6.3 Bending Light

> *When Lil's husband got demobbed, I said –*
> *I didn't mince my words, I said to her myself,*
> *HURRY UP PLEASE, IT'S TIME.*
> > T. S. Eliot, 1922

At the conclusion of his treatise on Opticks in 1704, the 62 year old Newton lamented that he could "not now think of taking these things into farther consideration", and contented himself with proposing a number of queries "in order to a farther search to be made by others". The very first of these was

> Do not Bodies act upon Light at a distance, and by their action bend its Rays, and is not this action strongest at the least distance?

From one point of view this may seem like a very natural suggestion. Newton's theory of universal gravitation already predicted that the path of any material particle (regardless of its composition) moving at a finite speed is affected by the pull of gravity. However, the finite speed of light was not well established in Newton's time, as discussed in Section 3.3, and it was far from clear that light consists of material particles. These uncertainties precluded Newton from making any definite prediction about whether and how light is affected by gravity. By the late 18th century the finite speed of light was well established so, although the constitution of light was still unknown, it was possible to apply Newton's law to compute the deflection of light by gravity – under the assumption that a pulse of light responds to gravitational attraction as does a particle of matter moving at the same speed. If the Sun's mass is located at the origin of xy coordinates then a particle moving with speed c = 1 along a nearly straight ray $y = r_0$ is subject to acceleration m/r^2 where $r^2 = r_0^2 + x^2$ is the distance to the Sun. Multiplying this by r_0/r gives the component of acceleration $a = (m/r^2)(r_0/r)$ transverse to the ray. Over any small segment we have $y = (1/2)ax^2$ (since x = t for a ray of light) up to a constant, and hence the angle of the ray is $dy/dx = ax = \tan(q)$ ~ q. From this we have $dq/dx = a = mr_0/r^3$, and integrating over the range x = -8 to +8 gives, to the lowest order of approximation, the total Newtonian angular deflection, $2m/r_0$. Of course, this crude derivation assumes a constant speed of light and a virtually straight path. Around 1784 Cavendish reached the same result by a more rigorous calculation, analyzing the actual hyperbolic path with varying speed, and in 1804 Soldner published the details of such an analysis.

The rectilinear coordinates x,y of a particle are related to the polar coordinates r,q by $x = r \cos(q)$ and $y = r \sin(q)$. Differentiating these expressions with respect to time, we have the components of the velocity

$$\dot{x} = -r\omega \sin(\theta) + \dot{r}\cos(\theta) \qquad \dot{y} = r\omega \cos(\theta) + \dot{r}\sin(\theta)$$

where dots signify time derivatives. Differentiating again, we get the components of the acceleration

$$\ddot{x} = \cos(\theta)(\ddot{r} - r\omega^2) - \sin(\theta)(r\dot{\omega} + 2\dot{r}\omega)$$
$$\ddot{y} = \sin(\theta)(\ddot{r} - r\omega^2) + \cos(\theta)(r\dot{\omega} + 2\dot{r}\omega)$$

Any vector whose x and y components are proportional to cos(q) and sin(q) respectively is parallel to r, and any vector whose x and y components are proportional to –sin(q) and cos(q) respectively is perpendicular to r (in the counter-clockwise direction), so it follows from the above expressions that the acceleration vector has the radial and tangential components

$$a_{radial} = \ddot{r} - r\omega^2 \qquad\qquad a_{tangential} = r\dot{\omega} + 2\dot{r}\omega$$

According to Newton's theory, a large gravitating body of mass m, located at the origin, exerts a purely radial force of magnitude $-m/r^2$ on a test particle at a distance r from the origin (in geometrical units with $G = c = 1$), so the Newtonian equations of motion for a test particle are

$$\ddot{r} - r\omega^2 = -\frac{m}{r^2} \qquad\qquad r\dot{\omega} + 2\dot{r}\omega = 0$$

The right hand equation is equivalent to $d(r^2 w)/dt = 0$, and hence the quantity $h = r^2 w$ is a constant. Letting u denote the reciprocal of r, we have $r = u^{-1}$ and $h/w = u^{-2}$, from which it follows that

$$\dot{r} = -h\frac{du}{d\theta} \qquad\qquad \ddot{r} = -h^2 u^2 \frac{d^2 u}{d\theta^2}$$

where we have used the fact that $w = dq/dt$. Substituting for w and r in the radial equation of motion, we get

$$-h^2 u^2 \frac{d^2 u}{d\theta^2} - u^{-1}\left(hu^2\right)^2 = -mu^2$$

Simplifying, we arrive at the familiar equation for the path of a test particle in a stationary spherical gravitational field according to Newtonian theory

$$\frac{d^2 u}{d\theta^2} + u = \frac{m}{h^2}$$

The general solution of this equation is

$$u(\theta) = A\cos(\theta) + \frac{m}{h^2}$$

Notice that if the speed of the particle is infinite, then h is infinite, and the equation of motion reduces to $r(q) = 1/A\cos(q)$, which is simply the equation of a straight line, so there would be no gravitational deflection. However, given the finite speed of light, and hence a finite value of h, we can solve this equation to determine the predicted gravitational deflection based on our Newtonian assumptions. Letting r_0 denote the distance of the pulse of light at its closest approach to the gravitating body, at which point $du/dq = 0$, the constant of integration can be written as $A = 1/r_0 - m/h^2$. Inserting this into the above equation, we get

$$u(\theta) = \frac{m}{h^2}\left[1 + \left(\frac{h^2}{mr_0} - 1\right)\cos(\theta)\right]$$

Reverting back to the radial coordinate r, we get the general equation for the path of a test particle in a spherical stationary gravitational field according to Newtonian theory

$$r(\theta) = \frac{h^2/m}{1 + \left(\frac{h^2}{mr_0} - 1\right)\cos(\theta)} \qquad (1)$$

This is the equation of a conic section, which is an ellipse, a hyperbola, or a parabola, accordingly as the value of the eccentricity parameter

$$\varepsilon = \frac{h^2}{mr_0} - 1$$

is less than, greater than, or equal to 1. Now, assuming a light pulse has the speed c at the perigee, we have $h = c\,r_0$. Choosing units so that $c = 1$, this is written as $h = r_0$, and

$$r(\theta) = \frac{r_0^2/m}{1 + \left(\frac{r_0}{m} - 1\right)\cos(\theta)}$$

The eccentricity is extremely large (since r_0 is much greater than m), so the path is a hyperbola that differs only slightly from a straight line, as indicated in the figure below.
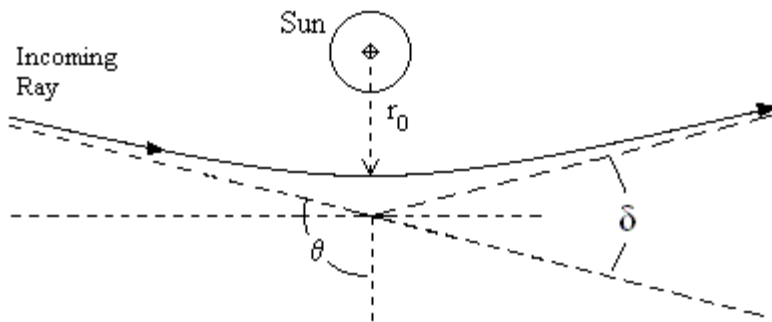
The asymptotes of this hyperbola occur at the angles where r goes to infinity, so we need only determine the angles q such that the denominator of the expression for r vanishes. Thus the two asymptotic angles are

$$\theta = a\cos\left(-\frac{m}{r_0 - m}\right) = \pm\left(\frac{\pi}{2} + \frac{m}{r_0} + \left(\frac{m}{r_0}\right)^2 + \frac{7}{6}\left(\frac{m}{r_0}\right)^3 \ldots\right)$$

To total deflection angle d equals the amount by which the difference between these two angles exceeds p, so we have

$$\delta_{Newtonian} = 2\left(\frac{\pi}{2} + \frac{m}{r_0} + \ldots\right) - \pi \approx 2\frac{m}{r_0}$$

In this derivation we assumed (with Soldner) that the speed of the light corpuscle (treated as a material particle) is c at the perigee, which according to Newtonian mechanics implies that it must be significantly less than c when the corpuscle is far from the Sun, since it is accelerated as it approaches. As an alternative, we might postulate that the speed is c at infinity, in which case Newtonian mechanics implies that the speed v_0 at perigee must be significantly greater than c. The change in potential energy of a particle of unit mass from infinity to the perigee is m/r_0, which must equal the change in kinetic energy given by $v_0^2/2 - c^2/2$. Again taking c = 1 this gives $v_0^2 = 1 + 2m/r_0$, from which we get $h^2 = r_0^2 + 2mr_0$. Substituting this into equation (1), we get

$$r(\theta) = \frac{\frac{r_0^2}{m}\left(1 + \frac{2m}{r_0}\right)}{1 + \left(\frac{r_0}{m} + 1\right)\cos(\theta)}$$

Thus the half-angle is acos($-m/(r_0+m)$), which again leads to the first-order result $2m/r_0$ for the deflection. Hence, regardless of whether we assume the speed of light is c at the perigee or at infinite, this is the amount of deflection predicted by Newton's theory of gravity, on the assumption that a pulse of light behaves like an ordinary material particle.

The best natural opportunity to observe this deflection would be to look at the stars near the perimeter of the Sun during a solar eclipse. The mass of the Sun in gravitational units is about m = 1475 meters, and a beam of light just skimming past the Sun would have a closest distance equal to the Sun's radius, r = $(6.95)10^8$ meters. Therefore, the Newtonian prediction would be 0.000004245 radians, which equals 0.875 seconds of arc. (There are 2p radians per 360 degrees, each of degree representing 60 minutes of arc, and each minute represents 60 seconds of arc.)

However, there is a problematical aspect to this "Newtonian" prediction, because it's based on the assumption that particles of light can be accelerated and decelerated just like ordinary matter, and yet if this were the case, it would be difficult to explain why (in non-relativistic absolute space and time) all the light that we observe is traveling at a single characteristic speed. Admittedly if we posit that the rest mass of a particle of light is extremely small, it might be impossible to interact with such a particle without imparting to it a very high velocity, but this doesn't explain why all light seems to have precisely the same speed, as if this particular speed is somehow a characteristic property of light. As a result of these considerations, especially as the wave conception of light began to supersede the corpuscular theory, the idea that gravity might bend light rays was largely discounted in Newtonian physics. (The same fate befell the idea of black holes, originally proposed by Mitchell based on the Newtonian escape velocity for light. Laplace also mentioned the idea in his Celestial Mechanics, but deleted it in the third edition, possibly because of the conceptual difficulties discussed here.)

The idea of bending light was revived in Einstein's 1911 paper "On the Influence of Gravitation on the Propagation of Light". Oddly enough, the quantitative prediction given in this paper for the amount of deflection of light passing near a large mass was identical to the old Newtonian prediction, $d = 2m/r_0$. There were several attempts to measure the deflection of starlight passing close by the Sun during solar eclipses to test Einstein's prediction in the years between 1911 and 1915, but all these attempts were thwarted by cloudy skies, logistical problems, the First World War, etc. Einstein became very

exasperated over the repeated failures of the experimentalists to gather any useful data, because he was eager to see his prediction corroborated, which he was certain it would be. Ironically, if any of those early experimental efforts had succeeded in collecting useful data, they would have proven Einstein *wrong*! It wasn't until late in 1915, as he completed the general theory, that Einstein realized his earlier prediction was incorrect, and the angular deflection should actually be *twice* the size he predicted in 1911. Had the World War not intervened, it's likely that Einstein would never have been able to claim the bending of light (at twice the Newtonian value) as a prediction of general relativity. At best he would have been forced to explain, after the fact, why the observed deflection was actually consistent with the completed general theory. Luckily for Einstein, he corrected the light-bending prediction before any expeditions succeeded in making useful observations. In 1919, after the war had ended, scientific expeditions were sent to Sobral in South America and Principe in West Africa to make observations of the solar eclipse. (Was the specific location of Principe chosen for its name, as a subliminal tribute to Newton's Principia?) The reported results were angular deflections of 1.98 ± 0.16 and 1.61 ± 0.40 seconds of arc, respectively, which was taken as clear confirmation of general relativity's prediction of 1.75 seconds of arc. This success, combined with the esoteric appeal of bending light and the romantic adventure of the eclipse expeditions themselves contributed enormously to making Einstein a world celebrity.

One other intriguing aspect of the story, in retrospect, is the fact that there is serious doubt about whether the measurement techniques used by the 1919 expeditions were robust enough to have legitimately detected the deflections which were reported. Experimentalists must always be wary of the "Ouija board" effect, with their hands on the instruments, knowing what results they want or expect. It's interesting to speculate on what values would have been recorded if they had managed to take readings in 1914, when the expected deflection was still just 0.875 seconds of arc. (It should be mentioned that many subsequent observations, summarized below, have independently confirmed the angular deflection predicted by general relativity, i.e., twice the "Newtonian" value.)

To determine the relativistic prediction for the bending of light past the Sun, perhaps the most direct approach is to set h equal to infinity in equation (6) of Section 6.2, because dt = 0 for a light path, and then solve for the nearly hyperbolic path. Another approach is to simply evaluate the solution of the four geodesic equations presented in Chapter 5.2, but this involves a three-dimensional manifold, with a large number of Christoffel symbols, etc. A more efficient variation of this method is to consider the problem from a two-dimensional standpoint. Recall the Schwarzschild metric in the usual polar coordinates

$$(d\tau)^2 = \left(\frac{r-2m}{r}\right)(dt)^2 - \left(\frac{r}{r-2m}\right)(dr)^2 - r^2(d\theta)^2 - r^2\sin(\theta)^2(d\phi)^2$$

We'll restrict our attention to a single plane through the center of mass by setting f = 0, and since light travels along null paths, we set dt = 0, which allows us to write the remaining terms in the form

$$(dt)^2 = \left(\frac{r}{r-2m}\right)^2(dr)^2 + \frac{r^3}{r-2m}(d\theta)^2 \qquad (2)$$

This can be regarded as the (positive-definite) line element of a two-dimensional surface (r, q), with the parameter t serving as the metrical distance. The null paths satisfying the complete spacetime metric with dt = 0 are stationary if and only if they are stationary with

respect to (2). This implies Fermat's Principle of "least time", i.e., light follows paths that minimize the integrated time of flight, or, more generally, paths for which the elapsed Schwarzschild coordinate time is stationary, as discussed in Chapter 3.5. (Equivalently, we have an angular analog of Fermat's Principle, i.e., light follows paths that make the angular displacement dq stationary, because the coefficients of (2) are independent of both t and q.) Therefore, we need only determine the geodesic paths on this surface. The covariant and contravariant metric tensors are simply

$$g_{\mu\nu} = \begin{bmatrix} \left(\dfrac{r}{r-2m}\right)^2 & 0 \\ 0 & \dfrac{r^3}{r-2m} \end{bmatrix} \qquad g^{\mu\nu} = \begin{bmatrix} \left(\dfrac{r-2m}{r}\right)^2 & 0 \\ 0 & \dfrac{r-2m}{r^3} \end{bmatrix}$$

and the only non-zero partial derivatives of the components of g_{mn} are

$$\frac{\partial g_{rr}}{\partial r} = \frac{-4mr}{(r-2m)^3} \qquad\qquad \frac{\partial g_{\theta\theta}}{\partial r} = 2r^2 \frac{r-3m}{(r-2m)^2}$$

so the non-zero Christoffel symbols are

$$\Gamma^r{}_{rr} = \frac{-2m}{r(r-2m)} \qquad \Gamma^r{}_{\theta\theta} = -(r-3m) \qquad \Gamma^\theta{}_{r\theta} = \Gamma^\theta{}_{\theta r} = \frac{r-3m}{r(r-2m)}$$

Taking the coordinate time t as the path parameter (since it plays the role of the metrical distance in this geometry), the two equations for geodesic paths on the (r, q) surface are

$$\frac{d^2 r}{dt^2} = \frac{2m}{r(r-2m)}\left(\frac{dr}{dt}\right)^2 + (r-3m)\left(\frac{d\theta}{dt}\right)^2$$

$$\frac{d^2\theta}{dt^2} = -2\frac{(r-3m)}{r(r-2m)}\left(\frac{dr}{dt}\right)\left(\frac{d\theta}{dt}\right)$$

(3)

These equations of motion describe the paths of light rays in a spherically symmetrical gravitational field. The figure below shows the paths of a set of parallel incoming rays.
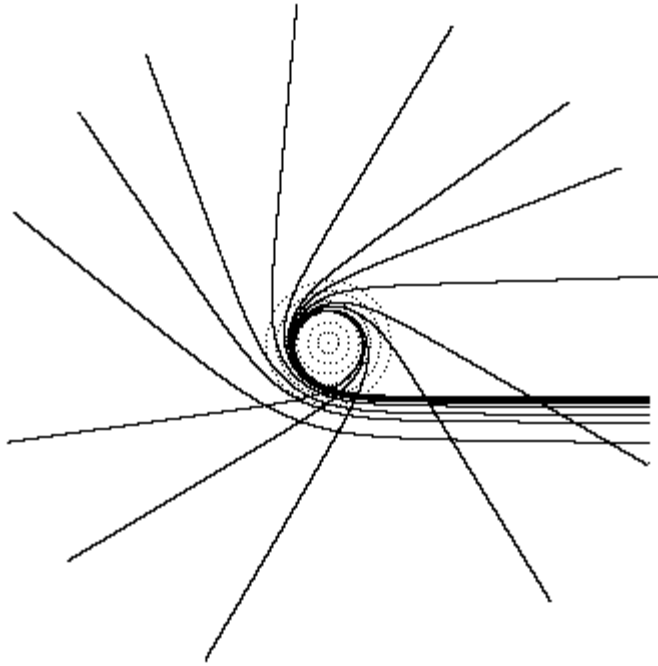
The dotted circles indicate radii of m, 2m, ..., 6m from the mass center. Needless to say, a typical star's physical radius is much greater than it's gravitational radius m, so we will not find such severe deflection of light rays, even for rays grazing the surface of the star. However, for a "black hole" we can theoretically have rays of light passing at values of r on the same order of magnitude as m, resulting in the paths shown in this figure. Interestingly, a significant fraction of the oblique incoming rays are "scattered" back out, with a loop at r = 3m, which is the "light radius". As a consequence, if we shine a broad light on a black hole, we would expect to see a "halo" of back-scattered light outlining a circle with a radius of 3m.

To quantitatively assess the angular deflection of a ray of light passing near a large gravitating body, note that in terms of the variable u = dq/dt the second geodesic equation (3) has the form (1/u)du = -[(2/r)(r-3m)/(r-2m)]dr, which can be integrated immediately to give ln(u) = ln(r-2m) - 3ln(r) + C, so we have

$$\frac{d\theta}{dt} = K\left(\frac{r-2m}{r^3}\right)$$

To determine the value of K, we divide the metric equation (2) by $(dt)^2$ and evaluate it at the perihelion $r = r_0$, where dr/dt = 0. This gives

$$1 = \frac{r_0^3}{r_0 - 2m}\left(\frac{d\theta}{dt}\right)^2$$

Substituting into the previous equation we find $K^2 = r_0^3/(r_0 - 2m)$, so we have

$$\frac{d\theta}{dt} = \sqrt{\frac{r_0^3}{r_0 - 2m}}\left(\frac{r-2m}{r^3}\right)$$

Now we can substitute this into the metric equation divided by $(dt)^2$ and solve for dr/dt to

give

$$\frac{dr}{dt} = \frac{r-2m}{r}\sqrt{1-\left(\frac{r_0}{r}\right)^3\frac{r-2m}{r_0-2m}}$$

Dividing dq/dt by dr/dt then gives

$$\frac{d\theta}{dr} = \frac{\sqrt{\dfrac{r_0^3}{r_0-2m}}}{\sqrt{1-\left(\dfrac{r_0}{r}\right)^3\dfrac{r-2m}{r_0-2m}}}\frac{1}{r^2}$$

Integrating this from $r = r_0$ to $¥$ gives the mass-centered angle swept out by a photon as it moves from the perihelion out to an infinite distance. If we define $r = r_0/r$ the above equation can be written in the form

$$d\theta = \int_{\rho=0}^{1}\frac{1}{\sqrt{1-\rho^2}}\frac{1}{\sqrt{1-2\left(\dfrac{1-\rho^3}{1-\rho^2}\right)\dfrac{m}{r_0}}}\,d\rho$$

The magnitude of the second term in the right-hand square root is always less than 1 provided r_0 is greater than 3m (which is the radius of light-like circular orbits, as discussed further in Section 6.5), so we can expand the square root into a power series in that quantity. The result is

$$d\theta = \int_{\rho=0}^{1}\frac{1}{\sqrt{1-\rho^2}}\left(1+\frac{1}{2}\left[2\frac{1-\rho^3}{1-\rho^2}\frac{m}{r_0}\right]+\frac{3}{8}\left[2\frac{1-\rho^3}{1-\rho^2}\frac{m}{r_0}\right]^2+\frac{5}{16}\left[2\frac{1-\rho^3}{1-\rho^2}\frac{m}{r_0}\right]^3+\cdots\right)d\rho$$

This can be analytically integrated term by term. The integral of the first term is just p/2, as we would expect, since with a mass of $m = 0$ the photon would travel in a straight line, sweeping out a right angle as it moves from the perihelion to infinity. The remaining terms supply the "excess angle", which represents the angular deflection of the light ray. If m/r_0 is small, only the first-order term is significant. Of course, the path of light is symmetrical about the perihelion, so the total angular deflection between the asymptotes of the incoming and outgoing rays is *twice* the excess of the above integral beyond p/2. Focusing on just the first order term, the deflection is therefore

$$\delta = 2\frac{m}{r_0}\int_{\rho=0}^{1}\frac{1-\rho^3}{\left(1-\rho^2\right)^{3/2}}\,d\rho$$

Evaluating the integral

$$\int\frac{1-\rho^3}{\left(1-\rho^2\right)^{3/2}}\,d\rho = -\sqrt{\frac{1-\rho}{1+\rho}}-\sqrt{(1-\rho)(1+\rho)}$$

from $r = 0$ to 1 gives the constant factor 2, so the first-order deflection is $d = 4m/r_0$. This gives the relativistic value of 1.75 seconds of arc, which is twice the Newtonian value. To higher orders in m/r_0 we have

$$\delta = 4\left(\frac{m}{r_0}\right) + \left(\frac{15}{4}\pi - 4\right)\left(\frac{m}{r_0}\right)^2 + \left(-\frac{15}{2}\pi + \frac{122}{3}\right)\left(\frac{m}{r_0}\right)^3 + \cdots$$

The difficulty of performing precise measurements of optical starlight deflection during an eclipse can be gathered from the following list of results:

Optical Deflection of Starlight During Eclipses

Date	Location	arc secs
29 May 1919	Sobral	1.98 ± 0.16
	Principe	1.16 ± 0.40
21 Sep 1922	Australia	1.77 ± 0.40
		1.42 to 2.16
		1.72 ± 0.15
		1.82 ± 0.20
9 May 1929	Sumatra	2.24 ± 0.10
19 June 1936	USSR	2.73 ± 0.31
	Japan	1.28 to 2.13
20 May 1947	Brazil	2.01 ± 0.27
25 Feb 1952	Sudan	1.70 ± 0.10
30 Jun 1973	Mauritania	1.66 ± 0.19

Fortunately, much more accurate measurements can now be made in the radio wavelengths, especially of quasars, since such measurements can be made from observatories with the best equipment and careful preparation (rather than hurriedly in a remote location during a total eclipse). In particular, the use of Very Long Baseline Interferometry (VBLI), combining signals from widely separate observatories, gives a tremendous improvement in resolving power. With these techniques it's now possible to precisely measure the deflection (due to the Sun's gravitational field) of electromagnetic waves from stars at great angular distances from the Sun. According to Will, an analysis in 2004 of over 2 million VBLI observations has shown that the ratio of the actual observed deflections to the deflections predicted by general relativity is 0.99992 ± 0.00023. Thus the dramatic announcement of 1919 has been retro-actively justified.

The first news of the results of Eddington's expedition reached Einstein by way of Lorentz, who on September 22 sent the telegram quoted at the beginning of this chapter. On the 7th of October Lorentz followed with a letter, providing details of Eddington's presentation to the "British Association at Bournemouth". Oddly enough, at this meeting Eddington reported that "one can say with certainty that the effect (at the solar limb) lies between 0.87" and 1.74", although he qualified this by saying the plates had been measured only preliminarily, and the final value was still to be determined. In any case, Lorentz's letter also included a rough analysis of the amount of deflection that would be expected due to ordinary refraction in the gas surrounding the Sun. His calculations indicated that a suitably chosen gas density at the Sun's surface could indeed produce a deflection on the order of 1", but for any realistic density profile the effect would drop off very rapidly for rays passing just slightly further from the Sun. Thus the effect of refraction, if there was any, would be easily distinguishable from the relativistic effect. He concluded

We may surely believe (in view of the magnitude of the detected deflection) that, in reality, refraction is

not involved at all, and your effect alone has been observed. This is certainly one of the finest results that science has ever accomplished, and we may be very pleased about it.

[Return to Table of Contents](#)