Protogravity: a quantum-theoretic precursor to gravity

Daniel Shanahan
Mariner’s Reach, Newstead, Queensland 4006, Australia
March 9, 2023

Abstract
As a consequence of the twin effect - the slower aging of the twin who makes a return trip compared with that of her brother who stays at home - an orbiting mass has a reduced proper time and thus a binding energy. Referring to this binding effect as protogravity, I argue that it was sufficient in itself to explain the preference of matter for bound rather than free motion in the early universe. I then show, from a consideration of the constraints imposed by laws of conservation and the principle of relativity, that this protogravity is also able to explain the important Schwarzschild metric, and thus the effects of gravity so far as these can be determined with accuracy from the Earth. I argue therefore that gravity need not involve a constraining geometry or any other extraneous force or effect, but is adequately explained as an emergent consequence of the manner in which elementary particles adapt to a change of inertial frame. Because the twin effect may be explained from the evolution of phase described by the de Broglie wave, this interpretation of gravity would provide gravity and quantum mechanics with a common origin in the wave-like nature of the elementary particles.

Keywords
emergent gravity · the twin effect · Schwarzschild metric · relativity of simultaneity · de Broglie wave · principle of relativity

It is said that more than 200 theories of gravity have been put forward ...
Sir Arthur Eddington, writing in 1920 [1], p. 64

1 Introduction
I explore the possibility that gravity has its origin with quantum mechanics in the manner in which the wave structure of an elementary particle must adapt to a change of inertial frame of reference.
A particle experiences a range of changes as it changes inertial frame. For a massive particle, these are the changes in length, time and simultaneity described by the Lorentz transformation and, at what will be regarded here as the more fundamental level, the corresponding variations in wavelength, frequency and phase defined for quantum mechanics by the de Broglie wave.

Whether deduced from the Lorentz transformation or from the de Broglie wave, these changes are the source of what has been called “the twin effect” — the slower aging of the twin who makes a return trip compared with that of her brother who stays at home. To a observer who is stationary with respect to an orbiting particle and is also beyond the reach of gravity (the notional observer at infinity), the particle has, from this dilation of time\(^1\), a reduced frequency \(\omega_E\) (its Einstein frequency) and, from the Planck-Einstein relation,

\[
E = h\omega_E, \tag{1}
\]

a correspondingly reduced energy \(E\) (where \(h\) is Planck’s constant).

This loss of proper time is empirically well-established. It was demonstrated by the Earth-circling clocks of the Hafele-Keating experiment [2] and is observed in the enhanced half lives of cosmic rays and accelerated muons [3]. It is evidenced every minute of every day, and to a high degree of accuracy, by the atomic clocks of global positioning systems\(^2\).

It follows from the twin effect that an orbiting particle, and consequently any orbiting object, has a binding energy, and it is this binding effect that I have referred to above as protogravity. If there were no other form of gravitational attraction in the universe, a system of mutually orbiting objects would have, from their orbital motion alone, a binding energy sufficient to hold those objects to their paths.

As thus described, this protogravity is not yet the gravity that is actually experienced. A stationary object also feels the effects of gravity, while a moving object experiences a dilation of time, not only from its movement relative to a gravitating mass, but from its proximity to that mass. Nor is the binding force due to the twin effect usually thought of as a form of gravity. For instance, in discussions of global positioning systems, a distinction is drawn between the dilation of time due to the orbital motion of the satellite, which is taken to be a consequence of special relativity, and the dilation due to the depth of the orbit within the gravitational influence of the Earth, as for example in Ashby [4].

Yet this protogravity contributes a binding force that would have been sufficient in itself to explain the tendency of matter to favour bound rather than

\(^1\)Taking a cosmic stance, time is lost. But taking the view that with less time expended, there is more remaining (for life or half-life), the commoner usage is that time becomes dilated (expands).

\(^2\)There are over a hundred GPS satellites orbiting the Earth. The United States, Russia, China and the European Union maintain international systems, while Japan and India have local systems.
unbound motion in the early universe. And as I will show, from a consideration of the further constraints imposed by the principle of relativity and the conservation of energy and angular momentum, this binding force is able to explain the important Schwarzschild metric, and in so doing, the effects of gravity so far as these can be ascertained with reasonable certainty from the Earth.

The twin effect will provide the underlying mechanism of gravitational attraction, while the principle of relativity and laws of conservation will determine the relative strength of this mechanism from one situation to another.

The dilation of time experienced by an object that is stationary with respect to a gravitating mass will be explained by the loss of momentum and thus energy that such an object experiences when it is brought to rest after falling from infinity, a loss which in accordance with the Planck-Einstein relation (Eqn. (1) above) is accompanied by a loss of frequency and thus of time. The further dilation experienced by an object that is moving within the influence of gravity will then follow from the stipulation, pursuant to the principle of relativity, that moving and stationary particle have the same interactions and dynamic relationships within that influence as they do when beyond it.

Why should this protogravity be deemed, as advertised above, a quantum-theoretic precursor of actual gravity? When the twin effect is explained (see Sect. 2 below) from the loss of phase in the direction of travel defined by the de Broglie wave, it acquires a common origin with quantum mechanics. It was de Broglie’s prediction of this “matter wave” in his famous thesis of 1923 that allowed a quantum mechanics in which massive particles are treated in terms of evolving wave characteristics. The Schrödinger equation and other equations of quantum mechanics for massive particles, including the Dirac and Pauli equations, were originally conceived as equations for the de Broglie wave, see Bloch [5] and Dirac [6]. Were it not for the de Broglie wave, there would not be a quantum mechanics, not at least a quantum mechanics for massive particles.

Even so, the explanation of gravity that I will offer in this paper could be presented with no mention at all of quantum theory or wave mechanics beyond the reference to the Planck-Einstein relation already made above. And this is the way in which I will initially present the argument. But this interpretation of gravity was initially conceived from a consideration of wave structure. Moreover, it will become apparent as the argument proceeds that in this interpretation, gravity is not strictly speaking a fundamental effect, but emergent from other laws of Nature, namely the Lorentz transformation, laws of conservation, the principle of relativity and the Planck-Einstein relation. These laws and principles are well-established, but are ultimately empirical and thus brute and unexplained, as also, I suggest, is the mysteriously superluminal de Broglie wave. In the concluding sections of the paper, I will endeavour to show that some at least of this miscellaneous collection of laws and phenomena have a common origin with gravity and quantum mechanics in the underlying wave structure of matter and radiation.
In this paper, I will not take gravity beyond the Schwarzschild metric. It seems unlikely that this “bootstrap” approach to gravity, in which orbital motion is self sustaining, would replicate general relativity in all possible situations. And that perhaps is reason in itself for pursuing this proposal. It is in relation to orbital motion, notably in its failure to explain galactic rotation curves and the origin of the angular momenta of the galaxies, that Einstein’s theory seems to require further investigation.

The twin effect is counterintuitive and as a preliminary step I will discuss in the next section how this dilation of time is related to the relativity of simultaneity.

2 The twin effect

The twin effect will be derived in three ways, these being in the historical order of their origination: (a) from the Lorentz transformation; (b) in Minkowski spacetime; and (c) as a consequence of the dephasing described by the de Broglie wave.

What is counterintuitive here is that an object should have an energy that is less when it is moving than when it is stationary. And of course, to an observer that a massive particle is passing, that particle does have an increased frequency and correspondingly increased energy,

$$E = \gamma E_o = \gamma \hbar \omega_o,$$

where $E_o$ and $\omega_o$ are, respectively, the energy and frequency of the particle in its rest frame, while $\gamma$ (in units in which $c = 1$) is the Lorentz factor,

$$\gamma = \left(1 - v^2\right)^{-1/2},$$

and it is this enhanced energy and its associated momentum that a moving object would bring to a collision with that observer.

But as an orbiting particle continues on its way, it is also experiencing the failure of simultaneity described by the Lorentz transformation, and this accumulates as a slowing of time and consequent decrease in energy per orbit that has (approximately at non-relativistic velocities) twice the magnitude of the increase in energy described by Eqn. (2).

(a) from the Lorentz transformation: These competing effects are combined in the time component,

$$dt' = \gamma (dt - v ds),$$

of the Lorentz transformation, which by the substitution,

$$ds = v dt,$$
describes a reduced proper time,
\[ d\tau = dt' = (1 - v^2)^{\frac{1}{2}} dt, \]
giving for a complete orbit,
\[ \tau = \int (1 - v^2)^{\frac{1}{2}} dt \]
and thus, for an orbiting particle of rest mass \( m_0 \), a reduced energy,
\[ m_0 P^{-1} \int (1 - v^2)^{\frac{1}{2}} dt, \]
where \( P \) is the orbital period as considered by the notional observer at infinity.

For the purpose of comparison, it will be more convenient to use from this point the energy per unit mass (referred to in astrodynamics as the specific mechanical energy \( \varepsilon \), see Bate et al. [7], at p. 15), which for the twin effect will be designated,
\[ \varepsilon_{\text{twin}} = P^{-1} \int (1 - v^2)^{\frac{1}{2}} dt, \]
where the binding energy (or mass defect), which is expressed in the negative, is,
\[ \Delta \varepsilon_{\text{twin}} = P^{-1} \int (1 - v^2)^{\frac{1}{2}} dt - 1. \]

**b) in Minkowski spacetime:** In a spacetime diagram (Minkowski [8]), the proper time \( \tau \) of a test particle following a time-like trajectory may be written,
\[ \tau = \int (ds^2)^{\frac{1}{2}} = \int (g_{uv} dx^u dx^v)^{\frac{1}{2}}, \]
where \( g_{uv} \) is the relevant metric, and \( u \) and \( v \) signify four-coordinates.

For the Minkowski metric, \( g_{uv} \) is the metric tensor,
\[ \eta_{uv} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \]
so that Eqn. (6) becomes, in differential form, the invariant interval,
\[ -d\tau^2 = ds^2 = -dt^2 + dx^2 + dy^2 + dz^2. \]

In the spacetime diagram of Fig. 1, a particle follows a curved (and thus accelerated) path between events \( A \) and \( B \). Because proper time is a scalar
Figure 1: The proper time experienced by an orbiting particle is less than that for a stationary particle.

invariant and thus has a magnitude on which all observers must agree, the axes of the diagram can be chosen (as they have been in the drawing) so that A and B lie conveniently on the t axis, which is thus the world line of an observer who in this frame or reference is stationary at \( x = 0 \).

In its own co-moving inertial frame, the particle is stationary. Thus, for any infinitesimal interval \( ds \) along its worldline,

\[
dx = dy = dz = 0,
\]

and Eqn. (7) reduces to,

\[
d\tau = dt.
\]

But to the observer at \( x = 0 \), for whom the particle is moving at the velocity \( v(t) \), Eqn. (7) becomes, after division by \( dt \),

\[
\frac{d\tau}{dt} = (1 - v^2)^{\frac{1}{2}} ,
\]

whereupon for the complete worldline of the orbit (along the curved path between A and B), we again have Eqn. (4), that is,

\[
\tau = \int (1 - v^2)^{\frac{1}{2}} dt
\]
as was deduced from the Lorentz transformation, and which as we have seen leads to the binding energy \( \Delta\varepsilon_{\text{twin}} \) of Eqn. (5).
(c) from the de Broglie wave: A massive particle has from the Planck-Einstein relation (Eqn. (1) above), an associated frequency (the Einstein frequency $\omega_E$), and from the de Broglie relation,

$$p = h\kappa_{dB},$$

a wave number $\kappa_{dB}$ (the de Broglie wave number), where $E$ and $p$ are, respectively, the energy and the momentum of the moving particle.

Frequency $\omega_E$ and wave number $\kappa_{dB}$ define for the moving particle, its de Broglie wave,

$$\psi_{dB} = e^{i(\omega_E t - \kappa_{dB} r)},$$

from which the evolution of phase per orbit is,

$$\varphi_{orbit} = \int [\omega_E dt - \kappa_{dB} ds],$$

which on making the substitution,

$$ds = v dt,$$

and using relations (1) and (9) becomes,

$$\varphi_{orbit} = \omega_0 \int (1 - v^2)^{\frac{1}{2}} dt,$$

so that the energy per unit mass is,

$$\varepsilon_{twin} = \frac{\omega_0 \int (1 - v^2)^{\frac{1}{2}} dt}{\omega_0 P} = P^{-1} \int (1 - v^2)^{\frac{1}{2}} dt,$$

giving for the binding energy per unit mass,

$$\Delta \varepsilon_{twin} = P^{-1} \int (1 - v^2)^{\frac{1}{2}} dt - 1,$$

which is the result obtained above as Eqn. (5) from the Lorentz transformation, but derived now from the evolving wave characteristics of the particle.

Thus all three derivations of the twin effect lead to the same dilation of time and binding energy $\Delta \varepsilon_{twin}$.

3 Newtonian gravity

As a further step toward the Schwarzschild metric, I will establish in this section that for the closed elliptical orbits of Newtonian gravity, binding energies derived from the twin effect correspond exactly with those deduced from the centripetal force supposed by Newton’s universal law of gravitation.

From observations recorded by Tycho Brahe, Kepler had deduced that the planets move in accordance with three laws:
1. The orbit of a planet is an ellipse with the sun at one focus;
2. A line drawn from a planet to the sun sweeps out equal areas in equal times; and
3. The square of the period of a planet is proportional to the cube of its mean distance from the sun.

Newton then showed in the *Principia* [9] that if angular momentum is conserved in accordance with the second of Kepler’s laws (the law of areas), the elliptical paths described by the first of those laws are consistent with the existence of a centripetal force acting directly between massive objects and having a strength varying inversely with the square of the distance between those masses. This is Newton’s universal law of gravitation,

\[ F = G \frac{m_1 m_2}{r^2}, \]

where \( G \) is the universal gravitational constant, \( m_1 \) and \( m_2 \) are the masses, and \( r \) is the distance between them.\(^3\)

In the *Principia*, Newton offered no explanation for this attractive force. He famously declared “hypotheses non fingo”, as in the recent translation by Cohen and Whitman [9], at p. 276:

I have not as yet been able to deduce from phenomena the reason for these properties of gravity, and I do not feign hypotheses. For whatever is not deduced from the phenomena must be called a hypothesis; and hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy. In this experimental philosophy, propositions are deduced from the phenomena and are made general by induction. The impenetrability, mobility, and impetus of bodies, and the laws of motion and the law of gravity have been found by this method. And it is enough that gravity really exists and acts according to the laws that we have set forth and is sufficient to explain all the motions of the heavenly bodies and of our sea.

In other writings, Newton described the notion that a force could act at a distance “without the mediation of any thing else” as “so great an absurdity that .... no man who has in philosophical matters a competent faculty of thinking can ever fall into it” [10]. And outside the *Principia*, he did consider possible explanations for this force, though seems ultimately to have seen in the orbits of the planets, the guiding hand of a divine providence.

\(^3\)To glimpse the enormity of Newton’s achievement, it is necessary to visit the Principia, and to understand that to pursue his many proofs and theorems, he had first to invent the calculus as well as the notion of mass as a measure of substance [9], Def. 1. And all this at a time when it was still possible for the rival theory of Descartes to explain the motions of the planets by circulating fluxes of a mysterious fluid.
It is a simple matter to show that the binding energies predicted by Newton’s central force correspond exactly with those obtained from the twin effect, that is,

\[ \Delta \varepsilon_{\text{Newton}} = \Delta \varepsilon_{\text{twin}}. \]

I will begin with a circular orbit, and then consider the more general case of an elliptical orbit.

From Eqn. (5) of the preceding section, the binding energy per unit mass due to the twin effect is,

\[ \Delta \varepsilon_{\text{twin}} = P^{-1} \int (1 - v^2)^{1/2} \, dt - 1, \]

which in the Newtonian approximation where \( v << c \) becomes,

\[ \Delta \varepsilon_{\text{twin}} = -P^{-1} \int \frac{1}{2} v^2 \, dt. \] (10)

For a circular orbit, \( v \) is constant, so that,

\[ \Delta \varepsilon_{\text{twin}} = -\frac{1}{2} v^2. \] (11)

Turning now to Newtonian gravity, the binding energy per unit mass, \( \Delta \varepsilon_{\text{Newton}} \), is the sum of the object’s kinetic and potential energies, the latter being taken to be zero at infinity. Thus,

\[ \Delta \varepsilon_{\text{Newton}} = T + V = \frac{1}{2} v^2 - \frac{GM}{h}, \] (12)

where \( h \) is the distance of the unit mass from the centre of the central mass, \( r \) being reserved here for the coordinate distance, as in the Schwarzschild metric, that is to say the distance observed by the notional observer at infinity).

For a circular orbit (Bate et al [7], at p. 34)

\[ v = \left( \frac{GM}{h} \right)^{1/2}, \]

and Eqn. (12) becomes,

\[ \Delta \varepsilon_{\text{Newton}} = \frac{GM}{2h} - \frac{GM}{h} = -\frac{GM}{2h}. \] (13)

and thus for a circular orbit it follows from Eqns. (11) and (13) that as required,

\[ \Delta \varepsilon_{\text{Newton}} = \Delta \varepsilon_{\text{twin}} = -\frac{1}{2} v^2 = -\frac{GM}{h}. \]
For an elliptical orbit (see, for instance, Logsdon [11], at p. 30), the velocity is given by Newton’s *vis-viva* formula,

\[ v = \left[ GM \left( \frac{2}{h} - \frac{1}{a} \right) \right]^{\frac{1}{2}}, \]

where \( a \) is the semi-major axis of the ellipse.

From Eqns. (12) and (14)

\[ \Delta \varepsilon_{\text{Newton}} = \frac{1}{2} v^2 - \frac{GM}{h}, \]

\[ = \frac{GM}{h} - \frac{GM}{2a} - \frac{GM}{h}, \]

\[ = - \frac{GM}{2a}. \]

Thus the binding energy is independent of the eccentricity \( e \) of the ellipse, depending only for a central mass \( M \) on the magnitude of the semi-major axis \( a \), and this is so for the limiting case of a straight line orbit (where the maxima are at \( h = \pm 2a \)), and that of a circular orbit (where \( h = a \)).

For an elliptical orbit, we have from Eqns. (10) and (14),

\[ \Delta \varepsilon_{\text{twin}} = P^{-1} \int \left[ \frac{GM}{h} - \frac{GM}{2a} \right] dt, \]

but it follows from the virial theorem (see Goldstein [12], at p. 85) that for orbits consistent with an inverse square law, the mean value of the kinetic energy of an orbiting object is half that of the mean value of its potential energy, that is,

\[ \langle T \rangle = \frac{1}{2} \langle V \rangle, \]

from which it follows that,

\[ \int \frac{GM}{h} dt = \int \frac{GM}{a} dt, \]

giving again as required,

\[ \Delta \varepsilon_{\text{twin}} = \Delta \varepsilon_{\text{Newton}} = -\frac{GM}{2a}. \]

### 4 The Schwarzschild metric

The Schwarzschild metric is a spherically symmetric and time-independent solution to Einstein’s field equation in a vacuum. It may be written (see, for example, Misner et al [13], at p. 607),

\[ ds^2 = -\left(1 - \frac{2M}{r}\right) c^2 dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2). \]
\[ ds^2 = -dr^2 = - \left( 1 - \frac{2GM}{r} \right) dt^2 + \left( 1 - \frac{2GM}{r} \right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \]  

(15)

where,

- \( ds \) is the invariant interval,
- \( d\tau \) is proper time - the time actually experienced within the metric,
- \( dt \) is coordinate time - the time experienced by a notional observer at infinity,
- \( r, \theta \) and \( \phi \) are spherical coordinates,
- \( G \) is again the universal gravitational constant, and
- \( M \) is a central mass.

The successes of the Schwarzschild metric include the anomalous precession of the perihelion of Mercury, the gravitational deflection and lensing of light, the redshift of light, and the Shapiro delay, see generally Will [14].

As discussed in Sect. 1, a distinction is commonly drawn in discussions of global positioning systems between time dilations due to special relativity (the twin effect) and those attributed to general relativity. Drawing the same distinction here, the contribution from special relativity can be isolated by ignoring temporarily the central mass \( M \), whereupon the metric (15) becomes,

\[ dr^2 = dt^2 - dr^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\phi^2, \]  

which is the metric of Minkowski spacetime expressed in spherical coordinates.

For an object making a return trip, this metric must induce the twin effect, as can be verified by expressing velocity in those same spherical coordinates, that is,

\[ v_r = \frac{dr}{dt}, \]
\[ v_\theta = r \frac{d\theta}{dt}, \]
\[ v_\phi = r \sin \theta \frac{d\phi}{dt}. \]

and using these expressions to eliminate \( dr, d\theta \) and \( d\phi \) from Eqn. (16), which becomes,

\[ dr^2 = dt^2 \left( 1 - v^2 - v^2 - v^2 \right) = dt^2 \left( 1 - v^2 \right), \]

giving for a complete orbit, as expected, the twin effect,

\[ \tau = \int (1 - v^2)^{1/2} dt. \]  

(17)

\(^4\)When considered from beyond the effects of gravity, these dilations are cumulative. But when considered from the surface of the Earth, the dilation due to gravity is less at the satellites than on the ground. The net effect is that the clocks of the satellites run faster than those on the ground, see Ashby [4].
Eqn. (17) will provide the dilation of time due to the twin effect for any return trip at all, including a meandering (and thus powered) excursion. But for a system of mutually orbiting objects, specifically here a test particle orbiting a much larger central mass, the trajectories of interest are those in which there is no variation in either the energy or the angular momentum of the orbiting masses.

In such a system, orbital motion once induced must endure, and my objective in what follows will be to show that in accommodating the further constraints imposed by conservation and the principle of relativity, the flat space metric of Eqn. (16) becomes the Schwarzschild metric of Eqn. (15). I will consider those constraints in the next section, postponing to Sect. 8, the consideration of why the predictions of the metric should approximate those of an inverse square law.\textsuperscript{5}

5 The metric components $g_{tt}$ and $g_{rr}$

As can be seen by comparing Eqns. (15) and (16), it is only in the tensor components,

\begin{align*}
g_{tt} & = \left( 1 - \frac{2GM}{r} \right), \quad \text{and} \\
g_{rr} & = \left( 1 - \frac{2GM}{r} \right)^{-1},
\end{align*}

(which describe, respectively, a dilation of time and an expansion of radial distance), that the Schwarzschild metric differs from that of Minkowski.

That some such dilation of time should be expected may be inferred by noticing that since an object has a diminished energy in a gravitational potential, it must also have the correspondingly reduced frequency contemplated by the Planck-Einstein relation (1). But before giving further consideration to the value of $g_{tt}$, it will be helpful to discuss the constraints imposed by the principle of relativity on the relationship between $g_{tt}$ and $g_{rr}$.

Notice firstly that if a particle experiences a reduction of frequency as a result of its depth within a gravitational potential, so also in the same degree must every other particle at the same depth, including every photon that is emitted at that depth when a system transitions from one state to another. If that were not so, inter-particle interactions would not be consistently the same at that depth as stipulated by the principle of relativity.\textsuperscript{5}

\textsuperscript{5}Nor do I consider here why $G$ has the value it has, other than to suggest that if gravity is a consequence of the twin effect, $G$ may not be fundamental, but a measure of the degree to which matter is gravitationally bound in the present epoch. In that case, $G$ would be akin to an intensive thermodynamic parameter, its apparently unchanging current value having been determined by the circumstances of the early universe.
Figure 2: The worldlines of successive crests of an electromagnetic wave emitted at $r_A$ and received at $r_B$. The coordinate time $\Delta t$ between departures from $r_A$ must be the same as that between corresponding arrivals at $r_B$. But the proper times will differ.

As frequencies are reduced and time becomes dilated in the same degree for all physical processes, including biophysical and mental processes, time itself runs more slowly within the gravitational field than it does outside it (or at least will seem to do so).

What then of wavelengths? Consider Fig. 2, which is a spacetime diagram of a kind commonly found in discussions of gravitational redshifts, see for instance Zee [15], at p. 304, and Moore [16], at p. 109. The diagram depicts the worldlines of successive crests of an electromagnetic wave emitted at $r_A$ and received at $r_B$. Because these successive paths from $r_A$ to $r_B$ are congruent, the coordinate time $\Delta t$ between departures from $r_A$ must be the same as that between corresponding arrivals at $r_B$. From the standpoint of a notional observer beyond the influence of the gravitating mass (the notional observer at infinity) the wave will thus have the frequency at $r_B$ that it had at $r_A$.

However the corresponding proper times will differ. The time actually experienced by a particle or observer at $r_A$ will be be less than at $r_B$. An incoming photon from a transition at $r_A$ will have a frequency less than that of a photon emitted from the corresponding transition at $r_B$.

Now consider what this implies for wavelengths. If, as assumed by special relativity, the photon from $r_A$ is to be observed at $r_B$ to have the velocity $c$, it follows from the relation,

$$\frac{\omega}{k} = c,$$

$^6$A t-meter, may aid the understanding here, see Moore [16], at p. 108.
that it must have a correspondingly longer wavelength,
\[
\lambda = \frac{2\pi}{\kappa},
\]
than the photon emitted from the same process at \( r_B \), which means in effect that lengths must increase as times become dilated, that is,
\[
g_{rr} = g_{tt}^{-1}.
\] (18)

It is commonly, but rather loosely, said in discussions of the gravitational redshift, (see, for instance Wald [17], at p. 137) that a photon suffers a loss of energy and reduction of frequency as it rises from a gravitating mass (the redshift) and gains energy and increases in frequency as it falls toward the gravitating mass (the blueshift). But as can be seen from the discussion above, the incoming photon is observed to be redshifted or blueshifted, as the case may be, because it was created in a reference frame in which frequencies, energies and wavelengths differ from those in the reference frame in which it is subsequently measured. If the photon were in fact to gain or lose energy, this would be contrary to the law of conservation of energy.

\( g_{tt} \) can now be deduced by considering the energy lost by a test particle that, after falling vertically from infinity, is brought to a stop at a distance \( r \) from the centre of the mass \( M \). This loss must be evaluated in the \( r \) and \( t \) coordinates of the Schwarzschild metric, and I will assume here (and discuss further in Sect. 8) that it is also in these coordinates that the inverse square law applies, this being why in this metric,
\[
V(r) = -\frac{GM}{r},
\]
(see Moore [16], at p. 117). On that assumption, a particle that has fallen from infinity has at \( r \) the velocity,
\[
v(r, t) = \frac{dr}{dt} = \left(\frac{2GM}{r}\right)^\frac{1}{2}.
\] (19)

As the particle falls, and for as long as its fall remains uninterrupted, its energy remains constant, retaining the value, \( E_0 = m_0 \), that it had at infinity. It thus follows from the relativistic equation of motion,
\[
E^2 - p^2 = m^2,
\] (20)
that as the momentum \( p \) of the particle increases, its relativistic (or effective) mass must decrease\(^7\). Let us suppose then that \( m \) decreases with \( r \) in accordance

\(^7\)The notion of a varying relativistic mass seems indispensible here, see generally Petkov [18], Chap. 9.
with an as yet unknown function $f(r)$, so that we can write,

$$
\begin{align*}
    m & \rightarrow m_0 f, \\
    p & \rightarrow \frac{m_0 f v}{(1 - v^2)^{1/2}},
\end{align*}
$$

and the relativistic equation of motion becomes,

$$(m_0)^2 - \frac{(m_0 f v)^2}{1 - v^2} = (m_0 f)^2,$$

whereupon, on solving for $f$, we have,

$$f = (1 - v^2)^{1/2}.$$ 

When brought to a stop at $r$, the particle loses its momentum and is left with the energy,

$$E = (1 - v^2)^{1/2} E_0,$$

so that from Eqn. (19), we have, as required,

$$g_{tt} = \left(\frac{dr}{dt}\right)^2 = (1 - \frac{2GM}{r}),$$

while from Eqn. (18),

$$g_{rr} = (1 - \frac{2GM}{r})^{-1}.$$

One question remains: why should a moving particle experience, both the dilation of time defined for a moving particle by the twin effect, as well as the dilation of time experienced by a stationary particle at the same position? The answer, shortly stated, is that the principle of relativity demands it. The laws of physics must hold in the same manner at $r$ as they do for a particle that is beyond the reach of gravity, which could not be the case if a particle moving at $r$ had a rest mass that differed from that of the stationary particle at the same position.

### 6 The de Broglie wave

The Schwarzschild metric has thus been explained from the twin effect, which as shown in Sect. 2 can itself be explained from the evolution of phase defined by the de Broglie wave. As discussed in Sect. 1, this same evolution of phase provided the basis for quantum mechanics, which raises the possibility that gravity and quantum mechanics have a common origin in wave structure.
But the de Broglie wave is superluminal and not as yet a suitable candidate for reconciling these two theories. What must first be provided is a physically reasonable provenance for the de Broglie wave itself.

I will argue that the de Broglie wave is better understood, not as an independent wave, but as the modulation (or dephasing or beating, see, for instance, Feynman et al [19], Vol. I, Chap. 48), of an underlying wave structure that is itself evolving through space at the subluminal velocity $v$ of the particle. I will base this argument on two empirically well-established laws of physics, the Lorentz transformation and the Planck-Einstein relation (Eqn. (1) above).

Before presenting this argument, I should first explain that this way of understanding the de Broglie wave is not at all novel. The de Broglie wave can be seen to emerge in the manner to be now described in two of the three demonstrations of the de Broglie wave in de Broglie’s own thesis of 1923 [20], one being a treatment in Minkowski spacetime [20], Chap. I, Sect. III, and the other, an intuitively more accessible mechanical model comprising an array of oscillating springs [20], Chap. I, Sect. II. This interpretation of the de Broglie wave was subsequently noticed by Mellen [21], and discussed at length by Wolff [22]. It has now acquired a “literature”, as listed in Shanahan [23], and see also Shanahan [24] to [26].

On the evidence of the Lorentz transformation, a massive particle comprises in its entirety underlying influences evolving at the velocity $c$ of light. If that were not so - if there were some other velocity having the same fundamental significance as $c$ - such a velocity would have its own Lorentz factor $\gamma$ (see Eqn. (3) ), and its own Lorentz transformation based on that factor $\gamma$, and the laws of physics would not then be the same from one inertial frame to the next. While a particle comprising underlying influences of more than one fundamental velocity might be stable in one inertial frame of reference, it could not retain the stability of its characteristic structure in any other inertial frame.

There are, of course, velocities in Nature that differ from $c$, those for instance of massive particles, sound waves, and refracted light. But, as Einstein explained in 1905 [27], such velocities transform in accordance with the relativistic formula for the composition of velocities. To explain the all-encompassing ambit of the Lorentz transformation, these other velocities must be regarded, not as fundamental, but as existentially dependent on $c$, that is to say, as the net effect of underlying influences that do evolve at velocity $c$.

On the evidence of the Planck-Einstein relation (Eqn. (1) ), it should also be assumed that these underlying influences of velocity $c$ are wave-like in nature, and that in its rest frame, a massive particle comprises wave-like influences having the characteristic frequency $\omega_0$ of the species of particle in question. In consequence, the particle would then have a characteristic wave number $\kappa_0$, satisfying the relation,

$$\frac{\omega_0}{\kappa_0} = c,$$

and thus a corresponding wavelength,
\[ \lambda_0 = \frac{2\pi}{\kappa_0}, \]

thereby according physical meaning to the Compton wavelength,

\[ \lambda_c = \frac{2\pi}{k_0} = \frac{\hbar}{mc}. \]

There is a wealth of corroborating evidence for this wave-based understanding of the nature of solid matter, but the item of evidence that seems particularly compelling is (as I will now show) the origin it provides for the de Broglie wave.

If, as argued above, a massive particle comprises underlying wave-like influences of velocity \( c \), it must comprise in its rest frame some form of standing or stationary wave. It is not necessary to consider the details of such a structure for it is easily shown that every standing or stationary waveform gives rise to a de Broglie wave when considered from another inertial frame of reference.

Consider the standing wave,

\[ R(x, y, z) e^{i\omega t}, \] (21)

which is evolving in time at some frequency \( \omega \), but for which no assumption has been made as to its manner of spatial variation. Under a relativistic boost in the \( x \)-direction, this waveform becomes the moving wave,

\[ R(\gamma (x - vt), y, z) e^{i\omega\gamma(t-vx)}, \] (22)

in which the spatial factor \( R(x, y, z) \) of standing wave (21) has become the carrier wave,

\[ R(\gamma (x - vt), y, z), \] (23)

which is evidently moving through space at the velocity \( v \) and, as indicated by the presence of the Lorentz factor \( \gamma \), is exhibiting the contraction of length predicted by special relativity.

The second factor,

\[ e^{i\omega\gamma(t-vx)}, \] (24)

in wave (22) is a transverse plane wave, which (in these units where \( c = 1 \) and \( v < c \)) is evolving through the carrier wave (23) at the superluminal velocity \( v^{-1} \). If the frequency \( \omega \) is now identified as the natural frequency \( \omega_0 \) of a massive particle, wave factor (24) can be rewritten in terms of the Einstein frequency,

\[ \omega_E = \frac{E}{\hbar} \gamma \omega_0, \] (25)

and de Broglie wave number,

\[ \kappa_{dB} = \frac{p}{\hbar} = \gamma \omega_0 v, \]
as,

$$e^{i(\omega_E t - \kappa dB x)},$$

and is now recognizable as the de Broglie wave, no longer an independent wave, but a modulation or dephasing of the underlying carrier wave. From Eqns. (22) and (26), the full composite particle wave structure is then,

$$R(\gamma (x - vt), y, z) e^{i(\omega_E t - \kappa dB x)}.$$  \hfill (27)

In this interpretation, this otherwise anomalous superluminal phenomenon achieves consistency with special relativity. Its superluminal velocity is no longer an embarrassment as the velocity of a modulation is not that of energy or information transport. And unlike the de Broglie wave considered alone, the full modulated wave structure (27) is a manifestly covariant relativistic object, capable in principle of taking its place in the tensor equations of relativistic physics. The Fitzgerald-Lorentz contraction appears in the carrier wave (23), while the dilation of time and failure of simultaneity are described by the modulation, that is to say, by the de Broglie wave (26).

The effect of the modulation is that the various parts of the moving wave are no longer cresting in unison as they had been in the standing wave, but in sequence, those ahead lagging in phase (and thus time) those behind. The de Broglie wave describes a progressive loss of phase in the direction of travel corresponding exactly in effect to the failure of simultaneity in that direction predicted by the Lorentz transformation.

Once the existence of the underlying wave structure is recognized, several mysteries become resolved of which I will mention below only those of relevance to the reconciliation of gravity and quantum mechanics.

7 Quantum mechanics from particle wave structure

If a massive particle were some form of tiny solid object, it would be exceedingly curious that the energy $E$ and momentum $p$ of this object should be associated with the wave characteristics, $\omega$ and $\kappa$ respectively, of a superluminal wave with which it would seem to have no physical nexus. Yet it was essentially on the basis of that association that wave mechanics was originated by Schrödinger and has since developed.

In constructing a wave equation that would have solutions consistent with the Planck-Einstein and de Broglie relations, Eqns. (1) and (9) respectively,
Schrödinger made the substitutions,

\[ p \rightarrow i\hbar \frac{\partial}{\partial x}, \quad \text{and} \]
\[ E \rightarrow i\hbar \frac{\partial}{\partial t}, \]

in the non-relativistic equation of motion,

\[ E^2 = \frac{p^2}{2} + V, \]

to obtain the non-relativistic Schrödinger equation,

\[ i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V \psi, \]

as he had likewise done in the corresponding relativistic equation of motion to obtain the corresponding relativistic wave equation (now called the Klein-Gordon equation).

These equations have owed their utility to the correspondence between, on the one hand, the frequency and wave number of the de Broglie wave, and on the other, the energy and momentum, respectively, of the associated particle. But it is the existence of the underlying carrier wave that makes sense of that correspondence, and it is the full wave (27), rather than the de Broglie wave considered alone, that provides an understanding of the nature of mass, energy, momentum and inertia.

In order to show that this is so, it will be helpful to have before us a model that, unlike the model described by Eqn. (27), displays the underlying velocity \( c \). I will take as a suitable model,

\[ \psi (r, t) = \frac{1}{2} |r|^{-1} [e^{i(\omega_o t - \kappa_o r)} - e^{i(\omega_o t + \kappa_o r)}], \quad \text{(28)} \]

which has the idealized form of a spherical standing wave, centred at \( r = 0 \) and constructed from incoming and outgoing waves of velocity \( c \), where,

\[ \frac{\omega_o}{\kappa_o} = c, \]

\( \kappa_o \) being here the Compton wave number.

Model wave (28) has a singularity at the origin and is itself unphysical, but will illustrate how the dynamic properties of a massive particle might originate from a thoroughly wave-theoretic treatment of matter.

On a boost in the \( x \)-direction, wave (28) becomes (on taking real parts),

\[ \Psi (x, y, z, t) = \sin \kappa_o \sqrt{\gamma^2 (x - vt)^2 + y^2 + z^2} \cos (\omega_E t - \kappa_d B x), \quad \text{(29)} \]

(where an amplitude factor has been omitted).
As in the case of travelling wave (27), wave (29) comprises, as one factor, a carrier wave of velocity $v$,
\[ \sin \kappa_0 \sqrt{\gamma^2(x - vt)^2 + y^2 + z^2}, \quad (30) \]
and as modulating factor, the de Broglie wave,
\[ \cos(\omega_B t - \kappa_B x), \]
which is of planar form and is evolving through the carrier wave at the superluminal velocity $v^{-1}$.

To show how this modulated wave structure is related to dynamic changes in the particle, it will suffice to consider rays passing through the centre of the waveform and moving forwardly and rearwardly along the direction of travel. In the rest frame of the particle, the superposition of these rays produces the one-dimensional standing wave,
\[ \Psi(x, t) = \frac{1}{2} \bigl[ e^{i(\omega_0 t - \kappa_0 x)} - e^{i(\omega_0 t + \kappa_0 x)} \bigr], \quad (31) \]
but when observed from a frame in which the particle is moving at velocity $v$, these forwardly and rearwardly moving rays, to be now labelled 1 and 2 respectively, transform as,
\[ e^{i(\omega_0 t - \kappa_0 x)} \rightarrow e^{i(\omega_1 t - \kappa_1 x)}, \]
\[ e^{i(\omega_0 t + \kappa_0 x)} \rightarrow e^{i(\omega_2 t + \kappa_2 x)}, \]
where, in accordance with the Doppler effect,
\[ \omega_1 = \gamma \omega_0 (1 + v), \quad \omega_2 = \gamma \omega_0 (1 - v), \quad (32) \]
\[ \kappa_1 = \gamma \kappa_0 (1 + v), \quad \kappa_2 = \gamma \kappa_0 (1 - v), \quad (33) \]
so that standing wave (31) becomes,
\[ \Psi(x, t) = \frac{1}{2} \bigl[ e^{i(\omega_1 t - \kappa_1 x)} - e^{i(\omega_2 t + \kappa_2 x)} \bigr], \]
which can also be written,
\[ \Psi(x, t) = \sin\left(\frac{\omega_1 - \omega_2}{2} t - \frac{\kappa_1 + \kappa_2}{2} x\right) \cos\left(\frac{\omega_1 + \omega_2}{2} t - \frac{\kappa_1 - \kappa_2}{2} x\right). \quad (34) \]

In wave (34), the second factor,
\[ \cos\left(\frac{\omega_1 + \omega_2}{2} t - \frac{\kappa_1 - \kappa_2}{2} x\right), \]

---

8 For a consideration of rays in other directions, see Shanahan [24].
is the de Broglie wave, from which the Einstein frequency and de Broglie wave number are therefore, respectively,

\[ \omega_E = \frac{\omega_1 + \omega_2}{2}, \]

and,

\[ \kappa_{dB} = \frac{\kappa_1 - \kappa_2}{2}. \]

In natural units in which \( \hbar = c = 1 \), the Planck–Einstein relation (1) is thus,

\[ E = \frac{\omega_1 + \omega_2}{2}, \tag{35} \]

while the de Broglie relation (9) is simply,

\[ p = \frac{\omega_1 - \omega_2}{2}. \tag{36} \]

The energy and momentum of the particle have thus been expressed in a simple and intuitive way as, respectively, the sum of and the difference between, the energies of forwardly and rearwardly moving waves.

In the same natural units, it follows from Eqns. (32) and (33) that,

\[ m = \omega_0 = \sqrt{\omega_1 \omega_2}, \tag{37} \]

while the relativistic equation of motion (Eqn. (20), that is,

\[ E^2 - p^2 = m^2, \]

can be seen to be the equality,

\[ \left( \frac{\omega_1 + \omega_2}{2} \right)^2 - \left( \frac{\omega_1 - \omega_2}{2} \right)^2 = \omega_0^2. \tag{38} \]

If inertia is now interpreted, not simply as the resistance of a massive particle to changes in its state of motion, but at a more fundamental level, as the resistance of a wave to changes in its oscillatory state, we have in Eqns. (35) to (38), a consistent scheme for the treatment in terms of wave characteristics of the energy, momentum, inertia and mass of a massive particle.

In summary, it has been argued in this and the preceding section: that on the evidence of the Lorentz transformation and Planck-Einstein relation, the elementary articles comprise underlying wave-like influences of velocity \( c \); that a massive particle must therefore comprise in its rest frame some form of standing wave; that when observed from another inertial frame this standing wave becomes a travelling wave from which the de Broglie wave emerges as a modulation; and that when considered in this way, the dynamic properties of a massive particle become the properties of a waveform.
Figure 3: The wave structure of a massive particle $P$ depicted in a symbolic manner at three positions along its orbit about a central mass $O$.

8 Gravity from particle wave structure

With the de Broglie wave explained, not as an improbable wave of superluminal velocity, but in a manner well-known from wave-theory (see again Feynman et al [19], Vol. I, Chap. 48), it is now possible to present a wave-theoretic explanation of the twin effect in which gravity emerges from the manner in which this wave structure must adapt to a change of inertial frame.

In Fig. 3, a test particle $P$ is depicted at three locations along its elliptical path about a central body $O$. To a co-moving observer, $P$ would have the form of a standing wave attenuated in intensity in accordance with an inverse square law. But to a notional observer at $O$, the wave structure of $P$ is continually (and continuously) adapting to its orbital path in the manner described for model particle (28) by Eqn. (29). As observed from $O$, the underlying carrier wave is contracted in the direction of motion, whilst its modulation (the de Broglie wave), which has been represented in the drawing by parallel transverse lines, is evolving through the ellipsoidally-contracted wavefronts of the carrier wave at superluminal velocity.

Considered, not as a wave in its own right, but as the modulation of an underlying wave structure, it is no longer puzzling that the superluminal de Broglie wave does not fly off at a tangent from its orbital path, a difficulty that confounded attempts, notably by Schrödinger (see Dorling [28]), to fit this mysterious “wave” to the orbits of atomic electrons. Nor should it now seem mysterious that the de Broglie wave appears to be “piloting” the subluminal
particle along its path, and yet at the same time, is continually overtaking but never leaving the moving particle. In effect, carrier wave and de Broglie modulation are wave factors in the one integral whole.

But the effect of relevance to gravity here is that, as a consequence of the modulation, the orbiting particle experiences the loss of phase in the direction of travel that results in the reduction of energy per orbit that was shown above to reproduce the binding energies of Newtonian gravity and the Schwarzschild metric.

While Newton was able to show that the closed elliptical orbits described by Kepler imply a force acting directly between one mass and another, he was unable to explain the origin of such a force. Nor should it be thought that the curvature of spacetime supposed by general relativity is an adequate explanation of gravity. There is, of course, a pleasing symmetry in Wheeler’s aphorism that,

*Space acts on matter, telling it how to move. In turn, matter reacts back on space, telling it how to curve* (Misner et al [13], p. 5).

But the geometric approach provides no apparent basis for this reciprocity of cause and effect. While the geodesic equation does explain how matter would move in a curved spacetime, general relativity is silent as to how matter causes this curvature.

On the other hand, in the interpretation of gravity proposed in this paper, the movement of matter provides its own explanation of why an object must persist in a bound state. Gravity becomes, in this sense, a form of inertia. Einstein stressed the close relationship between gravity and inertia, and on occasion referred to gravity as being a form of inertia, see Lehmkuhl [30]. In the scheme proposed here, it is likewise the inertia of a system of mutually orbiting objects - the tendency of the system to persist in its present state - that locks in the mass defect of the system and ensures that, failing the loss or input of further energy, the system retains that state.

But why should this binding effect follow (as it does exactly in Newtonian gravity) an inverse square law. Such a law is commonly explained from the geometric dilution of an effect that radiates outwardly from a point (see, for instance Wikipedia [29]). A similar dilution must occur in the wave-based explanation of matter described above, where the amplitude of a particle wave structure must decrease inversely with radial distance. Were it not to do so there would be a discontinuity in the movement of energy inwardly and outwardly through the wave structure.

Assuming that this same requirement of continuity constrains the composite wave structure of an aggregation of particles, this would explain the inverse square law governing Newtonian gravity. It would also be consistent with the apparent departure from the predictions of that law in the Schwarzschild metric, where the effective rest mass of the particle decreases with depth within the potential.
9 Concluding remarks

It would be only natural to suggest that I have the cart before the horse here - that in the ordering of explanatory priority, it is not the twin effect that explains gravity, but gravity that is the cause of orbital motion and is thus the explanation of the twin effect.

But it is not gravity, but the failure of simultaneity described by the Lorentz transformation that is the source of the twin effect, and that transformation applies to all matter, whether moving rectilinearly, following an orbital path, or moving in any other way. I stress again that even if there were no gravity in the usually recognized sense, the Lorentz transformation would induce the twin effect and an orbiting object would experience a dilation of time, and have nonetheless, a binding energy.

There is also here the promise of explanatory unification. On the evidence of the Planck-Einstein relation and Lorentz transformation, and with the corroborations of the de Broglie wave, I have argued for a unified wave-theoretic understanding of matter and radiation, implying in turn a common origin in wave structure for gravity and quantum mechanics.

There is thus, I suggest, ample reason to decide the issue of explanatory priority in favour of the twin effect, or to take this to a more fundamental level, in favour of a thoroughly wave-based explanation of matter and energy.

References


[23] D. Shanahan, Reverse Engineering the de Broglie Wave IJQF 9, 44-63 (2023)


