2009

An Optical Scalar Approach to Weak Gravitation Lensing

Louis Bianchini

Follow this and additional works at: http://vc.bridgew.edu/undergrad_rev

Part of the Other Physics Commons

Recommended Citation
Available at: http://vc.bridgew.edu/undergrad_rev/vol5/iss1/14

This item is available as part of Virtual Commons, the open-access institutional repository of Bridgewater State University, Bridgewater, Massachusetts. Copyright © 2009 Louis Bianchini
An Optical Scalar Approach to Weak Gravitation Lensing

LOUIS BIANCHINI

Louis will graduate summa cum laude in May of 2009, with a Bachelors of Science in Physics and Mathematics. Louis conducted this research, which was funded by the Adrian Tinsley summer program, with his mentor Dr. Thomas Kling of the Physics Department. This work was presented at the 2009 National Conference of Undergraduate Research. Louis will be entering graduate school in 2009 to pursue a doctorate in Physics.

We assume a thick gravitational lens governed by a Baltz \( n=1 \) matter density model. From the gravitational potential we then derive expressions for a Weyl tensor component and Ricci tensor component. A perturbative approach is taken to solve for the convergence and shear as given by Sach’s equation. By applying the geodesic deviation equation to a bundle of light rays with our expressions for the optical scalars we are able to derive expressions for the image shape and size at any point along the path from source to observer, in the case of an axially symmetric lens and source.

I. Introduction

The pre-relativity idea of gravitation was that objects with mass would interact with each other directly. The strength of this gravitational interaction is dependent on the mass of both objects, as well as the distance between them. However, with relativity came a new interpretation. The first change is that our use of only three coordinates to describe events is insufficient, since the inclusion of a time coordinate became necessary. Similarly, this changes the distance between objects as now we need to account for a change in time. Formally, this requires the use of tensors and a metric that accounts for gravitational effects when computing the distance or time between two events. Qualitatively, we say that space and time are connected to form a space-time. Moreover, the gravitational force an object of relatively small mass will experience as it moves through space-time is not best thought of as an interaction between it and all other massive objects. Instead, the force such an object experiences is caused by its movement through a curved space-time. The curvature of space-time, in turn, is caused by other massive objects.[1,2]

To work in relativistic physics, it is necessary to introduce four component vectors. This fact alone does not require any change in our geometry, however. It is the concept of curvature which requires the change in geometry. In a flat geometry, the metric is given by \( ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \), in coordinates of \((ct,x,y,z)\). Another representation of the metric is as a tensor, denoted by \( g \). A tensor, of rank \((k,l)\), is a multi-linear map taking \( k \) dual vectors and \( l \) vectors into a real number. There are often different interpretations, for example a \((1,1)\) tensor can take a dual and ordinary vector into a real number, or it can take an ordinary vector into an ordinary vector. For the flat
In curved space, however, this does not hold. In general, $g_{ab} \neq g^{ab}$. This curved space is best handled in a new type of geometry, based on the concept of Riemann manifolds.\(^{[1,6]}\)

Simply put, a manifold is a space which operates similar to flat geometry, in a local region, but whose global structure is allowed to deviate from Euclidean space. An example of this is the Earth, which we can associate a flat Euclidean geometry to, on a local scale (such as a room or a town), but the Earth is clearly curved overall. A Riemann manifold is an extension of a differentiable manifold, which is a manifold where Calculus works as expected, and allows an inner product to be defined in each tangent space of the manifold. In turn, a tangent space is a real vector space, which has the simplest mathematical properties. Whether done consciously or not, pre-relativity physics was performed in a vector space.

We may also note that one great benefit of using tensors is that any tensor equation which is true for one coordinate system is necessarily true for every coordinate system. There are some expressions which depend on what coordinates are used in pre-relativity physics; they would not be tensor equations. The usefulness of tensors is derived in part from the fact they are universal statements, applicable to any choice of coordinates. This allows us to exploit whatever symmetry is present and to choose coordinates that simplify the geometry of that particular problem.

### II. Space-Time Curvature

Imagine taking a particular vector associated with a point and moving it around some arbitrary closed loop. By intuition, we expect to get the same vector we started with as our result. However, this only holds for flat geometries\(^{[1]}\). In the space-time manifold there are indeed regions where this effect fails. The Riemann curvature tensor (denoted $R_{abcd}$) directly measures how much our choice for the arbitrary closed loop matters when we take our vector along the loop. The loop that we choose also depends on the underlying space-time curvature, which explains why for a flat space-time this effect is not seen, while curved space-times have this effect.

It turns out that we may decompose the Riemann tensor. In general relativity, we take a notation of raised and lowered indices, which mean different things. If we take a Riemann tensor, $R_{abcd}$, by use of the raised and lowered index $b$, this is equal to the Ricci tensor $R_{ab}$. Interestingly, if we take the trace (by letting $a=c$) of the Ricci tensor, we get $R_{ab} = R$. This expression, $R$, tells us how much the region of the space-time manifold we are investigating differs from a flat (Euclidean) geometry. Then the Ricci tensor can be used to get a direct measurement of the curvature of space-time. The Ricci tensor, however, is only half of the decomposition of the Riemann tensor. The other half is called the Weyl tensor, denoted $C_{abcd}$. The Weyl tensor is also called the conformal tensor, since as the name suggests, it is invariant to conformal changes (a mapping which preserves angles) to the metric. In practice, this means that if the Riemann manifold is to be conformally flat, then the Weyl tensor must vanish under contraction. The simplest way to ensure any tensor is to vanish under contraction is to make it strictly trace-free, as contraction will lead to terms that are part of the trace.\(^{[1,3]}\)

It is now necessary to introduce the idea of a null vector, and a null tetrad. In Euclidean geometry, the expression $\bar{A} \cdot \bar{A}$ computes the square of the length of the vector $A$. In the same manner, with a four dimensional vector and a given metric, we may compute the dot product of any vector with itself. This calculation is performed, assuming the Einstein summation notation, as $\bar{A} \cdot \bar{A} = g_{ab} A^a A^b$. Since in our vector we have three spatial and one time component, it is possible that this result is 0, meaning that the vector is null. If two events were related by such a vector, they would be on the same light cone. A null tetrad is then defined to be a set of four null vectors, $\{l^a, n^a, m^a, \bar{m}^a\}$, where $l^a, n^a \in \mathbb{R}$, $m^a \in \mathbb{C}$, and $\bar{m}$ is the complex conjugate of $m$.\(^{[1,2,6]}\) The four null vectors are chosen to be combinations of vectors that are part of the original vector space defining the metric. This forms the core of the Newman-Penrose spin coefficients. In this formalism, the Weyl tensor now has 10 independent components, one of which we are interested in. Similarly, the Ricci tensor can be represented by a set of scalars, four real and three complex. The two components of interest to us are given in equations (1) and (2)\(^{[5]}\).

\[
\Psi_0 = -C_{abcd} l^a m^b l^c m^d \quad (1)
\]

\[
\Phi_{00} = -\frac{1}{2} R_{ab} l^a l^b \quad (2)
\]
III. Calculating $\Psi_0$ of the Weyl tensor

We started with a Baltz $n = 1$ mass density model, given by\(^\text{(4)}\)

$$\rho_{\text{mass}} = \frac{M_0}{4\pi r (r + r_s)^2} \left( \frac{r_t^2}{r^2 + r_t^2} \right)$$  \hspace{1cm} (3)$$

In this model, $r_s$ is the scale radius for the mass distribution, and $r_t$ is the tidal radius. The value $M_0$ is a constant, with units of mass given by the formula $M_0 = 4\pi r_t^2 \rho_c$, where $\rho_c$ is the critical density, dependent only upon the area. The choice of this model is based on the success of the Navarro-Frank-White (NFW) dark matter halo distribution, which is the result of running multiple numeric simulations of gravitationally interacting particles. The NFW model has the problem of resulting in a divergent total mass; the Baltz model is essentially a truncation of the NFW model past the tidal radius. As such, this model has a finite total mass and is thus more realistic. Of importance to the calculation for gravitational potential ($\phi$) is the fact that this density function is symmetric with respect to the angular coordinates ($\theta\phi$). By solving for the potential ($\phi$) in spherical coordinates, by integrating twice, as given by Poisson’s equation we obtained the potential given in equation. We may note that the partial derivatives which appear in equation are actually total derivatives due to symmetry.

$$\nabla^2 \phi = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \phi}{\partial r} \right) = 4\pi G \rho_{\text{mass}}$$  \hspace{1cm} (4)$$

Here we chose the constants of integration, $C_1, C_2$, such that the potential is defined at $r = 0$, and that $\phi \to 0$ as $r \to \infty$. As such, the last two terms in equation are constants.

Since we are interested in obtaining expressions for the shear and convergence, we must first solve for the first order Weyl tensor component, $\Psi_0$ given in equation (8)\(^\text{(5)}\).

$$\Psi_0 = \frac{1}{2} (\phi_{xx} - \phi_{yy} - 2i\phi_{xy})$$  \hspace{1cm} (8)$$

We used MAPLE to perform the partial differentiation, which left us with a result of approximately thirty terms. Since our mass distribution was symmetric with respect to the angular coordinates our gravitational potential is also only dependent on the radial coordinate. Since the radial component is related to Cartesian coordinates by $r^2 = x^2 + y^2 + z^2$, and the Weyl tensor requires partial derivatives with respect to Cartesian coordinates, we expected to and obtained symmetric results for the real terms. The imaginary term which had mixed derivatives turned out to be similar in functional form. Note

$$\frac{d\phi}{dr} = \frac{M_0 G r_t}{(r_s^2 + r_t^2)^2} \left( r_s (r_s^2 + r_t^2) + \frac{r_s^2 - r_t^2}{r^2} \ln (r + r_s) - \frac{1}{2} \ln (r^2 + r_s^2) \right) + \frac{2r_t}{r^2} \tan^{-1} \left( \frac{r}{r_t} \right) - \frac{C_1}{r^2}$$  \hspace{1cm} (5)$$

$$\phi = \frac{M_0 G r_t^2}{(r_s^2 + r_t^2)^2} \left[ \tan^{-1} \left( \frac{r}{r_t} \right) - \frac{2r_t}{r} \tan^{-1} \left( \frac{r}{r_t} \right) \right] + \ln \left( \frac{r_s^2 + r_t^2}{r_s^2 (r + r_s)^2} \right) \left( \frac{\alpha}{2r} - \frac{\pi\alpha}{2r_t} - 2 \ln \left( \frac{r_t}{r_s} \right) \right)$$  \hspace{1cm} (6)$$
that in equation (9), the square of \((x-iy)\) is not to be performed by complex conjugation. Further, the units of each \(F_s(r)\) are per volume.

\[
\Psi_0 = \frac{GM_0 r_i^2}{2(r_i^2 + r_s^2)}(x - iy)^2 \left[ F_1(r) + F_2(r) + F_3(r) \right] 
\]

(9)

\[
F_1(r) = \frac{6r_s^{r_i^2}}{r^2(r^2 + r_s^2)} - \frac{6\beta}{r^2(r^2 + r_s^2)} - \frac{2r_s}{r^2(r + r_i)} - \frac{2r_s}{r^2(r + r_i)} 
\]

(10)

\[
F_2(r) = \frac{3\beta}{r^2(2 + r_i)} + \frac{\beta}{r^2(r + r_i) - \frac{2r_s}{r^2(r + r_i)} - \frac{2r_s}{r^2(r + r_i)}} 
\]

(11)

\[
F_3(r) = \frac{3}{2} \ln \left( \frac{r_s^2}{r^2(r + r_i)^2} \right) \beta - 6 \tan^{-1} \left( \frac{r_i F_s}{r_s^2} \right) 
\]

(12)

IV. Shear and Convergence

The shear and convergence are two important quantities in gravitational lensing. The shear measures how stretched the image of a distant light source appears. This effect would make a circular light source appear to be elliptical. This can be calculated out to any distance from the lensing source (mass distribution) so that we can understand how the image is distorted along its path to an observer. The convergence, on the other hand, shows how quickly the image is being focused. Much like a magnifying glass, a mass distribution can actually enhance the brightness of a distant object by gravitational lensing.

The equations governing shear and convergence are found in matrix form as

\[
DP = Q + P^2 
\]

(13)

\[
Q = \begin{bmatrix} \Phi_0 & \Psi_0 \\ \Psi_0 & \Phi_\infty \end{bmatrix} 
\]

(14)

\[
P = \begin{bmatrix} \rho & \bar{\sigma} \\ \sigma & \bar{\rho} \end{bmatrix} 
\]

(15)

Here, \(\rho\) is convergence, \(\sigma\) is shear, and \[D = \frac{1}{\sqrt{2}} \frac{d}{d\lambda} \]

is a parameter which measures the distance along the light ray from the observer. Since \(\Psi_0\) has already been calculated, the next step is to calculate \(\Phi_\infty\):

\[
\Phi_\infty = \frac{1}{2} \nabla^2 \varphi = 2\pi G \rho_{mass} = \frac{GM_0}{2r(r + r_i)^2} \left( \frac{r_i^2}{r^2 + r_i^2} \right) 
\]

(16)

This calculation required by equation (13) is difficult, and is treated perturbatively. To begin the perturbation, suppose that no mass distribution is present. This is called the zeroth order approximation, and accounts for light from a distant source arriving at the observer.

\[
D\rho_0 = \rho_0^2 + \sigma_0 \bar{\sigma}_0 \\
D\sigma_0 = 2\rho_0 \sigma_0 
\]

(17)

The zeroth order coupled differential equations (17) given by equation have the solution \[\rho_0 = \frac{-1}{\sqrt{2\lambda}}, \sigma_0 = 0 \]

The zeroth order answers tell us that images do not shear in free space and they appear dimmer further away. To the first order in potential, where the mass distribution is now being taken into consideration, but we assume that only linear terms of potential account for this effect, we obtain the following first order inhomogeneous linear differential equations

\[
D\rho_1 - 2\rho_0 \rho_1 = \Phi_\infty \\
D\sigma_1 - 2\rho_0 \sigma_1 = \Psi_0 
\]

(18)

These equations therefore have known solutions, given by

\[
\rho_1 = \frac{\sqrt{2}}{\lambda^2} \int_0^\lambda \lambda' \Phi_\infty (\lambda') d\lambda' 
\]

(19)

\[
\sigma_1 = \frac{\sqrt{2}}{\lambda^2} \int_0^\lambda \lambda' \Psi_0 (\lambda') d\lambda' 
\]

(20)

Here, the constant of integration which normally appears in solutions to linear differential equations has been taken into account via the use of a definite integral. This is especially useful in equation (20), which has no exact solution, however the definite integral may be evaluated by numerical techniques.
Equation (27) now represents a set of coupled differential equations. The simplest way to decouple these equations would be if $\sigma \in \mathbb{I}$, as it would naturally separate. We can do this by proper choice of $\Psi_0$, namely if we let $y=0$ be the path along our zeroth order light ray, with any choice of $x$.

In the case that $y=0$, we may still observe some interesting conclusions. Note that since $\sigma \in \mathbb{I}$, it is understood that $\text{Im}(\sigma)-0$, and that $p=p_0 \cdot p_1$ by perturbation. Then equation (27) turns into a set of decoupled separable first order homogeneous differential equations given as

$$D_a = -(\rho_0 + \rho_1) a - \sigma_i a$$  \hspace{1cm} (28)

$$D_b = -(\rho_0 + \rho_1) b + \sigma_i b$$  \hspace{1cm} (29)

Since these equations are separable, we are free to integrate the first order terms as definite integrals, while dealing with the zeroth order term indefinitely. We choose a constant of integration such that, in zeroth order, $a(\lambda+L)=1$. Here, we are choosing $L$ to be the overall distance from source to observer, whereas $\lambda$ measures the distance along the path. As we will see, this implies that $a(\lambda+0)=0$, which physically means that the light rays are coming from the same point source. The integration of equation (28) yields

$$\frac{1}{\sqrt{2}} \frac{da}{d\lambda} = \frac{1}{\sqrt{2}} \lambda^{-1} - a - (\sigma_i + \rho_1) a$$  \hspace{1cm} (30)

$$\frac{da}{a} = \frac{d\lambda}{\lambda} - \sqrt{2} (\sigma_i + \rho_1) d\lambda$$  \hspace{1cm} (31)

$$\ln a = \ln \lambda + \ln C - \sqrt{2} \int_0^\lambda (\sigma_i + \rho_1) d\lambda'$$  \hspace{1cm} (32)

Then the choice of $C$ is clear, $C = \frac{1}{L}$. Taking the exponential of both sides, we obtain a direct result for $a$ as:

$$a = \frac{\lambda}{L} e^{-\sqrt{2} \int_0^\lambda (\sigma_i + \rho_1) d\lambda'}$$  \hspace{1cm} (33)

The result for $b$ follows a similar procedure, and we ultimately obtained

$$b = \frac{\lambda}{L} e^{\sqrt{2} \int_0^\lambda (\sigma_i + \rho_1) d\lambda'}$$  \hspace{1cm} (34)
Given these two values we can obtain expressions for the area of the image, as well as the ratio of axes, which gives direct information about the shape of the image. To this end,

\[ A = \pi ab = \pi \frac{\lambda^2}{L^2} \exp \left( -2\sqrt{2} \int_0^\lambda \rho_1 d\lambda' \right) \]  

\[ \frac{b}{a} = \exp \left( 2\sqrt{2} \int_0^\lambda \sigma_1 d\lambda' \right) \]  

Not only do we have expressions for the area and shape of the image of light rays, we also can make an interesting observation due to their functional form. We already know there are points along the path of a light ray where the image collapses to zero area. These points are called conjugate points, and they are the location of caustics. By looking at the expression for \( A \), we can see that there is no term capable of making \( A = 0 \), aside from \( \lambda = 0 \). In this process, the possibility for caustics has been lost. This is actually due to the choice of using a perturbation theory approach.

**VI. Data Results**

For this section we have used the following parameters to correspond to a plausible lensing scenario in units scaled by the age of the Universe and where \( G = 1; z_0 = -0.41718, L = 0.67514, M_0 = 5.227 \times 10^{-9}, r_t = 0.00136, r_s = 1.21 \times 10^{-5} \). The angle \( \alpha \) from figure 1 is allowed to vary and essentially controls the path which the light takes to reach the observer. The first quantities we need to evaluate are \( \Psi_0 \), \( \Phi_{00} \), \( \rho_1 \), and \( \sigma_1 \). We chose to evaluate \( \rho_1 \) as given by (19), by use of a Riemann sum. By l'Hôpital's rule both \( \sigma_1(0) \) and \( \rho_1(0) \) evaluate to zero, and as such we use a right-handed approximation of \( n \)-intervals. In this case the approximation is given by

\[ \int_0^\Delta \! f(\lambda') d\lambda' \approx \sum_{k=1}^n f(k\Delta') \Delta' \]  

Here we used a fixed interval, which causes \( \Delta \lambda' \) to be constant, and thus can come outside the summation. Further, since we started the integral at 0, \( \Delta \lambda' = \lambda/n \).

\[ \rho_1 = \frac{\sqrt{2}}{\lambda^2} \left[ \Delta \lambda' \! \left( \sum_{k=1}^n (k\Delta')^2 \Phi_{00} \left( k\lambda' \right) \right) \right] \]  

\[ \rho_1 = \frac{\sqrt{2}}{\lambda^2} \left[ \Delta \lambda' \! \left( \sum_{k=1}^n k^2 \Phi_{00} \left( \frac{k\lambda}{n} \right) \right) \right] \]  

Since equation (20) has no known integral it is necessary to evaluate \( \sigma_1 \) by numeric methods. We are free to apply the same technique, and obtained a similar result given by equation (41).

\[ \sigma_1 = \frac{\sqrt{2} \lambda^2}{n^2} \sum_{k=1}^n k^2 \Phi_{00} \left( \frac{k\lambda}{n} \right) \]  

In figure 2 we have the graph of \( \Phi_{00} \) versus \( \lambda \) from 0 to \( L \). From the parameters chosen we may see that the lensing galaxy is located at \( \lambda = 0.25796 \), which corresponds to the large spike that appears on the graph. This suggests that the effect of the gravitational lens is concentrated close to the galaxy, as expected. By not specifying any angle in the graph, this is actually taken along the light path which goes through the middle of the gravitational lens. Away from the lens, the effect is several orders of magnitude lower. In figure 3, the graph of \( \sigma_1 \) is presented, for the same parameters as figure 2, with the exception of the angle. As \( \sigma_1 \) is defined as an integral of \( \Phi_{00} \), we would expect that the abrupt spike seen in figure 2 would also appear in figure 3. Indeed, this is the case. However, in figure 3 it is not appropriate to go through the middle of the lens, as this is where the perturbative technique is least applicable. Instead, we choose three angles, \( 500'' \), \( 750'' \), and \( 1000'' \). These effectively control the path the light takes from source to observer. By doing so, our light paths do not go through the middle of the lens and are thus in the region where the perturbation applies. Up to the lens, there is very little convergence as shown in figure 3. We would expect this, as there is effectively no lens between the source and values of \( \lambda < 0.25796 \).

Figure 4 is a graph of \( \Psi_0 \) versus \( \lambda \) from 0 to \( L \). As with figure 2, this graph is not made with a specific angle and thus goes through the middle of the gravitational lens. Figure 4 also has the same abrupt spike at the lens, except here it is downward. In figure 5, a graph of \( \sigma_1 \) versus \( \lambda \), this manifests itself in a similar manner as figure 3. We chose the same angles in figure 5 as were used in figure 3, for the same perturbative reasons. As with figure 3, there is almost no shearing up to the gravitational lens.

In figure 6, the exponential part of equation (35) is graphed. While not showing explicitly the size of the bundle of light rays as it passes from the source to the observer, this is a necessary portion of that ultimate result. As is expected from figure...
3, we observe that there is a magnifying effect by the galaxy which becomes noticeably apparent at the lens. The amount of magnification is relatively small, well under 1%. To calculate the actual size, an additional factor of $\frac{\pi \lambda^2}{L^2}$ would need to be introduced to this graph. However, the dominant effect seen on such a graph is $\lambda^2$, which makes it difficult to determine what effect the lens plays. In addition, the outside factor merely tells us how light expands in the absence of a gravitational lens.

VII. Conclusion
In this research, we have created the formalism for a new approach to gravitational lensing in situations where a thin lens approximation is not appropriate. Beginning with a fully three-dimensional mass distribution model, and solving for the gravitational potential, we derived expressions for the optical scalars, convergence, and shear, by taking a perturbative approach to the Sach's equation. With this, the geodesic deviation equation was applied to a bundle of light rays to determine the shape parameters and size. Graphs of the optical scalars as well as the Weyl tensor component and Ricci tensor component which controls the optical scalars were presented. In addition, the effect of the gravitational lens on the size of a bundle of light rays was shown graphically to cause magnification of less than 1% for reasonable lensing scenarios. Further research is expected to be conducted to obtain further graphs and to extend the formalism.

References