Photon theory of gravity – an advance from Einstein's relativity

Xianming Meng
Research School of Physics, Australian National University, Canberra, ACT 2601
email: xianming.meng@anu.edu.au

Abstract

Based on a postulate that photons of low frequencies (undetectable by current technology) are the gravity force carrier, the paper derives quantitative results that are the same as or very similar to those derived in the special and general relativity theories and explains experiments and observations better. These quantitative results are the mass-energy formula, the energy momentum equation, and those for relative mass, the transverse Doppler effect, gravitational red shift, planetary precession, the deflection angle of light in gravitational lensing, the orbits around a black hole, and the strength and direction of gravitational waves (orbit decay of pulsars). At the detailed level, the results and explanations are different from some of those achieved using Einstein's relativity theory, such as the explanation of the null experimental result for the Doppler effect of electromagnetic waves reflected from a transversely moving surface, the reason for gravitational red shift, and the size of the light sphere around a black hole. The paper claims that both the high-order Doppler effect and the gravitational red shift occur only at the point of photon emission. The paper also explains why the predicted deflection angle of gravitational lensing is slightly greater than the measured deflection angle and why the predicted pulsar orbit decay is close but differs from calculations based on observations.

Key words: photon density; mass-energy equation; general relativity theory; black hole; gravitational waves; gravitational lensing

1. Introduction

This paper is an attempt to advance Einstein's theories of special and general relativity. Einstein was a great imaginer, analogist, and tradition-breaker. His postulate of the constant speed of light, relativistic velocity addition formula and the concept of spacetime shocked the world. His numerous thought experiments and bold claims that were based on analogies of different reference frames and electromagnetic fields were proved correct, so his abstract and counterintuitive theory eventually conquered the world.

However, most people would agree that these theories are fairly difficult to understand due to the involvement of transformation of reference framse, the interaction of spacetime and matter-energy, the Reimann geometry, and higher dimensional tensors. Furthermore, it can be argued that the analogical approach based on the postulates of Einstein may not pinpoint the cause of the relativistic phenomena. By proposing that photons carry the gravitational force, this paper provides a straightforward explanation for the proven relativistic phenomena identified or predicted by Einstein. The paper also explains some experiments and phenomena that Einstein's theory cannot explain in a rigorous fashion, and provides new predictions to test the theory.

The paper is organized as follows. Section 2 shows the advances made in relation to special relativity. Applying a few assumptions, we will derive results consistent with those achieved using the special relativity theory, and we will also show that the new theory can better explain the mixed experimental results on the transverse Doppler effect. Building upon the previous section and with new assumptions centred on low frequency photons being the carrier of the gravity force, Section 3 obtains the results the same as or similar to those in the general relativity theory and explains the differences at the detailed level. Section 4 concludes the paper.

2. Advancing from special relativity

This section starts with the statements and explanations of the key assumptions of the new theory, which are the basis for deriving the formulas for the key physical quantities. The section also applies the new theory to explain the mixed experimental results on transverse Doppler effects.

2.1. Assumptions

The proposed photon density theory requires a few intuitive assumptions. We call them assumptions because while they are not yet proven, they are consistent with, or extended from, physical observations in an inertial frame of reference. The fundamental assumption is that light consists of periodically emitted photons, so light is basically particles with the periodicity that give the wave property. As a result, this assumption does not conflict with the belief by most physicists that light is electromagnetic waves. There are four additional assumptions that are applied to this new theory.

(1) Light (photon) speed is independent of the speed of the source.

This assumption is also found in the special relativity theory. The assumption is supported by observations and experiments such as stellar aberration, the Doppler effect of light, observation of the movement of binary stars, the speed of γ rays from mesons, and the one-way Michelson and Morley experiments.

(2) All material objects emit photons evenly in all directions (isotropic emissions), and the number of photons emitted per second is proportional to the magnitude of the mass at rest.

The isotropic emission assumption suggests that the number of photons emitted in each direction for any period of time is equal. Photon emission is a common phenomenon and we observe that the emission intensity of a light source is the same in all directions, so isotropic emission is plausible. All materials emit light at high temperature, but it seems that most objects do not emit photons at low temperature. This may be because the emission frequencies at low temperature are too low to be detected by current technology. Other things being equal (e.g. temperature and volume), an object with a larger mass tends to produce a proportionally higher light intensity, which indicates that the photon numbers per emission are proportional to the magnitude of the rest mass.

Given that matter emits photons evenly in all directions, we focus on the photon emission rates in one representative direction. The number of photons emitted per second in any direction can be calculated as:

$$N=f*e*m_0 \tag{1}$$

where m_0 is mass at rest, e is the number of photons per emission by one unit of mass for a given direction, f is emission frequency (the number of emissions per second), and N is the number of photons per second emitted by the total mass m_0.

If the first photon emitted by an object travels a distance of s in t seconds, given N is the photon numbers emitted per second, there are $N*t$ number of photons covering the distance s, so the line density can be expressed as:

$$d=Nt/s=f*e*m_0\ t/s$$

Apparently, line density d is also proportional to rest mass m_0.

(3) The inertia of matter is proportional to its photon density.

This assumption is a natural extension of the common wisdom that mass at rest is the measurement of the inertia of a matter. Based on assumption 2, the density of photons is proportional to the magnitude of the mass at rest, so the density can serve as a measurement of mass and thus a measurement of inertia. One may further argue that photon density may result in the inertia of the emitter through photon pressure or photon matter interaction. This argument implies that photon density may be the cause of inertia.

When a photon emitter moves, the photon density changes: it increases in front of the emitter and decreases behind the emitter. This change in density structure may change the average density and thus affect the level of inertia.

We use the concept of relative mass as the indicator for the changing inertia of the object. For an object at rest, its relative mass equals its rest mass, m_0, and the corresponding photon density is denoted d_0. When the object starts to move, the relative mass m is different from m_0, and photon density d is different from d_0. Since we assume inertia or relative mass is proportional to photon density, we can write:

$$m = k * d \quad \text{and} \quad m_0 = k * d_0\ , \quad \text{where } k \text{ is a constant.}$$

By utilizing the above equations, assumption (3) can be crystalized as:

$$m = m_0 d/d_0 \tag{2}$$

(4) Emission frequency is proportional to the inverse of emitted photon density.

The pressure of emitted photons can adversely affect the further emission of photons from the emitter, so photon density d emitted by an object and its emission frequency f are inversely related to each other. This assumption can be expressed as:

$$f \propto \frac{1}{d} \text{ or } fd = constand$$

If the rest mass m_0 has an emission frequency of f_0 and a photon density of d_0, the assumption can be further expressed as:

$$fd = f_0 d_0 \qquad or \quad f = f_0 d_0 / d \tag{3}$$

2.2. Basic physical quantities and concepts

Using the four assumptions in the previous section, we can derive important physical quantities for stationary and moving objects.

(1) Mass

We formally define rest mass as the mass measured when an object is stationary in the measurement frame, denoted as m_0. When the object is moving, the measured mass is called relative mass, denoted as m. Given the emission frequency f_0 for the object at rest, its photon emission rate (photon number per second) N_0 can be obtained by applying equation (1) to the case:

$$N_0 = f_0 e m_0 \tag{1'}$$

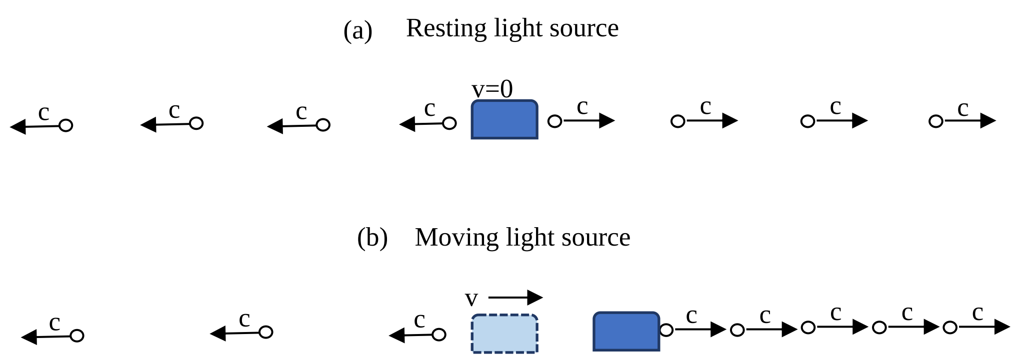
Before we proceed further, we need to define photon density more specifically. Due to the symmetry of a stationary point emitter, one may tend to measure the photon density in a volume of a sphere centred on the emitter. However, because photon density decreases as the radius of the sphere increases, this is not a good way to measure the overall photon density. If the emitter starts to move, the measurement becomes even more troublesome.

Considering the fact that a moving emitter does not affect the density in the directions perpendicular to the movement, we can use the formula for cylinder volume along the direction of movement to measure overall photon density. This approach (see appendix) shows that the impact of the moving emitter on the average photon density is the same as the photon density along the direction of the movement, so we consider photon density in axis of moving emitter only, as shown in Fig. 1.

Denoting c as the speed of light in a vacuum, as shown in panel (a) of Fig. 1, we can infer that if N_0 photons are emitted in 1 second and cover the distance of c in any direction, the line density of photons can be calculated for the object at rest:

$$d_0=N_0/c \tag{4}$$

Fig.1 Speed of light source, photon density, and relative mass



When the object is moving, the structure of the photon density will change. As shown in panel (b) of Fig. 1, a light source with speed v will have denser photons in front of the source in the direction of movement and less dense photons behind the source. The change in photon density structure may also affect the average photon density and thus affect the photon emission frequency. Based on equations (1) and (1'), we have:

$$N=fem_0=(f/f_0)*(f_0em_0)=N_0f/f_0 \tag{5}$$

The photon density ahead and behind the source can be calculated respectively as:

$$d_1=N/(c-v) \quad \text{and} \qquad d_2=N/(c+v)$$

As such, the average line density of the light source can be expressed as:

$$d=(d_1+d_2)/2=Nc/(c^2-v^2) \tag{6}$$

One may argue that the average density can be calculated by 2*N* photons (*N* photons ahead and *N* photons behind the object) covering a distance of (*c*+*v*)+(*c*-*v*). This approach is invalid or inaccurate because it gives more weighting (because of the longer distance) to the photons behind the source.

Plugging equations (4) and (5) into equation (6), we have:

$$d=(N_0 f/f_0)c/(c^2-v^2) = d_0(f/f_0)c^2/(c^2-v^2) \tag{7}$$

Plugging equation (3) into equation (7), we have:

$$d^2=d_0^2c^2/(c^2-v^2)$$

Solving the above equation and using equation (3) again, we obtain:

$$d = d_0/\sqrt{1-\frac{v^2}{c^2}} \tag{8}$$

$$f = f_0\sqrt{1-\frac{v^2}{c^2}} \tag{9}$$

Using equations (2) and (8), we have relative mass:

$$m = m_0 d/d_0 = m_0/\sqrt{1-\frac{v^2}{c^2}} \tag{10}$$

The difference between relative mass and resting mass can be referred to as kinetic mass, which can be calculated as:

$$m_k = m - m_0 = m_0/\sqrt{1-\frac{v^2}{c^2}} - m_0 \tag{11}$$

A word of caveat is necessary here. Fig. 1 is very similar to a graph that explains the longitude or ordinary Doppler effect, while the resulting equation (9) is similar to the transverse or relativistic Doppler effect. This may cause confusion for some readers. It is important to highlight that here we are dealing with photon density rather than perceived photon frequency. Although the ordinary Doppler effect or the PERCEIVED frequency/wavelength change at the observation point can be demonstrated by a similar graph, we concern in Fig. 1 and the resulting equations (8) and (9) only the resultant change in photon density, which is objective regardless of the state of the observer. Later, we will show that this objective impact is the source of the so-called relativistic Doppler effect. We will also show that if the observer is not moving at the same speed as the light source, the ordinary Doppler effect will be added on the top of the frequency change shown by equation (9).

(2) Momentum

Since relative mass changes with speed, the definition and the formula for momentum in classical physics needs to be upgraded to reflect the changing relative mass. The general formula for momentum can be expressed as:

$$p=\int Fdt=\int d(mv)=\int (mdv+vdm)=mv$$

Using equation (10) and integration by parts, one can easily verify that this general formula holds for relative mass:

$$p = \int (mdv + vdm) = \int \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} dv + \int vd \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$= \int \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} dv + v \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} - \int \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} dv = mv$$

(3) Energy

The calculation of kinetic energy also needs to be based on relative mass:

$$K=\int Fdx=\int (dx/dt)Fdt=\int vd(mv)=\int (mvdv+v^2dm) \tag{12}$$

Equation (10) can be rewritten as:

$$m^2(c^2-v^2)=m_0^2c^2$$

Differentiating the above equation, we have:

$$c^2dm-v^2dm-mvdv=0$$

or

$$v^2dm+mvdv=c^2dm \tag{13}$$

Plugging equation (13) into equation (12), we have:

$$K = \int c^2 dm$$

Equation (10) shows that when the speed increases from zero to v, the mass increases from m_0 to m. The above calculus in this range can be evaluated as:

$$K=mc^2-m_0c^2$$

The term m_0c^2 indicates the amount of energy related to rest mass, so we call it rest energy or internal energy E_0:

$$E_0=m_0c^2 \tag{14}$$

Similarly, the term mc^2 indicates the energy related to relative mass, so we call it relative energy:

$$E=mc^2 \tag{15}$$

The difference between relative and rest energy is kinetic energy, which is related to kinetic mass:

$$K=m_kc^2=(m-m_0)c^2=E-E_0$$

Alternatively, we have:

$$\mathrm{E}=K+E_0=K\ m_0c^2=mc^2$$

From the mass-energy equation and the definition of momentum we can easily derive the energy-momentum relationship. Squaring the equation for momentum definition we have:

$$p^2 = (mv)^2 = \frac{m_0^2 v^2}{1 - \frac{v^2}{c^2}}$$

We can solve the above equation for v to obtain:

$$v^2 = \frac{p^2 c^2}{p^2 + m_0^2 c^2}$$

Plugging this into equation (10) and squaring both sides of equation (15), we have:

$$E^2 = (mc^2)^2 = \frac{m_0^2c^4}{1-\frac{v^2}{c^2}} = \frac{m_0^2c^4}{1-\frac{p^2}{p^2+m_0^2c^2}} = (pc)^2 + (m_0c^2)^2 = (pc)^2 + E_0^2$$

This is the energy-momentum equation.

2.3. Explaining experimental results on transverse Doppler effect

From the time dilation effect, special relativity suggests that there would be a relativistic or higher-order Doppler effect. If the light source and the observer move in a direction perpendicular to the light ray, there will be no ordinary Doppler effect, but there will still be a red shift of light frequency to be observed, namely the transverse Doppler effect. Einstein suggested experimental observations of this effect to confirm his theory. Ives and Stilwell (1938[i], 1941[ii]) confirmed the higher-order Doppler effect by measuring the frequency shift of light emitted from atoms moving at high velocities. The emission experiments by some researchers (e.g. Hay et al, 1960[iii]; Kundig, 1963[iv]; Kaivola et al., 1985[v] ; Klein et al., 1992[vi]) have directly confirmed the transverse Doppler effect. However, when Jennison and Davies (1974[vii], 1975[viii]) attempted to measure the transverse Doppler effect using a rotating mirror, they got a null result. With the advancement of technology, Thim (2003[ix]) tried to accurately measure the Doppler shift of microwaves reflected from transversely moving/rotating antennas but also reported a null result. This section explains these mixed experimental results.

We start with the formula for the ordinary Doppler effect:

$$\frac{\lambda}{\lambda_0} = \frac{c - v * cos\theta}{c} = 1 - \frac{v}{c}cos\theta \tag{16}$$

where λ and λ_0 are the perceived and original wavelength of light, respectively; c is the speed of light, v the speed of the moving light source and θ the angle between v and the light ray towards the observer.

The Lorentz transformation necessitates that for a moving object, time slows down by the amount of the Lorentz factor $\gamma = 1/\sqrt{1 - v^2/c^2}$, so the relativity theory predicts a high-order Doppler effect due to time dilation. Incorporating this high-order effect into the ordinary Doppler effect, we have the formula for the full Doppler effect:

$$\frac{\lambda'}{\lambda_0} = \gamma(1 - \frac{v}{c}cos\theta) = \frac{1-\frac{v}{c}cos\theta}{\sqrt{1-\frac{v^2}{c^2}}} \tag{17}$$

Using our photon density theory, we can derive the same formula as above. As discussed in Section 2.2, the photon density around a moving light source can cause pressure and thus negatively affect the photon emission frequency. The photon emission frequency of a moving light source is described by equation (9). In terms of wavelength, the equation can be written as:

$$\lambda_M = \lambda_0/\sqrt{1 - \frac{v^2}{c^2}} \tag{18}$$

where λ_0 is the wavelength from a stationary light source and λ_M is the wavelength from a moving light source.

The change in wavelength described by equation (18) will also be perceived by the observer, so the ordinary Doppler effect in equation (16) should be upgraded to:

$$\lambda' = \frac{\lambda}{\lambda_0} * \lambda_M = \frac{c-v*cos\theta}{c}\frac{\lambda_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1-\frac{v}{c}*cos\theta}{\sqrt{1-\frac{v^2}{c^2}}}\lambda_0 \tag{19}$$

This equation is exactly the same as equation (17), which was derived based on time dilation. Using photon density, we can explain with ease the mixed experimental results on the high-order Doppler effect. During the emission experiments (Ives and Stilwell, 1938, 1941; Hay et al, 1960; Kundig, 1963; Kaivola et al., 1985; Klein et al., 1992), the emitter is moving[1], so the frequency of emitted light is reduced by the increased photon density. In the reflection experiments carried out by Jennison and Davies (1974, 1975) and Thim (2002), the emitter is stationary, so there is no change in the photon density around the emitter and thus no change in the frequency of emission. This leads to the experimental null results.

The special relativity theory can easily explain the positive experimental results on the transverse Doppler effect, but it cannot logically explain the null results of the reflection experiments. As

[1] Some may think the emitters in experiment by Hay et al. (1960) and Kundig (1963) are stationary. Careful reading of their experiment reveals that the light frequency they measured is from the second emitter, which is moving with respect to the counter.

the reflector (i.e. the transverse moving mirror/surface) and the light source are moving relatively, the time would be dilated when the photons arrive at the reflector, which means the light frequency should decrease. Using the same reasoning, there is relative movement between the reflector and the observer, so there must be another time dilation effect when the photons move from the reflector to the observer. This would cause another red shift. Adding two red shifts, the observed transverse Doppler effect would be twice as much as that observed in the emission experiments. This is at odds with the reflection experiments.

Due to the significant success of using relativity to predict other phenomena, the null experimental results of the transverse Doppler effect from a rotating mirror have largely been neglected. The only discussion about them is an explanation by Sfarti (2010[x]), who states that the null result in Thim's experiment is expected because both the light source and the detector are stationary. Mathematically, Sfarti's explanation can be expressed by a Lorentz transformation and an inverse-Lorentz transformation.

Let the reflector move on the x axis while the photons travel on the y axis. For photons to move from light source A to reflector A', the following 2D Lorentz transformation formula is used to transform the stationary frame of A (*txy*) to the moving frame of A' (*t'x'y'*):

$$x' = \gamma(x - vt),\ y' = y,\ t' = \gamma\left(t - \frac{vx}{c^2}\right),\ \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{20}$$

From the above formulas we can calculate the time *t'* in the new reference frame A' as:

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) = \frac{t}{\gamma} - \frac{vx\prime}{c^2} \tag{21}$$

Using A' itself as the new reference frame, point A' is stationary, so x' is constant for point A and we have $\Delta t'=\gamma\Delta t$. Since v is less than c, γ>1, so $\Delta t'=\gamma\Delta t<\Delta t$. This means a smaller amount of change in *t'* is equivalent to more change in *t*,(i.e. the time in reference frame (*t'x'y'*) elapses more slowly than in the frame (*txy*). This is the time dilation effect claimed in the special relativity theory.

When the photons move from reflector A' to observer A'', an inverse-Lorentz transformation seems applicable because observer A'' and the light source are both in the stationary frame. Applying the following inverse Lorentz transformation:

$$x = \gamma(yx' + vt'),\ y = y',\ t = \gamma\left(t' + \frac{vx\prime}{c^2}\right),\ \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{22}$$

we can obtain:

$$t = \beta\left(t' + \frac{vy\prime}{c^2}\right) = \frac{t\prime}{\beta} + \frac{vy}{c^2} \tag{23}$$

Plugging the *t'* in equation (20) into equation (23), we have a null transverse Doppler effect of reflected photons:

$$t \quad = \beta\left(t' + \frac{vy\prime}{c^2}\right) = \frac{t\prime}{\beta} + \frac{vy}{c^2} = t - \frac{vy}{c^2} + \frac{vy}{c^2} = t$$

Since *t* has not changed, it seems that the special relativity theory provides a perfect explanation for the null results of the reflection experiments.

A similar but more advanced Lorentz transformation is to use a four-dimensional Lorentz boost to the energy momentum of the photon $(E,p_x,p_y,p_z) = (1,0,1,0)$:

$$\begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma \\ -\gamma v/c \\ 1 \\ 0 \end{pmatrix}$$

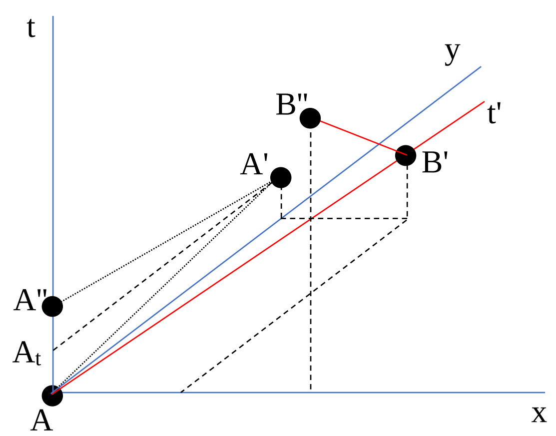
The results show that the photon's energy and x-direction momentum change but the momentum/speed in the y direction does not change. When the photon is reflected back from the mirror, the y-direction momentum changes sign. Using an inverse-Lorentz boost, we obtain:

$$\begin{pmatrix} \gamma & \gamma v/c & 0 & 0 \\ \gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \gamma \\ -\gamma v/c \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

It is apparent that the energy and momentum are unchanged, except for a sign change for the momentum in the photon travel axis. This also seems to be consistent with the null results of the reflection experiments.

However, the above explanation involves an inappropriate application of the inverse Lorentz transformation. The inverse transformation formula is used to transform the same event back to the original reference frame, but they apply this formula to a new event. To use the (inverse) Lorentz transformation, one must satisfy a condition that the two reference frames have a common initial point. This can be seen from equations 20 and 22. If we let t=x=0, we have t'=x'=0, and vice versa. A new event like the reflected photon does not satisfy this condition. This can be demonstrated by a Minkowsky spacetime diagram (Fig.2):

Fig. 2 Lorentz transformation for photon reflection experiments



In Fig. 2, the stationary lab reference frame is indicated by spacetime (t,x,y). The reflection experiment in this reference frame can be described as an incoming light from A to A' and then a reflected light from A' to A''. All events occur in the yt plane because photon reflection occurs only on the y axis. In the frame of the moving mirror (t', x', y'), the reflection experiment can be described as an incoming light from A to B' and then a reflected light from B' to B''. The Lorentz transformation can transform event AA' from the lab reference frame to the AB' in the moving mirror frame and the inverse Lorentz transformation can transform AB' back to AA'. However, one cannot transform the new event B'B'' back to A'A'' by inverse Lorentz transformation: the starting points of the new events (A’ and B’) are different and are not at the origin of two reference frames, thus the (inverse) Lorentz transformation is not applicable. An appropriate Lorentz transformation can be used only when one set the initial point (i.e. the reflection point) as the common origin of the two reference frames involved.

The other possible argument in defending the time dilation explanation of this case may be that the spinning mirror/plate is not an initial frame so one should apply the general relativity rather than the special relativity. This argument is not valid for the reflection experiments. Because the reflected photons (or electromagnetic waves) contact only the surface of the mirror/plate, they have not entered the spinning system and have experienced no attractive force from the system. The part of touching surface moves at an almost constant speed in the transverse direction, so it should be viewed as an inertial frame.

The null results of the transverse Doppler effect from a reflected surface can be viewed as indirect evidence for the claim that a frequency change only occurs to the photon emissions from the moving light source. Direct evidence may be obtained by an experiment in which the light source is stationary while the detector is moving transversely and rapidly. The photon density theory expects a null result for the Doppler effect while Einstein’s relativity theory expects a positive result. This type of experiment can shed light on what causes the transverse Doppler effect; however, a rapid-moving detector may affect the stability and accuracy of the detection. One needs suitable technology to overcome this problem.

3. Advancing from general relativity

The general relativity theory provides a number of predictions that are quantitatively confirmed by observations. Using the photon density approach, this section examines these predictions. It begins with some additional assumptions, then examines Newton’s gravitational formula and the predictions from the general relativity theory.

3.1 Assumptions

As well as the assumptions listed in Section 2.1, the following assumptions are necessary to explain phenomena related to gravity.

(G1) All matter emits low frequency photons that are attracted by other matter.

(G2) The gravity force is transferred first by the attraction between the first photon and the mass and then by the attraction between the photons in a ray.

(G2) The attraction force between two objects is proportional to the density of the photons or the relative mass of the objects.

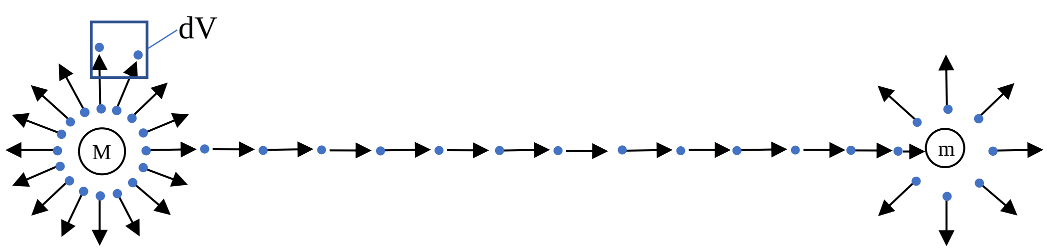
3.2 Results and rationale

With the above assumptions, we can explain the relativistic phenomena in the gravity fields.

(1) Deriving and upgrading Newtonian gravitational formula.

We start the derivation from Fig. 3, which shows mass *M* radiates photons in all directions (Assumption 2). As the photons in the ray from left to right attract each other, mass *M* can attract mass *m* through a great distance (Assumption G2). However, as the distance increases, the photon density decreases and the attraction between *M* and *m* decrease significantly (Assumption G3).

Fig. 3. Photon density and gravitational force



The number of photons emitted by mass *M* in one second is proportional to the frequency of emission *f* and to the amount of mass, $n=k_1f*M$, so the total number of photons emitted in time period dt is $dN=ndt= k_1fMdt$.

Given that the photons travel at the speed of light *c*, the distance the photons cover in time *dt* is *cdt*. The sphere volume from *r* to *r+cdt* (i.e. *dr=cdt*) is:

$$dV= \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} dv= \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} r^2 \sin\theta dr d\theta d\varphi=4\pi r^2 cdt.$$

The density of fundamental particles at radius *r* is $D=dN/dV=k_1fM/(4\pi cr^2)$. Based on Assumption G3, the gravity force can be expressed as:

$$F=k_2*D*m=k_1k_2fMm/(4\pi cr^2)$$

If we calibrate k_1 and k_2 to satisfy $G=k_1k_2f/(4\pi c)$, the above equation becomes Newton's gravity equation. We can rewrite Newton's formula in a different form:

$$F = 4\pi Gm\left(\frac{M}{4\pi r^2}\right)$$

The term in the brackets is a measurement of the area density of photons emitted by *M*, letting the area density be:

$$\rho_S = \frac{M}{4\pi r^2}$$

We can write Newton's formula as:

$$F = 4\pi \mathrm{Gm}\rho_S \tag{24}$$

In considering the assumptions in Section 3.1, a number of amendments should be made for Newton's gravitational law. First, the gravitational force is not instant. Since the gravity is transferred through a chain of photons, the gravitational force is realized only after the first photon has arrived at mass *m*.

Second, the mass in the gravity equation should be relative mass. Based on our derivation, photon density is the determinant for gravity force. From our discussions in Section 2.2, we know that when an object starts to move, its average photon density changes and this leads to a change in relative mass. Relative mass is more accurate than rest mass as a measure of photon density and thus as a determinant of gravitational force. The necessity to replace rest mass with relative mass is manifested by the precession of planetary orbits. As will be shown in the next section, the precession of Mercury can be explained well by the change in relative mass of the planet.

Third, there may be other factors involving gravitational force. One example is temperature. Since temperature can change the photon emission frequency and intensity, it directly affects photon density. If our assumptions of the role of photon density in gravity are proven to be correct, a change in temperature should change the size of the gravitational force. However, due to the weakness of the gravitational force and the rareness of dramatic temperature changes in large masses in daily life, the effect of temperature on gravity is hard to quantify. Notwithstanding, the successful detection of gravitational waves seems to be indirect supporting evidence of this. In ordinary situations, gravitational waves are weak and stable so they are hard to detect. In the event of a supernova or black hole collision, temperature surges and large number of photons are emitted in a very short period of time, which causes a sudden surge in the gravitational force and makes the gravitational wave detectable. Future observations may provide more data regarding any relationship between temperature and gravitational force.

(2) Explaining planetary precession

For an orbiting mass *m*, its speed *v* has two parts. One part is related to the radial of the orbit, $v_1=dr/dt$, and the other is related to the angular change, $v_2=dr/d\theta * d\theta/dt = r * d\theta/dt$. Since any orbits in the field must satisfy conservation of total energy *E* and angular momentum $L=mv_2$, we can express total energy at any point on an orbit as:

$$E = \frac{1}{2}m\left(\frac{dr}{dt}\right)^2 + \frac{1}{2}m\left(r\frac{d\theta}{dt}\right)^2 - \frac{GMm}{r} = \frac{1}{2}m\left(\frac{dr}{dt}\right)^2 + \frac{1}{2}m\left(\frac{L}{mr}\right)^2 - \frac{GMm}{r}$$

The first term is the kinetic energy, the second term is angular potential energy, and the third term is the potential energy of the gravitational field. The total potential (including angular potential) of the orbit can be expressed as:

$$\Phi = -\frac{GM}{r} + \frac{1}{2}\frac{L^2}{m^2r^2} \qquad (25)$$

However, this is a Newtonian potential as it assumes that the mass of the particle does not change with speed. Considering the relative mass of the particle, the Newtonian field potential should be updated to:

$$\Phi = -\frac{GM}{r} + \frac{1}{2}\frac{L^2}{{m_0}^2 r^2}\left(1 - \frac{v^2}{c^2}\right) \tag{26}$$

The total speed *v* can be obtained from energy conservation:

$$E = \frac{1}{2}mv^2 - \frac{GMm}{r} \tag{27}$$

For point mass in an elliptical orbit, the amount of total energy is fixed by:

$$E = -\frac{GMm}{2a} \tag{28}$$

Where *a* is the semi-major of the orbit.

Using equation (28), then Obtaining *v* from equation (27) and substituting into equation (26), we have:

$$\Phi = -\frac{GM}{r} + \frac{1}{2}\left(1 + \frac{GM}{c^2 a}\right)\frac{L^2}{{m_0}^2 r^2} - \frac{1}{2}\frac{L^2}{{m_0}^2 r^2}\frac{2GM}{c^2 r} \tag{29}$$

The added term $\frac{GM}{c^2 a}$ will change the orbit marginally. For Mercury, $\frac{GM}{c^2 a} = 2.5455 * 10^{-8}$. The last term is the relativistic correction related to r^{-3}, which is the same as that derived in general relativity. With this correction term, we can derive the same disturbance to the orbit and thus a result for the precession of Mercury's orbit, which is very similar to that based on general relativity, so we can explain the unexplained 43'' per century orbit precession as well as general relativity does. This type of derivation is demonstrated in textbooks (e.g. Cheng, 2005; Schutz, 2009[xi]), so it is not necessary to repeat them here.

(3) Gravitational red shift

Based on assumption G3, the attraction force between photons and mass causes the photons to accelerate towards the centre of the mass. The acceleration rate (i.e. gravitational force for photons) is $g=GM/r^2$. When free photons at speed *c* travel from position A at infinity distance from the centre of the mass to position B at distance R from the centre of the mass, their speed should increase to *c'*:

$$c' = c + \int_0^{t1} g dt = c + \int_\infty^R g d(\frac{r}{c}) = c + \int_\infty^R \frac{GM}{cr^2} dr = c + \frac{GM}{cR} = c(1 + \frac{GM}{c^2R}).$$

An alternative approach is to use energy conservation. Although a photon does not have mass, it has energy and momentum, just like a massive particle. The kinetic energy of a photon can be expressed in the usual way but let the mass variable m=1, or simply omit the mass variable. The work done on a photon is the force F=g (i.e. acceleration rate for photons) multiplied by distance. As a free photon falling into a gravitational field at position B, the increased speed *c'* should satisfy:

$$\frac{1}{2}c'^2 - \int_\infty^R F dr = \frac{1}{2}c'^2 - \int_\infty^R \frac{GM}{r^2} dr = \frac{1}{2}c'^2 - \frac{GM}{R} = \frac{1}{2}c^2$$

From this equation, we can have:

$$c' = \left(c^2 + \frac{2GM}{r}\right)^{\frac{1}{2}} = c\left(1 + \frac{2GM}{c^2R}\right)^{\frac{1}{2}} \approx c + \frac{GM}{cR} \qquad (30)$$

As the above equation shows, the kinetic energy transformed from the gravity potential should increase the speed of the photons entering the gravity field. However, these photons will eventually have the same speed as locally emitted photons. This can be explained by the higher pressure caused by the higher photon density when getting closer to the gravity centre. This pressure adversely affects photon emissions and thus can be inversely related to the emission frequency.

The effect of photon pressure caused by a gravitational field can be described by how much the light speed being depressed by the pressure, or the ratio of the would-be speed *c'* calculated in equation (30) to the actual speed *c* in the gravitational field:

$$\frac{c'}{c} \approx 1 + \frac{GM}{c^2R}$$

Assuming the photon pressure indicated by the above equation is inversely related to the photon emission frequency, the emission frequency *f'* in a gravitational field will be less than the frequency *f* outside the gravitational field (infinitely away from the mass centre) and can be expressed as the following ratio:

$$\frac{f'}{f} = \frac{c}{c'} = 1/(1 + \frac{GM}{c^2R}) \approx 1 - \frac{GM}{c^2R} \qquad (31)$$

The equation shows that the locally emitted photons appear to be red shifted by $\frac{GM}{c^2R}$. This result is exactly the same as the gravitational red shift based on Einstein's equal principle.

Despite having the same result, the different causes behind it lead to different predictions. Since Einstein attributed the red shift to the time dilation caused by gravity, his theory necessitates that the frequency of the same light be measured differently in gravitational fields of different strengths, so a high frequency light far away from the gravity centre gets red shifted when getting close to the centre. On the other hand, the present study suggests that the red shift is caused by the higher photon density pressure in a stronger gravitational field, which can prevent the photon speed from increasing in the gravitational field, and thus reduce the frequency of emissions. As such, if the photons are already emitted by a stationary light source in a weaker gravity field and travel to a stronger gravity field, their frequency will not change as the time standard does not change. In other words, the source far away from the mass centre emits higher frequency light than the same type of source close to the mass centre, but the light frequency will not change once the photons are emitted.

The different predictions can be tested by measuring the frequency of the same light source in gravitational fields of different strengths. According to Einstein's theory, the measured emission frequency should be the same, as the time for the measurement is dilated to the same degree as the time for the light source. Our study suggests otherwise, because the photon emission frequency changes by different degrees in gravitational fields of different strengths while the non-emitting measurement instrument (e.g. a detector) does not affected by gravity strength.

The work of Chou et al (2010[xii]) can categorically differentiate the two hypotheses. Chou et al examined the time dilation effect from both the speed of the light source and from the strength of gravity field. They built two nearly identical optical clocks and compared the clock tick frequency under two scenarios: (1) setting the ions in one clock in harmonic motion by a RF electrical field while keeping the other on stationary at the same elevation, (2) raising one clock to a higher elevation (about 33cm) than the other. The experiment shows that the frequency of the clock at higher elevation is greater than that of the other clock while the clock in motion ticks slower than the stationary one. The amount of tick frequency change was consistent with Einstein's prediction.

Since the clock frequency indicates time, it seems to suggest that the time is dilated for the moving clock and for the clock at a position closer to ground. The experiment results seem to support Einstein's time dilation, however, if we go to the details of the experiment, we will find it actually supports the emission frequency change hypothesis.

Optical clock has three major components. First, a highly stable reference frequency or 'clock transition', which is provided by optical absorption of transition of different states of atoms or ions. Second, a laser or local oscillator that can stabilize its frequency to the clock transition, third, a femtosecond comb that can count the frequency of local oscillator, or clock ticks. Based on Rosenband et al (2008)[xiii], Chou et al (2010a, b[xiv]), the transition frequency of their atom clocks is the $^1S0 - {}^3P_0$ transition frequency of Al^+ ions in a trap. The probing laser (local oscillator) utilizes a reference laser transported to the ion traps through optical fiber. The clock signals from the two atomic clocks then transmitted through optical fibers to femtosecond comb for comparison.

Since the measured effect of moving ions in Chou et al (2010) is similar to that of Ives and Stilwell (1938), we focus on the effect of different elevations. If the expectation from the general relativity is correct that the time in a weaker gravity field dilates less, the 'clock transition' frequency at the higher elevation should be higher. Meanwhile, the frequency of the fiber-transmitted probing laser should also increase when the laser reaches the higher elevation. The degree of change in two frequencies should be the same because time dilation is equal to all components at the same height. As such, the less dilated probing laser will match the less dilated clock transition and no effort is needed to change the probing laser frequency (i.e., no error signal is produced to the laser lock). Once the less dilated (and thus higher frequency) laser at higher elevation is transmitted to femtosecond comb at lower elevation for comparison, the higher frequency will be dilated back to the original frequency thanks to the increased gravity strength. As a result, there should be no frequency difference for two clocks. This expectation based on time dilation conflicts with experimental outcomes.

On the other hand, if the gravity causes only an emission frequency change, the experimental result is very easy to explain. Since the gravity field at higher elevation increases only the light frequency of emitter, the probing laser frequency has to be increased by the experiment system in order to match the clock transition frequency. When the probing laser signals are transmitted to the femtosecond comb at lower elevation, the increased frequency does not change and thus is recorded. Compared with the clock frequency at the lower elevation, we find the frequency at higher elevation is blue-shifted. This is what observed in Chou et al (2010).

In short, the results of Chou et al is at odd with the time dilation hypothesis, but can be easily explained by an emission frequency change induced by gravity strength. A more clear-cut experiment can be done by checking the frequency of same type of light source in a space station and on earth. If the frequencies measured at two locations are the same, the experiment supports time dilation hypothesis. Otherwise, it rejects the time dilation and supports the emission frequency change hypothesis.

(4) Gravitational lensing

The gravitational deflection effect was proposed by classical physicists long before Einstein; however, the classical approach has two drawbacks. One is that photons are supposed to be massless so they should not be affected by gravity. The other is that the deflection angle calculated by classical physicists is only half of what has been proven by observations. Using the equal principle, Einstein initially derived the same amount of deflection angle as the classical physicists did, but using the general relativity theory he obtained double the angle. The observation of a deflection angle of this amount played a key role in the acceptance of the general relativity theory. In this section, we explain quantitatively the measured deflection angle.

Our assumption is that photons are gravitational force carriers, so the classical explanation that photons are attracted to mass is plausible. Based on the photon speed in the gravitational field of *c'*, the classical physicists calculated the index of refraction n_R as:

$$n_R = c'/c = 1 + \frac{GM}{c^2 R}$$

This refraction index gives a deflection angle of $\frac{2GM}{c^2 R}$, which is half that of the result from the general relativity theory. The analogy of photon deflection in the gravitational field to the refraction of light between two media is reasonable at the qualitative level, but it is not suitable at the quantitative level. During a refraction, the light travels in a straight line, pivoting at the interface of two media. However, in a gravitational field, light travels along a curve. The trajectories of photons have a significant impact on the apparent light path and on the calculation of the deflection angle.

The trajectory of a photon in the gravitational field can be obtained by using Newton's second law and the conservation of angular momentum formula. Again, for a photon, we need to omit the mass in these laws:

$$F = \frac{d^2 r}{dt^2}$$

$$L = v \times r = r^2 \frac{d\theta}{dt}$$

In the polar coordinates, the acceleration rate in the radial direction (a_R) and in the transverse direction (a_T) can be expressed as:

$$a_R = \frac{d^2 r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2, \qquad a_T = \frac{1}{r}\frac{d}{dt}\left(r^2 \frac{d\theta}{dt}\right) = \frac{1}{r}\frac{dL}{dt}$$

On an orbit in a gravitational field, there is a force on the radial direction but no force on the transverse direction, so a_T=0, which implies there is no change in angular momentum, i.e. conservation of angular momentum. The acceleration in the radial direction is caused by the gravitational force. Applying Newton's second law of motion, the radial acceleration rate must equal the gravitational accelerate rate:

$$\frac{d^2r}{dt^2} - r\frac{d^2\theta}{dt^2} = -\frac{GM}{r^2}$$

Applying conservation of angular momentum, we have:

$$\frac{d^2r}{dt^2} - r\left(\frac{L}{r^2}\right)^2 = -\frac{GM}{r^2}$$

By variable transformation, $u=1/r$ and $d/dt=d/d\theta*d\theta/dt=(L/r_2)*d/d\theta$, the above equation can be transformed to:

$$\frac{d^2u}{d\theta^2} + u = \frac{GM}{L^2}$$

The solution for this harmonic is:

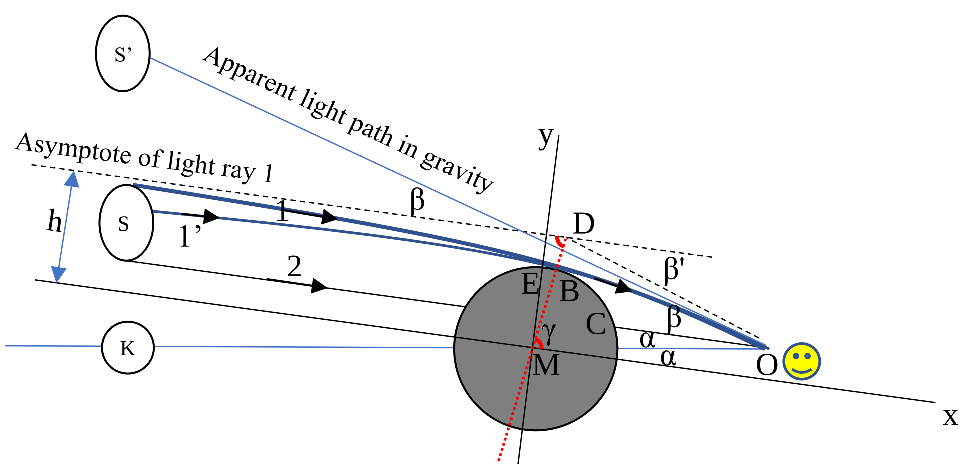
$$u = \frac{1}{r} = \frac{GM}{L^2} + Acos\theta = \frac{a}{b^2}(1 + ecos\theta)$$

This is the photon trajectory in gravity caused by mass *M*, where *e* is the eccentricity of the orbit and *a* and *b* are the vertex and co-vertex, respectively, of a hyperbola. The parameter *A* is to be calibrated by the initial condition.

Next, we use the photon trajectory to calculate the deflection angle observed during a sun eclipse. Fig. 4 shows the deflection of light from star *S* during a sun eclipse. When the sun *M* is absent (i.e., is not between star *S* and the observer on the earth), light ray 2 from the start comes to the observer, who records an observation angle α. When sun *M* is in the position shown in Fig. 4, light ray 2 is blocked by the sun but light ray 1 from star *S*, which does not come to the observer when the sun is absent, is deflected and comes to the observer. Based on experience, the observer assumes a straight light path and finds that the star's position appears to have changed during a sun eclipse. Due to the high speed of the photons, the relatively small mass of the sun (e.g., compared with black holes), and the short distance between the earth and the sun, the light deflection in the gravity of the sun is small, so the deflected light from star *K* on the extension line from observer *O* to the centre of the sun (zero observation angle) cannot be seen from the earth.

For convenience, we use the line passing through the centre of the sun and parallel with light ray 2 as the x axis, and the perpendicular line passing through the centre of the sun as the y axis. The direction of light ray 1 depends on the distance from star *S* to the sun. If star *S* is infinitely far away from the sun, light ray 1 should be parallel to light ray 2. This is shown by the asymptote line (the dashed line) of incoming light ray 1, which is parallel to light ray 2, while light ray 1 is bent by the gravity of the sun. In this case, the observed deflection angle is β. If star *S* is a finite distance from the sun, light ray 1' forms an angle with light ray 2. Based on this light path, the deflection angle is greater than the measured angle β. For simplicity, we assume that star *S* is very far away from the sun compared with the distance between the earth and the sun, so light ray 1 is used in our analysis.

Fig. 4 Calculating deflection angle in a gravity field



Assuming that the speed of the photon in incoming light ray 1 is *c* with energy $E = \frac{1}{2}c^2$ and that the perpendicular distance from the sun to the asymptote of incoming light ray 1 is *h*, the angular momentum of the photon is *L*=*c***h*. When the photon is deflected from point B on the edge of the sun (*MB*=*R*, the radius of the sun), the conservation of energy necessitates:

$$E = -\frac{GM}{R} + \frac{L^2}{2R^2} = \frac{1}{2}c^2$$

This relates angular momentum to the radius of the sun:

$$L^2 = c^2R^2 + 2GMR$$

Since angular momentum is conserved, using the initial angular momentum of the photon we have:

$$h = (R^2 + 2GMR/c^2)^{1/2}$$

This is the requirement for a light ray passing thought the edge of the sun. If *h* is smaller, the photon will hit the sun.

As the total energy *E*>0, the deflection has an unbounded hyperbolic trajectory. The *L*-formula (momentum conservation) and the *E*-formula (energy conservation) require that vertex *a* and co-vertex *b* must satisfy:

$$L^2 = c^2h^2 = \frac{GMb^2}{a} and \; E = \frac{1}{2}c^2 = \frac{GM}{2a}$$

From these two equations we can calculate:

$$\frac{b}{a} = \frac{c^2h}{GM}$$

The deflection angle is the angle at point D formed by the two asymptote lines of light ray 1 (the two dashed lines), namely (π-β')=2γ. The semi-deflection angle γ is the angle between one of the asymptote lines and the line perpendicular to light ray 1 at point B (the red dotted line), or between the red dotted line and the x axis. Based on the geometry of conics, this semi-deflection angle must satisfy:

$$tg\,\gamma = \frac{b}{a} = \frac{c^2h}{GM}$$

Alternatively,

$$\tan\left(\frac{\beta'}{2}\right) = \tan\left(\frac{\pi}{2} - \gamma\right) = cot\,\gamma = \frac{GM}{c^2h} \approx \frac{GM}{c^2R}$$

This is the classical result or the result based on the equal principle, which is only half of the value confirmed by observations. This result is correct for one photon or unrelated photons, but not for a light ray. A light ray consists of continuously emitted inter-attracted photons, and it is this inter-attraction between photons that transfers the gravitational force. If a light ray goes straight towards a gravity centre, the gravitational force is felt by the first photon and then transferred through to the other photons, i.e. the head photon shields other photons from direct interaction with mass. However, if the light ray does not point to the mass centre, as usual the photons in the ray feel the gravity from the previous photon through the chain effect, but also are exposed to the gravity field directly, i.e. the photons experience a double attraction force. As such, the orbit equation should be:

$$\frac{1}{r} = \frac{2GM}{L^2} + Acos\theta = \frac{a}{b^2}(1 + ecos\theta)$$

From this orbit equation, the *GM* in the previous *L*-formula and *E*-formula should be replaced by 2*GM*. As a result, we have:

$$\tan\left(\frac{\beta'}{2}\right) = cot\,\gamma = \frac{2GM}{c^2h} \approx \frac{2GM}{c^2R}$$

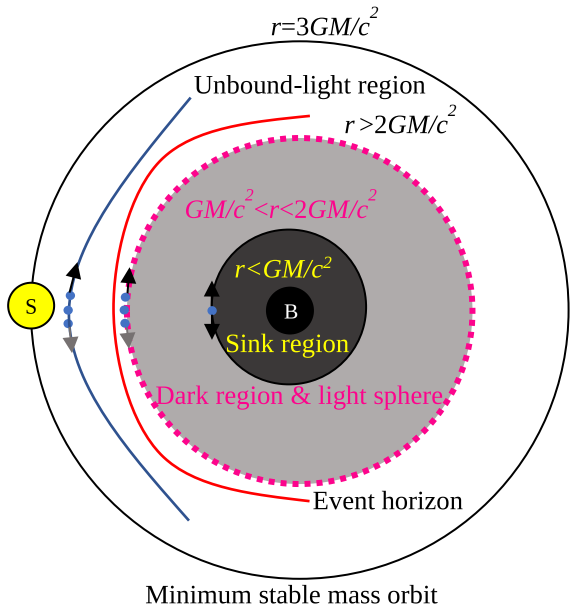
This result corresponds to that derived from the general relativity theory. As shown in Fig. 4, the deflection angle based on the asymptotes is generally greater than that measured on the earth. This is also consistent with the observations (e.g. Dyson et al., 1920[xv], Wang et al. 2000[xvi]).

(5) Black holes

If the mass of an object is very large, the gravitational force can overpower the pressure of nuclear reactions, the status of matter may change, and the object does not emit high frequency photons. As current technology has difficulty detecting the low frequency photons emitted by this type of object, the object appears black. Also, the enormous gravitational force can suck in any matter and photons that are in its vicinity, so the area becomes a sinking black hole.

However, a black hole does not suck in the photons emitted by itself because it relies on them to enact the gravitational force.

Fig. 5 Regions of a black hole



As the distance to the centre of a black hole increases, the impact of the black hole changes and thus we can divide the black hole area into different regions. As shown in Fig. 5, for a radius of less than GM/c^2, any photons from other objects will be absorbed by the black hole, and it is more so for very large particles. Thus, this region can be called a sink region. If the radius is greater than GM/c^2, the photons moving perpendicularly or outwards will not be sucked into the black hole; however, the gravitational force of the black hole is still large enough to suck in any light rays as they consist of inter-attracted photons and experience double gravity force when the ray is not pointing towards the center of blackhole. If the distance from the centre of the black hole is $2GM/c^2$, the light rays will move around the black hole in circular orbits, forming a light circle or light sphere. For $GM/c^2<r<2GM/c^2$, no light escapes from the black hole, so this region cannot be seen from outside world. We can call it a dark region. Any events that occur in this dark region are unknown to the outside world. If $r>2GM/c^2$, light can escape the black hole and be detected by the outside world. As a result, $r=2GM/c^2$ becomes an event horizon.

All the regions in Fig. 5 are enclosed by a circle with a radius of $3GM/c^2$, indicating that, within this circle, no massive particle can exist without being sucked in by the black hole. By intuition, if an object can achieve a very high speed, e.g. close to the light speed *c*, it could orbit the black hole between the two circles ($2GM/c^2<r<3GM/c^2$) in Fig. 5. However, at very high speed, the relative mass increases substantially and thus causes more gravitational force, so the orbit may not be stable. We can examine this quantitatively by using the orbit potential derived previously to explain the precession of Mercury.

For simplicity, we consider a circular orbit, so the semi-major '*a*' in equation (29) becomes the radius of orbit *r*. Letting $r^*=2GM/c^2$, equation (29) can be rewritten as:

$$\Phi = -\frac{GM}{r} + \frac{1}{2}\frac{L^2}{{m_0}^2 r^2} - \frac{1}{4}\frac{L^2}{{m_0}^2 r^2}\frac{r^*}{r} \qquad (32)$$

The last term is the potential energy caused by increased relative mass. A stable orbit must have the lowest energy. We can find the lowest energy orbit by differentiating Φ with respect to *r*:

$$\frac{d\Phi}{dr} = \frac{GM}{r^2} - \frac{L^2}{{m_0}^2 r^3} + \frac{3}{4}\frac{L^2 r^*}{{m_0}^2 r^4} = 0 \qquad (33)$$

Solving this equation for *r*, we have:

$$r = \frac{L^2}{2GM{m_0}^2}\left(1 \pm \left(1 - \frac{3GM{m_0}^2}{L^2} r^*\right)\right)^{1/2} \qquad (34)$$

The solution with the negative sign in equation (34) leads to a lower *r*, which is actually an energy maximum. At a smaller *r*, the contribution of the last term in equation (33) dominates, so an decrease in *r* leads to an increase in potential energy, indicating a local energy maximum. The higher *r* solution with the positive sign is an energy minimum orbit, so it is a stable orbit. The smallest stable orbit is achieved by setting:

$$1 - \frac{3GM{m_0}^2}{L^2} r^* = 0 \text{ , or } r^* = \frac{L^2}{3GM{m_0}^2}$$

Thus, the smallest radius of a stable orbit is:

$$r_0 = \frac{L^2}{2GM{m_0}^2} = \frac{3}{2}r^* = \frac{3GM}{c^2}$$

The results derived in this section differ from those in general relativity, which claims that the photon sphere is r= 3GM/c^2 while the minimum stable mass orbit is r= 6GM/c^2. Our results seem more reasonable than those in general relativity. For example, the circular light orbit of $r^*=2GM/c^2$ in Fig. 5 coincides with the event horizon, which is a consistent result. On the other hand, by assuming a zero length of the geodesic tangent vector for light, the general relativity claims a photon sphere of $r=3GM/c^2$, that is greater than the event horizon. If the light at $r=1.5r^*$ (outside of the event horizon) travels in a circle, it is bounded and thus cannot be seen from the far distance. As a result, no information within this radius can be revealed to the outside world, so $r=1.5r^*$ would become the event horizon. This is in conflict with the event horizon r^* derived from the same theory.

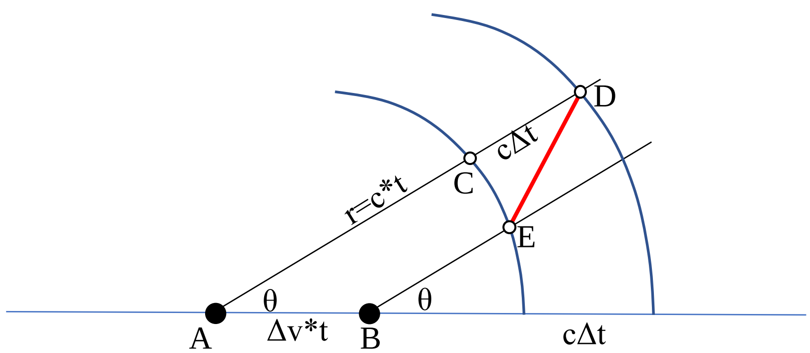
The general relativity theory explains a black hole as a singularity in the Schwarzchild metric and a tipping over of light cones, which suggests that black holes have no features other than mass, angular momentum and charge – the so-called 'black holes have no hair'. Based on quantum energy fluctuation, Hawking speculated that a black hole can emit particles to the outside region while also having an increase in negative energy within. This kind of analogy has some merit and the predictions may be useful for future research. However, with very limited information being able to be obtained from a black hole, it is difficult to know what is really going on inside the black hole, so the existing theories cannot be verified. When the technology advances enough to be able to detect the low frequency photons, we may gain more knowledge about black holes.

(6) Gravitational waves

Given that emitted photons are the carriers of gravity, it is easy to explain gravitational waves. Since there are attractions between photons and between the mass and the photons, an acceleration or oscillation of a mass will cause a subsequent change in photon density, and this density change will transmit from near to far away, just like a sound wave or an electromagnetic wave. Using this reasoning we can explain quantitatively the orbit decay of binary pulsars.

First, using Purcell's approach, we can derive the Larmor formula in the gravity field. As shown in Fig. 6, the stationary mass M at point A emits photons, but it suddenly accelerates and moves to point B in time t, so AB=$\Delta v*t$. Meanwhile, the photons emitted at point A travel to point C, AC=$r=c*t$. When the photon emitted at B travels to point E in time Δt, the photon at point C travels to point D, and CD=$c*\Delta t$. If mass M did not move, its gravity strength F should be measured by the distance between consecutive photons at C and D. With the acceleration, the distance between the consecutive photons is DE, so the acceleration causes a change in gravity strength F', which can be indicated by the perpendicular distance CE=AB*$sin\theta=\Delta v*t*sin\theta$.

Fig. 6 Gravity field change caused by motion of mass



Based on Newton's gravitational formula, the gravity strength without motion is $F=GM/r^2$, which is indicated by the length of CD. The change in gravity strength caused by motion can be calculated as:

$$F'=\text{CE/CD}*F=\Delta v*t*sin\theta/(c*\Delta t)*GM/r^2.$$

Letting $a=\Delta v/\Delta t$ be the acceleration rate and noticing $r=ct$, we have:

$$F'=a*t*sin\theta*GM/(c/r^2)=GM*a*r*sin\theta/(c^2r^2)=(GM*a/(c^2r))*sin\theta \qquad (35)$$

The formula shows that an acceleration of a mass can generate an additional gravitational field that has a strength that is only a small fraction (a/c^2) of the stationary gravitational force, but reduces slowly at the rate of $1/r$. The term $sin\theta$ gives the component of the acceleration that is perpendicular to the light ray.

Recall that based on equation (24), gravitational force is related to the area of photon density ρ_S by $F=4\pi G^*\rho_S$. Equation (35) shows that the area of the photon density involved in acceleration-caused gravity strength is:

$$\rho_S=M^*a^*sin\theta/(4\pi c^2 r)$$

Using ρ_S and F', we calculate the rate of energy loss. Since the energy loss rate is defined as energy loss per second, that is the work done by force F' pushing the photons in a unit of area to travel for 1 second. Given the number of photons in a unit area ρ_S and the speed of light c, the work done in 1 second or energy loss rate is:

$$\Delta E = W = F' * \rho_S * c = \frac{4\pi G(M*a*\sin\theta)^2*c}{(4\pi c^2 r)^2} = \frac{GM^2a^2sin^2\theta}{4\pi c^3r^2} \tag{36}$$

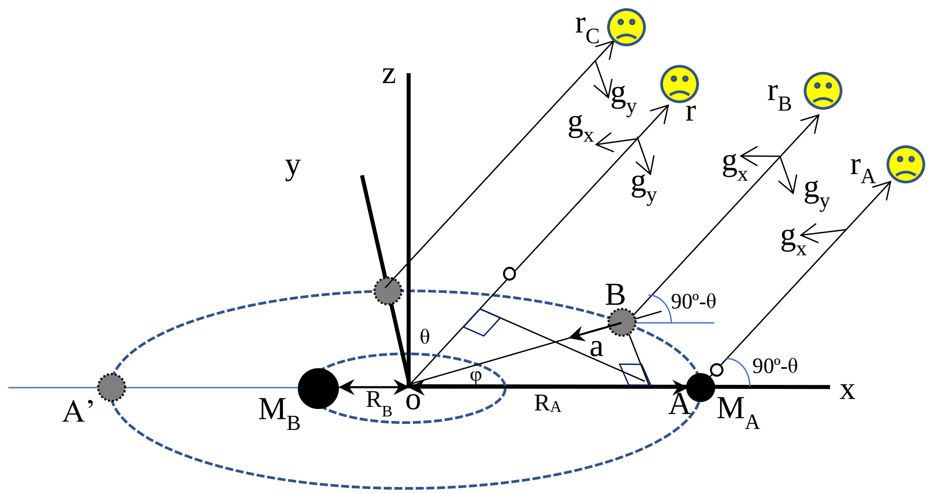
This is the energy loss rate at given angle θ and radius r. To calculate the total energy loss rate due to acceleration, we integrate ΔE over the sphere of radius r to calculate the energy flux passing through the whole sphere:

$$\frac{dE}{dt} = \int_0^{2\pi} d\varphi \int_0^{\pi} Wr^2 sin\theta d\theta = \int_0^{2\pi} d\varphi \int_0^{\pi} \frac{GM^2a^2sin^3\theta}{4\pi c^3r^2} r^2 d\theta = \frac{2GM^2a^2}{3c^3} \tag{37}$$

Thus, we have derived the Larmor formula for the gravitational field without invoking the Poyting vector (with the analogical approach, one can derive the same formula from the Poyting vector).

Second, we apply the same approach to binary pulsars. The orbits of binary pulsars are often ellipses. To simplify the case, we assume two stars orbiting in circles of different sizes around the common centre of the mass, as shown in Fig. 7.

Fig. 7. Gravity field change caused by binary stars



Since the two stars move constantly, we use their centre of mass as the origin of a cartesian coordinate system xyz. The two stars of masses M_A and M_B orbit in circles with radii of R_A and R_B, respectively, in the xy plane. The observer is in the xz plane (i.e. azimuth angle φ=0) and the observation direction has a polar angle θ. Since the observer is very far away from the pulsar, in order to simplify the analysis, the paths that the light travels from the centre of the mass and from the different positions of the star are drawn as parallel lines, e.g. the rays of r, r_A, r_B and r_C.

For a binary system, the two stars are always on the opposite side of the orbits and move in the opposite direction, so they have the same angular speed: $\omega_A=\omega_B=\omega$. Their mass and radii satisfy $M_AR_A=M_BR_B$. Letting $R=R_A+R_B$, $M=M_A+M_B$, we have $R_A=M_BR/M$, $R_B=M_AR/M$.

If both stars are on the x axis (e.g. one star at point A), their (radial) acceleration forces are on the x axis, and the situation is quite similar to what we discussed in Fig. 6, so the acceleration will generate a gravity strength in the direction of g_x, which is perpendicular to the direction of r_A. If these stars are on the y axis, the acceleration rates will generate a gravity strength in the direction of g_y, which is also perpendicular to the direction of r_C. For a general case at position B, the acceleration will generate gravity strength in both perpendicular directions, g_x and g_y. These two directions together with direction r_B form another cartesian coordinates system.

Since their acceleration rates depend on their position in the orbits, using position B as an illustration, we can express the acceleration rate for the star of mass M_A as:

$$\overrightarrow{a_A} = \overrightarrow{a_{Ax}} + \overrightarrow{a_{Ay}} = a_A \vec{x} cos\varphi + a_A \vec{y} sin\varphi = R_A \omega^2 \left(\vec{x} cos\omega t_A + \vec{y} sin\omega t_A \right)$$

Here t_A is the retarded time at M_A, which is the observation time t minus the time the light covers r_A, namely $t_A=t-r_A/c$. Based on equation (35), we can calculate the gravitational field generated by this acceleration. Noticing that the x component of acceleration and the observation ray forms an angle of 90º-θ and the y component is perpendicular to the observation ray, we have:

$$\overrightarrow{F_A} = \frac{GM_A}{c^2 r_A} * \left(\overrightarrow{g_x} a_{Ax} \sin\left(\frac{\pi}{2} - \theta\right) + \overrightarrow{g_y} a_{Ay} sin\omega t_A \sin\frac{\pi}{2} \right)$$

$$= \frac{GM_A \omega^2 R_A}{c^2 r_A} \left(\overrightarrow{g_x} cos\omega t_A \sin\left(\frac{\pi}{2} - \theta\right) + \overrightarrow{g_y} sin\omega t_A \sin\frac{\pi}{2} \right)$$

$$= \frac{GM_A \omega^2 R_A}{c^2 r_A} \left(\overrightarrow{g_x} cos\omega t_A cos\theta + \overrightarrow{g_y} sin\omega t_A \right)$$

Since the star moves constantly, t_A and r_A will keep changing and make the problem complicated. With some correction terms, we can change the measurement of t_A and r_A to t_R and r related to the centre of mass O, where t_R is the retarded time at O: $t_R=t–r/c$.

$$\overrightarrow{F_A} = \frac{GM_A \omega^2 R_A}{c^2 r} \left(\overrightarrow{g_x} cos\omega t_R cos\theta + \overrightarrow{g_y} sin\omega t_R \right)$$

This equation is correct if M_A has the above-described acceleration and stays at point O, i.e. a two dimensional (2D) harmonic oscillator at point O. Since this 2D harmonic oscillator is actually orbiting in the orbit, two corrections are necessary to approximate this circulation movement.

The first approximation of the orbiting movement is to separate the 2D oscillator to two 1D oscillators and let them oscillate on the x and y axes simultaneously. Compared with the stationary oscillator at the centre, the impact of the position variation on the x and y axes should be accounted for. The projections of the two positions of M_A on the opposite side of the orbit (e.g. B and B') on the x axis are $-R_A cos\varphi$ and $R_A cos\varphi$, respectively. This gives a length variation of $2R_A cos\varphi$, which corresponds to the length variation on the g_x axis of $2R_A cos\varphi cos\theta$ or a time difference of light travel of $\Delta t_{A1x}=2cos\varphi cos\theta\, R_A/c$. During this time interval, the gravitational force F_A will generate a momentum P_{A1}:

$$\overrightarrow{\mathrm{p_{A1x}}} = \overrightarrow{\mathrm{F_A}}\Delta\mathrm{t_{A1x}} = \frac{2GM_A\omega^2R_A^2}{c^3r}\left(\overrightarrow{\mathrm{g_x}}cos\omega\mathrm{t_R}cos\theta + \overrightarrow{\mathrm{g_y}}sin\omega\mathrm{t_R}\right)cos\omega\mathrm{t_R}cos\theta$$

Similarly, the projections of the two positions of M_A on the opposite side of the orbit (e.g. B and B') on the y axis (and thus on the g_y axis) are $-R_Asin\varphi$ and $R_Asin\varphi$, respectively. This gives a length variation of $2R_Asin\varphi$, which corresponds to a time difference of light travel of $\Delta t_{A1y}=2sin\varphi R_A/c$. During this time interval, the gravitational force F_A will generate a momentum P_{A1y}:

$$\overrightarrow{\mathrm{p_{A1y}}} = \overrightarrow{\mathrm{F_A}}\Delta\mathrm{t_{A1y}} = \frac{2GM_A\omega^2R_A^2}{c^3r}\left(\overrightarrow{\mathrm{g_x}}cos\omega\mathrm{t_R}cos\theta + \overrightarrow{\mathrm{g_y}}sin\omega\mathrm{t_R}\right)sin\omega\mathrm{t_R}$$

The total momentum accounting for the length variation due to the oscillation on the x and y axes is:

$$\overrightarrow{\mathrm{p_{A1}}} = \overrightarrow{\mathrm{p_{A1x}}} + \overrightarrow{\mathrm{p_{A1y}}} = \frac{2GM_A\omega^2R_A^2}{c^3r}\left(\overrightarrow{\mathrm{g_x}}cos\omega\mathrm{t_R}cos\theta + \overrightarrow{\mathrm{g_y}}sin\omega\mathrm{t_R}\right)(cos\omega\mathrm{t_R}cos\theta + sin\omega\mathrm{t_R})$$

Using the radius and mass relations of binary stars, $R_A=M_BR/(M_A+M_B)$, and denoting $\mu=M_AM_B/(M_A+M_B)^2$, the above equation can be rewritten as:

$$\overrightarrow{\mathrm{p_{A1}}} = \overrightarrow{\mathrm{p_{A1x}}} + \overrightarrow{\mathrm{p_{A1y}}} = \frac{2G\mu M_B\omega^2R^2}{c^3r}\left(\overrightarrow{\mathrm{g_x}}cos\omega\mathrm{t_R}cos\theta + \overrightarrow{\mathrm{g_y}}sin\omega\mathrm{t_R}\right)(cos\omega\mathrm{t_R}cos\theta + sin\omega\mathrm{t_R})$$

By convention, μM_B is called the reduced mass of M_B. Similarly, for star B we have its momentum generated by acceleration, so we can derive a similar result for $\overrightarrow{\mathrm{p_B}}$. Summing up the momentum for both stars, we have the total momentum due to the accelerations:

$$\overrightarrow{\mathrm{p_1}} = \frac{2G\mu M\omega^2R^2}{c^3r}\left(\overrightarrow{\mathrm{g_x}}cos\omega\mathrm{t_R}cos\theta + \overrightarrow{\mathrm{g_y}}sin\omega\mathrm{t_R}\right)(cos\omega\mathrm{t_R}cos\theta + sin\omega\mathrm{t_R})$$

or

$$\overrightarrow{\mathrm{p_1}} = \frac{G\mu M\omega^2R^2}{c^3r}\Big(\overrightarrow{\mathrm{g_x}}((cos2\omega\mathrm{t_R} + 1)cos^2\theta + sin2\omega\mathrm{t_R}cos\theta) + \overrightarrow{\mathrm{g_y}}(sin2\omega\mathrm{t_R}cos\theta + cos2\omega\mathrm{t_R} - 1)\Big)$$

Differentiating the above equation with respect to t_R, we have momentum gradient or the gravitational force in the direction perpendicular to r:

$$\overrightarrow{\mathrm{F}_1} = \frac{2G\mu M\omega^3 R^2}{c^3 r}\left(\overrightarrow{\mathrm{g_x}}(-sin2\omega\mathrm{t_R}cos^2\theta + cos2\omega\mathrm{t_R}cos\theta) + \overrightarrow{\mathrm{g_y}}(cos2\omega\mathrm{t_R}cos\theta - sin2\omega\mathrm{t_R})\right)$$

Applying the same method as we used to derive equation (37), the Larmor formula for gravitational field, we can calculate the energy loss rate at the radius r and angle θ as:

$$\frac{dE1}{d\mathrm{t_R}} = \frac{4\pi G\mu^2 M^2\omega^6\mathrm{R}^4 c}{4\pi^2\mathrm{c}^6 r^2}((-sin2\omega\mathrm{t_R}cos^2\theta + cos2\omega\mathrm{t_R}cos\theta)^2 + (cos2\omega\mathrm{t_R}cos\theta - sin2\omega\mathrm{t_R})^2)$$

Using the fact that over one or many cycles the average values $<sin^2 2\omega t_R>=<cos^2 2\omega t_R>=0.5$, $<sin2\omega t_R cos2\omega t_R>=0$, we have:

$$\frac{dE1}{d\mathrm{t_R}} = \frac{G\mu^2 M^2\omega^6\mathrm{R}^4}{2\pi\mathrm{c}^5 r^2}(cos^4\theta + 2cos^2\theta + 1) = \frac{G\mu^2 M^2\omega^6\mathrm{R}^4}{2\pi\mathrm{c}^5 r^2}(cos^2\theta + 1)^2 \qquad (38)$$

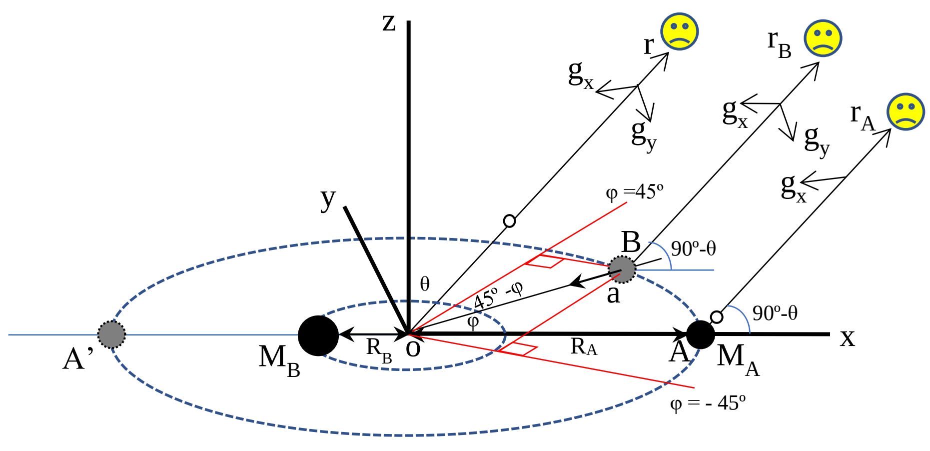
This energy loss corresponds to the oscillations of plus polarizations in general relativity:

$$h_+(t) = \frac{2\mathrm{G}^2\mathrm{M_A}\mathrm{M_B}}{\mathrm{c}^4 rR}(cos^2\theta + 1)cos2\omega\mathrm{t_R}$$

Our approximation so far has a shortcoming: the oscillations on the x and y axes do not have symmetrical results. For example, the impact of the oscillation on the y axis does not relate to the observation angle θ but the impact of the oscillation on the x axis does. To overcome this shortcoming, we add two additional oscillation directions at $\varphi= \pm\pi/4$ and covariant shift the observation angle to these axes, as shown in Fig. 8. The projections of positions B' and B on the $+\pi/4$ axis are $-R_A cos(\pi/4-\varphi)$ and $R_A cos(\pi/4-\varphi)$, respectively, and on the $-\pi/4$ axis are $-R_A sin(\pi/4-\varphi)$ and $R_A sin(\pi/4-\varphi)$, respectively. The variations in length on these axes are $2R_A cos(\pi/4-\varphi)$ and $2R_A sin(\pi/4-\varphi)$, respectively. Since the observation angles on both axes are $\pi/2-\theta$, the expected momentums due to the time variation of $\Delta t_{A2}=2(cos(\pi/4-\varphi)+sin(\pi/4-\varphi))R_A/c$ are:

$$\overrightarrow{\mathrm{p_{A2}}} = \overrightarrow{\mathrm{F_A}}\Delta\mathrm{t_{A2}} = \frac{2GM_A\omega^2 R_A^2}{c^3 r}\left(\left(\overrightarrow{\mathrm{g}_{+\frac{\pi}{4}}}\cos\left(\frac{\pi}{4}-\varphi\right)\sin\left(\frac{\pi}{4}-\theta\right)\right)\left(\cos\left(\frac{\pi}{4}-\varphi\right)+\sin\left(\frac{\pi}{4}-\varphi\right)\right)\right.$$
$$\left. + \left(\overrightarrow{\mathrm{g}_{-\frac{\pi}{4}}}sin\left(\frac{\pi}{4}-\varphi\right)\sin\left(\frac{\pi}{4}-\theta\right)\right)\left(\cos\left(\frac{\pi}{4}-\varphi\right) + \sin\left(\frac{\pi}{4}-\varphi\right)\right)\right)$$

Fig. 8. Gravity field change caused by binary stars



Similarly, for star B we can derive its momentum generated by acceleration. Summing up the momentum for both stars, we have the total momentum due to the accelerations:

$$\overrightarrow{p_2} = \frac{2G\mu M\omega^2 R^2}{c^3 r}\left(\left(\overrightarrow{g_{+\frac{\pi}{4}}}\cos\left(\frac{\pi}{4}-\varphi\right)cos\theta\right)\left(\cos\left(\frac{\pi}{4}-\varphi\right)+\sin\left(\frac{\pi}{4}-\varphi\right)\right.\right.$$
$$\left.+\left(\overrightarrow{g_{-\frac{\pi}{4}}}sin\left(\frac{\pi}{4}-\varphi\right)cos\theta\right)\left(\cos\left(\frac{\pi}{4}-\varphi\right)\ +\sin\left(\frac{\pi}{4}-\varphi\right)\right)\right)$$

Noting $\varphi=\omega t_R$, we have

$$\overrightarrow{p_2} = \frac{G\mu M\omega^2 R^2}{c^3 r}\left(\overrightarrow{g_{+\frac{\pi}{4}}}\left(cos2\left(\frac{\pi}{4}-\omega t_R\right)+1+sin2\left(\frac{\pi}{4}-\omega t_R\right)\right)cos\theta\right.$$
$$\left.+\overrightarrow{g_{-\frac{\pi}{4}}}\left(sin2\left(\frac{\pi}{4}-\omega t_R\right)+cos2\left(\frac{\pi}{4}-\omega t_R\right)+1\right)cos\theta\right)$$

Differentiating the above equation with respect to t_R, we have the momentum gradient or gravitational force in the perpendicular direction:

$$\overrightarrow{F_2} = \frac{2G\mu M\omega^3 R^2}{c^3 r}\left(\overrightarrow{g_{+\frac{\pi}{4}}}\left(-sin2\left(\frac{\pi}{4}-\omega t_R\right)+cos2\left(\frac{\pi}{4}-\omega t_R\right)\right)cos\theta\right.$$
$$\left.+\overrightarrow{g_{-\frac{\pi}{4}}}\left(cos2\left(\frac{\pi}{4}-\omega t_R\right)-sin2\left(\frac{\pi}{4}-\omega t_R\right)\right)cos\theta\right)$$

Applying the same method as we used to derive equation (37), we can calculate the energy loss rate at the radius r and angle θ as:

$$\frac{dE2}{d\mathrm{t_R}} = \frac{G\mu^2 M^2 \omega^6 \mathrm{R}^4}{\pi \mathrm{c}^5 r^2}\left(-sin2\left(\frac{\pi}{4} - \omega \mathrm{t_R}\right) + cos2\left(\frac{\pi}{4} - \omega \mathrm{t_R}\right)\right)^2 4cos^2\theta$$

Using the fact that over one or many cycles the average values $<sin^2 2\omega t_R> = <cos^2 2\omega t_R> = 0.5$, $<sin2\omega t_R cos2\omega t_R> = 0$, we have:

$$\frac{dE2}{d\mathrm{t_R}} = \frac{G\mu^2 M^2 \omega^6 \mathrm{R}^4}{2\pi \mathrm{c}^5 r^2} * 4cos^2\theta$$

This energy loss corresponds to the oscillations of the cross polarizations in general relativity:

$$h_\times(t) = \frac{4\mathrm{G}^2 \mathrm{M_A M_B}}{\mathrm{c}^4 rR} cos\theta sin2\omega \mathrm{t_R}$$

Adding the two parts of the energy loss rates together, we have:

$$\frac{dE}{d\mathrm{t_R}} = \frac{G\mu^2 M^2 \omega^6 \mathrm{R}^4}{2\pi \mathrm{c}^5 r^2}((cos^2\theta + 1)^2 + 4cos^2\theta)$$

Integrating over the sphere of radius r, we have:

$$\frac{dE}{d\mathrm{t_R}} = \frac{G\mu^2 M^2 \omega^6 \mathrm{R}^4}{2\pi \mathrm{c}^5} \iiint \frac{(cos^2\theta + 1)^2 + 4cos^2\theta}{r^2} r^2 sin\theta dr d\theta d\varphi$$

$$= \frac{G\mu^2 M^2 \omega^6 \mathrm{R}^4}{2\pi \mathrm{c}^5} \frac{64\pi^2}{5} = \frac{32\pi}{5} \frac{G\mu^2 M^2 \omega^6 \mathrm{R}^4}{\mathrm{c}^5}$$

This result is exactly the same as that derived from the general relativity theory. With this formula, one can derive the results for the orbit decay and the in-spiralling of the pulsar. The procedure and results should be the same as that achieved in previous research (e.g. Hulse and Taylor, 1975[xvii], Damour and Deruelle, 1986[xviii], Taylor and Weisberg, 1989[xix], Weisberg and Taylor, 2003[xx], Stair, Thorsett, Arzoumanian, 2004[xxi], Cheng, 2005[xxii]).

It is worth mentioning that by using an analogy of electromagnetic waves, Hilborn (2018)[xxiii] derived results for orbiting binaries that had the same form as that obtained from general relativity but with a coefficient that was 16 times smaller. The direction of the gravitational wave was also different from the results achieved in the general relativity theory. The excellent work of Hilborn shows the similarity between a gravitational wave and an electromagnetic wave. However, the difference in coefficients and in directions of gravitational waves also shows that the analogical approach does not reveal the mechanism behind the results so the results may not be correct in every aspect.

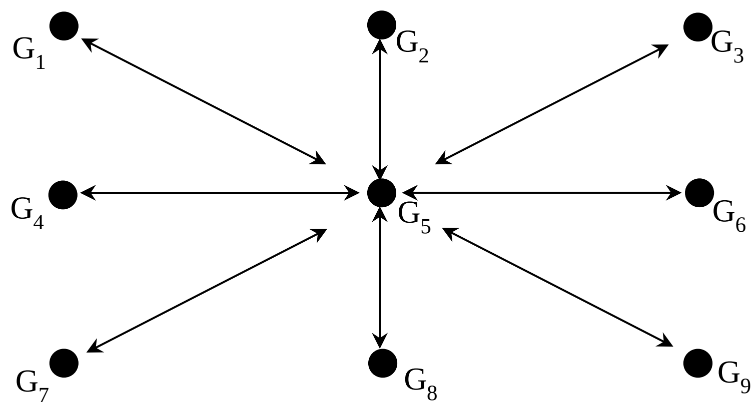
(7) Expanding homogeneous universe

Astronomical observation shows that the universe is expanding and is largely homogeneous. This does not seem to be compatible with any theory that uses Newtonian gravity. Our theory is an advance from both the Newtonian gravity theory and Einstein's relativity theory, so we need to explain the seeming incompatibility.

The common wisdom is that Newtonian gravity leads to a centralized universe. One of the reasons that Einstein was not happy with Newtonian gravity was that its centralized universe was not consistent with the almost homogeneous distributions of stars in the sky. The conclusion of the centralized Newtonian universe may come from the reasoning of the gravity pulling effect or from the analogy of solar systems, galaxies, and classes of galaxies, but it may not be valid if one considers the universal movements and vastness of space. If an object has low or no speed, it will be captured by a mass centre or enter an orbiting system. With the vast space and constant moving objects, a homogeneous static universe can be explained with ease.

In Fig. 9 we consider nine vastly separated galaxies. There are gravitational forces between any two galaxies. For example, galaxy G5 is attracted by eight other galaxies. If these forces are largely cancelled out, G5 will not move into any of the other galaxies. Even if these forces are not cancelled out and G5 moves around near its position, G5 may not be dragged into any other galaxies because these forces are very weak due to the vast distance between them. The reader may be quick to point out: how about the other eight galaxies at the edge? They would move to G5 because there are no other forces to balance the gravity from G5. This would happen if they are at the edge, i.e. if the universe has a limited size. As such, Newtonian gravity is consistent only with an infinite homogeneous universe. Is the universe finite? We cannot answer this question because current technology does not allow us to see the edge. Most people believe the universe is infinite.

Fig. 9. Homogeneous universe in an unlimited space



The acceleration of the expanding universe is currently explained by the repulsive force from dark matter. This explanation runs into problems with both Newtonian gravity and Einstein's general relativity as the repulsive force would dampen the gravitational force between matters. To explain the expanding universe, a cosmological tensor is added to the Einstein equation and a very small cosmological constant is assigned so that it will only affect the object on a cosmological scale. The same thing could also be done to the Newtonian gravity equation to explain the expanding universe. In short, the cause of the expanding universe is still a mystery that does not favour either Newtonian gravity or Einstein's general relativity.

4. Conclusions

The paper presents a new theory to explain relativistic phenomena in a straightforward way. Using a few plausible assumptions, the paper obtains the same or very similar relativistic results that explain the results predicted by Einstein, such as relative mass, the mass-energy formula, the energy momentum equation, the transverse Doppler effect, gravitational red shift, planetary precession, the deflection angle of light in gravitational lensing, the orbits around a black hole, and the strength and directions of gravitational waves (orbit decay of pulsars). The paper not only helps a wide audience to understand the relativistic phenomena, but also asks a deep question: what is the real cause of these phenomena?

The advance of the new theory from Einstein's relativity theory centred on a methodologic change. Using an analogical approach, Einstein captured the similarities of different objects, events and systems, so many of his predictions were proven correct. However, his approach is unable to pinpoint the cause of the phenomena, which may result in mild contradictions or inaccuracies. The new approach is to propose a plausible specific cause and mechanism for all relativistic events. This paper achieves this target and explains them logically and quantitatively. The paper also suggests experiments and observations that can test the new theory. If the tests support the new theory, we are confident that the causes and mechanisms proposed in the paper are likely to be true. If future experimental work contradicts our prediction, we can examine the gap between the theory and the experimental results and propose a new theory. Either way, our understanding of the relativistic phenomena will improve in our journey to find the truth.

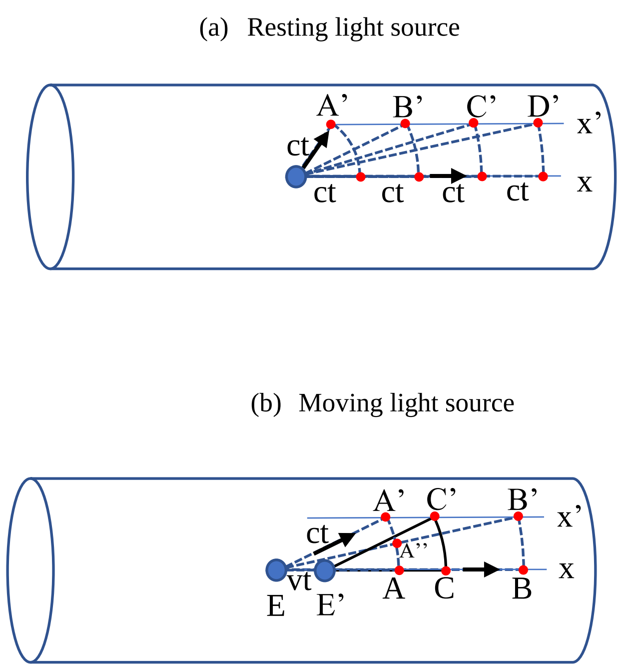
References:

Appendix: Measuring photon density of a moving emitter

A. Line density of photons

As explained in the text, the process for measuring sphere volume is not suitable for measuring the overall photon density of a moving emitter, so we use the process for measuring cylindric volume, as shown in Fig. A1. When the emitter moves along the axis of the cylinder, the photon density on the cross section is not affected (to be explained in section B), so we focus on line density of photons at (or parallel to) the direction of cylinder axis x.

Fig. A1 Longitude density of photons



Panel (a) in Fig. A1 shows the case of a resting light source. Assume an emission period of *t* (or emission frequency *f*=1/*t*) and photon speed of *c*. When the photons are travelling, such as in the direction indicated by the x axis, the photons are distributed evenly with the distance of the neighbouring photons being ct. On any line parallel to the x axis, e.g. the x’ line, we can find the same number of photons corresponding to those on the x axis by drawing the arcs of various radii. The photons on x’ are not distributed evenly; however, since the length of the cylinder can go to infinity, we can always find a length where x’ contains the same number of photons as those on the x axis. As such, the average distance of the neighbouring photons is also *ct*, i.e. <A’B’>=(A’B’ + B’B’ + C’D’ +…)/*n*=*ct.*

The situation of a moving emitter is shown in panel (b). For ease of demonstration, we only show two pairs of photons and increase the distance of the neighbouring photons. If the emitter is stationary, the first set of photons travel to B and B’ and the second set of photons travel to A and A’. As discussed above, AB=<A’B’>=*ct*.

Now we consider that, after emitting the first set of photons, the emitter starts to move to the right at speed v. When it starts to emit the second set of photons, the emitter has travelled a distance vt to E’ and the first set of photons have travelled to A and A’’. When the second set of photons travel to C and C’, the first set of photons arrive at B and B’. Because of the parallel shift of shape EAA’ to A’CC’, it is obvious that AC=A’C’=vt. As a result, the new distance of the neighbouring photons is BC=<B’C’> =ct-vt.

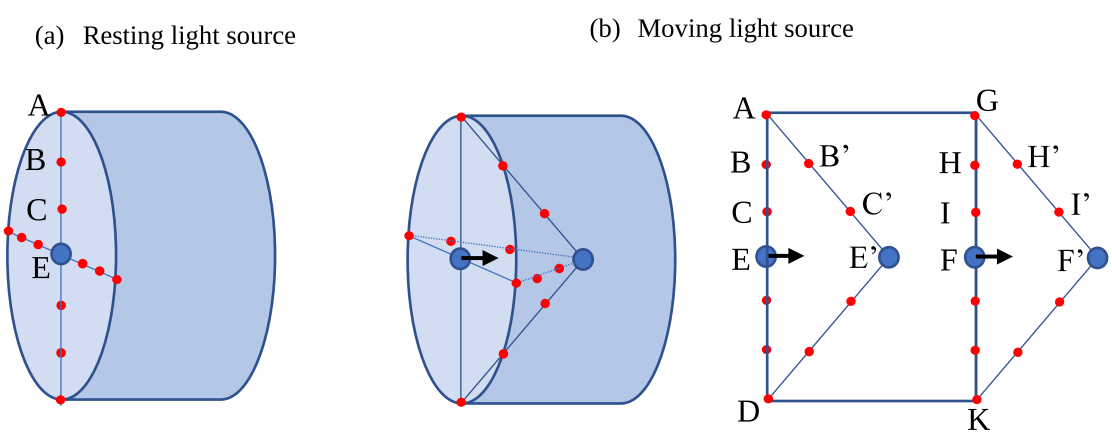
The line density of photons at the right of the emitter at E’ can be expressed as $d1=c/(ct-vt)=cf/(c-v)$. Similarly, the density at the left of E’ is $d2=c/(ct+vt)=cf/(c+v)$, and the average density is $d=(d1+d2)/2=c^2f/(c^2-v^2)$.

B. Photon density at the perpendicular cross section.

Since the cross section is perpendicular to the movement of the emitter, by intuition we know the photon density at the cross section should not be affected by the speed of the emitter. This section provides a discussion on this.

The case of the resting emitter is shown in panel (a) of Fig. A2. Although the photon density is higher in the area close to the emitter, the photon densities on any radius in the cross section are equal.

Fig. A2 Photon density at the cross section



When the emitter starts to move to the right, the photons with a speed perpendicular to the speed of the emitter will also move to the right of the cross section. The cross section deforms to a surface of a cone, as shown in panel (b). Although the line density of photons on the cone surface (e.g. density on AE') is less than that on the perpendicular cross section (e.g. density on AE) when the emitter is stationary, i.e. AB'>AB, B'C'>BC, etc., the projection of the photon density of the cone surface onto the cross section will be unchanged, e.g. the projection of the photons at B' and C' in panel (b) will be B and C. The longitudinal cross section view in panel (b) shows that the perpendicular cross sections AD and GK deform to cone surfaces ADE' and GKF', respectively. Considering a continuing process, when the new cone KGF' is created at the right, a cone of the same size will move into the cylinder ADKG from the left , so calculation of cross section photon density should use the projections of the photons in the cone surface onto the cross section. In other words, the average cross section density is the flattened density on the cone surface. As a result, the projected photon density on the perpendicular cross section and thus the cross-section density are not affected.

---

[i] Ives, H. E.; Stilwell, G. R., 1938, An experimental study of the rate of a moving atomic clock. Journal of the Optical Society of America. 28 (7): 215.

[ii] Ives, H. E.; Stilwell, G. R., 1941, An experimental study of the rate of a moving atomic clock II. Journal of the Optical Society of America. 31 (5): 369.

[iii] Hay, H.J., J. P. Schiffer, T. E. Cranshaw, and P. A. Egelstaff, 1960, Measurement of the Red Shift in an Accelerated System Using the Mössbauer Effect in Fe57. Physical Review Letters, 4, 165

[iv] Walter Kündig. 1963, Measurement of the Transverse Doppler Effect in an Accelerated System. Phys. Rev. 129, 2371

[v] Kaivola, M., Poulsen, O., Riis, E., Lee, S., 1985. Measurement of the relativistic Doppler shift in neon. Physical Review Letters. Volume 54, Issue 4. P. 255-258. 1079-7114 (electronic). 0031-9007 (printed). DOI: 10.1103/PhysRevLett.54.255.

[vi] Klein F, Laroche T, Cardenas ME, Hofmann JF, Schweizer D, Gasser SM, 1992, Localization of RAP1 and topoisomerase II in nuclei and meiotic chromosomes of yeast. The Journal of Cell Biology, 117(5):935-48

[vii] Jennison, R.C. and Davies, P.A., 1974, Reflection from a transversely moving mirror, Nature, 248: 660-661

[viii] Davies, P.A., and Jennison, R.C. 1975, Experiments involving mirror transponders in rotating frames, Journal of Physics A: Math, 8: 1390

[ix] Thim, H. W. 2003, Absence of the relativistic Doppler shift at microwave frequencies, IEEE Instrumentation and Measurement, 52(5): 1660–1664

[x] Sfarti, A. 2010, Comment on "The case of absence of transverse Doppler effect", IEEE Instrumentation and Measurement, 59(2): 494-495

[xi] Schutz, B., A first course in general relativity, 2nd edition, Cambridge university press

[xii] Chou, C. W. Hume, D. B. Rosenband, T. and Wineland, D. J. 2010, Optical Clocks and Relativity, Science, 329(5999): 1630-1633, DOI: 10.1126/science.1192720

[xiii] Rosenband, T., D. B. Hume, P. O. Schmidt, C. W. Chou, A. Brusch, L. Lorini, W. H. Oskay, R. E. Drullinger, T. M. Fortier, J. E. Stalnaker, S. A. Diddams, W. C. Swann, N. R. Newbury, W. M. Itano, D. J. Winelandand , J. C.

Bergquist, 2008, Frequency Ratio of Al+ and Hg+ Single-Ion Optical Clocks; Metrology at the 17th Decimal Place, Science, 319:1808-1812.

[xiv] Chou, C. W., D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband, 2010, Frequency Comparison of Two High-Accuracy Al+ Optical Clocks, Phys. Rev. Lett. 104, 070802

[xv] Dyson, F. W., A. S. Eddington, C. Davidson, 1920, A determination of the deflection of light by the Sun's gravitational field from observations made at the total eclipse of May 29, 1919, Phil. Trans. R. Soc. Lond. A 220, 291–333.

[xvi] Qian-shen Wang, Xin-she Yang, Chuan-zhen Wu, Hong-gang Guo, Hong-chen Liu, and Chang-chai Hua, 2000, Precise measurement of gravity variations during a total solar eclipse, Phys. Rev. D 62, 041101(R)

[xvii] Hulse, R. A., Taylor, J. H., 1975, Discovery of a pulsar in a binary system, Astrophysical Journal, Vol. 195, p. L51-L53

[xviii] Damour, T., Deruelle, N., 1986, General relativistic celestial mechanics of binary systems. II. The post-Newtonian timing formula, Ann. Inst. Henri Poincaré Phys. Théor., 44(3): 263 - 292

[xix] Taylor, J. H., Weisberg, J. M., 1989, Further Experimental Tests of Relativistic Gravity Using the Binary Pulsar PSR 1913+16, Astrophysical Journal 345:434

[xx] Weisberg, J., & Taylor, J. H. 2003, The relativistic binary pulsar B1913+16, in ASP Conf. Ser. 302, Radio Pulsars, ed. M. Bailes, D. J. Nice, & S. E. Thorsett (San Francisco: ASP), 93

[xxi] Stairs, I. H., S. E. Thorsett, and Z. Arzoumanian. Measurement of Gravitational Spin-Orbit Coupling in a Binary-Pulsar System, Phys. Rev. Lett. 93, 141101

[xxii] Cheng, T., 2005, Relativity, Gravitation and Cosmology, Oxford University Press.

[xxiii] Hilborn, R., 2018, Gravitational waves from orbiting binaries without general relativity, American Journal of Physics, 86, 186-197.