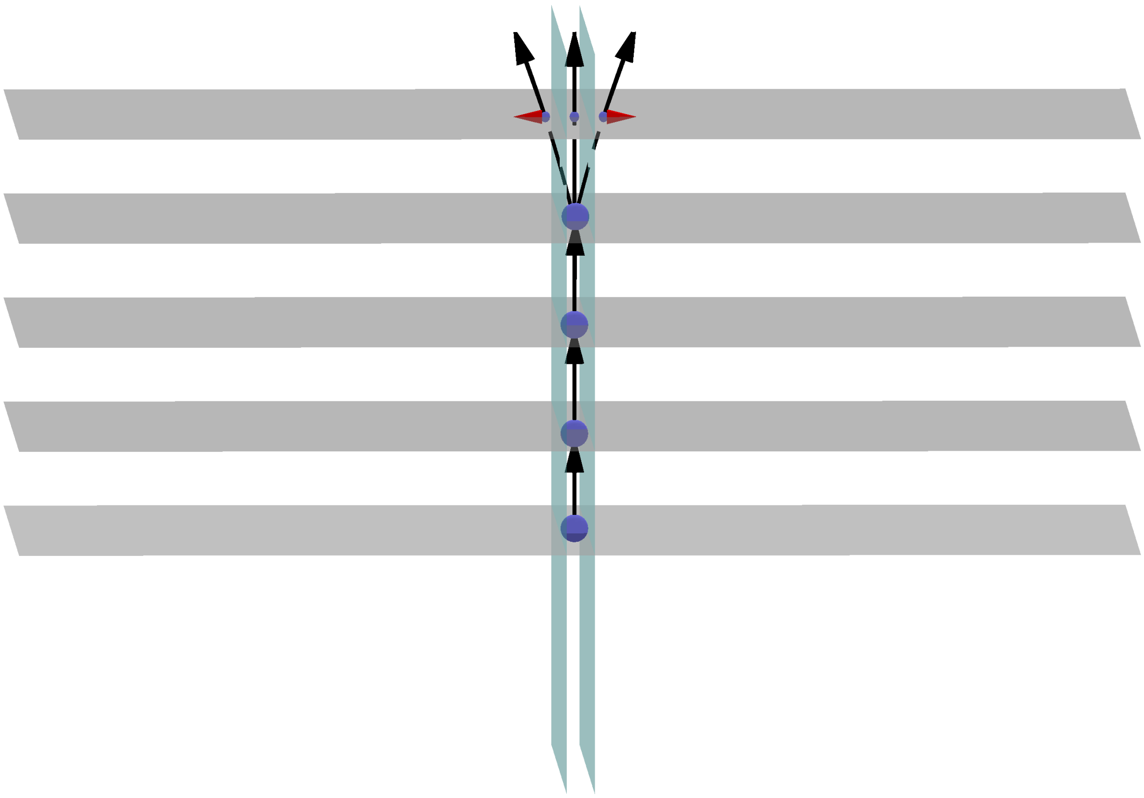

Figure 15: The above shows consistency with both the conservation law in flat spacetime and Newton's Third Law. Momentum density increases come from momentum fluxes, mass density changes come from momentum gradients, and the momentum of the center of mass is unchanged, i.e. it continues taking an autoparallel path in flat spacetime. The mass-energy equivalence demands that the increase in kinetic energy shown above must result in a decrease in the energy bound within the particles of the system, i.e. a decrease in their summed rest masses, as reflected by the smaller spheres on the topmost spatial hypersurface. In the case of gravitation, spacetime is non-flat, the conservation law becomes $T^{\mu\nu}_{;\nu} = 0$, and Newton's Third Law still remains as an independent postulate for non-gravitational interactions, with the consequence that the center of mass of a closed system still follows autoparallels, however these autoparallels are no longer embedded in flat spacetime.