A fully relative theory of gravitation

Peter R. Lamb

Abstract

A fully relative geometric theory of gravitation is proposed which retains Einstein’s formulation of a gravitational field as a distortion of space-time by matter. The distortion is judiciously revised to be proportional to the stored energy density of a local matter concentration relative to the background distortion. The resultant scalar theory therefore reproduces all the standard predictions of (tensor) general relativity theory for the present background potential (density of background stored energy). However, the magnitude of the distortion by local matter reduces as the total potential, due to all other matter, increases. This is shown to remove the inconsistency with quantum mechanics and the need to postulate dark matter and dark energy. The revised theory is Lorentz invariant and space is always locally flat but the magnitude of the space-time interval and the speed of light are no longer invariant. The theory reproduces the standard results of an invariant metric, but is consistent with Mach’s principle that inertia is due to the gravitational effect of the rest of the matter in the universe. The singularity at the centre of black holes is avoided and inflation is accommodated. The revised theory predicts a dependence of the shape of the light curves of supernovae on their galactic environment.

Keywords

Gravitation, Dark matter, Dark energy, Mach’s principle, General relativity

P. R. Lamb
Institute for Frontier Materials, Deakin University, Geelong Waurn Ponds 3216, Australia
e-mail: peter.lamb@deakin.edu.au
Introduction

General Relativity Theory (GRT), as proposed by Einstein, has survived for nearly 100 years and has had remarkable success in predicting observed effects, particularly those occurring in our solar system. These include the precession of the planet Mercury’s perihelion, the bending of starlight by the sun and the redshift of the frequency of light (and radio waves) in a gravitational field. The predicted effect of clocks running slower at higher gravitational potential has to be taken into account in order for the satellite-based global positioning system to give correct locations. More recent effects predicted and observed include the Shapiro time delay, the gravitational lensing of galaxies and galaxy clusters and the rate of energy loss of binary pulsar systems. An additional argument for a tensor gravitational theory of the form of GRT is that it is required by a graviton of spin 2. Currently, there appear to be no contradictions between observations and GRT.

However, there are some clouds on the horizon, although all of them have supporters and opponents over whether they are real problems. The first possible clouds are the need to postulate dark matter and dark energy. The observation of the lower than expected brightness of distant supernovae has been taken to imply that there has been an acceleration in the rate of expansion of the universe [1,2]. This in turn implies that some three-quarters of the universe consists of an unexplained dark energy. The rotation speed of stars in the disk portions of spiral galaxies is observed to be in poor agreement with that expected from Newtonian gravitation and the observed mass distribution, based on assumptions for the luminance to mass ratio of matter in the cores of galaxies. The rotation curves do not decrease in the expected inverse square root relationship but are nearly constant outside of the central bulge. This discrepancy is thought to betray the presence of a halo of dark matter. Diffuse dark matter haloes have also been put forward to explain the observed gravitational lensing of distant galaxies and galaxy clusters and the evidence for dark matter is considered by some to be compelling [3]. However, the proposed dark matter can neither absorb nor emit electromagnetic radiation and cannot be attributed to neutrinos. Despite extensive searches no candidates for this non-baryonic dark matter have yet been observed and none are predicted within the standard model of particle physics. Similarly, there is no persuasive theoretical explanation for the existence or magnitude of dark energy [4].

A second concern is that GRT does not appear consistent with Mach’s principle that inertia is a result of all the other matter in the universe. Mach’s principle reflects the fundamentally relativistic view that movement can only be specified relative to other objects and that therefore there can be no inertia if no other matter is present. This is encapsulated in Special Relativity which requires that the laws of physics are independent of the relative unaccelerated motion of the observers. GRT sought to extend this principle to include gravitation, however, it allows solutions where there is inertia but no other matter. Under GRT, if the only two masses in the universe are tied by string and rotate about their centre
then the string will be in tension. Mach’s principle, however, requires that this tension go to zero because there would be no other object that would allow the rotation to be defined.

For some critics the biggest concern is that, so far, it has not been possible to make GRT consistent with quantum mechanics because the strength of the gravitational potential varies inversely with the separation of the objects. Under GRT, energy corresponds to mass and, at the Planck scale, the available energy becomes enormous due to the Uncertainty Principle. In quantum electrodynamics the forces also tend to infinity at short distances but are brought under control by renormalization which can be seen as a shielding introduced by the polarization of the virtual particle/antiparticle pairs that arise as the available energy increases. However, the gravity of these particle/antiparticle pairs adds to the gravitational force under GRT and there seems to be no way of avoiding the infinities. Proponents of string theory see extra dimensions at the Planck scale offering a potential solution.

Another on-going area of controversy, if not inconsistency within GRT, is the apparently different behaviour of space at small and large scales. The expansion of the universe is compared with the stretching of the surface of a balloon under inflation. However, it is argued that the expansion applies only to truly cosmological distances, that is, the distances between galaxies or galaxy clusters. In the balloon analogy, the galaxies are represented by dots or small coins of fixed diameter stuck on the balloon surface, which is consistent with the lack of any observed expansion of galaxies. As the balloon inflates, the galaxies remain the same size, and the pattern of galaxies simply expands in size [5]. A corollary is that distant galaxies can recede at super-luminal velocities whereas under special relativity this would require an infinite amount of energy. It has been argued that bound systems such as atoms do not expand [6], but why do galaxies bound by gravity behave differently to the universe of galaxies where gravity is the only known binding force? Alternatively, it has been argued that there is no problem because “the expansion of space is neither more nor less than the increase over time of the distance between observers at rest with respect to the cosmic fluid” [7]. However, this seems problematic unless the cosmic fluid is different from space and an aether is being reintroduced? The expansion proposed by GRT takes place where there is little matter but is inhibited where there is matter.

A final concern of some is that, if not even light can escape from a black hole, then how is it that early in the big-bang universe when the density of matter was enormous that it too would not have behaved like a black hole? This conceptual problem appears to be related to the fact that our current theories predict a singularity at the centre of a black hole, which really means that the existing theories are no longer valid. This may also be related to the need for an explanation of the cause of inflation – the extremely rapid expansion of the early universe – that has been hypothesised to account for the observed uniformity of the cosmic ray background across regions that could not seem to have been causally connected.
A revised theory that could reproduce the predictions of GRT but throw light on any of these possible discrepancies or inconsistencies would be most helpful.

**Outline of the revised theory**

GRT is based on the invariant interval $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$ of special relativity, but allows a gravitational field to alter the dimensions of space and time while keeping $c$ and the interval invariant. The revised theory proposes that the scale of space ($R$) depends linearly on the stored energy density (which corresponds to a potential $\Phi$) while both $c$ and $t$ depend on the square root of the potential. This proposal automatically has that time is proportional to distance divided by the speed of light. In a region of different but constant potential there is still an interval $ds'^2 = -c'^2dt'^2 + dx'^2 + dy'^2 + dz'^2$, which is invariant, but only within that region i.e. $ds'^2 \neq ds^2$. A gravitational field now corresponds to a simple stretching of the spatial dimensions. The increased stretch then causes an increase in the speed of any wave travelling in that space according to the square root of the increase in stretch, which increases the time interval ($T$) by $R/c$ or the square root of the stretch. Time, interpreted as the relative rate at which events occur, appears slower by $1/\sqrt{R}$.

The fact that $c$ varies means that the energy stored in particles as mass, i.e. $m = E/c^2$, is not necessarily constant. An increase in the stretch of space, due to an increase in the presence of other matter, should lead to a decrease in the magnitude of the stretch due to a local particle of matter, by analogy with metal balls on a rubber sheet. Thus, it is proposed that the energy stored by matter, as mass, decreases as the local distortion due to other matter increases.

The energy stored by matter therefore decreases as that matter moves closer to other matter. This appears as kinetic energy of motion, that is, the matter will be accelerated and the magnitude of the force will depend on the amount of matter so that the local gravitational acceleration will be the same for different quantities of matter.

Note that this is a scalar theory. The forces arise only from changes in magnitude of a local energy density with position or time. However, it is different from other field theories in that the properties (the stored energy) of particles in the field are no longer constant, but depend on the total distortion due to all matter.

**Outline of potential problems and objections**

The potential problems or objections include:

i) how can the speed of light be allowed to vary as this should violate relativity?

ii) if $c$ and mass vary in this way, how is it that the variations have not been observed?

iii) time seems to be increasing as dimensions increase but the magnitude is only changing by the square root of that proposed by GRT?

iv) this is a scalar theory when a tensor theory has been shown to be essential?
v) the theory must lead to predicted consequences inconsistent with current observations?

As it turns out, all these concerns can be successfully addressed, as will be set out below. However, first some background arguments are presented on why a theory of the form suggested is required.

**Background**

GRT proposes a universal absolute space-time well away from the gravity of massive objects, an invariant metric, whose magnitude is not relative. This lack of relativity and the inconsistency with Mach’s principle can be traced to the assumption that a constant density of surrounding background matter has no effect. At first sight, this assumption seems plausible because the acceleration, due to a gravitational field, is observed to be proportional to the rate of change of gravitational potential. However, if the gravitational distortions of space from opposite directions add, rather than cancel, then the proportionality remains but the magnitude of the gravitational field reduces as the total potential increases. So the observer, immersed in a uniform density of background matter, sees gravitational forces that depend on the rate of change of gravitational potential due to the nearby matter, but is not aware that the magnitude of these forces depends on the density of background matter.

In GRT any spherically symmetric solution of the vacuum field equations must be stationary and asymptotically flat (Birkhoff’s Theorem). As a result, the metric of a vacuum inside a stationary spherical shell of matter must be the flat Minkowski metric. Thus, under the assumption that a flat metric corresponds to no gravitation, the gravitational field must vanish inside the sphere, which agrees with Newtonian gravitation, and the mass and radius of the spherical shell have no discernible effects upon the laws of physics as they are observed in the interior [8]. Interestingly, it has been shown recently that this does not mean that the shell of matter has no effect [9]. The outside mass distribution makes an interior clock run slower. Nevertheless, GRT requires that a surrounding non-rotating uniform shell has no effect on the observed gravitational behaviour and does not lead to any curvature of the spatial dimensions. This appears to be inconsistent with Mach’s principle, if inertia is due to gravitation, because it implies that the surrounding matter has no gravitational effect.

This discrepancy is still a highly contentious issue. Woodward and Mahood reviewed the gravitational explanation of the origin of inertia and concluded that GRT dictates that inertia is gravitationally induced [10]. However, they pointed out that GRT allows solutions in which test bodies in otherwise empty space experience inertial reaction forces. They also concluded that inertial reaction forces in GRT are produced by an instantaneous radiative interaction with chiefly the most distant matter in the universe; which violates special relativity.
GRT postulates that the speed of light is an invariant and that the cause of gravitation is a curvature in the underlying metric of space-time. A curvature requires distortions, which can only be expansions or contractions, of the dimensions. A spherical shell of matter can be envisaged as an infinite number of small pairs of equally distant masses in opposite directions, each of which is entirely symmetric perpendicular to the line joining them. So any unbalanced spatial distortion can only be an expansion or contraction in the radial direction. The lack of gravitational effects from a stationary spherical shell in GRT has been taken to imply that spatial distortions from opposite directions cancel, analogous to the cancellation of electric charges on a surrounding conducting sphere which ensures gauge invariance of electromagnetic interactions. This cancellation of curvature is implausible because symmetry demands that both masses contribute an expansion or both a contraction, so distortions from opposite directions must add rather than cancel.

The fractional effect of an asymmetry in the matter distribution will depend on the total distortion from all matter. Hence, if the total distortion is much larger than the additional distortion due to local matter then the additional distortion will appear on an approximately constant background. A common 2-D analogy is the way metal balls will distort a stretched rubber sheet. With rubber sheets, or springs, the initial distortion introduces a tension that makes further distortion more difficult. A more highly tensioned sheet is distorted less. If this is the case for space, then two predictions follow. Firstly, the effect of additional matter on the distortion of space should decrease when the overall density of matter increases. Secondly, the wave speed of a distortion should increase when more matter is present. A spherically uniform and stationary distribution of distant matter does not produce a gravitational acceleration. However, it should have an effect such that the magnitude of any acceleration, due to a local inhomogeneity of the matter distribution, is determined by the ratio of the imbalance in extension to the total extension. The revised theory is consistent with these requirements.

Relativity and a variable speed of light

The assumption of the constancy of the speed of light, in empty space, $c$ was taken over from the Maxwell-Lorentz theory of electromagnetism. The equivalence of inertial systems (the special principle of relativity) then leads to the result that time and distance are relative, that the interval $ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2$ is an invariant, and that the laws of nature are invariant with respect to Lorentz transformations. This, however, does not imply that the speed of light has to be constant in a different location where the gravitational potential is different. The problem actually traces back to the idea that a vacuum constitutes “empty space”. If the dimensions of space are determined by matter then it is never empty. The background matter determines both $c$ and the dimensions of space, and the time interval is always proportional to distance divided by the speed of light. In a region of different constant potential there is still an interval $ds'^2 = -c'^2 dt'^2 + dx'^2 + dy'^2 + dz'^2$, which
is invariant, but only within that region i.e. $ds'^2 \neq ds^2$. Lorentz invariance is always maintained.

In GRT the “effective velocity” of light is understood to change in the presence of a gravitational field because of the ways that distance and time transform in this field. It makes more sense to expect both the dimensions of space and the actual speed of light to vary which will then cause time to vary. If the scale of space is proportional to the total potential and both $c$ and the time interval proportional to the square root of the potential, then $cT$ is proportional to potential and $T = R/c$. Local Lorentz invariance can still be maintained even though the speed of light varies [10,11]. Here the interval is invariant (in an inertial frame) only within a region of constant potential. Travelling to a region of different matter content requires movement through a gravitational potential and hence through a non-inertial transformation.

It has been claimed that the postulate of special relativity (i.e. that all physical laws are valid in all inertial systems) implies that $c$ is the same in all inertial frames [12]. However, it only implies that $c$ must be the same within an inertial frame. The reason is that an inertial frame at a different constant gravitational potential (a different matter/energy density) can only be reached via an acceleration.

The result of ignoring the background potential in GRT is to remove any effect of the change in total potential that occurs when a body is in free fall. The Einstein equivalence principle states that in a freely falling reference frame the physics is the same as an inertial frame with no gravity. However, in general, the two frames are only instantaneously equivalent. If a body is in free fall towards the earth then the total potential is changing. In both frames, tidal effects are seen in any object of a finite size. In the freely falling frame the tidal forces are due to changes in magnitude and direction of the gravitational field with position and time, whereas in the inertial frame, the tidal forces are ascribed only to differences in position. In the freely falling frame the magnitude of the tidal forces will increase as the body moves closer to the centre of the earth. A formulation that reproduces the initially correct tidal forces, but does not allow their magnitude to change (i.e. GRT), will only be correct in the limit that any change in potential is small relative to the total potential.

GRT equates gravity with a distortion of a fixed interval of space-time, but an alternative is to have the distortion correspond to a change in the magnitude of the interval in space-time. The magnitude of $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = \eta_{\mu \nu}dx^\mu dx^\nu$ changes as well as its components rather than just the components. Thus it is proposed that the larger the background energy density due to matter or background potential $\Phi$ (related to the gravitational potential but increasing as two masses approach each other), the larger the dimensions of space and the higher the speed of light. The acceleration due to a given potential gradient will still reflect the rate of change of the distortion, that is, the slope of space and time. By analogy with the stretched rubber sheet or a spring, the rate of increase
in distortion (by local matter) will decrease with increasing distortion, but equally the rate of increase in the scale of space \( R \) (due to all matter) will decrease with increasing scale. The distortions and potentials in two regions of space are predicted to be related by 
\[
\frac{R_1}{R_2} = \frac{\Phi_1}{\Phi_2}.
\]
However, the revised theory has both \( c \) and \( T \) increasing in proportion to 
\[
\sqrt{\left(\frac{\Phi_1}{\Phi_2}\right)}
\]
which retains \( T = \text{distance} / c \).

In GRT the tidal forces, which for example cause separated objects in a lift in free fall to move closer together, are equated with curvature, which is the second derivative of the metric. The differences of the gravitational effects on neighbouring points are proportional to the derivative of the gravitational field, and the second derivative of the gravitational potential [13]. Hence, GRT equates gravitational effects with the curvature of a metric. The revised theory still has the gravitational force proportional to the rate of change of the potential and the differences on neighbouring points proportional to the second derivative. However, the potential is the total potential due to all matter and the distortion is a simple stretching of space and of the product \( cT \). As will be shown below, the distortion of a metric of fixed magnitude is, to first order, equivalent to changing the dimensions of space-time and \( c \) as proposed. Equating the size of the tidal effects with the observed strength of gravity and speed of light in our local region means that the magnitude of the effects of gravity will only be correct when the total potential \( \Phi \) is similar to that observed locally. In a location where the total potential is significantly different, for example, closer to the centre of our galaxy, or at an earlier epoch when the universe was denser, then the predictions of GRT will be in error.

If \( c \) varies, then the relationship \( E = mc^2 \) implies that the mass of a stationary object (of fixed energy) decreases in inverse proportion to the square of an increasing speed of light. This appears consistent with the analogy of a stretched rubber sheet where the distortion of the sheet by matter decreases as the total distortion increases. The equation indicates that \( c^2 \) is the conversion factor for the mass equivalent of the same amount of energy in different regions of space, but not necessarily the same amount of matter (number of particles). It is tempting to assume that the same number of elementary particles corresponds to the same energy in different regions of space but this would be equivalent to assuming that the energy of such stationary states is independent of the amount of other matter present. If time varies in different regions then it might be expected that the energy of the time stationary state (an elementary particle) might also vary.

In 1864 Maxwell pointed out the paradox that arose because gravitation was an attractive force: “the presence of dense bodies influences the medium so as to diminish this energy [of the medium] whenever there is an attractive force” [14]. If we demand conservation of energy and that the number of elementary particles is conserved when two bodies accelerate towards each other under gravitational attraction, then the medium or particles must give up energy. Hence, we are forced to conclude that the energy of the same
elementary particle (or the energy of the distortion of the medium) is lower in the presence of other matter. This should not be surprising if the dimensions of space are larger and the passage of time is slower. The question is: how do \( m, c \) and \( t \) vary with the potential?

From the rubber sheet analogy, it should be expected that mass (which should be related to the amount of distortion of the sheet by a given amount of energy) will decrease as the potential (due to other matter) becomes larger. The relationship \( E = mc^2 \) reflects the conservation of energy in special relativity and should be expected to hold for all regions, and hence all values of \( c \). If the number of particles \( N \) is constant, but the energy per particle decreases as the potential increases then mass must have a strong dependence on \( c \). Therefore, it is proposed that \( m \propto 1/c^4 \), so that the energy per particle decreases in proportion to \( 1/c^2 \). If the distortion by a particle decreases in inverse proportion to the increase in expansion of space then \( m \propto E/\Phi \) with \( \Phi \propto c^2 \). This is consistent with the analogy and with an attractive gravitational force because the stored energy of a particle decreases as it gets closer to other matter. The expression \( m_g \propto E/\Phi \) then explains gravitational mass as the potential energy (and distortion of space) of a particle relative to the potential (and distortion) due to all other particles. Hence, it is proposed that \( R_1/R_2 = \Phi_1/\Phi_2 = (c_1/c_2)^2 \), in which case \( c = R/T \) implies \( R_1/R_2 = (T_2/T_1)^2 \).

An increase in spacing of particles by \( R \) corresponds to a reduction in density by \( 1/R^3 \). If the mass, or stored energy, per particle increases by \( R^2 \) when the scale of space, and hence wavelength of light, decreases by \( 1/R \), then the energy density will decrease by \( R^2/R^3 \) or \( 1/R \). Thus the proposed dependence of mass on \( 1/c^4 \) is consistent with the scale of space being proportional to the energy density and the square of the speed of light.

The implications for conservation of momentum also need to be considered. Force is defined as the rate of change of momentum \( \vec{F} = d\vec{p}/dt \) where \( \vec{p} = (E/c^2)\vec{v} = m\vec{v} \). However, if special relativity is to hold, all velocities must be relative to the speed of light. The momentum-like quantity that should be preserved is therefore \( \vec{p}c = E(\vec{v}/c) = mc\vec{v} \). The force is then \( \vec{F} = d(\vec{p}c)/d(ct) \), which is consistent with time being distance divided by \( c \) and work (the change in energy) being force by distance. This is also consistent with \( E^2 = m^2c^4 + p^2c^2 \) being a conserved quantity and with gravitational and inertial mass being the same property and hence having a fixed ratio, as required by the Eötvös experiments.

**Time**

The apparent time interval \( dt' \) in a gravitational potential due to a massive spherical object is observed to good accuracy to have the form \( dt'^2 = (1+2\Delta\phi/c^2)dt^2 \) relative to a time interval far away from the massive object. The difference in potential is \( \Delta\phi = -GM/r \). The gravitational time dilation at distance \( r \) from a sphere of mass \( M \) is then \( dt'^2 = (1-2GM/rc^2)dt^2 \). This effect of time appearing to run faster in a higher gravitational
potential, e.g. as you move away from the Earth, is well established and is seen in the need to make appropriate corrections to the signals from the satellites of the gravitational positioning system. The time dilation can be derived under GRT using the equivalence principle and light being redshifted by gravity such that $\Delta \omega / \omega = \Delta \phi / c^2$, with $c$ constant and the observed frequency of light being inversely proportional to the local proper time $\tau$ [13]. It is found that $d\tau = \left[1 + (\phi_1 - \phi_2) / c^2\right]d\tau_2$ or $d\tau_1^2 \approx (1 + 2\Delta \phi / c^2)d\tau_2^2$.

However, under the proposed theory the background potential $\Phi$ should determine $c$ and $c^2 \propto \Phi$. (The current understanding is that the gravitational potential $\phi$ increases with distance from a massive object. However, here it is proposed that this potential is a measure of the energy stored as mass which increases with a decrease in the background potential. The total potential $\Phi$ is actually the background potential $\Phi_\text{b}$ minus the (conventional) gravitational potential $\phi$ and increases with the speed of light and with an increase in the density of matter.) We then have $T'^2 = (1 + \Delta \Phi / \Phi)T^2 \approx (\Phi' / \Phi)^2T^2$ where $\Phi' = \Phi_\text{b} - \Delta \phi$. However, $T'^2 = (\Phi' / \Phi)^2T^2$ implies $T' / T = \Phi' / \Phi = (c'/c)^2$, which conflicts with $R = cT$ and $\Phi / \Phi' = R / R'$ and with the proposal that $T$ increases in proportion to $\sqrt{(\Phi' / \Phi)}$. Thus the observed gravitational time dilation appears, at first sight, to conflict with the expectations of the proposed theory. This conflict can be resolved if i) the assumption is dropped that electromagnetic transitions of the same energy in different regions of space correspond to the same frequency or wavelength. It also leads to a revision in the nature of time.

The proposed revised notion of time is that it is the relative rate at which events of a particular energy occur in regions of different background potential. Thus, the number of oscillations of a photon as it traverses the same rod in a different location could be considered as the number of events, or ticks of the clock. If space (and the wavelength of the photon) expands in proportion to $\Phi$ but $c$ increases in proportion to $\sqrt{\Phi}$ then the length of the interval of time $T$, the spacing between ticks, will increase by $\sqrt{\Phi}$. The spacing between ticks will be larger when the potential due to matter is larger, so the relative rate $\nu'$ at which events in a region of higher potential occur will appear slower when viewed from a region of lower potential i.e. $(T / T')^2 = \Phi / \Phi' = (\nu' / \nu)^2 = (c / c')^2 = R / R'$. Time reflects the relative rates at which identical events occur in different regions. Time goes faster when the interval between events (ticks) is smaller.

The behaviour, in terms of time dilation, of any system having a given energy, momentum and angular momentum should be consistent under the changed distance scale and speed of light, if the system is located in a different region of space. So that, for example, one revolution of the same system, having the same angular momentum and energy, should appear to take relatively longer in proportion to $\sqrt{\Phi}$. The kinetic energy $\frac{1}{2}mv^2$ of mass $m$ will be preserved if the energy content of the mass is proportional to $1/c^2$ and the normalized

Original version submitted 14th August, 2013. Revision 1 submitted 9th September 2013
velocity $v/c = d/ct$ is used. The linear momentum variable that should remain constant for the same amount of energy is $\vec{p}c = E(\vec{v}/c) = mc^2 (\vec{v}/c)$. Hence, normalized momentum defined as $\vec{p}c = m_0 \vec{v}c$ with $m_0 = m_i$ should be expected to be the preserved quantity which matches energy. The period for the rotation of the same energy content ($E$) with the same momentum (hence $v'/c' = v/c$) will change according to $T'/T = \omega' \omega = (v/r)/(v'/r') = c'/c$ which is $(T'/T) = \sqrt{\Phi'/\Phi}$, as required. The apparent time dilation of events of the same energy in regions of different potential is consistent with the revised definition of time. It is also consistent with the conservation of the energy-momentum four vector $(E, \vec{p}c)$. The conservation of the four vector confirms that the formulation maintains general covariance and the theory is Lorentz invariant.

The above behaviour, the consistency of the apparent time dilation, applies to systems which involve small perturbations to the local potential on a large and approximately constant background potential. The apparent behaviour will be different if the background potential changes significantly. However, the equivalence principle should hold that the physical laws should be the same once the different background is taken into account. This is seen to be the case because an identical clock will have the same number of particles and so will have energy and normalized momentum that are both reduced by a factor of $\Phi$. The same behaviour is then observed if distance and the product $cT$ are also normalized by dividing by $\Phi$.

Time $t$ is now a quantity that indicates the relative rate at which events of the same energy, relative to the background energy, occur. It replaces the GRT concept of “proper time” that was absolute for all space well away from local gravitational distortions. This makes it clear that “proper time” is effectively based on the assumption that the gravitational potential of distant matter has no effect. GRT has actually introduced a sweepingly absolute concept: the invariant interval of absolute space-time. It is said that, partly for this reason, Einstein and other physicists preferred “invariance theory”, because it involved something that, at its core, was not relative [15].

In a region of constant potential, the energy of a photon of frequency $\nu$ is $E = h\nu$, from quantum mechanics. If it is assumed that $h$ is constant, independent of the potential, then the relationship indicates that the frequency of a distortion of space (how fast space “vibrates”) that carries the same energy is independent of the potential. However, this is inconsistent with time varying between regions of different potential. If it is demanded that energy is a meaningful property for both the gravitational potential (energy) of space and the energy of a photon then the value of $h$ must, in some way, reflect the value of the background potential. In which case, an electromagnetic transition of the same energy should not have the same frequency in a different region of space, as is normally assumed.

So how does frequency change and how is this revised concept of time, with $(t'/t)^2 = (T'/T)^2 = \Phi'/\Phi = (c'/c)^2$, consistent with the apparent time dilation seen in, for
example, the Pound-Rebka-Snider experiments? These experiments all appear to be explained by a gravitational frequency shift $\frac{\Delta \omega}{\omega} = \frac{\Delta \phi}{c^2}$, which corresponds to $t'/t = \Phi / \Phi'$, if $\Phi = c^2$. The answer is that the frequency for the same energy (same transition) is shifted lower by $\sqrt{(\Phi / \Phi')}$ in moving from a frame with higher acceleration to one with lower acceleration. There is also a shift in emission frequency but this is cancelled by the shift of $\sqrt{(\Phi / \Phi')}$ due to the time dilation between the two locations (the time dilation between inertial frames is due to the difference in energy density). This shift in the frequency due to time dilation is consistent with $E = h \nu$ and the wavelength varying in proportion to the scale of space, because $\nu = c / \lambda$, provided the energy of the photon is unchanged. This in turn implies that $h \propto c = \sqrt{\Phi}$ so that the product $hc$ is proportional to $\Phi$ and $\nu \propto E / c$. The value of the dimensionless fine structure constant $\alpha = e^2 / 4\pi\varepsilon_0 hc$ remains constant because the electric constant $\varepsilon_0$ has dimensions of farad/metre and distance (metres) scales as $R \propto \Phi$. So called, “natural units” ($h = c = 1$) can be used within a region of unchanged gravitational potential but the magnitude of the units will vary with the potential.

If the energy of an electromagnetic transition is the same in all inertial frames independent of the background energy density, then a photon emitted in one frame, that does not gain or lose energy in a gravitational field, will be resonantly absorbed in the other frame. This will occur even though the clock rates in the two frames are different. The shift in frequency seen in the Pound-Rebka-Snider experiments is just that due to the different accelerations of the two frames. This is the same gravitational redshift predicted by GRT and the two theories give the same prediction for the redshift. The time dilation has no effect on the energy of the photon otherwise the predictions would be different.

There is no shift in frequency from the photon losing energy in “escaping” from a gravitational field. This is a mistaken viewpoint [16]. The photon does not lose energy. Nor is it “attracted” by the gravitational field; its path is a straight line in the distorted space. The change in frequency is because of the change in the dimensions of space-time plus that due to any change in acceleration between the frames. No change will be observed in the wavelength of a transition of the same energy in inertial frames in different regions (if the total energy density is approximately constant). Under the proposed theory, the emitted wavelength changes as the ratio of the potentials. However, the wavelength then expands or contracts in the same proportion in going to a region of different potential. Thus the wavelength of light is not observed to change unless the background potential has changed between the emission and reception. For an approximately constant total potential, and inertial frames, the observer does not see any changes in wavelength with potential. However, the apparent duration of events of the same energy will increase by the square root of the ratio of the potentials. If the universe is expanding then a reduction in potential of $\Phi / \Phi'$ will be accompanied by an increase in (the rate of) time of $\sqrt{\Phi'} / \Phi$.

Original version submitted 14th August, 2013. Revision 1 submitted 9th September 2013
Scalar Theory

It has been shown that a graviton of spin 2 necessarily leads to a tensor theory of gravity of the form of GRT [17]. Yet here a scalar theory is being proposed, so how can this be resolved? Arguments for the spin of the graviton have been set out clearly by Hatfield [18]. As a force carrying particle it needs to carry integer rather than half-integer spin. Next, it is found that when the exchanged particle has odd integer spin then like charges repel and opposite charges attract. Given that gravity is universally attractive then it must correspond to even spin. If the exchanged particle has spin 0, then it does not couple to the spin 1 photon. This is because spin 0 corresponds to a scalar field and the calculation of the exchange amplitude shows that a spin 0 graviton only couples to the trace of the stress-energy tensor. However, the stress-energy tensor for the electromagnetic field in Minkowski space is traceless. Hence scalar gravitational fields do not couple gravity to light. It is then argued that since light “falls” in a gravitational field and that light is deflected by massive objects then light couples to the gravitational field and the graviton cannot be spin 0. Spin 2 is the next choice. Finally, it is found that a massless spin 2 graviton couples to the stress-energy tensor in exactly the same way as the gravitational field of general relativity theory (GRT). From this it is concluded that, since GRT is in agreement with observation, the graviton must be spin 2.

The understanding that photons do not lose energy (“fall”) in travelling between regions of different gravitational potential is actually part of standard GRT [16]. The time dilation causes the frequency to appear to be changed but the energy is unchanged. The resulting explanation of the gravitational redshift actually leads to the conclusion that the receiver atom is at a higher gravitational potential and therefore that all its energy levels are increased (blueshifted) [13,16,19]. The necessary conclusion is that particles lose gravitational energy when moving to a region with a higher density of matter. Thus the concept that particles, but not photons, must lose energy in moving to a region with a higher density of matter should not be novel. However, here it is proposed that the energy of an electromagnetic transition is also constant but the energy equivalence (the energy that can be stored) of the matter changes.

The photon does not lose or gain energy in a gravitational field. Nor is it deflected by the gravitational field; its path is a straight line in the distorted space. The observed changes in frequency are because of the changes in the dimensions of space-time. The photon does not couple to the stored energy of the gravitational field, i.e. to gravitons, or to other photons (except elastically via virtual particle/antiparticle pairs). Hence, the graviton should be expected to have spin 0 and a scalar theory of gravitation is required.

Brans and Dicke [8] sought to add a scalar component to allow variations in the gravitational “constant” but combined it with the existing tensor theory of GRT. Recently, Novello et al. [20] summarised the difficulties of previous scalar theories of gravity noting that the main handicap of such theories is that gravity does not couple to the electromagnetic field when

Original version submitted 14th August, 2013. Revision 1 submitted 9th September 2013
the source of gravity is the trace of the energy-momentum tensor. It is not a difficulty as it is what is observed. However, as they noted, it leads to a space-time dependent mass. It turns out that this is also consistent with observations and can explain why massive objects gain energy in a gravitational field.

Giulini [21] examined the arguments that Einstein gave on why a special-relativistic theory of gravity based on a massless scalar field could be ruled out merely on theoretical grounds. He concluded that the two main arguments are unsatisfactory and that such a theory seems formally perfectly viable, though in clear contradiction with experiments. However, the primary noted contradiction with experiment was that the theory predicted that a travelling (free) electromagnetic wave would not be bent in a gravitational field. The other noted contradiction was in the calculation of the precession of the planet Mercury, but this was based on the assumptions of GRT for the nature of mass and time.

Giulini also argued that trapped electromagnetic radiation in a box with mirrored walls will induce stresses in the box’s walls due to radiation pressure. This will increase the weight of the box by $\Delta m = E_{\text{rad}} / c^2$ and, in this sense, bound electromagnetic fields do carry weight (have mass). However, this is not a problem, it is entirely consistent with the revised theory that the total potential, and therefore stretching of space, is proportional to the energy density. It also throws light on the argument used by Feynman [18] to rule out a spin zero graviton. He argued that it was inconsistent with a volume of hot gas being more massive than the same gas when cold. This is because the hot gas needs additional pressure to maintain the same volume. This stored energy gives mass (but only if it is stored). It was then argued that this was inconsistent with lower mass corresponding to lower binding energy (of nuclei). The argument appears to be that lower mass corresponds to less gravitational attraction which contradicts energy being released when objects get more strongly bound in a gravitational field. However, in the proposed theory, if particles move closer together then the local energy density increases, but the energy stored per particle (mass) decreases, so that the amount of free energy released (the heat of the gas particles in the above argument) corresponds to the stored energy lost as mass. Thus the arguments for a spin 2 graviton are flawed and a spin 0 graviton, and hence a scalar theory, is demanded.

It is noted here that there seems no reason to assume that the gravitational interaction, or graviton, be quantised. The gravitational potential just reflects the relative location and movement of other massive objects taking into account that the information can only propagate at the speed of light.

**Consistency with GRT and Newtonian predictions**

The next question is whether and how the predictions of an invariant interval $ds'^2 = ds^2$ with $c$ constant and a distorted metric (i.e. GRT) can be reconciled with a variable $c$ and $ds'^2 = (\Phi' / \Phi)^2 ds^2$. Firstly, distance, which will vary with the scale of space $R$, is calculated...
from the return time of electromagnetic signals assuming $c$ is constant and that the local time is $t' = (\Phi / \Phi')t = (\phi' / \phi)t$. Time is assumed faster by just the amount to underestimate distance by $(\Phi / \Phi')$ so that the scale of space appears unchanged. Secondly, assuming $c$ constant, but that the “effective velocity” is distance divided by time, overestimates the actual velocity by $\sqrt{(\Phi / \Phi')}$ while assuming $t' = (\phi' / \phi)t$ overestimates $t$ by $\sqrt{(\Phi / \Phi')}$ so that $ct$ is overestimated by $(\Phi / \Phi')$, or underestimated by $(\Phi' / \Phi)$. Thirdly, the wavelength of light expands and contracts according to the local potential so the changes in scale of space are invisible. The result is that $ds'^2$ appears to be $D^2 - c^2t^2$, where $D^2 = x^2 + y^2 + z^2$, i.e. the magnitude of the interval appears invariant because of the assumptions made. However, the actual interval is $ds'^2 = (\Phi' / \Phi)^2D^2 - (\Phi' / \Phi)^2c^2t^2$. Assuming $ds'^2 = D^2 - c^2t^2$ and equating this with $g_{\mu\nu}dx^\mu dx^\nu$, under GRT, means that the observed changes in time $t' = (\phi' / \phi)t$ must be compensated by inverse distortions of distance so that the determinant of $g_{\mu\nu}$ is 1. For $D=0$, GRT has $g_{00} = -(1 + \phi / c^2)^2$. This corresponds to $-(\Phi' / \Phi)^2$ under the convention that the potential increases as a massive object is approached. The magnitude of the distance term of the apparent metric, that multiplies $D^2$, will be the inverse of the magnitude of $g_{00}$ i.e. $(\Phi' / \Phi)^2$. Hence, the predictions of the revised theory match those of a distorted metric, at least while the total potential is similar.

An invariant metric with $c$ constant reproduces the equivalent distortion of space-time as $(T / T')^2 = \Phi / \Phi' = (c / c')^2 = R / R'$ in the limit that the magnitude of the potential is similar to that observed locally, i.e. the value of the gravitational “constant” $G$ is the local value. This equivalence extends to the angular deflection, precessions and time delays predicted by GRT. The two theories are not easily distinguished locally unless the differences in total potential are large or a parameter such as $G$ or the speed of light can be measured with an accuracy better than $\Delta\phi / \Phi$. However, the magnitude of gravitational lensing, gravitational acceleration and clock rate will appear to be different in distant regions where the background potential is not the same as in the region of our solar system.

The revised theory will be in agreement with all the standard predictions of GRT in the region of our solar system because the apparent distortions of a metric are reproduced. The increase by a factor of two in the bending of light, over the pre-GRT calculation, is easily reproduced and has a clearer explanation. The standard GRT calculation of the bending of light is to use the Schwarzschild line element which has the form [5]:

$$ds^2 = \left[1 - \frac{2GM}{rc^2}\right]c^2dt^2 - \left[1 - \frac{2GM}{rc^2}\right]^{-1}dr^2 - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)$$
For the GRT calculation of the deflection of light by the sun and the calculation of the Shapiro time delay, the value of $ds^2$ is set to zero and the last term is ignored. The presence of two terms involving $2GM$ leads to an expression for the angle of deflection ($\delta\phi$) of:

$$
\delta\phi = \frac{4GM}{c^2r_{\text{min}}}, \text{ where } r_{\text{min}} \text{ is the distance of closest approach.}
$$

In the revised theory, a $2GM$ term in $dr^2$ or $c^2dt^2$ corresponds to $(\Phi' / \Phi)^2$. The $4GM$ dependence of the angle of deflection arises because there is an expansion of space of $\Phi' / \Phi = R' / R$ both along the path of the light ray and perpendicular to that path.

The terms involving $GM$ come from $\Delta\phi = -GM / r$ and so incorporate the local values of $M$ and $r$, which will depend on the magnitude of the local background potential. The value of $G$ in $\Delta\phi = -GM / r$ will vary in different regions of space with the value increasing away from concentrations of matter. Hence the Newtonian and GRT predictions hold in the limit that the scale of space is constant or, equivalently, that the background potential is enormous relative to the local variation in potential.

The revised theory is a geometric theory in that the geometry of space-time is altered by matter which then affects the motions and physical measurements of matter. However, the distortion of space is a simple stretching or contraction of a flat space, and of $cT$, rather than a complex curvature. The concept of time is changed to one of the relative rates at which events of the same energy occur when the background energy changes. There is now a reason for the acceleration of matter in a gravitational field in terms of a reduced storage of energy of the constituent particles as the amount of other matter increases. Mass reduces as the background potential increases. Energy and momentum are conserved.

The gravitational force is proportional to the change in stored energy as the total potential changes, and the total potential depends on all the other matter, and its relative motion, in the universe whose influence could have reached the local position. The energy that can be stored by a particle as mass $m = E / \Phi$ is inversely proportional to the background potential. The force that arises from the energy available for acceleration as the potential increases with a change in position is $F = -\partial E / \partial r = m\partial\Phi / \partial r$.

However, if mass, distance and speed of light vary in this way, why don’t we see it? Firstly, with distance, it needs to be appreciated that because electromagnetic signals are used to measure distances of faraway objects then differences in the speed of light and dimensions of space are hidden. Only differences in clock rates or differences in the magnitudes of gravitational effects can be observed. Secondly, if the differences in potential are small relative to the background potential then the predicted differences in clock rates are nearly the same. However, GRT and the revised theory will give different predictions for the magnitudes of angular deflections and apparent motion (velocities) of massive objects in regions where the total potential is significantly different.
So, how should the differences in mass, distance and the speed of light show up in observations of our solar system? For objects gravitationally bound to a large central mass we know that Newton’s inverse square law $F(r) = ma = m\ddot{\phi}/\dddot{r} = GMm/r^2$ holds accurately in the limit of low velocities. If the local variation in potential reflects the distortion due to the central object then it should vary according to $M/r$, although the scale of $r$ will change if the total potential changes (and the central amount of distortion by a given number of particles will be lower in a region of higher background potential). If acceleration reflects the slope of the distortion of space then, in any region, it should be proportional to the rate of change of the total distortion with distance. The force law appears to hold accurately because $m'/m = (c'/c)^2$ and $(R'/R)^2 = (c'/c)^2$ so that the radial contraction in the scale of space, $R'/R = \Phi'/\Phi$, which is not visible (but alters the value of $1/r^2$) is cancelled by the change in mass for an object with a constant number of particles. The acceleration is the rate of change of velocity (relative to $c$) with time which introduces a dependence of the acceleration on the potential but the effect is cancelled by using the apparent distance (rather than the contracted distance).

The field equation, which describes how an isolated free particle responds to changes in its environment, can be derived as follows. An increase in potential with distance will mean that the stored energy of the particle will decrease, this will manifest itself as a force in that direction. The force will result in a variation with time of the particle’s momentum. Therefore, $F_i = -\partial E / \partial x_i = dp_i / dt$. Special relativity has $\vec{p} = m_i\gamma \vec{v}$, where $m_i$ is the inertial mass, and it is proposed that the gravitational mass is given by $m_G = E / \Phi$. Hence, $m_idv_i / dt = -m_G \partial \Phi / \partial x_i$, where the $\gamma$ term has been set to 1 because the apparent behaviour for an observer travelling with the particle is being considered. The result is that $a = dv_i / dt = -(m_G / m_i) \partial \Phi / \partial x_i$. If $\Phi = \Phi_s + Gm_i / r$ then $a = \partial \Phi / \partial r = Gm_i / r^2$ which is Newton’s law of gravitation and it can be seen that it applies in the limit that the background potential is constant and the total potential is nearly constant so that variations in $c$, $t$ and distance scale can be ignored. The behaviour of non-relativistic massive objects in a region of a large and approximately constant background potential will be Newtonian except that the strength of the interaction will vary inversely as the background potential. If the total potential is not constant then the scale of distance, velocity and time all change with distance from the central object but these changes are not easily observed using electromagnetic radiation which also responds to the potential. This is accommodated in the predictions of GRT by assuming $c$ is independent of potential but that time is proportional to the potential.

The expected variation of $G$ inside a sphere containing different uniform densities of matter can be determined as follows. The $1/r$ dependence of the local gravitational potential, $\Delta \phi = GM / r$, means that an infinitesimally thin shell of matter, in an otherwise empty space, will produce a uniform constant expansion inside the sphere ($\Delta \Phi \approx 4\pi r \rho dr$). In the limit...
of a uniform density, a stretch proportional to \( r \) (i.e. an increase in \( r \) from \( R_1 \) to \( R_2 \)) reduces
the volume by \( 1/r^3 \), so that the ratio of the energy densities is
\((E_2 / E_1) / (R_2 / R_1)^3 = (M_2 / M_1) / (R_2 / R_1)^2\). This is consistent with the ratio of the potentials
being \( G_2 M_2 R_1^4 / G_1 M_1 R_1^2 \) if \( G_2 / G_1 = R_1 / R_2 \), i.e. \( G \) is proportional to \( 1/\Phi \).

A comment on the field equation is in order. In principle, a force should also arise if the
stored energy of the particle decreases with time, but does not move relative to other
matter. However, this means that the total potential has increased which corresponds to a
spherically symmetric decrease in the distance to other matter. This is a back-reacti-
on the particle on the field, which contradicts the assumption of an isolated free particle.

The revised theory has the dimensions of space, time, speed of light and mass dependen-
t on the amount of other matter. The change in size is real and should not be confused with the
apparent contraction of objects and dilation of time when an object is moving relative to the
observer. Special relativity still applies although the size of the contraction and dilation also
depend on the amount of matter because of the change in speed of light.

**Elaboration of the Theory**

The proposal is that space expands by an amount determined by the stored energy. It will
expand until the energy through a surface enclosing the stored energy becomes constant.
That is, until the energy from matter outside the volume (the background flux) has the same
magnitude. Particles are stationary states of stored energy, so the energy near them will be
larger and space more expanded (for a given background flux). However, the amount of
energy that can be stored in each particle will decrease as the background flux increases.

If the spacing between a homogeneous, uniform distribution of matter increases by a factor
\( R \), then the density of matter decreases by \( 1/R^3 \). If there is a decrease in stretch by \( 1/R \)
then there will be an increase in (the stored energy) mass per particle of \( R^2 \) and the total
stored energy in the expanded volume will increase by \( R^2 \) but the energy density will only
decrease by \( 1/R \), which is consistent with stretch being proportional to the energy density.
If the energy \( E \) stored by a particle increases by \( R^2 \) then the total flux through a surface
enclosing the stored energy will increase by \( R^2 \). However, if the scale of space decreases by
\( 1/R \), then the flux per unit area will remain constant.

The potential will still give rise to Poisson’s equation of Newtonian gravity, i.e. \( \nabla^2 \Phi = 4\pi G \rho \)
but \( \rho \) must be interpreted as the energy density \( E/c^2 \) with \( G \) and \( c \) as variables. The left
hand side of the equation then corresponds to the rate of change of total potential energy
for a small local change in energy. This will vary with the total potential so that the left hand
side of the equation will reflect the additional stretch produced by a given amount of stored
energy relative to the stretch produced by the total stored energy. Hence, if the background
stretch varies then \( \Delta R / R = 4\pi G / c^2 \).

Original version submitted 14th August, 2013. Revision 1 submitted 9th September 2013
Differences from standard GRT predictions

A. The avoidance of singularities and gravitational redshifts between inertial frames

The revised formulation relates the distortion in the dimensions of space to the amount of matter relative to the background density of matter. In the limit of a very high density of matter the gravitational mass and acceleration will tend to zero, which has the attraction, along with other variable mass theories [22], of removing the singularity at the centre of a black hole. In fact, it also removes the blackness of black holes because photons do not lose energy in a gravitational field. The wavelength is unchanged but time slows so that the number of photons per unit time tends to zero, but light can still escape. The frequency of light and passage of time will be slowed enormously but will not be stopped.

It is not being claimed that gravitational collapse does not occur, for example, when a neutron star exceeds a certain mass. However, at extreme densities gravitational interactions will tend to zero so the object will not collapse to a singularity. The nature and properties of the substance at such extreme energy and pressures is unclear. The observed wavelength of light should be shifted strongly to the red under GRT but not under the revised theory. There does not appear to be any evidence for gravitational redshifts in wavelength between inertial frames. All observed gravitational redshifts appear to be between frames in which at least one frame is subject to a force (gravitational acceleration) [23,24] where the theories give similar predictions for small changes in potential.

The cosmological singularity in the early universe is also avoided as mass tends to zero at high densities. This possibility for variable mass theories has been pointed out previously [25] and, more recently, a formulation based on GRT but with rest mass decreasing with gravitational potential has been used to suggest the non-existence of black holes as light can always escape [26].

B. No need for dark energy and inflation

The most obvious change in matter distribution with time is the expansion of the universe. As a result the speed of light will have slowed, the scale of space will have decreased and time will have increased as the universe has expanded. Currently, the observed redshift of the wavelength $\lambda$ of light from distant galaxies is interpreted as the apparent increase in dimension, or scale of space, of the universe between when the light was emitted $R$ and received $R_0$. This leads to the prediction that $R_0 / R = 1 + Z$, where $Z = (\lambda_{rec} - \lambda_{em}) / \lambda_{em}$.

Under the revised theory, changes in the scale of space with distance are invisible in terms of wavelength shifts, or apparent size, because the scale of space is expanding or contracting exactly in tune with the wavelength. Therefore, changes in wavelength must be due to relative motion of the emitter and receiver, or to an apparent change in distance with time due to a change in potential with time. However, changes in time between locations with different total potentials can be observed. If the potential (energy density) is
decreasing with time, then this will lead to a redshift in wavelength as a function of distance while the difference in clock rate will also reflect the difference in static potential.

If the total potential has decreased over the time, between emission and reception, by $R / R_0$, where $R$ is the scale of space in the past and $R_0$ is the scale now (and $Z = (R - R_0) / R_0$), then the speed of light will have been higher, $c = c_0 \sqrt{1 + Z}$, so the frequency for a given amount of radiated energy was $1 / \sqrt{1 + Z}$ different at the time of emission. This shift in frequency means that the emitted energy will be “cooled” by this factor and so the apparent energy will be reduced by $1 / \sqrt{1 + Z}$. The luminosity (energy per unit time) of a given star in a region of higher potential will be observed to be lower by $R / R_0$ because time was running slower and the frequency gets shifted lower. However, for a star to have the same temperature in a different region of space it would need to have $R_0 / R$ times the number of atoms, so it is not an identical star.

For type 1a supernovae, the total amount of energy released appears to be approximately constant. It has been found that the brighter supernovae wax and wane more slowly than fainter ones. However, when time scales of individual light curves are stretched to fit the norm and the brightness is scaled according to the stretch, then most light curves match. This can be explained under the revised theory. Most of the light emitted is due to the decay of radioactive nuclei synthesised in the explosion. The brightness is therefore dependent on the number of particles rather than the mass and the number of particles will have a dependence on the background potential in the region of the supernova (although the total potential is probably dominated by the matter in the white dwarf). However, the rate of decay of the light curve will reflect the difference in apparent clock rate due to the difference in potential between us and where the light was emitted. The brightness will then reflect the actual distance in terms of number of wavelengths (the variation in wavelength with potential hides any variation in the stationary scale of the space traversed by the light), but will be reduced by $1 / \sqrt{1 + Z}$ if the scale of space is decreasing with time.

This can be plotted against $Z = \Delta \lambda / \lambda = \Delta R / R$. The Union 2.1 data [27] for type 1a supernovae in terms of the distance calculated from the luminosity versus the redshift is given in Fig. 1. Firstly, the uncorrected distance is plotted (Fig. 1(a)) then the distance corrected for the decrease in scale of space with time (Fig. 1(b)). The linear fits, which take into account the errors, show that most of the change in slope has been removed. The $R^2$ value of the fits (not constrained to go through the origin) increases from 0.979 to 0.990.

There is now almost no evidence for a change in the rate of expansion (rate of change of energy density), so there is no need to propose dark energy. The fractional change in the expansion with distance is nearly constant (independent of distance). This would seem to be a consequence of an initially rapid expansion asymptotically approaching a steady value.
Fig. 1 Union data [27] for a) uncorrected, and b) corrected luminosity distance (Megaparsecs) of Type 1a supernovae versus redshift. (Linear fits shown by dashed lines.)
Normally, the supernovae data are plotted assuming a linear velocity-distance law. Such a law applies quite generally in expanding and isotropic models under GRT [28]. In this case, “... spatial homogeneity and isotropy imply a preferred (universal) space, and the time invariance of homogeneity and isotropy implies a preferred (cosmic) time. In the comoving frame, space is isotropic, receding bodies are at rest, and peculiar velocities have absolute values” [28]. Thus the linear velocity-distance law is based on the assumptions of GRT including constant $c$, and can lead to recession velocities that exceed the speed of light. The redshift-distance law, however, appears to apply to a homogeneous universe that expands and contracts to match the local energy density between regions.

WMAP data also provide information about dark matter and energy independent of supernovae results. The data are consistent with a flat universe to better than 1% [29]. A fit to the data using the ΛCDM model (based on GRT) then gives the percentage of baryonic matter as 4.56%, cold dark matter as 22.7%, and the rest as dark energy. The WMAP data suggests that the universe is spatially flat so the value of the Hubble parameter in the conventional formulation of general relativity has been used to determine that it has a critical density of $9.30 \times 10^{-27}$ kg/m$^3$, using $\rho_c = 3H_0^2 / 8\pi G$ [30]. However, under the revised theory the universe is necessarily flat and the apparent rate of expansion ($H_0$) appears to have a fixed relationship with the energy density.

A fit to the data of Figure 1b gives the current fractional rate of expansion $\Delta R / R = 7.18 \times 10^{-27}$. This is possibly related to the current strength of the gravitational field and speed of light, i.e. $\Delta R / R = 4\pi G / c^2$ or $9.33 \times 10^{-27}$. The similarity of the values is encouraging but may be a coincidence and needs further investigation. The current distance scale of the universe would have to be overestimated by about 23% to bring the numbers into line.

The proposed phenomenon of inflation, the extremely rapid exponential expansion of the early universe, was hypothesized to explain the uniformity of the cosmic microwave background radiation and flatness of space. This was necessary to explain why widely spaced regions could have previously been causally connected and in thermal equilibrium. Under the revised theory, gravity is weaker, time is slower and the speed of light faster, as you go back in time to when the universe was more dense. Therefore, the amount of matter which was causally connected increases in the past and the time available for it to occur also increases.

The predictions from baryon acoustic oscillations [31], periodic fluctuations in the density of the visible baryonic matter of the universe, need to be re-examined in light of the revised theory. However, since length scale and wavelength expand synchronously with gravitational potential, it seems likely that the agreement with observations of such a standard length ruler will be unaffected.

**C. No need for dark matter**

Original version submitted 14th August, 2013. Revision 1 submitted 9th September 2013
The initial evidence for dark matter was from the rotation curves of galaxies and then from gravitational lensing [3]. Calculations of the expected effects assume that the amount of gravitational acceleration and curvature are the same as locally observed, and together with the clock rate are independent of the amount of other matter present. None of these assumptions is correct. The matter to luminosity ratio should also increase as the amount of other matter present increases because the clock rate will decrease. The effects appear to be of the correct form to explain the anomalous rotation curves of galaxies and the amount of gravitational lensing of clusters of galaxies without the need to hypothesize dark matter halos.

According to Newtonian mechanics the velocity of rotation should decrease as \(1/\sqrt{r}\) in moving away from a large central concentration of matter. However, the measured velocity is calculated assuming that the distance and time scales are independent of the gravitational potential. The apparent velocity will then be the actual velocity multiplied by \(\sqrt{\Phi_B/(\Phi_B + a/r)}\), or approximately \(\sqrt{r}\) if the background potential is small. The net effect is that the velocity curve will appear to remain approximately constant until the potential approaches the background level. (Such large changes in gravitational potential should lead to large redshifts in the wavelength of light under GRT but none are predicted under the revised theory.)

It is observed that the rotational velocities of the stars and gas in spiral galaxies are constant out to great distances [32]. It is also observed that galaxies with a uniform distribution of luminous matter have a rotation curve which slopes up to the edge. These observations are consistent with the above explanation provided the background potential, away from galaxies, is small. In general, the dependence of \(G\) on the inverse of the total potential means that for all concentrations of matter a dark matter halo will need to be inferred to explain the rotation of galaxies or the amount of gravitational lensing under the assumption of constant \(G\). This might then explain the scaling relation between the size and density of galactic halos [33].

The distribution of radiation, visible matter and mass as deduced from gravitational lensing in colliding galaxy clusters has been used as strong evidence for the existence of dark matter [34,35]. This needs re-examination not only because the revised theory will yield a different determination of matter distribution but also because, in the revised theory, the number of particles for a given mass decreases with decreasing potential and so the estimates of the density of interacting gas may need revision. It appears that much more of the matter may reside in the galaxy cores than the gas, as indicated by the data.

Additional evidence of the need for dark matter has been that the estimated baryon density is insufficient to account for the observed elemental abundances of Big Bang nucleosynthesis and WMAP data [3]. However, both these calculations change as the masses and hence interactions are dependent on the background energy density.
D. Consistency with quantum mechanics

The revised theory removes the apparent inconsistency between quantum mechanics and general relativity at very small length scales. The inconsistency arises because the uncertainty principle indicates that, at very short distances and times, the available energy can be very large. Under GRT this energy corresponds to a large mass which induces such a severe warping of space that a “quantum foam” results [36]. However, if only stored energy corresponds to mass and if the stored energy increases then the dimensions of space also increase and mass and time decrease and the infinities do not occur. It needs to be show that the proposed scalar theory with energy density as the source of the distortion of space-time leads to Newton’s equation for gravitation in the limit of low velocities and is consistent with the Schrödinger equation and the Compton wavelength of massive objects.

E. Consistency with Mach’s principle

If there is no shell of surrounding matter, and no uniform background density of matter, then the distance scale goes to zero except near the only massive object(s). The distortion of space becomes enormous but only extends a tiny distance from the object. A system having the same angular momentum will rotate faster in proportion to $1/\sqrt{\Phi}$. Therefore, for the same speed of rotation, the angular momentum will tend to zero as the background potential tends to zero. So the energy and momentum for a given rotation speed of any system tends to zero as the amount of background matter tends to zero.

New Predictions

A. Supernovae light curves

An observable consequence of such significant changes in the total potential (and therefore matter to luminosity ratio) in the region of galaxies is in the light curve of type 1a supernovae (at all $Z$). The collapse from a white dwarf is understood to occur when the gravitational pressure exceeds the resistance from electron degeneracy. The total luminosity (energy released), should be determined by the available gravitational energy and so should be approximately constant, whereas the apparent clock rate (time) will be slower when the density of matter is higher. As pointed out earlier the number of particles and therefore brightness, due to the decay of radioactive species, should increase in proportion to the background potential and this is a potential explanation of why supernovae light curves can fit a distribution that normalizes the brightness according to the rate of decay [37,38]. The rate of decay should be proportional to the square root of the gravitational potential in the local environment of the supernova, with slower decaying supernovae lying closer to the centre of larger galaxies. Some evidence has been seen for the decay rate being correlated with the morphology of the galaxy [38], but it is not known if the decay rate has been plotted against an estimate of the local matter density (potential).
The decay rate should increase in proportion to $\sqrt{r}$ within the region of a galaxy for which the light curve is flat.

Cepheid variables might also allow a mapping of the gravitational potential. Their luminosity is closely related to their period of oscillation, with longer periods corresponding to higher luminosity. A larger gravitational potential will lead to a slower passage of time and so to a lower luminosity and increased period. However, the number of particles per unit of mass will also change and this prediction needs careful examination. The difference between the observed luminosity and the expected luminosity for the observed period should depend on the local gravitational potential. The effect should show up as a shift in the luminosity-distance relationship to lower brightness for Cepheids closer to their galactic centre but may be difficult to distinguish from effects due to increasing dust (reddening) and metallicity.

B. Change in distance scale due to the expansion of the universe

If the universe is expanding then the density of matter is decreasing and so the scale of space should be decreasing. The number of wavelengths to a distant stationary object will not change but the speed of light will slow so the elapsed time will increase and not as many wavefronts will return in a given time, so the apparent frequency (of a signal sent to the distant object and back) will decrease.

The rate of contraction of the distance scale can be calculated from a fit to Fig. 1(b). A linear least squares fit (weighted by the quoted error on each point) indicates that the distance travelled by light (at the current speed) in the time (in current units of time) that the dimensions of the universe have doubled ($Z=1$) was 4702 Megaparsecs. So the current fractional rate of expansion should be $c/(4702$ Megaparsecs) or $2.74\times10^{-18}$ per second. Therefore, if an electromagnetic signal is used to measure distance by a signal received and sent back at approximately the same frequency (locked to the received frequency) then the observed frequency should be uniformly decreasing by an amount $\Delta f/f$ of $2.74\times10^{-18}$. The corrected supernova data, however, do not exactly follow a linear fit. The low $Z$ values indicate a slightly faster rate of expansion ($c/(4511$ Megaparsecs) or $2.86\times10^{-18}$ per second), so the current rate of change of frequency may be slightly higher.

The Pioneer spacecraft have been observed to have an anomalous deceleration towards the sun of approximately $8\times10^{-10}$ m/s$^2$ based on a steady downwards drift in frequency of approximately $6\times10^{-9}$ Hz/s or $1.51\$Hz in 2.11 GHz in 8 years or $2.84\times10^{-18}\$s/s [39]. A more recent analysis has suggested that the anomalous deceleration has been decreasing with time [40], and that the deceleration can be explained by the selective radiation of heat energy in the direction away from the sun [41]. However, the remarkable closeness of the observed drift in frequency with that predicted here suggests that the change of the scale of space with time is the more probable explanation for most of the constant component.

C. Variation in the value of the gravitational constant
The value of the so-called gravitational constant will vary according to \((\Phi / \Phi')\). Hence, the value of \(G\) will increase by a factor of approximately \(2GM(1/r_i - 1/r_f) / c^2\) in moving away from a massive object. This implies a fractional increase of \(2 \times 10^{-8}\) in moving from the Earth to far away from the Sun. The current fractional error on the value of \(G\) is about \(1 \times 10^{-4}\), so there is little hope of detecting this change.

\[D. \text{ Gravitational wave detection}\]

The revised understanding of the nature of gravitational distortions is that the wavelength of light and the dimensions of space change in unison. A consequence is that the interference of light waves cannot be used to detect a change in path length due to a travelling variation in gravitational potential. Methods, such as the LIGO detectors \([42]\) which examine the interference pattern of light of fixed frequency sent down two paths, should not see any change in the relative phase. There would be shifts in frequency of the light emitted at differences in time, but the interference is between light which has been generated at the same time (if the two paths are equal in length) and averaged over a time interval corresponding to the total path length of the multiple passages in each arm.

\[\text{Discussion}\]

The apparent inconsistency between GRT and Mach’s principle can be traced to the assumption that the outcome of any local non-gravitational experiment is independent of where and when in the universe it is performed. The weak equivalence principle is the equivalence of inertial and gravitational mass so that the trajectory of any freely falling body does not depend on its internal structure, mass or composition \([43]\). The so-called Einstein equivalence principle states that the weak equivalence principle holds and, in addition, that: the outcome of any local non-gravitational experiment in a freely falling laboratory is independent of the velocity of the laboratory and its location in space-time (local Lorentz invariance), and independent of where and when in the universe it is performed (local position invariance). Local position invariance requires or assumes that the background distribution of matter has no effect (in an inertial frame). This has been supported by the argument that a null-geodesic can be used to travel from one point in space-time to another with a different background without any observable consequences. The argument, however, breaks down because it inherently assumes that the sum of infinitesimal changes, from changes in the distribution of external matter, tends to zero. That is equivalent to assuming that the initial and final distributions of external matter are the same. However, if the associated distortions do not cancel, then differences in the mean density of background matter will affect the outcome of experiments that involve mass and the properties of space (i.e. its space-time dimensions). Thus the proposed theory is consistent with the weak equivalence principle but not the stronger Einstein equivalence principle which includes positional invariance. The underlying equivalence principle, that experiments give the same
result independent of location, velocity and acceleration, requires that the distortion of space is correctly taken into account.

The revised theory does not have an absolute space-time but produces the same predictions as an invariant metric if the background distribution of matter is fixed. It therefore reproduces the standard results of general relativity, but is consistent with Mach’s principle. The theory is truly relative, whereas GRT has a space-time of fixed magnitude. Dimensions, time, the speed of light, wavelength and mass are all relative to the amount of other matter present and so the revised theory obeys Mach’s principle. The presence of matter gives space dimensions but the increase in dimensions for a given increase in matter decreases as the density of matter increases. Time lengthens and the speed of light increases with density. A consequence is that the energy of an elementary particle, a stationary state, will decrease as the gravitational potential (density of matter) increases. Therefore, the stored energy decreases as objects move closer together which provides an explanation for gravitational attraction.

The revised theory has general covariance with energy and normalized momentum conserved. It reproduces all the standard predictions of a constant metric, because it gives the same relationship between distance, speed of light and time, provided the change in potential is much smaller than the total potential. The discrepancies between GRT and observations, i.e. dark matter and dark energy, and the inconsistency with quantum mechanics and presence of singularities, all occur where the total potential is significantly different. It appears that all the discrepancies can be avoided by the revised theory. However, this needs careful examination and the further implications, such as for the predicted evolution of galaxies, also need investigation.

The Sachs-Wolfe effect [44], which gives variations in the wavelength of the cosmic microwave background based on the perturbed Einstein Field Equations, cannot be attributed to losses or gains in energy of photons as they move through gravitational potentials. Such changes in energy do not occur. Instead the variations in wavelength must reflect different amounts of expansion due either to different times of emission or to changes in the rate of expansion of regions of different gravitational potential.

The observation of gravitational radiation from binary neutron stars moving at relativistic speeds would also seem to be in need of review. However, it is likely that the revised theory will generate equivalent behaviour to GRT. A change in potential can only propagate at the speed of light and so the local potential will not reflect the position of a source mass when the change in potential influences the receiving mass. Thus the change in direction of motion will not be along the line to the source mass when it is observed there. This is analogous to a distortion of space not centred on the source masses and in that sense can be seen as a gravitational wave.
If a photon has energy but does not have any net energy stored as mass but can split into a particle/antiparticle pair, then it can be concluded that particles and antiparticles have the opposite effect on the dimensions of space. Hence, antiparticles must induce a contraction in spatial dimensions in the local region of the universe. Antiparticles will still not be able to store as much energy when the background potential increases and so will be gravitationally attracted to a concentration of particles. However, symmetry suggests that in a region where antiparticles dominate it will be antiparticles that induce an expansion and particles will induce a contraction. Particles near a boundary between regions will be attracted to the region where the like kind of particle dominates. This appears to be a potential explanation of where the initially equal quantity of antimatter may have gone.

Quantum electrodynamics is a gauge invariant theory. An arbitrary constant value can be added to the potential without affecting the results. However, this is not so for this revised theory of gravitation. It is not gauge invariant and the masses of all particles decreases with increasing background potential and so with moving back in time when the density of the universe was larger. This can be compared with the Higgs mechanism, which corresponds to a spontaneous breaking of gauge invariance, and the Higgs boson or Higgs field which interacts with all particles to give them mass. The mechanism and the revised theory of gravitation must be related. In the limit that the background potential is enormous, a constant small potential can be added without affecting the results. So gauge invariance is approached as the masses of particles tend to zero. It would appear that the Higgs field that gives particles mass must be equated with the field due to all other matter in the universe. However, this field acts somewhat differently as the amount of stored energy (mass) per particle is reduced as the density of stored energy (the field) increases.

Summary and Conclusions

A revised theory of gravitation in terms of a stretching of space, and the product $cT$, dependent on the background density of matter is proposed. The underlying difference from most conventional theories of gravitation is that the distortions of space from opposite directions add rather than cancel. The key consequence is that the magnitude of gravitational interactions is no longer independent of any surrounding shell of matter. Gravitational interactions are therefore not gauge invariant. The theory also has the speed of light and mass being dependent on the background potential but energy and normalized momentum are conserved. It is proposed that particles cannot store as much energy when the density of other matter increases. The resultant release of energy gives rise to the kinetic energy of gravitational acceleration. The theory leads to a revised concept of time and makes it clear that photons do not lose energy in escaping a gravitational field. The necessary consequence is that the exchange “quanta” of the gravitational field have zero spin and the requirement is for a scalar field theory, as opposed to GRT which is a tensor theory where the graviton has spin 2. The proposed theory is shown to reproduce the standard predictions of a curved space-time with an invariant metric for the local (constant)
background potential. However, it removes the inconsistency of current general relativity theory with quantum mechanics, removes gravitational singularities and appears to avoid the need to hypothesize dark matter or dark energy. The theory appears to be fully relative in that it is consistent with Mach’s principle. A search for correlations between the light curves of supernovae and position relative to galactic centres may help confirm the revised theory.

Acknowledgements

The author would like to thank Xiujuan (Jane) Dai for her support and encouragement, A. G. (Tony) Klein for pointing to the lack of need for quantisation and Paul Lasky for his scepticism and criticism which helped in improving the arguments.

References


Original version submitted 14th August, 2013. Revision 1 submitted 9th September 2013


