The Deflection of Light, Einstein's Equivalence Principle

and

the Implications of Maxwell-Newton Approximation

C. Y. Lo

Applied & Pure Research Institute

17 Newcastle Dr., Nashua, NH 03060, U.S.A.

March 2003

Abstract

It is shown that, just as the gravitational red shifts, the bending of light-rays can be considered as a consequence of the equivalence principle. This is achieved by deriving the Maxwell-Newton Approximation for massive matter from the equivalence principle together with related physical requirements such as special relativity and Newton's theory as a first approximation of the time-time component of the space-time metric, and thus beyond Schiff's approach. This first order approximation of the space-time metric, which is covariant with respect to the Lorentz transformation, is the foundation of Einstein's radiation formula that is supported by the Hulse-Taylor experiments on the binary pulsars. Concurrently, it is shown that the Maxwell-Newton Approximation is generally a first order approximation of an Einstein equation that is compatible with the 1995 update equation. Thus, although the 1915 Einstein equation, which is compatible with the Maxwell-Newton Approximation for a static problem, has no physical solution for a dynamic situation, general relativity remains a viable theory. Moreover, since the Maxwell-Newton Approximation is derived from general physical considerations, further confirmation by Stanford Gravity Probe-B is expected.

1. Introduction

Einstein [1,2] proposed three testable consequences of general relativity, namely: a) the gravitational red shifts; b) bending of light-rays; c) the perihelion of Mercury. The perihelion of Mercury is not a prediction but a confirmation because the perihelion had been known from observations. The prediction of gravitational red shifts is directly related to Einstein's equivalence principle only, but not his field equation. However, the bending of light-rays was considered as a confirmation of Einstein's field equation because this calculation used a space-time metric obtained from the Einstein equation.

In 1911, Einstein [3] failed to show that the equivalence principle is sufficient to derive the right amount of light bending. Recently, Moreau et al. [4] interpret this as "the failure of the equivalence principle to fully account for the total deflection of electromagnetic waves in the sun's gravitational field." They believe that "the total rate of deflection includes the effect of spacetime curvature". It seems, their interpretation is due to carelessness and negligence the details of Einstein's work. As Eddington [5,6] observed, that few understand general relativity, in particular Einstein's equivalence principle, adequately.

Einstein [1,2] has made clear that the deflection rate of light depends on the first order derivatives of the metric elements. On the other hand, the spacetime curvature depends on the second order derivatives [2]. Both the Ricci spacetime curvatures of the Schwarzschild metric and the flat metric of special relativity, are zero in vacuum. However, the curvature of Einstein's approximate metric is not zero, but the derived bending rate is essentially identical to the bending rate due to the Schwarzschild metric [7]. These all show that the deflection rate is not directly related to the spacetime curvature (see also Section 3).

Thus, such an interpretation of light bending is clearly incorrect. In 1915 Einstein [1,2] has already made clear that it is necessary to use the equivalence principle to calculate the deflection rate. From Einstein's calculation [2], the 1911 failure is only on accounting for the full effects of the equivalence principle [1)]. It will be shown that the bending of light-rays can be considered as a consequence of the equivalence principle, just as the gravitational red shift is currently considered.

In this paper, it will be shown that the equation of "linearized gravity" can actually be considered as a direct consequence of the equivalence principle (Section 4). Moreover, this derivation is independence of the 1915 Einstein equation and the linearized harmonic gauge condition, and thus beyond Schiff's approach (see § 5). Such an independency is necessary because the 1915 Einstein equation has been found incompatible with the weak field equation for a dynamic situation [8]. In addition, this independency is consistent with the fact that a physical gauge is unique for a given frame of reference. The resulting equation is called the Maxwell-Newton Approximation because this is a set of equations similar to Maxwell's and is based on Newton's theory as a first order approximation, and thus this set of equations may not be valid if the source is not massive matter. To this end, let us first discuss the relationship between Einstein equation and the weak field equation obtained by linearization.

2. The Questions of Linearized Gravity for Weak Gravitation

In physics, a solution of a field equation for an isolated physical system should be finite and bounded [2)] in amplitude although the non-zero domain of the field may not be bounded. For instance, a reason is that the field energy density is usually associated with to the amplitude of the field. For a weak source of gravity, as suggested by Einstein [1,2], general relativity is not an exception. The principle of causality[3)] [9,10] requires that a weak source should produce weak gravity and therefore a bounded (in amplitude) metric. Mathematically, however, one does not know whether there is any solution of weak gravity for a weak source because the non-linear field equation may not be valid. Although the Einstein equation has bounded static solutions, it may not have bounded (in amplitude) dynamic solutions since there is no bounded (in amplitude) plane-wave solution [10].

For weak gravity, the terms in the Ricci curvature tensor would have a definite order of deviations from the flat metric η_{ab}. Linearized gravity is based on the assumption that higher order terms are negligible when compare with the sum of the first order terms. Thus, even the solution is bounded and weak, the linearized field equation may not relate to such an approximation at all [11]. For a nonlinear differential equation, the second order terms can play a decisive physical role [8]. If the solution is not bounded, then it is meaningless to consider the order of deviations and the scheme of linearization breaks down completely.

Thus, to see the validity of this ordering scheme, it is extremely important to check whether the resulting linear field equation is compatible with the Einstein equation. If they are incompatible, this means at least one of the equations is invalid. Based on existing evidences, this can be the linearized equation or the Einstein equation [8,11].

To discuss linearized gravity and related problems, it would be useful to outline the linearization steps:

1) The non-linear Einstein's field equation of 1915 version [1,2] is

$$G_{ab} \equiv R_{ab} - ½ R g_{ab} = -KT(m)_{ab} \qquad (1a)$$

where its source $T(m)_{ab}$ is the energy-stress tensor for massive matter and can depend on the space-time metric g_{ab} and R_{ab} is the Ricci curvature tensor. Note that for G_{at} (a = x, y, z, t), there is no second order time derivative [7,12]. Thus, for eq. (1a), the initial condition of a Cauchy problem is restricted by four constraints. Also, the source has been formally extended to a general energy-stress tensor T_{ab} [13].

Due to a certain freedom of choice of coordinates, a gauge condition must be added. The harmonic gauge condition is

$$\frac{\partial}{\partial x^a}(|g|^{1/2}g^{ab}) = 0 \qquad (1b)$$

where g is the determinant of metric g_{ab}. Note that Eddington [6] rejected the notion of an arbitrary gauge, which can lead to the acceptance of unphysical solutions [9,14,15] and a violation of the equivalence principle[4)].

2) Linearized gravity is a result of neglecting all terms of explicit second order of deviations—a problematic approach since a first order term can be zero in a finite region [8]. Then, eq. (1a) and eq. (1b) are respectively linearized to be

$$G^{(1)}{}_{ab} = -K\,T_{ab}, \qquad \text{where } G^{(1)}{}_{ab} \equiv G_{ab} - G^{(2)}{}_{ab} = \frac{1}{2}\partial^c\partial_c\bar{\gamma}_{ab} + H^{(1)}{}_{ab}, \tag{2a}$$

where

$$H^{(1)}{}_{ab} \equiv -\frac{1}{2}\partial^c[\partial_b\bar{\gamma}_{ac} + \partial_a\bar{\gamma}_{bc}] + \frac{1}{2}\eta_{ab}\partial^c\partial^d\bar{\gamma}_{cd};$$

and

$$\partial^c\bar{\gamma}_{ac} = 0, \tag{2b}$$

where $\bar{\gamma}_{ab} = \gamma_{ab} - \frac{1}{2}\eta_{ab}(\eta^{cd}\gamma_{cd})$ and $\gamma_{ab} = g_{ab} - \eta_{ab}$. Validity of (2b) was implicitly questioned by Einstein [2].

3) The gauge (2b) implies that $H^{(1)}{}_{ab} = 0$ and that eq. (2a) is reduced to (see also [6] & [8])

$$\frac{1}{2}\partial^c\partial_c\bar{\gamma}_{ab} = -K\,T_{ab}, \tag{3a}$$

Eq. (3a) is the well-known linearized field equation, whose structure is similar to that of Maxwell's equation. Thus, its Cauchy initial condition can be arbitrary. This manifests that eq. (3a) and eq. (1a) may not be compatible for dynamic problems.

An implicit assumption is the existence of bounded solution for weak gravity. This is a common implicit assumption in physics, and is generally true for a linear equation. However, for a non-linear equation, this assumption must be verified. For the case of static field, this assumption is valid as shown by the harmonic solution and the isotropic solution [7]. However, for a dynamic case, such an assumption is incorrect [8]. This is also the reason that no bounded dynamic solution has ever been found.

An inhomogeneous solution of eq. (3a) is

$$\bar{\gamma}_{ab}(x^i, t) = -\frac{K}{2\pi}\int\frac{1}{R}T_{ab}[y^i, (t - R)]d^3y, \quad \text{where } R^2 = \sum_{i=1}^{3}(x^i - y^i)^2. \tag{3b}$$

Solution (3b) can reproduce Newtonian gravity.

In linearized gravity, there are two independent equations namely: the linear field equation (3a) and the linearized gauge condition (2b). It should be noted that, however, these equations together require

$$\partial^a T_{ab} = 0, \tag{4}$$

the linearized conservation law that implies no radiation [16,17]. Since eq. (4) was used in deriving Einstein's radiation formula, there is an apparent inconsistency. But, one can use the conservation law,

$$\nabla^a T_{ab} \equiv \partial^a T_{ab} + \Gamma_{ac}{}^a T^c{}_b - \Gamma_{cb}{}^a T_a{}^c = 0 \tag{5}$$

instead of eq. (4). However, since eq. (2a) directly implies eq. (4), there is little hope that Einstein's radiation formula can be justified with linearized gravity. Also, when the source is not a massive energy tensor [18], eq. (3a) may be incompatible with weak gravity. However, since eqs. (3a) and (5) imply $H^{(1)}{}_{ab}$ is of the second order [6], one might hope that eq. (3a) would give the first order approximation of eq. (1a). But, eq. (3a) has been proven to be not a dynamic approximation [8] for eq. (1a).

This incompatibility means that eq. (1a) is invalid for dynamic problems if eq. (3a) is a first order approximation. But, eq. (3a) is supported by all observations [8,19]. On the other hand, validity of eq. (1) for a dynamic situation has been questioned by Gullstrand [20,21] as early as 1921. Also, from studying the gravity of electromagnetic waves that Einstein equation must be modified because of the necessity of having an anti-gravity coupling to accommodate the gravitational wave [18].

Historically, Fock [22] pointed out that, for harmonic coordinates, there are divergent logarithmic deviations from expected linearized behavior of the radiation. But, since some vacuum solutions are not logarithmic divergent [23], Fock's result was regarded as merely due to the method used. However, the correct interpretation should be that eq. (1) has no bounded dynamic solutions [8]. In 1936 Einstein and Rosen [24,25] are the first to discover the excluding of the gravitational wave solution. However, this was not generally accepted because unbounded "time-dependent" solutions unrelated to a dynamic source, had been considered as valid in physics because the principle of causality [18] was not recognized. In concord with Einstein et al., it was proven in 1991 [9,26] that there is no bounded gravitational plane-wave solution.

Moreover, Hu, Zhang & Ding [12] found that a perturbative calculation on radiation depends on the approach chosen. Mathematically, this manifests that such a dynamic solution is unbounded. Also, the post-Newtonian approaches (1/c-expansions) are fraught with serious internal consistency problems [27] because they often lead, in higher approximations, to divergent integrals. Being an extension of the linearization, the post-Minkowskian approach (K-expansions) has logarithmic divergence.

Nevertheless, this incompatibility may mean, as Einstein [2] remarked, only that the source of eq. (1a) is subjected to modification. Although Einstein has claimed that his theory is logically complete [21], Klein [28] pointed out that there is no satisfactory proof of rigorous validity of eq. (1a). Also, gauge eq. (2b) is not an integral part of general relativity [6,15] since Newton's theory was derived with just static weak gravity. Hogarth [29] conjectured that the source tensor should be non-zero in vacuum to account for gravitational radiation. In other words, general relativity would still be valid (see Section 4).

3. Implicit Assumptions in Deriving the Gravitational Red Shifts and the Bending of Light.

Based on eq. (3) and that the static massive energy-stress tensor $T(m)_{ab}$,

$$T(m)_{tt} = \sigma \quad \text{and} \quad T(m)_{ab} = 0 \text{ otherwise,} \tag{6a}$$

where σ is the mass density, Einstein obtained the static spacetime metric,

$$ds^2 = c^2\left(1 - \frac{K}{4\pi}\int dV_0 \frac{\sigma}{R}\right)dt^2 - \left(1 + \frac{K}{4\pi}\int dV_0 \frac{\sigma}{R}\right)(dx^2 + dy^2 + dz^2), \tag{6b}$$

If Einstein's equivalence principle is valid, under the influence of only gravity, for an observer P at point (x, y, z, t) with an arbitrary velocity v, there is a co-moving local Minkowski coordinate system (X, Y, Z, T) [5)],

$$ds^2 = c^2dT^2 - dX^2 - dY^2 - dZ^2. \tag{7}$$

Consider the case of v = 0, then for observer P one has

$$c^2dT^2 = c^2\left(1 - \frac{K}{4\pi}\int dV_0 \frac{\sigma}{r}\right)dt^2 ; \tag{8a}$$

and thus

$$(dX^2 + dY^2 + dZ^2) = \left(1 + \frac{K}{4\pi}\int dV_0 \frac{\sigma}{r}\right)(dx^2 + dy^2 + dz^2), \tag{8b}$$

It follows from eq. (8a), the red shift for a frequency, which would have been ν_0 in a Minkowski space

$$\nu = \nu_0(g_{tt})^{1/2} = \nu_0\left(1 - \frac{K}{4\pi}\int dV_0 \frac{\sigma}{r}\right) \tag{9a}$$

and

$$L = \frac{d\ell}{dt} = c(1 - \frac{K}{4\pi}\int dV_0 \frac{\sigma}{r}), \qquad \text{where} \qquad d\ell^2 = (dx^2 + dy^2 + dz^2), \tag{9b}$$

is the light speed of the space at (x, y, z, t) since $(dX^2 + dY^2 + dZ^2)/dT^2 = c^2$.

If we imagine the sun, of mass M, concentrated at the origin of our system of coordinates, then a ray of light, traveling parallel to the x_3-axis, in the x_1-x_3 plane, at a distance Δ from the origin, will be deflected, in all by an amount [1,2]

$$\alpha = \int_{-\infty}^{+\infty} \frac{1}{L}\frac{\partial L}{\partial x_1} dx_3 = \frac{KM}{2\pi\Delta} \tag{9c}$$

where $M = \int \sigma dV_0$ (see also Appendix). Thus, the red shifts, the light speeds, and the light deflection are consequences eq. (3). It should be noted that Δ is also measured "in the sense of Euclidean geometry [1]." In fact, all his predictions are expressed in the sense of Euclidean geometry. This is due to that Einstein's Riemannian space necessarily has a Euclidean-like structure.

To understand the Euclidean-like structure, we must first clarify what "measure" means in relation to Einstein's equivalence principle. In Einstein's theory, the measuring instruments are resting but in a free fall state [1,2]. From Einstein's equivalence principle, time dilation and space contraction are obtained. Based on such measurements, Einstein believed, "In the general theory of relativity, space and time cannot be defined in such a way that differences of the spatial coordinates can be directly measured by the unit measuring-rod, or differences in the time coordinates by a standard clock". However, if the measuring instruments are attached to the frame of reference, since the measuring instruments and the coordinates being measured are under the same influence of gravity, a Euclidean-like structure emerges as if gravity did not exist. From metric (6), it is clear that this operational defined Euclidean-like structure is a necessary complimentary structure of Einstein's physical space. In other words, the physical meaning of space-time coordinates has already been provided in the theoretical framework of general relativity.

On the other hand, eq. (9a) has been accurately approximated from the equivalence principle with Newtonian gravity [3]. An implicit assumption is that the Newtonian potential provides a first order approximation of the exact g_{tt} In other words; the Poisson equation would give a first order approximation. It will be shown in next Section that this assumption together with special relativity and the equivalence principle would lead to eq. (3), and this is independent of the Einstein equation (1a). Then, eq. (9b) and therefore (9c) can also be considered as directly from the equivalence principle.

Some theorists considered Einstein's equivalence Principle as non-essential since his predictions can be obtained without its explicit usage [7]. Another problem is that the same predictions can be obtained from different metrics such as the Schwarzschild solution and the harmonic solution [30]. Currently, this fact is considered as due to the gauge invariance. However, this is incorrect since the second order effect of gravitational red shifts is different from these metrics. This would means that Einstein's equivalence principle is needed to find out the realistic metric from those that can pass the four standard tests [7,19].

A realistic gauge is unique for a given frame because the time dilation and space contractions are measurable according to equations (7) and (8). Therefore, there must be some ways to identify the correct solution although the precise methods are not yet known. (Note also that the criticisms of Whitehead [31] to Einstein's understanding are actually not valid for general relativity.) Since the Maxwell-Newton Approximation can be derived independent of the harmonic gauge condition, it is a new criterion for selecting a physically valid metric for the case of static gravity due to a spherical symmetric massive source.

4. The Equivalence Principle, the Maxwell-Newton Approximation, and the Field Equation

Moreover, there is a need to build solution (6b) on a solid foundation. First, the bending of light is only approximately a static problem. Second, since eq. (1a) has no bounded dynamic solutions, it is incompatible with the principle of causality, which requires the existence of a solution for weak gravity. Since linear eq. (3a) is supported by experiments, to reaffirm the validity of general relativity, one must show that eq. (3a) is compatible with the theoretical framework of relativity. This need is independent of whether eq. (3a) is an approximation of a modified Einstein equation.

In general relativity [2], there are three basic assumptions namely: 1) the principle of equivalence; 2) the principle of covariance; (Note, as rectified by Lo [15] that covariance, in accord with Einstein [2], is restricted to space-time coordinate systems which are compatible with the equivalence principle. Although a tensor equation is mathematically covariant, a tensor equation alone has no meaning in physics unless the space-time coordinates are valid in physics.) and 3) the field equation whose source can be modified. Thus, it is possible that eq. (3a) can be an approximation of a modified Einstein equation.

It will be shown first that eq. (3a) is compatible with the other two assumptions. Note that eq. (3a) is invariant with respect to the Lorentz transformations. Moreover, solution (3b) implies that eq. (3a) is compatible with the principle of causallity and Einstein's notion of weak gravity. Thus, eq. (3a) as an approximation for the specified coordinate system is compatible with the requirement of covariance. It remains to show that eq. (3a) is compatible with the equivalence principle.

The equivalence principle implies that the geodesic equation,

$$\frac{d^2x^\mu}{ds^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0, \tag{10a}$$

where

$$\Gamma^\mu{}_{\alpha\beta} = (\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta})g^{\mu\nu}/2, \quad \text{and} \quad ds^2 = g_{\mu\nu}dx^\mu dx^\nu. \tag{10b}$$

is the equation of motion for a neutral particle. In comparison with Newtonian theory, one obtains [1,2] that

$$\Phi \approx c^2 g_{00}/2. \tag{11}$$

Since the gravitational potential Φ satisfies the Poisson equation $\Delta\Phi = 4\pi\kappa\rho$, according to the correspondence principle, one has the field equation, $\Delta g_{00} \approx 8\pi\kappa\, c^{-2}\, T_{00}$, where $T_{00} \approx \rho$, the mass density and κ is the coupling constant.

Then, according to special relativity and the Lorentz invariance, one has

$$\frac{1}{2}\partial_c\partial^c g_{ab} = \frac{1}{2}\partial_c\partial^c \gamma_{ab} = -4\pi\kappa c^{-2}[\alpha T(m)_{ab} + \beta Z(m)\eta_{ab}], \tag{12a}$$

where

$$Z(m) = \eta^{cd}T(m)_{cd}\,, \qquad \alpha + \beta = 1, \tag{12b}$$

α and β are constants, and $T(m)_{ab}$ is for massive matter. Eq. (12) is a field equation for the first order approximation (as assumed) for weak gravity of moving particles. An implicit condition is that the flat metric η_{ab} is the asymptotic limit at infinity.

To have the exact equation, since the left hand side of eq. (12a) does not satisfy the covariance principle, one must search for a tensor whose difference from $\partial_c\partial^c\gamma_{ab}/2$ is of second order in κc^{-2}.

In Riemannian geometry, it has been proven [7] that the curvature tensor "$R^{\lambda}{}_{\mu\nu\kappa}$ is the only tensor that can be constructed from the metric tensor and its first and second derivatives, and is linear in the second derivatives." Thus, Einstein identified the Ricci curvature tensor R_{ab} ($\equiv R^{\lambda}{}_{a\lambda b}$) as the required tensor; and if R_{ab} includes no first order term other than $\partial_c\partial^c\gamma_{ab}/2$, the exact field equation would be

$$R_{ab} = X^{(2)}{}_{ab} - 4\pi\kappa c^{-2}\,[\alpha\, T(m)_{ab} + \beta\, T(m)g_{ab}], \quad \text{where } \; T(m) = g^{cd}T(m)_{cd} \tag{13a}$$

is the trace of $T(m)_{ab}$, $X^{(2)}{}_{ab}$ is a second order unknown tensor chosen by Einstein to be zero. However, a non-zero $X^{(2)}{}_{ab}$ may be needed to ensure eq. (12) as an approximation of eq. (13a) [8].

Now, let us examine R_{ab} further whether the above physical requirement can be valid. Let us decompose

$$R_{ab} = R^{(1)}{}_{ab} + R^{(2)}{}_{ab}, \tag{14a}$$

where

$$R^{(1)}_{ab} = \frac{1}{2}\partial_c\partial^c\gamma_{ab} - \frac{1}{2}\partial^c[\partial_b\gamma_{ac} + \partial_a\gamma_{bc}] + \frac{1}{2}\partial_a\partial_b\gamma\ , \tag{14b}$$

and $R^{(2)}_{ab}$ consists of higher order terms. To be compatible with the requirement that eq. (12) provides the first order approximation, the sum of other linear terms must be of second order. This is feasible because a second order term can be obtained by a suitable linear combination of $\partial^c\gamma_{cb}$ and $\partial_b\gamma$. From eq. (12a), one obtains

$$\partial_c\partial^c(\partial^a\gamma_{ab}) = -8\pi\kappa c^{-2}[\alpha\partial^a T(m)_{ab} + \beta\partial_b Z(m)]\ . \tag{15a}$$

It is clear that $4\pi\kappa c^{-2}\partial^a T(m)_{ab}$ is of second order but $4\pi\kappa c^{-2}\partial_b Z(m)$ is not. Now, from (15a), one obtains

$$\frac{1}{2}\partial_c\partial^c(\partial^a\gamma_{ab} + C\partial_b\gamma) = -4\pi\kappa c^{-2}[\alpha\partial^a T(m)_{ab} + (\beta + 4C\beta + C\alpha)\partial_b Z(m)] \tag{15b}$$

It follows eq. (14b) and eq. (12b) that, for the other terms to be of second order, one must have

$$\beta + 4C\beta + C\alpha = 0, \qquad 2C + 1 = 0, \qquad \text{and} \qquad \alpha + \beta = 1. \tag{15c}$$

The solution of eq. (15c) is $C = -1/2$, $\alpha = 2$, and $\beta = -1$. Thus,

$$\frac{1}{2}\partial_c\partial^c\gamma_{ab} = -8\pi\kappa c^{-2}[T(m)_{ab} + \frac{1}{2}Z(m)\eta_{ab}]\ . \tag{12c}$$

for the first order approximation, is determined to be the field equation of massive matter. Eq. (12c) is independent of the exact form of the unknown second order term $X^{(2)}_{ab}$. Then, it is possible to obtain from eq. (13a) an equation different from eq. (1a),

$$G_{ab} \equiv R_{ab} - \frac{1}{2}g_{ab}R = -K[T(m)_{ab} - Y^{(1)}_{ab}], \qquad K = 8\pi\kappa c^{-2} \tag{13b}$$

where

$$KY^{(1)}_{ab} = X^{(2)}_{ab} - \frac{1}{2}g_{ab}\{X^{(2)}_{cd}\,g^{cd}\},$$

is of second order. The conservation law $\nabla^c T(m)_{cb} = 0$ and $\nabla^c G_{cb} \equiv 0$ implies also

$$\nabla^a Y^{(1)}_{ab} = 0. \tag{13c}$$

$Y^{(1)}{}_{ab}$ has been proven as the gravitational energy tensor of $t(g)_{ab}$ although its exact form is not clear [8].

Note that Einstein obtained the same values for α and β by considering eq. (13a) after assuming $X^{(2)}{}_{ab} = 0$. An advantage of the approach of considering eq. (12) and eq. (14) is that the assumption $X^{(2)}{}_{ab} = 0$ is not needed. The anti-gravity coupling of $t(g)_{ab}$ is due to the radiation formula [8] and explains the failure of eq. (1a). Note that the existence of such an anti-gravity coupling implies that the energy conditions of the singularity theorems [32], is unrealistic.

In short, the field equation of first order approximation is eq. (3a) if the source is restricted to massive matter. This means

$$\frac{1}{2}\partial^c\partial_c \bar{\gamma}_{ab} = -KT(m)_{ab} . \tag{12d}$$

An implicit condition is that the flat metric η_{ab} is the asymptotic limit. However, the problematic gauge (2b) is not used even just formally. The solution of eq. (12) is compatible with the equivalence principle as demonstrated by Einstein [2] in his calculation of the bending of light. Thus, eq. (12) is compatible with eq. (6), and the derivation is self-consistent. Moreover, since this derivation of (12) is independent of the harmonic gauge, this approximation is a criterion that all the static solution must be satisfied. *Accordingly, the long-standing Schwarzschild solution is unfortunately rejected as a physically valid solution.*

One might argue that Einstein equation (1a) could be "derived" from a linear equation more general than eq. (12a), if one regards the gravitational field as a spin 2 field coupled to the energy tensor [19,33]. However, such a "pure" theoretical approach is not really consistent with Newton's theory and related observations because Newton's theory is not gauge invariant. More important, in such a "proof", the existence of bounded dynamic solutions for eq. (1a) must be invalidly assumed.

While gravitational red shifts support any eq. (12a), the bending of light supports only eq. (12c). Thus, this choice can be considered as required also by observations. One may object this on the ground that an unverified assumption that the gravitational effect due to light is negligible has been used in the calculation. Nevertheless, such an assumption can be considered as justified from observation. In fact, the theoretical justification of this negligence has been proven to be valid [18]. Since eq. (12) as a first order approximation is proven on the basis of the equivalence principle, the experimental supports of this equation are verifications of this principle. Now, it is clear that general relativity can be derived logically, in spite of Klein's [28] criticism.

Note that eq. (12) need not be valid if the source is not massive [9,10]. On the other hand, theoretical consistency requires the source of eq. (13) must be extended to include the energy-stress tensor of electromagnetism [3]. Since the source of eq. (12) is limited to massive energy-stress tensor, it is more appropriate to call eq. (12) as the Maxwell-Newton Approximation [8].

5. Comments on Comparison with Schiff's Approach

In 1960, Schiff [33], based on also the equivalence principle and special relativity, concluded that it is actually possible to obtain the light deflection, as well as the red shift, in a valid manner without using the full theory. To avoid confusion among some readers, perhaps it is necessary to clarify the differences from Schiff's approach.

Schiff's approach utilized the time dilation and spatial contraction in special relativity. Since spatial contraction in special relativity is one dimensional, his method cannot handle a multi-dimensional contraction, which solution (6a) represents. Essentially, Schiff's work is a reinterpretation of the Schwarzschild solution, which has only a one-dimensional spatial contraction. Consequently, Schiff's result has not reached the level of a field equation for a first order approximation. Moreover, since Schiff's extension was based on the source less cases, it needs not be valid for gravity with a source. This derivation of the Maxwell-Newton Approximation shows that Schiff's extension is actually not generally valid.

Schiff's objective is on experimental tests of a static case of the general theory of relativity because he did not know that the 1915 Einstein equation does not have a dynamic solution [8]. He also was unaware of that a physical gauge is unique for a given frame. On the other hand, in order to interpret the binary pulse experiments [8], our objective is to justify the Einstein equation of 1995 update and in particular the related Einstein radiation formula, and thus general relativity is still a viable theory.

To this end, one must justify a first order approximation, which is covariant with respect to the Lorentz transformation, because Einstein's radiation formula must deal with particles with changing speeds. Because the geodesic equation is considered, we actually use the infinitesimal version of Einstein's equivalence principle [1,2]. Consequently, our results are far richer than Schiff's, and include exposing the shortcoming of Einstein's derivation of his equation. Not only the field equation for such a first order approximation is obtained, but also it has been shown that the 1995 update equation is completely justified within the theoretical framework of general relativity.

6. Discussions and Conclusions

The starting point of this analysis is to see the full effects of the equivalence principle in connection with the bending of light-rays. Another problem is that since the 1915 Einstein equation (1a) does not have bounded dynamic solutions, one cannot assume that the Einstein equation gives a solution at the static limit. In other words, Gullstrand's criticism [20] on the calculation of Mercury perihelion turns out to be completely justified. Therefore, it is not clear whether the linearized equation (3a), is compatible with general relativity. Furthermore, since this linear equation, as a first order approximation is supported by experiments, it is not clear that general relativity is still a viable theory. In fact, some believe that general relativity is not valid [35].

Based on Einstein's equivalence principle, we obtain the Maxwell-Newton Approximation eq. (12) within the theoretical framework that leads to general relativity, and thus show that the bending of light-rays, just as the gravitational shifts, is a consequence of the equivalence principle. Moreover, general relativity has been proven remaining a viable theory since the Maxwell-Newton Approximation can be derived from the theoretical framework of general relativity. Concurrently, it is shown that the Maxwell-Newton approximation is a first order approximation of the 1995 Einstein equation update (13), and therefore a modification of Einstein equation is justified within general relativity although the exact form remains to be clarified.

Moreover, eq. (12) and the update eq. (13) are compatible with the principle of causality, and the fact that gravitational waves carry energy. Physically, this explains why eq. (1a) is dynamically incompatible with the notion of weak gravity, and cannot have a dynamic solution. These two equations form the basis of Einstein's radiation formula, which requires a dynamic solution of the first order [8]. However, the question of a dynamic Einstein equation to justify the calculation for the perihelion of Mercury remains unsolved because the second order terms of a dynamic spacetime metric are required.

Nevertheless, some difficult theoretical problems have been solved. For instance, although eq. (3a) as a first order approximation is supported by all experiments so far, validity of eq. (3a) may still be questioned. Some considered [12] the wave component in g_{at} (for a = x, y, z, or t) as artificially induced by the harmonic gauge since eq. (3a) can be incompatible with eq. (1a). Now, eq. (12) for weak gravity is justified independently by physical principles that lead to general relativity. Thus it is on the most solid theoretical ground possible within the theoretical framework of general relativity. No longer anyone can maintain the hope that eq. (1a) may have a dynamic solution.

In spite of these, Christodoulou and Klainerman [36] claimed to have constructed bounded gravitational (unverified) waves. In addition to their claims being obviously incompatible with the findings of others, as Perlick [37] pointed out, validity of their mathematics is dubious. Further investigation shows that their proofs are mathematically incomplete and therefore invalid [38]. Furthermore, their presumed initial conditions are incompatible with Einstein's radiation formula and are unrelated to dynamic sources. In short, they simply have mistaken an invalid assumption (which does not satisfy physical requirements) as a wave.

Thus, assuming the existence of bounded dynamic solutions for eq. (1a), though prevailing [39-41], is actually invalid. Furthermore, the derivation of the Einstein field equation is intimately related to experiments, and one no longer can maintain an illusion that general relativity is a consequence of pure thought. On the other hand, Klein's criticism [28] on a lack of satisfactory proof is no longer valid. Also, this theoretical development would be devastating to the viewpoint [19,30,42] that accepted Einstein's equation but rejected Einstein's principles. It is hoped that those theorists can [6)] adjust to the new reality.

There are differences between eq. (12) and eq. (13). Eq. (13), which includes implicitly the equation of motion for particles, describes the interaction between gravity and matter. However, eq. (12) provides only an approximate influence of matter to the gravitational field. There are important differences in physics between eq. (3a) and eq. (12) although they have the identical form in mathematics. Eq. (12) is explicitly for a massive source only, but eq. (3a) is also not valid if the source is an electromagnetic wave [9]. Whereas eq. (3a) is derived with a certain chosen gauge for eq. (la), eq. (12) is completely general. This is consistent with the fact that a physical gauge is unique for a given frame. This generality has two important implications. 1) The Maxwell-Newton Approximation eq. (12) will be further confirmed by the Gravity Probe-B gyroscopes experiment [43]. 2) The Schwarzschild solution would also be rejected by this experiment.

7. Acknowledgments.

The author gratefully acknowledges stimulating discussion with Professors J. E. Hogarth, P. Morrison, C. Wong, K. Young, and X. Yu. The author would like to thank the referees for useful suggestions, and valuable comments. This work is supported in part by Innotec Design, Inc., California, USA.

Appendix: The Deflection of Light and Theoretical Consistency in General Relativity

In the calculations of the star light bending, both wave and particle approaches [1,7] give the same deflection. This shows that Huyghen's principle is compatible with Einstein's equivalence principle from which the null geodesic equation for a light is derived. This manifests also that the particle-wave duality is not necessarily excluded from a classical theory. However, there are theoretical problems to be settled due to the requirement of theoretical self-consistent.

Note that almost all authors consider the light as consisting of massless particles, photons, which obey the null geodesic equation [7,16,19,36]. But, Einstein [2] alone maintained the derivation of light bending through the Huyghen's principle, which is not directly related to the motion of a particle (although he did adapt the method of calculating the perihelion advance by others.) Perhaps, Einstein had already aware of the theoretical problem that the Einstein-Maxwell equations cannot produce the geodesic equation for the light [18], although his field equation would include the equation of motion for massive matter.

A geodesic equation cannot be generated from the electromagnetic energy tensor, since it generates the Lorentz force. Nevertheless, if non-electromagnetic radiation is included in the light, its energy tensor can generate a geodesic equation. Thus, in general relativity, the light bending by following a geodesic necessitates that the energy-stress tensor of light (i.e., the totality of related free radiations) must have a form, which is different from that of the electromagnetic energy tensor [18].

Experimentally electromagnetic waves and photons are inseparable although particles and waves are conceptually distinct [44]. Quantum phenomena show conclusively that the light is distinct from just a classical electromagnetic wave. On the other hand, it has never been shown conclusively that the light is identical to an electromagnetic wave in Maxwell's theory. What has been shown is only that the light is inseparably associated with an electromagnetic wave. Moreover, there is no physical principle, which dictates that the distinction between the light and an electromagnetic wave is limited to quantum phenomena.

The bending of light is an example since it is independent of its frequency, this makes it possible to treat this problem classically that there are non-quantum aspects of the light, which are different from the classical electromagnetic wave. A distinct energy-stress tensor for photons would support that there are intrinsic connections between general relativity and quantum theory. In spite of the fundamental differences, as Bohr [45] argued, there are possibly intrinsic connections among them. Note that both theories have their foundations on different aspects of a common physical phenomenon -- the velocity and quantum of light.

Moreover, an implicit assumption in calculating the light bending is that the gravity due to the light itself is negligible [2,7]. Although this is believable, to justify the calculation as a prediction of the theory, such an assumption must be proven self-consistently within general relativity. On the other hand, if an electromagnetic wave and the related photons are distinct objects as concluded, without a distinct photonic energy-stress tensor, one would encounter that Einstein's equation cannot have a physical solution [9]. Therefore, the question of a distinct photonic energy-stress tensor is actually an integral part of general relativity since the electromagnetic energy-stress tensor must be included in the source [32].

In short, it is necessary to have a distinct photonic energy-stress tensor such that the geodesic equation for photons is derived and that the gravity due to light is negligible. Since it has been proven that such an energy-stress tensor for photons can be found [18], it becomes justified to use solution (6b) and $ds^2 = 0$ to calculate the bending of light

For a static problem, eq. (3a) can be a first order approximation of Einstein equation (1a). To clarify this, let us consider the isotropic solution and the Schwarzschild external solution [7], which is

$$ds^2 = (1 - KMG/4\pi\rho)c^2dt^2 - (1 - KMG/4\pi\rho)^{-1}d\rho^2 - \rho^2 d\theta^2 - \rho^2 \sin^2\theta d\varphi^2, \quad \text{for} \quad \rho > KMG/4\pi \qquad \text{(A1)}$$

where

$$M = \int \sigma dV_0, \qquad \text{and} \qquad c^2KM/4\pi\rho = \kappa M/\rho \qquad \text{(A2)}$$

is the Newtonian potential for a spherical symmetric mass distribution. And the isotropic form is

$$ds^2 = [(1 - KMG/16\pi r)^2/(1 + KMG/16\pi r)^2]dt^2 - (1 + KMG/16\pi r)^4 (dx^2 + dy^2 + dz^2) \qquad \text{(A3)}$$

where

$$\rho = r(1 + KMG/16\pi r)^2, \quad \text{for} \quad \rho > KMG/4\pi \quad \text{and} \quad r = [x^2 + y^2 + z^2]^{1/2} \qquad \text{(A4)}$$

Clearly,

$$ds^2 = c^2 (1 - \frac{K}{4\pi}\int dV_0 \frac{\sigma}{R})dt^2 - (1 + \frac{K}{4\pi}\int dV_0 \frac{\sigma}{R})(dx^2 + dy^2 + dz^2), \qquad (6b)$$

where

$$R = [(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]^{1/2}, \qquad \text{and } \sigma = \sigma (x_0, y_0, z_0)$$

is the first approximation of (A3) since $\int dV_0\sigma / R$ = M/r, but is not the first order approximation of the Schwarzschild solution (A1). If (A1) is used to calculate the geodesic and to obtain the light deflection, to the first order approximation, one obtains [7]

$$\alpha' = KM/2\pi\rho_0 \qquad (A5)$$

where ρ_0 is the impact parameter. Since ρ_0 must be larger than the radius of the sun, ρ_0 is approximately equal to the corresponding Δ in eq. (9c). On the other hand, according to Huyghen's principle, if n is a direction perpendicular to the propagation of the light speed γ, the light-ray envisaged in the plane (γ , n) has the curvature - $\partial\gamma/\partial n$. It follows that a ray of light, traveling parallel to the x_3-axis in the x_1-x_3 plane, at a distance Δ from the origin, will be deflected, in all, by an amount [1,2]

$$\alpha = \int_{-\infty}^{+\infty} \frac{1}{L}\frac{\partial L}{\partial x_1} dx_3, \qquad \text{where} \quad L = c(1 - \frac{KM}{4\pi r}) \qquad (A6)$$

It follows [46]

$$\alpha \approx \int_{\theta=-\pi/2}^{\theta=\pi/2} \frac{1}{r^2}\cos\theta dx_3 = \frac{KM}{2\pi\Delta} \qquad \text{where} \quad \cos\theta = \Delta/r \qquad (A7)$$

(A4) also implies that $\alpha' \approx \alpha$. Of course, one can directly use (6b) for the geodesic equation, and get the same result, but there is little point to do a new calculation since existing results derived by Fock [30] can be used readily.

Thus, for these two different metrics, the two different approaches get the same approximation. The parameters ρ and r should be physical the same since these metrics have the same frame. However, the diffeomorphism (A4) requires that they are distinct.

ENDNOTES

1) This is supported by the fact that the Fermat's principle can be derived from the geodesic equation as shown by Pauli [47].
2) In this paper, the boundedness is always addressed to the amplitude unless otherwise specified.
3) The principle of causality [18] assumes that the causes of phenomena are identifiable. This principle is commonly used in symmetry consideration in electrodynamics. In general relativity, Einstein and subsequent theorists have used this principle implicitly on symmetry considerations [1,7]. Thus, in practice, the physical meaning of space coordinates has been used not

only subsequently, but also right at the beginning. Other consequences are that parameters unrelated to any physical cause in a solution is not allowed and that a dynamic solution must be related to a dynamic source.

4) A satisfaction of the equivalence principle requires that a time-like geodesic must represent a physical free falling. This means that the manifold is a physical space (-time), which models reality, and all (not just some) physical requirements are sufficiently satisfied. Thus, the equivalence principle may not be satisfied in a Lorentz Manifold [15] which implies only the necessary condition that the mathematical existence of a co-moving local Minkowski space.

5) Some theorists believed that Einstein's equivalence principle does not imply the necessary physical existence of a local Minkowski space (7) although this is clearly stated in eq. (106) in Einstein's book "The Meaning of Relativity". This misinterpretation [48] comes from in Pauli's version [47, p. 145], in which the existence of a local Minkowski space is only a mathematical possibility. Experimentally, it is known that the local Minkowski space exists in a spacecraft under only gravity. On the other hand, there is no theoretical or observational support for the substitution of Einstein's equivalence principle with Pauli's version. Currently, the fact that some predictions can be obtained from different metrics is considered as due to the gauge invariance. However, this is incorrect since the second order effect of gravitational red shifts is different from these metrics. This means that Pauli's version is inadequate, and Einstein's equivalence principle is needed to find out the realistic metric from those that can pass the four standard tests.

6) Max Planck once remarked, "A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it." Fortunately, it seems, mathematics is an exception to his rule.

REFERENCES

1. A. Einstein, Ann. Phys. (Leipig) **49**, 769-822 (1916).

2. A. Einstein, ***The Meaning of Relativity*** (Princeton Univ. Press, 1954), p. 63, p. 87 & p. 93.

3. A. Einstein, Ann. Phys. (Leipzig) **35**, 898-908 (1911).

4. W. Moreau, C. Purdie, and J. Wood, Am. J. Phys. **67** (3), 241-251, March (1999).

5. K. S. Thorne, ***Black Holes & Time Warps*** (Norton, New York, 1994).

6. A. S. Eddington, ***The Mathematical Theory of Relativity*** (Chelsa, New York, 1975), p. 129.

7. S. Weinberg, ***Gravitation and Cosmology*** (John Wiley, New York, 1972).

8. C. Y. Lo, Phys. Essays, 7 (4), 453 (1994).

9. C. Y. Lo, Phys. Essays, **10** (3), 424-436 (1997).

10. C. Y. Lo, in Proc. Seventh Marcel Grossmann Meeting on Gen. Relativ., Stanford Univ., 25-29 July 1994, ed. R. Jantzen & M. Keiser, Ser. ed. R. Ruffini, 525 (World Sci., Singapore, 1996).

11. C. Y. Lo, Astrophys. J., **455**, 421 (Dec. 20, 1995).

12. N. Hu, D.-H. Zhang, & H. G. Ding, Acta Phys. Sinica, **30** (8), 1003-1010 (Aug., 1981).

13. A. Einstein, 'Physics and Reality (1936)' in ***Ideas and Opinions*** (Crown, New York, 1954), p. 311.

14. D. Kramer, H. Stephani, E. Herlt, & M. MacCallum, ***Exact Solutions of Einstein's Field Equations***, ed. E. Schmutzer (Cambridge Univ. Press, Cambridge, 1980).

15. C. Y. Lo, Phys. Essays, **11** (2), 264-272 (1998).

16. R. M. Wald, ***General Relativity*** (The Univ. of Chicago Press, Chicago, 1984), p. 84-88.

17. X. Yu, Astrophysics and Space Science, **194**: 159-163 (1992).

18. C. Y. Lo, Phys. Essays, **12** (2), 226 (June, 1999).

19. H. C. Ohanian & R. Ruffini, ***Gravitation and Spacetime*** (Norton, New York, 1994).

20. A. Gullstrand, Ark. Mat. Astr. Fys. **16**, No.8 (1921); ibid. Ark. Mat. Astr. Fys. **17**, No.3 (1922).

21. A. Pais, ***'Subtle is the Lord ...'*** (Oxford University Press, New York, 1996).

22. V. A. Fock, Rev. Mod. Phys. **29**, 325 (1957).

23. H. Bondi, F. A. E. Pirani, and I. Robinson, Proc. R. Soc. London A **251**, 56-533 (1959).

24. A. Einstein & N. Rosen, J. Franklin Inst. **223**, 43 (1937).

25. ***The Born-Einstein Letters***, commentary by Max Born (Walker, New York, 1968), p. 125.

26. C. Y. Lo, in Proc. Sixth Marcel Grossmann Meeting On General Relativity, 1991, ed. H. Sato & T. Nakamura, 1496 (World Sci., Singapore, 1992).

27. L. Blanchet and T. Damour, Phil. Trans. R. Soc. Lond. A, **320**, 379-430 (1986).

28. O. Klein, Z. F. Physik **37**, 895 (1926).

29. J. E. Hogarth, "Particles, Fields, and Rigid Bodies in the Formulation of Relativity Theories", Ph. D. thesis 1953, Dept. of Math., Royal Holloway College, University of London (1953), p. 6.

30. V. A. Fock, ***The Theory of Space Time and Gravitation***, translated by N. Kemmer (Pergamon Press, 1964).

31. A. N. Whitehead, ***The Principle of Relativity*** (Cambridge Univ. Press, Cambridge, 1922).

32. S. W. Hawking and R. Penrose, Proc. Roy. Soc. London, A **314**, 529-548 (1970).

33. C. Y. Lo, Phys. Essays, **10** (4), 540-545 (1997); ibid., Astrophys. J. **477:** 700 (March 10, 1997).

34. L. I. Schiff, Am. J. Phys. **28** (4), 340-343 (1960).

35. H. Yilmaz, Nu. Cim. 107 B, 941 (1992).

36. D. Christodoulou and S. Klainerman, ***The Global Nonlinear Stability of the Minkowski Space*** (Princeton University Press, 1993).

37. Volker Perlick (republished with an editorial comment), Gen. Relat. Grav., 32, 761-763 (2000).

38. C. Y. Lo, Phys. Essays, 13 (1), 109-120 (March, 2000).

39. C. W. Misner, K. S. Thorne, & J. A. Wheeler, ***Gravitation*** (Freeman, San Francisco, 1973).

40. T. Damour, in ***300 Years of Gravitation*** (Cambridge Univ. Press., Cambridge, 1987).

41. T. Damour & B. Schmidt, J. Math. Phys. **31** (10), 2441-2453 (1990).

42. J. L. Synge, ***Relativity*** (North-Holland, Amsterdam, 1956).

43. C. W. F. Everitt et al. in Proc. Seventh Marcel Grossmann Meeting on Gen. Relativ., Stanford, July 1994, ed. R. Jantzen & M. Keiser, Ser. ed. R. Ruffini, 1533 (World Sci., Singapore, 1996).

44. V. F. Weisskopf, ***The Privilege of Being a Physicist*** (Freeman, San Francisco, 1988), p. 129.

45. P. A. Schilpp (ed.), ***Albert Einstein as Philosopher and Scientist*** (Evanston 1949), p. 152.

46. A. Einstein, H. A. Lorentz, H. Weyl, and H. Minkowski, ***The Principle of Relativity*** (Dover, New York, 1952).

47. W. Pauli, ***Theory of Relativity*** (Pergamon Press, London, 1958), p. 154, & p. 163.

48. J. Norton, "What was Einstein's Principle of Equivalence?" in Einstein's Studies Vol. 1: ***Einstein and the History of General Relativity,*** eds. D. Howard & J. Stachel (Birkhäuser, 1989).