Phase velocity and light bending in a gravitational potential

José-Philippe Pérez and Brahim Lamine

Université de Toulouse, UPS-OMP, IRAP and CNRS, IRAP, 14, Avenue Edouard Belin, F-31400 Toulouse, France

E-mail: jose-philippe.perez@irap.omp.eu and brahim.lamine@irap.omp.eu

Received 9 March 2018, revised 31 May 2018
Accepted for publication 6 June 2018
Published 5 July 2018

Abstract

In this paper we review the derivation of the light bending equation, obtained before the discovery of general relativity (GR). This is intended for students learning about GR, and for specialists who will shed new light on, and make new connections between, these historic derivations. Since 1915 it has been well known that the observed bending of light has two contributory factors: the first is directly deduced from the equivalence principle alone and was obtained by Einstein in 1911; the second comes from the spatial curvature of spacetime. In GR, these two components are equal, but other relativistic theories of gravitation can give different values to those contributions. In this paper we give a simple explanation, based on the wave-particle picture, of why the first term, which relies on the equivalence principle, is identical to the one obtained by a purely Newtonian analysis. In this context of wave analysis, we emphasize that the dependency of the velocity of light on the gravitational potential, as deduced by Einstein, concerns the phase velocity. Then we wonder whether Einstein could have envisaged already, in 1911, the second contribution, and therefore the correct result. We argue that considering a length contraction in the radial direction, along with the time dilation implied by the equivalence principle, could have led Einstein to the correct result.

Keywords: light bending, general relativity, equivalence principle

1. Introduction

The Newtonian theory of light bending was published in 1801 by the German physicist Soldner [1, 2]. The author develops Kepler’s classical motion of a particle of light, of mass...
submitted to the gravitational force exerted by a mass $M$ with spherical symmetry. He obtained the usual hyperbolic motion and computed the deflection angle $\chi_N$ of the trajectory in the Newtonian approximation. By applying this analysis to a particle of light grazing the Sun, he found the value $\chi_N \approx 0.87\text{ as}$, which is exactly half the experimental value measured in 1919 [3]. In the first section we review Soldner’s computation with modern notations.

In 1911 [4], Einstein proposed a new analysis of light bending, based on the equivalence principle alone. He was led to the conclusion that a dilation of duration is produced by a gravitational potential. This leads to the conclusion that a certain velocity of light should depend on the gravitational potential $\Phi$.

$$c_{\Phi} = c \left(1 + \frac{\Phi}{c^2}\right)$$

(1)

This velocity is smaller than $c$, the value in the absence of potential ($\Phi = -GM/r < 0$). In his original paper of 1911 [4], Einstein does not give a real physical interpretation of this velocity, but simply speaks of speed of light. Using the Huygens–Fresnel principle, he deduced the trajectory of a light ray by requiring that they are normal to wavefront. Curiously, he found the same expression as Soldner’s Newtonian result. In the second section, we review the Einstein argument in a slightly different way, which sheds new light on the Einstein derivation. In particular, we show that the velocity obtained by Einstein has to be interpreted as a phase velocity, not the light speed (which remains a fundamental constant). We then argue that the de Broglie wave transposition of Soldner’s analysis explains the identical result obtained by Einstein.

Only a few years later, as part of the complete theory of general relativity (GR) [5, 6], Einstein obtained the correct value of this deviation, i.e. the double of the previous result. Many authors have discussed the reason for the doubling of the Newtonian result in GR [7].

In the third part of this paper, we propose a new light to interpret this doubling. For this, we propose a generalization of the physical analysis of Einstein, accompanying the time dilation due to the gravitational potential, by a concomitant contraction of the radial lengths (see [8] where this idea has already been proposed, though with a different approach to ours). This derivation is an intuition that Einstein could have had, more than a formal proof, because it is already known that the correct result cannot be recovered simply from the equivalence principle and the Newton’s limit alone [9].

2. Soldner’s Newton theory

In this section, we briefly summarize how Soldner computed the bending of light by a massive body using a Newtonian approach. For a complete historical perspective on the Newtonian influence of gravitation on light, see [10]. For this computation, Soldner hypothesized that light is made of material particles, for which it is possible to apply Newton’s laws in order to obtain the trajectory. To justify his hypothesis, he added, in the part relating to the possible opposition he may face, that light should be considered as matter:

*Hopefully, no one would find it objectionable that I treat a light ray as a heavy body. That light rays have all the absolute [basic] properties of matter one can see from the phenomenon of aberration which is possible only because light rays are truly material. And furthermore, one cannot think of a thing which exists and works on our senses that would not have the property of matter.*
Soldner’s computation was prior to Maxwell’s theory, in which the speed of light is a constant\(^2\). From Soldner’s perspective, the speed of a particle of light is not a constant, but varies along the path around the massive body, just like an ordinary material particle. In his publication, there is therefore a free parameter, which he took as being the speed of light measured at the level of perihelion \(P\); in the following, we will note this velocity as \(v_P\).

The trajectory of a particle of light \(A\) can be deduced from the conservation of the massic mechanical energy \(e_m = v^2/2 - GM/r\) and the massic angular momentum \(\ell = r^2 v e\). In these expressions, \(r\) and \(\varphi\) are the polar coordinates of \(A\), in the plane of motion defined by \(O\) and the normal vector \(\ell\) (figure 1). Combining these two expressions gives

\[
\ell^2 + \frac{\ell^2}{r^2} - 2\frac{GM}{r} = 2e_m
\]

(2)

The mass \(m\) of the particle of light, which was unknown to Soldner in 1801, does not appear in this equation. This observation is simply a reformulation, in the case of light, of the underlying hypothesis of the equality of the gravitational mass, which appears in gravitational energy, and of the inertial mass, present in the angular momentum. This hypothesis was postulated earlier by Galileo and then tested experimentally, with a relative precision of \(10^{-3}\), by Newton using pendulums made of different materials.

Following Soldner, the constant \(\ell\) can be expressed with respect to the speed of light \(v_P\) at the perihelion, \(v_P = v_P\). Likewise, one can also introduce the impact parameter \(b\), so that \(\ell = r v_P = b v_c\) (see figure 1 for notation). Equation (2) can be rewritten using dimensionless quantities. Introducing \(\rho \equiv r/r_P\) (as Soldner did), expression (2) is written more conveniently, if we introduce the gravitational potential at the perihelion \(\Phi_P \equiv -GM/r_P\), as

\[
r_P^2 \left( \frac{d\rho}{v_P \, df} \right)^2 + \frac{1}{\rho^2} + \frac{2}{\rho} \frac{\Phi_P}{v_P^2} = \frac{2e_m}{v_P^2} = \left( \frac{v_c}{v_P} \right)^2
\]

(3)

This equation is identical to the one obtained by Soldner. He solved equation (3) with lengthy calculations because the usual Binet change of variable was not yet known. Using the reduced Binet variable \(u \equiv 1/\rho = r_P/r\), one finds from (3):

\[
\frac{d^2u}{d\varphi^2} + u = -\frac{\Phi_P}{v_P^2} \equiv \frac{r_P}{p}
\]

(4)

If the light is grazing on the surface of the attractive body, \(r_P = R\), with \(R\) the radius of the massive body. The dimensionless quantity \(-\Phi_P/v_P^2\) is positive and reduces to the compactness \(C \equiv GM/(Rv_P^2)\) of the object, which physically represents the ratio between the gravitational energy and the mass energy. For objects like planets or stars, the compactness is very small compared to unity, so that the right hand side of equation (4) is very small and the

\[^2\text{In particular, it is independent of the gravitational potential, because gravitation and electromagnetism are not coupled in this theory.}\]
solution is nearly the usual Newton solution. For the Sun and the Earth, we find respectively (taking \(v_p \approx c\)):

\[ C_\odot \approx 2 \times 10^{-6} \quad \text{and} \quad C_m \approx 7 \times 10^{-10} \quad (5) \]

The solution of equation (4) is given by \(u(\varphi) = A \cos(\varphi - \varphi_0) + n_p/p\). We then determine the constant \(A\) using the condition on the perihelion, \(u = 1\) when \(\varphi = \varphi_0\), which gives \(1 = A - \Phi_p/v_p^2\). Thus, the solution for \(r\) is a conic of parameter \(p\) and eccentricity \(e\):

\[ r = \frac{p}{1 + e \cos(\varphi - \varphi_0)} \quad \text{with} \quad p = r_p \left(\frac{v_p^2}{-\Phi_p}\right) \quad \text{and} \quad e = \left(\frac{v_p^2}{-\Phi_p}\right) - 1 \quad (6) \]

We can also relate \(e\) to the massic mechanical energy:

\[ e_m = \frac{v_p^2}{2} \left(1 + \frac{2\Phi_p}{v_p^2}\right) = \frac{v_p^2}{2} \left(\frac{e - 1}{e + 1}\right) = \frac{v_p^2}{2} \left(\frac{\Phi_p}{v_p^2}\right)^2 (e^2 - 1) \quad (7) \]

The previous expression allows us to study the type of trajectories as a function of the value of the eccentricity: hyperbolic motion for \(e_m > 0\) \((e > 1)\) parabolic motion for \(e_m = 0\) \((e = 1)\) and elliptic motion for \(e_m < 0\) \((e < 1)\). Soldner found that in practice \(e_m > 0\), because the condition \(-\Phi_p/v_p^2 \ll 1\) was satisfied for the stars known at that time. Therefore \(e \gg 1\) according to equation (6) and the corresponding trajectories of the particles of light are hyperbolic ones.

Soldner briefly evoked the existence of bounded solutions, characterized by \(e_m < 0\), i.e. \(GM/(n_p v_p^2) > 1/2\). He added, however, that this condition was not realistic, or in any case it did not correspond to any known object at that time. Indeed, the stars seen in the sky were already considered as sun-like, whose mass and radius were known with sufficient precision. The compactness should be of the same order of magnitude as \(C_\odot\), and therefore very small (see equation (5)).

The Newtonian deviation angle \(\chi_N\) is easily obtained by writing the asymptotic condition \(r \to \infty\), i.e. \(\cos(\varphi_m - \varphi_0) = -1/e\). By choosing \(\varphi_m = 0\) for the direction of the incident ray, the ray emerges asymptotically in \(\varphi = \pi + \chi_N\), so that \(\cos \varphi_0 = \cos(\pi + \chi_N - \varphi_0) = -1/e\). Hence \(\chi_N = 2\varphi_0 - \pi\) and therefore \(\tan \varphi_0 = \tan(\chi_N/2 + \pi/2) = -\tan^{-1}(\chi_N/2)\). Since \(\cos \varphi_0 = -1/e\), \(\tan \varphi_0 = -(e^2 - 1)^{1/2}\), and we find the following result obtained by Soldner:

\[ \tan \left(\frac{\chi_N}{2}\right) = \frac{1}{(e^2 - 1)^{1/2}} = \frac{-\Phi_p/v_p^2}{(1 + 2\Phi_p/v_p^2)^{1/2}} \quad (8) \]

This Newtonian result is an exact result, which does not rely on any assumption. In the limit \(-\Phi_p/v_p^2 \ll 1\), it gives \(\chi_N \approx 2GM/(n_p v_p^2)\). Or, since \(n_p v_p = b v_\infty\) and \(r_p \approx b\) (at lowest order in \(-\Phi_p/v_p^2\)),

\(^3\) Soldner wrote, ‘Since it does not matter how much mass it would be so great that it could produce such an acceleration gravity, a light ray describes, in the world known to us, always hyperbola.’ We will discover much later that such objects, for which the trajectory of light realizes \(e_m < 0\), do exist in nature; for example, black holes. Note that Michell already considered the bounded trajectory of light, but in a rather different situation: he considered radial trajectory of light from massive objects, from which the escape velocity would be greater than the speed of light [10].
\[ \chi_N \approx \frac{2GM}{b \nu_c^2} = \frac{r_S}{b \nu_c} \]  
where  \[ r_S = \frac{2GM}{c^2} \]  
(9)
is the Schwarzschild radius. With the intention to estimate the orders of magnitude, Soldner used the speed of light measured by Bradley in 1729, using the aberration of stars [11]  
4. The result obtained by Soldner is half that predicted by GR in 1915 [5]. Moreover, its expression (8) is not universal, because it involves the speed of light at the perihelion (or equivalently, \( \nu_c \)), the latter being not considered, at the time of Soldner, as a universal constant. However, Soldner seems to suppose that this speed, which is much greater than the speed of celestial objects (planets, stars), must be, according to the law of Galilean composition of velocities, quite close to the value which he used in its numerical applications (see also the discussion in [10] which takes up the argument of Michell about the variation of the speed of light in a gravitation field). Assuming that \( \nu_c \approx c \), one obtains, if the light is grazing, for the Sun and the Earth respectively:  
\[ \chi_{N,\odot} \approx 0.87 \text{ as} \quad \text{and} \quad \chi_{N,\oplus} \approx 0.28 \times 10^{-3} \text{ as} \]
Soldner deduced from these numerical results that the deviation of light near the Sun was too small to be measured at his time². He (unknowingly) announced a result that would be tested experimentally more than a century later [3]. It is interesting to note that he published the result of his analysis, even though the conclusion was that the effect was not observable⁶.

3. Einstein’s relativistic theory of 1911

Einstein had already noticed in 1907, in his review article on special relativity, that, according to the principle of equivalence, a light ray has to be bent by gravitation [12]. In 1911 he carefully studied the influence of a gravitational potential \( \Phi \) on the propagation of light in vacuum. For a review of the original derivation, see [13]. He based its arguments on two pillars:

(i) Special relativity, including Maxwell’s theory of electromagnetism. It contains in particular the universal character of the speed of light in vacuum and the Doppler–Fizeau effect.

(ii) The equivalence principle he developed to build the theory of GR; this principle affirms the equivalence between an observer at rest in a uniform gravitational field and an observer uniformly accelerated in the absence of gravitation (see [14] for philosophical considerations concerning the principle of equivalence).

Inspired by Einstein’s reasoning, let us consider two observers, each one with an identical clock. These two observers are assumed to have a uniform acceleration \( a \); for example, by being both in the same rocket subjected to this acceleration. These two observers exchange photons, from the emitter \( E \) to the receiver \( R \) located at a distance \( H \) (figure 2 on the left). Due to the Doppler–Fizeau effect, the frequency \( \nu_e \) of the electromagnetic wave received by \( R \) differs from the frequency \( \nu_e \) of the wave emitted by \( E \). At lowest order (ignoring relativistic corrections which would produce a negligible second-order effect here), the photon is received by \( R \) after a time interval \( H/c \). The velocity of \( E \) is then \( v = aH/c \). As a result,

⁴ Note that Bradley obtained this speed, in units of speed of the Earth around the Sun, the latter being poorly known at the time.

⁵ He concludes with this sentence: ‘So it is clear that nothing is necessary, at least in the present state of practical astronomy, that one should take into account the disturbance of light rays by attracting celestial bodies.’

⁶ He even adds in its conclusion: ‘At any rate, I do not believe that there is any need on my part to apologize for having published the present essay just because the result is that all perturbations are unobservable.’
according to the Doppler–Fizeau effect, the relation between \( \nu_t \) and \( \nu_e \) is (still at lowest order):

\[
\nu_t = \nu_e \left( 1 + \frac{v}{c} \right) = \nu_e \left( 1 + \frac{aH}{c^2} \right)
\]

(10)

Because of the equivalence principle, the situation in an accelerated rocket is physically equivalent to that which one at rest observes in a uniform gravitational field \( G = G_0 \, e_z \), such that \( G_0 = a \) (figure 2 on the right). We remind the reader that \( G \) is such that the Newtonian gravitational force \( F \) exerted on a mass \( m \) submitted to the gravitational field is \( F = mG \). Introducing now the gravitational potential \( \Phi \), one has \( \Phi_e - \Phi_t = G_0H > 0 \). The gravitational potential is related, up to a constant, to the gravitational potential energy of a mass \( m \) in the gravitational field by the relation \( E_p = m\Phi \). Thus

\[
\nu_t = \nu_e \left( 1 + \frac{\Phi_e - \Phi_t}{c^2} \right) \quad \text{or} \quad \nu_t \left( 1 + \frac{\Phi_t}{c^2} \right) = \nu_e \left( 1 + \frac{\Phi_t}{c^2} \right)
\]

(11)

to first order [15].

This theoretical prediction of Einstein’s was tested experimentally for the first time by Pound and Rebka in 1960 [16]. In his article written in 1911, Einstein proposed to measure this effect using the shift of the spectral lines of the Sun, while emphasizing that the effect was very small since \( C_0 \approx 2 \times 10^{-6} \).

According to Einstein, equation (11) does not express just a simple Doppler–Fizeau effect on an electromagnetic wave, but more fundamentally an influence of the gravitational potential on time. To reach this conclusion, one can argue that the number of oscillation cycles in a wave packet exchanged between \( E \) and \( R \) must be preserved\(^7\). Therefore, introducing the proper durations \( \tau_e \) and \( \tau_t \) measured by clocks in \( E \) and \( R \), one has \( \nu_t \, d\tau_t = \nu_e \, d\tau_e \), that is to say \( \nu_0 \, d\tau_0 \) is constant or equivalently:

\[
\frac{d\tau_0}{1 + \Phi/c^2} = d\tau_0
\]

(12)

\( \tau_0 \) being the proper duration measured by a distant observer, located at a point for which \( \Phi \approx 0 \) (typically at infinity). What is true for the photon frequency must be true for all other fields: in other words it is the proper duration \( \tau_0 \) that flows differently for \( E \) and for \( R \).

\(^7\) Likewise, Einstein argued that the number of nodes and antinodes between \( E \) and \( R \), when a standing wave is established between the transmitter and the receiver, has to be constant, otherwise we would be in the presence of a nonstationary process, which is excluded.
The dependency of $τ_Φ$ on the gravitational potential has of course to remain compatible with the foundations of the special relativity and the equivalence principle. It implies, in particular, that the speed of light, as measured by an observer at the point where he stands (this precision is important), has to stay equal to $c$,

$$\frac{dr}{dτ_0} = c \quad \text{which implies} \quad \frac{1}{1 + \frac{Φ}{c^2}} \frac{dr}{dτ_0} = c$$

(13)

Hence the speed of light $c_{p,Φ}$ measured by a distant observer (with proper time $τ_0$), who observes the propagation of the latter in the vicinity of a massive star, will be\(^8\)

$$c_{p,Φ} = \frac{dr}{dτ_0} = c \left( 1 + \frac{Φ}{c^2} \right)$$

(14)

Einstein obtained this expression in 1911 [4] with a different argument. Nevertheless, in his paper, he was not clear about the physical interpretation of this velocity. In particular, he was a little bit embarrassed about the fact that special relativity and the equivalence principle have to imply a constancy of the speed of light, while its result shows, on the contrary, a dependency on the gravitational potential. He even wrote that ‘the principle of the constancy of the speed of light is not valid in the sense that serves as a basis for the usual theory of relativity’. In fact, there is no inconsistency with special relativity and the key point here is that this velocity $c_{p,Φ}$ is relevant only for a distant observer. An observer at the level of the perihelion would indeed measure that the speed of light is equal to $c$ at this point, and this is not in contradiction with equation (14). The second key ingredient is that this velocity is in fact a phase velocity. Einstein did not mention this term in his paper of 1911, where he used the generic term speed of light without distinguishing between phase or group velocity. If this velocity is interpreted as a group velocity, it would imply that light would be bent in the opposite direction, that is to say outwards instead of towards the central body!

Hopefully, Einstein used a wave analysis of the bending, and therefore arrived at a bending towards the central mass. To do this, he considered the propagation of a wavefront propagating at velocity $c_{p,Φ}$ in a nonuniform gravitational potential, and deduced the trajectory of light through the Malus theorem. We adopt here another approach, based on the eikonal equation.

Indeed, the dependency of the phase velocity of light on a gravitational potential $Φ$ can also be interpreted in terms of an effective refraction index $n_Φ$ of the (empty) medium in which light propagates, according to

$$n_Φ = \frac{c}{c_{p,Φ}} = \frac{1}{1 + \frac{Φ}{c^2}} \approx 1 - \frac{Φ}{c^2} > 1$$

(15)

In order to determine the trajectory, one can now use the eikonal equation in the approximation of the geometrical optics. Introducing the Frenet base $(e_t, e_n)$ and the curvilinear abscisse $s$ along the trajectory, the equation of the light ray is given by [17]

$$\frac{d}{ds} (n_Φ e_t) = \text{grad} \ n_Φ$$

(16)

Multiplying this equation by $e_n$ and introducing the elementary deflection angle of the path, $dχ = -de_n \cdot e_t$, one gets

\(^8\) Note that this relation, relativistic in essence, supposes that the gravitational potential $Φ$ is defined without additive constant; in Newtonian mechanics, the effect of the constant is neutralized by the infinite value of the speed of propagation of light.
Since $d\chi / ds$ is of order 1, we can take $n_{\phi} \approx 1$ at zeroth order. We then find the integral expression of the deflection angle $\chi_{E,11}$ obtained by Einstein in 1911:

$$n_{\phi} \frac{d\chi}{ds} = -\frac{1}{c^2} \nabla \Phi \cdot e_n$$

The minus sign indicates a deviation towards the massive object. Treating $\nabla \Phi$ as a small perturbation, the previous integral (17) can be computed on a straight line rather than the actual curved trajectory. If we denote by $x$ the coordinate of the current point $A$ on the trajectory, one gets, using $\Phi = -GM/r$ (figure 3):

$$\chi_{E,11} \approx -\frac{1}{c^2} \int \nabla \Phi \cdot e_n \, ds$$

(17)

where $e_r \cdot e_n = -\cos \theta$; the angle $\theta$ varies from $-\pi/2$ to $\pi/2$ when $x$ moves between $-\infty$ to $\infty$. Since $r = b / \cos \theta$ and $x = b \tan \theta$, with $b$ the impact parameter, $dx = b \, d\theta / \cos^2 \theta$ and therefore

$$\chi_{E,11} = \frac{GM}{bc^2} \int_{-\pi/2}^{\pi/2} \cos \theta \, d\theta = \frac{r_s}{b}$$

(19)

This result is identical to Soldner’s, although the approaches adopted are substantially different. To understand the reason, let us use the wave aspect of any physical object, based on the Hamilton–Jacobi formalism and the link between the action $S$ associated to a particle and the phase $\varphi = S/h$ of the associated wave [18]. The velocity of the particle of light is given by $v^2 = v_{\infty}^2 - 2\Phi$, so that

$$v = v_{\infty} \left(1 - \frac{2\Phi}{v_{\infty}^2}\right)^{1/2} \approx v_{\infty} \left(1 - \frac{\Phi}{v_{\infty}^2}\right) > v_{\infty}$$

(20)

As already mentioned, should this velocity be interpreted as a phase velocity, it would give an effective refractive index $n_{\phi} < 1$ (see equation (15)), and therefore an opposite light bending compared to observations. In order to determine the phase velocity $c_{p,\Phi}$, we can consider the displacement of a wavefront ($\varphi = \text{cte}$) between $t$ and $t + \, dt$:

$$0 = \frac{d\varphi}{dt} = \frac{\partial \varphi}{\partial t} + c_{p,\Phi} \cdot \nabla \varphi$$

(21)

The displacement being perpendicular to the wavefront, $c_{p,\Phi}$ is colinear to $\nabla \varphi$. Finally, replacing $\varphi$ with $S/h$, we deduce
\[
\frac{\partial S}{\partial t} + c_{p,\phi} |\text{grad} S| = 0
\]  
(22)

This equation is analogous to the Hamilton–Jacobi equation [19], provided that \( c_{p,\phi} \) is expressed as a function of the generalized momentum. Then, one can use the fact that the time derivative of the action is equal to the opposite of the Hamiltonian, \( \frac{\partial S}{\partial t} = -H \). And because the Hamiltonian does not depend explicitly on time, it is a constant \( H = \mathcal{E} \) so that

\[
c_{p,\phi} = \frac{E}{|\text{grad} S|} = \frac{\mathcal{E}}{p}
\]  
(23)

where the generalized momentum \( p = \gamma mv \) [20] is identified with \( \text{grad} S \) in the Hamilton–Jacobi formalism \( (p_i = \partial S/\partial q_i, [19]) \). Combining the previous equations finally gives

\[
v \times c_{p,\phi} = \frac{E}{\gamma m} \approx c^2
\]  
(24)

since \( \mathcal{E} = \gamma mc^2 + m\Phi \approx \gamma mc^2 \). It can be seen that the mass of the particle disappears and that this last relation is also valid for relativistic particles. It leads to the following relation between the phase velocity in the presence of a gravitational potential, and the phase velocity in its absence:

\[
c_{p,\phi} \approx \frac{c^2}{v_v} \approx \frac{c^2}{v_v(1 - \Phi/v_v^2)} \approx c \left( 1 + \frac{\Phi}{c^2} \right) < c
\]  
(25)

where we used \(-\Phi/v_v^2 \ll 1\) and \( v_v \approx c \). This expression of the phase velocity is exactly the same as the one obtained by Einstein in 1911. As shown above, the wave associated to the particle of light is the fundamental ingredient in understanding the identical results obtained by Soldner and Einstein. Note, nevertheless, that the Einstein result is more universal because the speed of light \( c \) is a real constant of nature \( (v_v \) is not).

4. Einstein’s relativistic theory of 1915

In 1915, Einstein reanalyzed, in the framework of his theory of GR, the deviation of a light ray by a mass distribution with spherical symmetry. He obtained a result which is the double of what he initially published in 1911. In this new result, a first contribution is attributed to the influence of the gravitational potential on time (it is exactly the effect computed in 1911), and a second contribution, of the same magnitude, is related to the deformation of space (spatial curvature). As already mentioned, this new result was confirmed experimentally in 1919 [3].

In this last section, we wonder whether Einstein could have already obtained the right answer in 1911. We first explain why the formal answer is no, and then propose a guess that could have led Einstein to GR before 1915.

To begin with, Einstein could not have established rigorously the correct expression until he had completed the theory of GR. The reason is that there are several possible relativistic theories of gravitation, which are all in agreement with the equivalence principle (see [21] for a review), but differ from Einstein’s GR. Also, different attempts have been made to simply recover the Schwarzschild metric from the equivalence principle and the Newtonian limit alone, but none succeeded [9]. Only experiments finally made it possible to decide in favor of Einstein’s theory. All these relativistic theories of gravitation predict a first contribution identical to the one obtained by the Newton approach (cf equation (19)). In GR, as already shown, this contribution is understood to stem from a curvature of time. The difference lies in

---

9 The velocity \( v_v \) can be interpreted as the group velocity of the electromagnetic light wave, which allows us to recover the well-known relation (24) on the product between the group and phase velocity.
the second contribution, which physically depends on the way space is curved by energy. For example, in Nordström’s theory of gravitation of 1913 [22], the two previous contributions precisely cancel each other and give a deviation of light which is identically zero\(^{10}\), in contradiction with the experiment of 1919 [3]. Nevertheless, Nordström’s theory is theoretically viable, fully relativistic and in accordance with the equivalence principle.

However, one of the lessons of special relativity is that space and time are profoundly linked into a spacetime concept. Therefore, it seems natural to apply to space what has been observed with time: if duration depends on the gravitational field, length should also depend on gravitational field. The question is to know what modification should be made to length. Going back to equation (12), we can write the relation between \(d\tau\) and \(d\tau_0\) as a time dilation relation. Indeed, by posing \(\Phi = -2\Phi_0\), one has

\[
d\tau_0 = \gamma_\phi \, d\tau_0 \quad \text{with} \quad \gamma_\phi = \left(1 + \frac{2\Phi}{c^2}\right)^{-1/2} = \left(1 - \frac{\Phi_0}{c^2}\right)^{-1/2} \geq 1 \quad (26)
\]

The duration in a distant observer is dilated. One can try a contraction of length in the radial direction, that is to say in the direction in which the gravitational potential varies. We would then have

\[
dr_0 = \frac{dr_\phi}{\gamma_\phi} \quad (27)
\]

with \(dr_\phi\) the length travelled during time \(d\tau_0\) at the level of the particle of light, while \(dr_0\) is the length as seen by a distant observer. Then, instead of starting from (13), we have to require, because of the equivalence principle,

\[
c = \frac{dr_\phi}{d\tau_0} \quad (28)
\]

So that the phase velocity would be given by

\[
c_{\Phi,\Phi} = \frac{dr_0}{d\tau_0} = \frac{1}{\gamma_\phi} \frac{dr_\phi}{d\tau_0} = \frac{c}{\gamma_\phi} = c\left(1 + \frac{2\Phi}{c^2}\right) \quad (29)
\]

This is the new phase velocity measured by a distant observer. We obtain the same relation as equation (14), simply replacing \(\Phi\) with \(2\Phi\). It is worth noting that the radial contraction of equation (27) is nothing else than a physical manifestation of the curvature of space. This contraction also defines the right direction of the parallel transport of the photon [23].

The previous result is retrieved, in a more modern way, by considering the following modification of the square of the interval:

\[
ds^2 = \left(1 - \frac{r_s}{r}\right) c^2 \, dr^2 - \frac{1}{1 - r_s/r} \, dr^2 \quad \text{with} \quad \frac{r_s}{r} = \frac{2GM}{rc^2} = -\frac{2\Phi}{c^2} \quad (30)
\]

We recover the spacetime interval of the Schwarzschild metric proposed by the latter in 1916 [24]. The trajectory of the light can be obtained according to \(ds^2\) and therefore

\(^{10}\) From a modern point of view, this is due to the fact that Nordström’s theory is a scalar theory \(\phi\), and that the coupling Lagrangian should be \(\phi \, T\) with \(T\) the trace of the energy-momentum tensor. For an electromagnetic field, this trace is zero, and therefore light cannot be coupled to a scalar.
\[ c_\Phi = \frac{dr}{dt} = \frac{c}{\gamma_\Phi} = c \left( 1 + \frac{2\Phi}{c^2} \right) = c \left( 1 - \frac{\mathfrak{r}_2}{r} \right) \]  \quad (31)

This expression for \( c_\Phi \) looks like the one obtained initially by Einstein in 1911, with the factor 2 which affects the gravitational potential. It is then sufficient to use Einstein’s wave reasoning to obtain a double deviation angle, in accordance with the observations [3].

Notice that other choices were a priori admissible. For example, in the Nordström theory, this choice would be not to contract the radial lengths, but on the contrary, to expand them, \( dr_0 = \gamma_\Phi dr_\Phi \). This amounts to treating space and time with the same factor. In modern language, this means that the metric is conformally flat, that is to say, \( ds^2 = f(\Phi) (c^2 dr^2 - dr^2) \). This would give \( c_{\mu,\Phi} = c \). Therefore, in this theory, because the phase velocity is a constant, light is not bent. Physically, there is a perfect compensation between the effect on time and the effect on space.

5. Conclusions

Let us remember the two essential points:

(i) From the Newtonian perspective, Soldner showed as early as 1801 that light should be deflected by a spherical mass. This deviation is identical (at lowest order) to the one obtained by Einstein in 1911, although their approaches differ substantially. The equality of both results comes from the principle of equivalence and the link between the velocity of a physical object and the velocity of the associated de Broglie wave.

(ii) Intuitive reasoning based on the effect of a gravitational potential on radial lengths and thus on the curvature of space could have put Einstein on the track of GR as early as 1911; he would then have found the right result for light deviation, and at the same time the Schwarzschild metric (see also [8]).

ORCID iDs

Brahim Lamine © https://orcid.org/0000-0002-9416-2320

References

[1] Soldner J 1801 Ueber die Ablenkung eines Lichtstrals von seiner geradlinigen Bewegung, durch die Attraktion eines Weltkörpers, an welchen er nahe vorbei geht Astronomisches Jahrbuch für das Jahr 1804 161–172

English translation in Jaki S 1978 Johann Georg von Soldner and the gravitational bending of light Foundations of Physics 8 11–2

[2] Hadley H G 1966 Predictions of the gravitational bending of light before einstein Am. J. Phys. 34 162

[3] Dyson F, Eddington A and Davidson C 1920 A Determination of Defection of Light by the Sun’s Gravitational Field from Observations Made at the Total Eclipse of May 29, 1919 Phil. Trans. A CCXX-A 579 291–332

[4] Einstein A 1911 Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes Ann. Phys., Lpz. XXXV 898–908

[5] Einstein A 1915 Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie Preussische Akademie der Wissenschaften, Sitzungsberichte 831–9

[6] Einstein A 1916 Die Grundlage der allgemeinen Relativitätstheorie Ann. Phys., Lpz. XLIX 769–882

[7] Lerner L 1997 A simple calculation of the deflection of light in a Schwarzschild gravitational field Am. J. Phys. 65 1194
[8] Provost J-P and Bracco C 2016 The 1895 Lorentz transformations: historical issues and present teaching Eur. J. Phys. 37 045601–11
[9] Gruber R P et al 1988 The impossibility of a simple derivation of the Schwarzschild metric Am. J. Phys. 56 265
[10] Eisenstaedt J 1991 De l’influence de la gravitation sur la propagation de la lumière en théorie newtonienne. L’archéologie des trous noirs Arch. Hist. Exact Sci. 42 315–86
[11] Bradley J 1729 A letter to E. Halley giving an account of a new discovered Motion of the Fix’d Stars Physical Transaction of the Royal Society 35 637–61
[12] Pais A 1993 Albert Einstein, Life and Work InterEditions 539 177
[13] Darrigol O 2015 Mesh and measure in early general relativity Studies in History and Philosophy of Modern Physics 52 163–87
[14] Brown H R and Read J 2016 Clarifying possible misconceptions in the foundations of general relativity Am. J. Phys. 84 327–34
[15] Pérez J-Ph 2016 Redshift Formulas and the Doppler–Fizeau Effect Eur. J. Phys. Vol. 37 015604 (8 pp)
[16] Pound R and Rebka G 1959 Gravitational red-shift in nuclear resonance Phys. Rev. Lett. 3 439–41
[17] Pérez J-P 2013 Optique, Fondements et Applications 698 7th edn (Dunod) p 204, 262
[18] Pérez J-P, Carles R and Pujol O 2013 Quantique, fondements et applications by de Boeck p 175
[19] Landau L and Lifchitz E 1969 Mechanics Mir 228 191
[20] Pérez J-P 2016 Relativité, fondements et applications 439 3rd edn (Paris: Dunod) p 51
[21] Norton J D 1992 Einstein, Nordström and the early demise of scalar, Lorentz-covariant theories of gravitation Arch. Hist. Exact Sci. 45 17
[22] Nordström G 1913 Zur Theorie der Gravitation vom Standpunkt des Relativitätprinzips Ann. der Phys. 42 533
[23] Ferraro R 2003 The equivalence principle and the bending of light Am. J. Phys. 71 168
[24] Schwarzchild K 1916 Über das Gravitationsfeld eines Masspunktes nach der Einsteinschen Theorie Sitzber. Deut. Akad. Wiss. Berlin Maths-Phys. K1 189–96 Tech.