Massachusetts Institute of Technology
Department of Physics

Physics 8.962 Spring 1999

Gravitational Lensing from Hamiltonian Dynamics

©1999 Edmund Bertschinger and Rennan Barkana. All rights reserved.

1 Introduction

The deflection of light by massive bodies is an old problem having few pedagogical treatments. The full machinery of general relativity seems like a sledge hammer when applied to weak gravitational fields. On the other hand, photons are relativistic particles and their propagation over cosmological distances demands more than Newtonian dynamics. In fact, for weak gravitational fields or for small perturbations of a simple cosmological model, it is possible to discuss gravitational lensing in a weak-field limit similar to Newtonian dynamics, albeit with light being deflected twice as much by gravity as a nonrelativistic particle.

The most common formalism for deriving the equations of gravitational lensing is based on Fermat's principle: light follows paths that minimize the time of arrival (Scheider et al. 1992). As we will show, light is deflected by weak static gravitational fields as though it travels in a medium with variable index of refraction $n = 1 - 2\phi$ where ϕ is the dimensionless gravitational potential.

With the framework of Hamiltonian dynamics given in the notes *Hamiltonian Dynamics of Particle Motion*, here we present a synopsis of the theory of gravitational lensing. The Hamiltonian formulation begins with general relativity and makes clear the approximations which are made at each step. It allows us to derive Fermat's least time principle in a weak gravitational field and to calculate the relative time delay when lensing produces multiple images. It is easily applied to lensing in cosmology, including a correct treatment of the inhomogeneity along the line of sight, by taking advantage of the standard formalism for perturbed cosmological models.

Portions of these notes are based on a chapter in the PhD thesis of Barkana (1997).

2 Hamiltonian Dynamics of Light

Starting from the notes *Hamiltonian Dynamics of Particle Motion* (Bertschinger 1999), we recall that geodesic motion of a particle of mass m in a metric $g_{\mu\nu}$ is equivalent to Hamiltonian motion in $3+1$ spacetime with Hamiltonian

$$H(p_i, x^j, t) = -p_0 = \frac{g^{0i}p_i}{g^{00}} + \left[\frac{(g^{ij}p_ip_j + m^2)}{-g^{00}} + \left(\frac{g^{0i}p_i}{g^{00}}\right)^2\right]^{1/2} . \tag{1}$$

This Hamiltonian is obtained by solving $g^{\mu\nu}p_\mu p_\nu = -m^2$ for p_0. The spacetime coordinates $x^\mu = (t, x^i)$ are arbitrary aside from the requirement that $g_{00} < 0$ so that t is timelike and is therefore a good parameter for timelike and null curves. The canonical momenta are the spatial components of the 4-momentum one-form p_μ. The inverse metric components $g^{\mu\nu}$ are, in general, functions of x^i and t. With this Hamiltonian, the exact spacetime geodesics are given by the solutions of Hamilton's equations

$$\frac{dx^i}{dt} = \frac{\partial H}{\partial p_i} , \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x^i} . \tag{2}$$

Our next step is to determine the Hamiltonian for the problem at hand, which requires specifying a metric. Because we haven't yet derived the Einstein field equations, all we can do is to pick an ad hoc metric. In order to obtain useful results, we will choose a physical metric representing a realistic cosmological model, an expanding Big Bang cosmology (a Robertson-Walker spacetime) superposed with small-amplitude spacetime curvature fluctuations arising from spatial variations in the matter density. For now, the reader will have to accept the exact form of the metric without proof.

The line element for our metric is

$$ds^2 = a^2(t)\left[-(1+2\phi)dt^2 + (1-2\phi)\gamma_{ij}dx^i dx^j\right] . \tag{3}$$

In the literature, t is called "conformal" time and x^i are "comoving" spatial coordinates. The cosmic expansion scale factor is $a(t)$ and is related to the redshift of light emitted at time t by $a(t) = 1/(1+z)$. To get the non-cosmological limit (weak gravitational fields in Minkowski spacetime), one simply sets $a = 1$. The Newtonian gravitational potential $\phi(x^i, t)$ obeys (to a good approximation) the Poisson equation. (In cosmology, the source for ϕ is not ρ but rather $\rho - \bar{\rho}$ where $\bar{\rho}$ is the mean mass density; we will show this in more detail later in the course.) We assume $|\phi| \ll 1$ which is consistent with cosmological observations implying $\phi \sim 10^{-5}$.

In equation (3) we write $\gamma_{ij}(x^k)$ as the 3-metric of spatial hypersurfaces in the unperturbed Robertson-Walker space. For a flat space (the most popular model with theorists, and consistent with observations to date), we could adopt Cartesian coordinates for

which $\gamma_{ij} = \delta_{ij}$. However, to allow for easy generalization to nonflat spaces as well as non-Cartesian coordinates in flat space we shall leave γ_{ij} unspecified for the moment.

Substituting the metric implied by equation (3) into equation (1) with $m = 0$ yields the Hamiltonian for a photon:

$$H(p_i, x^j, t) = p(1 + 2\phi) \ , \quad p \equiv \left(\gamma^{ij} p_i p_j\right)^{1/2} \ . \tag{4}$$

We have neglected all terms of higher order than linear in ϕ. Not surprisingly, in a perturbed spacetime the Hamiltonian equals the momentum plus a small correction for gravity. However, it differs from the proper energy measured by a stationary observer, $E = -V^\mu p_\mu$, because the 4-velocity of such an observer is $V^\mu = (a(1-\phi), 0, 0, 0)$ (since $g_{\mu\nu} V^\mu V^\nu = -1$) so that $E = a^{-1} p(1 + \phi)$. The latter expression is easy to understand because a^{-1} converts comoving to proper energy (the cosmological redshift) and in the Newtonian limit ϕ is the gravitational energy per unit mass (energy).

Why is the Hamiltonian not equal to the energy? The answer is because it is conjugate to the time coordinate t which does not measure proper time. The job of the Hamiltonian is to provide the equations of motion and not to equal the energy. The factor of 2 in equation (4) is important — it is responsible for the fact that light is deflected twice as much as nonrelativistic particles in a gravitational field.

To first order in ϕ, Hamilton's equations applied to equation (4) yield

$$\frac{dx^i}{dt} = n^i(1 + 2\phi) \ , \quad \frac{dp_i}{dt} = -2p\nabla_i\phi + \gamma^k{}_{ij} p_k n^j (1 + 2\phi) \ , \quad n^i \equiv \frac{\gamma^{ij} p_j}{p} \ . \tag{5}$$

We will drop terms $O(\phi^2)$ throughout. We have defined a unit three-vector n^i in the photon's direction of motion (normalized so that $\gamma_{ij} n^i n^j = 1$). The symbol $\gamma^k{}_{ij} = \frac{1}{2}\gamma^{kl}(\partial_i \gamma_{jl} + \partial_j \gamma_{il} - \partial_l \gamma_{ij})$ is a connection coefficient for the spatial metric that vanishes if we are in flat space and use Cartesian coordinates. Beware that ∇_i is the covariant derivative with respect to the 3-metric γ_{ij} and not the covariant derivative with respect to $\gamma_{\mu\nu}$, although there is no difference for a spatial scalar field: $\nabla_i \phi = \partial_i \phi$.

Note that the cosmological expansion factor has dropped out of equations (5). These equations are identical to what would be obtained for the deflection of light in a perturbed Minkowski spacetime. The reason for this is that the metric of equation (3) differs from the non-cosmological one solely by the factor $a^2(t)$ multiplying every term. This is called a conformal factor because it leaves angles invariant. In particular, it leaves null cones invariant, and therefore is absent from the equations of motion for massless particles.

In the following sections we shall represent three-vectors (and two-vectors) in the 3-space with metric γ_{ij} using arrows above the symbol. To lowest order in ϕ, we may interpret these formulae as giving the deflection of light in an unperturbed spacetime due to gravitational forces, just as in Newtonian mechanics. The difference is that our results are fully consistent with general relativity.

3 Fermat's Principle

When $\partial_t \phi = 0$, the Hamiltonian (eq. 4) is conserved along phase space trajectories and the equations of motion follow from an alternative variational principle, Maupertuis' principle (Bertschinger 1999). Maupertuis' principle states that if $\partial H(p_i, q^j, t)/\partial t = 0$, then the solution trajectories of the full Hamiltonian evolution are given by extrema of the reduced action $\int p_i \, dq^i$ with fixed endpoints. This occurs because

$$\int p_i \, dq^i - H \, dt = \int p_i \, dq^i - d(Ht) + t \, dH \ . \tag{6}$$

The Ht term, being a total derivative, vanishes for variations with fixed endpoints. The $t \, dH$ term vanishes for trajectories that satisfy energy conservation, and we already know (from the Hamilton's equations of the full action) that only such trajectories need be considered when $\partial H/\partial t = 0$. Thus, the condition $\delta \int p_i \, dq^i = 0$, when supplemented by conservation of H, is equivalent to the original action principle.

Expressing p_i in terms of dx^i/dt using Hamilton's equations (5) in the full phase space for the Hamiltonian of equation (4), the reduced action becomes

$$p_i \, dx^i = p \gamma_{ij} n^j \, dx^i = H(1 - 2\phi) \gamma_{ij} n^i \, dx^j = H \, dt \ . \tag{7}$$

Using $H = \text{constant} \equiv h$, Mauptertuis' principle yields Fermat's principle of least time,

$$\delta \int dt = \delta \int [1 - 2\phi(x)] \left(\gamma_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} \right)^{1/2} ds = 0 \tag{8}$$

for light paths parameterized by s. We leave it as an exercise for the reader to show, using the Euler-Lagrange equations, that if s measures path length, equation (8) yields equations (5) exactly (to lowest order in ϕ) when $\partial_t \phi = 0$. In comparing with equation (5), one must be careful to note that there the trajectory is parameterized by $dt = (1 - 2\phi) ds$ so that $\vec{n} = d\vec{x}/ds$ is a unit vector.

Thus, for a static potential ϕ (even in a non-static cosmological model with expansion factor $a(t)$), light travels along paths that minimize travel time but not path length (as measured by the spatial metric γ_{ij}). The null geodesics behave as though traveling through a medium with index of refraction $1 - 2\phi$. To minimize travel time, light rays will tend to avoid regions of negative ϕ; therefore light will be deflected around massive bodies.

Fermat's principle is exact for gravitational lensing only with static potentials. In most astrophysical applications, the potentials are sufficiently relaxed so that $\partial_t \phi$ may be neglected relative to $n^i \nabla_i \phi$ and Fermat's principle still applies. The one notable exception is microlensing, where the lensing is caused by stars (or other condensed objects) moving across the line of sight. In this case, one may still apply Fermat's principle after boosting to the rest frame of the lens.

4 Reduction to the Image Plane

In equation (8), the action is invariant under an arbitrary change of parameter, $s \to s'(s)$ with $ds'/ds > 0$. This is not a physical symmetry of the dynamics, and as a consequence we may eliminate a degree of freedom by using one of the coordinates to parameterize the trajectories. A similar procedure was used to eliminate t in going from equation (6) to equation (8). Here, as there, the Lagrangian is independent of the time parameter, enabling a reduction of order. However, for reasons that will soon become clear, this reduction cannot be done using the reduced action (Maupertuis' principle) but instead follows from reparameterization of the Lagrangian.

To clarify the steps, we start with

$$L_3(x^i, dx^j/ds) = [1 - 2\phi(x)] \left(\gamma_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} \right)^{1/2} \tag{9}$$

for the Lagrangian in the three-dimensional configuration space (eq. 8). Because the Lagrangian does not depend explicitly on s, the Hamiltonian is conserved and we may attempt to reduce the order as in the previous section. The first step is to construct the Hamiltonian. Under a Legendre transformation, $L_3 \to H_3(p_i, x^j, s) = p_i(dx^i/ds) - L_3$ where $p_i = \partial L_3/\partial(dx^i/ds)$ is the momentum conjugate to x^i. But we quickly run into trouble: as the reader may easily show, H_3 vanishes identically.

What causes this horror? The answer is that L_3 is homogeneous of first degree in the coordinate velocity dx^i/ds, which is equivalent to the statement that the action of equation (8) is invariant under reparameterization. Physically, the Hamiltonian vanishes because of the extra symmetry of the Lagrangian, which is unrelated to the dynamics. The physical Hamiltonian should include only the physical degrees of freedom, so we must eliminate the reparameterization-invariance if we are to use Hamiltonian methods.

This is done very simply by rewriting the action (eq. 8) using one of the coordinates as the parameter. The radial distance from the observer is a good choice: for small deflections of rays traveling nearly in the radial direction toward the observer, r will be single-valued along a trajectory.

To fix the parameterization we must write the spatial line element in a Robertson-Walker space in terms of r and two angular coordinates:

$$dl^2 \equiv \gamma_{ij} dx^i dx^j = dr^2 + R^2(r)\gamma_{ab}(\xi) d\xi^a d\xi^b \ . \tag{10}$$

Here $1 \leq a, b \leq 2$ and γ_{ab} is the metric of a unit 2-sphere. The coordinates ξ^a are angles and are dimensionless. Note that r measures radial distance ($\gamma_{rr} = 1$) and $R(r)$ measures angular distance. We will not give the exact form of $R(r)$ here except to note that for a flat space, $R(r) = r$. In the standard spherical coordinates, $\gamma_{\theta\theta} = 1$ and $\gamma_{\phi\phi} = \sin^2\theta$. We will leave the coordinates in the sphere arbitrary for the moment, and use γ_{ab} and its inverse γ^{ab} to lower and raise indices of two-vectors and one-forms in the sphere.

Our action, equation (8), is the total elapsed light-travel time t (using our original spacetime coordinates, eq. 3). The reparameterization means that now we express the action as a functional of the two-dimensional trajectory $\xi^a(r)$:

$$t[\xi^a(r)] = \int_0^{r_S} [1 - 2\phi(\xi, r)] \left[1 + R^2(r)\gamma_{ab}\frac{dx^a}{dr}\frac{dx^b}{dr}\right]^{1/2} dr \ . \tag{11}$$

This action is to be varied subject $\delta\xi^a = 0$ at $r = 0$ (the observer) and $r = r_S$ (the source).

In writing equation (11), we have neglected $\partial\phi/\partial t$ and we have neglected terms $O(\phi^2)$ (weak-field approximation). As we will see, the angular term inside the Lagrangian is small when the potential is small, and therefore we can expand the square root, dropping all but the lowest-order terms. To the same order of approximation, we may neglect the curvature of the unit sphere, and set $\gamma_{ab} = \delta_{ab}$. (We can always orient spherical coordinates so that $\gamma_{ab} = \delta_{ab}$ plus second-order corrections in ξ.) These approximations together constitute the small-angle approximation. In practice it is well satisfied; observed angular deflections of astrophysical lenses are much less than 10^{-3}.

With the weak-field and small-angle approximations, the action becomes

$$t[\xi^a(r)] = r_S + \int_0^{r_S} L_2\, dr \ , \quad L_2\left(\xi^a, \frac{d\xi^b}{dr}, r\right) = \frac{1}{2}\, R^2(r)\delta_{ab}\frac{d\xi^a}{dr}\frac{d\xi^b}{dr} - 2\phi(\xi^a, r) \ . \tag{12}$$

Note that the Lagrangian now depends on the "time" parameter, so we have eliminated the parameterization-invariance.

To get a Hamiltonian system, we make the Legendre transformation of the Lagrangian L_2. The conjugate momentum is $p_a = R^2(r)\delta_{ab}d\xi^b/dr$. The Hamiltonian becomes

$$H(p_a, \xi^b, r) = \frac{p^2}{2R^2(r)} + 2\phi(\vec{\xi}, r) \ . \tag{13}$$

On account of the small-angle approximation, $\vec{p}$ and $\vec{\xi}$ are two-dimensional vectors in Euclidean space ($p^2 \equiv \delta^{ab}p_a p_b$). Noting that r plays the role of time, this Hamiltonian represents two-dimensional motion with a time-varying mass $R^2(r)$ and a time-dependent potential 2ϕ.

With the Hamiltonian of equation (13), Hamilton's equations give

$$\frac{d\vec{\xi}}{dr} = \frac{\vec{p}}{R^2(r)} \ , \quad \frac{d\vec{p}}{dr} = -2\frac{\partial\phi}{\partial\vec{\xi}} \ . \tag{14}$$

These equations and the action may be integrated subject to the "initial" conditions $\xi = \xi_0$, $\vec{p} = 0$ and $t = t_0$ at the observer, $r = 0$:

$$
\begin{aligned}
\vec{\xi}(r) &= \vec{\xi}_0 - \frac{2}{R(r)} \int_0^r \frac{R(r-r')}{R(r')} \frac{\partial \phi}{\partial \vec{\xi}}(\vec{\xi}(r'), r')\, dr' \\
\vec{p}(r) &= -2 \int_0^r \frac{\partial \phi}{\partial \vec{\xi}}(\vec{\xi}(r'), r')\, dr' \\
t(r) &= t_0 - r - \int_0^r \left[\frac{p^2(r')}{R^2(r')} - 2\phi(\vec{\xi}(r'), r') \right] dr' \ .
\end{aligned}
\tag{15}
$$

Note that here t is the coordinate time along the past light cone; the elapsed time (the action) is $t_0 - t$. The two terms in the time delay integral arise from geometric path length (the p^2 term) and gravity. Half of the gravitational potential part comes from the slowing down of clocks in a gravitational field (gravitational redshift) and the other half comes from the extra proper distance caused by the gravitational distortion of space.

Equations (15) provide only a formal solution, since ϕ is evaluated on the unknown path $\vec{\xi}(r')$. The reader may verify the solution by inserting into equations (14). One needs the following identity for the angular distance in a Robertson-Walker space, which we present without proof:

$$
\frac{\partial}{\partial r} \left[\frac{R(r-r')}{R(r)R(r')} \right] = \frac{1}{R^2(r)} \ . \tag{16}
$$

It is easy to verify this for the flat case $R(r) = r$.

When the potential varies with time, we cannot use Fermat's principle or the further reduction achieved in this section. Instead, one has to integrate the original equations of motion (5). It can be shown (Barkana 1997) that, under the small-angle approximation, these equations also have the formal solution given by equation (15), with the single change that ϕ also becomes a function of t and that t must be evaluated along the trajectory: $\phi(\vec{\xi}(r'), r', t(r'))$. Thus, we obtain the physical result that the potential is to be evaluated along the backward light cone.

5 Astrophysical Gravitational Lensing

The astrophysical application of gravitational lensing is based on the following considerations. Given an observed image position $\vec{\xi}_0$, we wish to deduce the source position $\vec{\xi}_S \equiv \vec{\xi}(r_S)$ using equation (15) to relate $\vec{\xi}(r_S)$ to $\vec{\xi}_0$. The result is a mapping from the image plane $\vec{\xi}_0$ to the source plane $\vec{\xi}_S$). This mapping is called the lens equation.

By integrating the deflection $\vec{\xi}_S - \vec{\xi}_0$ for a given distribution of mass (hence potential) along the line of sight from the observer, and for a given cosmological model (hence angular distance $R(r)$), one can compute the source plane positions for the observed images.

In practice, we wish to solve the inverse problem, namely to deduce properties of the mass and spatial geometry along the line of sight from observed lens systems. How can this be done if we know only the image positions but not the source positions?

There are several methods that can be used to deduce astrophysical information from gravitational lenses (Blandford and Narayan 1992). First, the lens mapping $\vec{\xi}_S(\vec{\xi}_0)$ can become multivalued so that a given source produces multiple images. In this case, the images provide constraints on lensing potential and geometry because all the ray paths must coincide in the source plane. This method can strongly constrain the mass of a lens, especially when the symmetry is high so that an Einstein ring or arc is produced.

Another method uses information from $t(r)$. If the source is time-varying and produces multiple images, then each image must undergo the same time variation, offset by the $t - t_0 + r$ integral in equation (15). Because this method involves measurement of a physical length scale (the time delay between images, multiplied by the speed of light), it offers the prospect of measuring cosmological distances in physical units, from which one can determine the Hubble constant. This is a favorite technique with MIT astrophysicists.

Another way to get a timescale occurs if the lens moves across the line of sight, in the phenomenon called microlensing. Gravitational lensing magnifies the image according to the determinant of the (inverse) magnification matrix $\frac{\partial \xi_S^a}{\partial \xi_0^b}$. If the angular position of the lens is close to ξ_S so that the rays pass close to the lens, the magnification can be substantial (e.g. a factor of ten). A lens moving transverse to the line of sight will therefore cause a systematic increase, then decrease, of the total flux from a source. From a statistical analysis of the event rates, magnifications and durations, it is possible to deduce some of the properties of a class of lensing objects, such as dim stars (or stellar remnants) in the halo surrounding our galaxy (more colorfully known as MACHOs for "MAssive Compact Halo Objects").

A fourth method, called weak lensing, uses statistical information about image distortions for the case where the deflections are not large enough to produce multiple images, but are large enough to produce detectable distortion. This method can provide statistical information about the lensing potential. It is a favorite method for trying to deduce the spectrum of dark matter density fluctuations.

There are many other applications of gravitational lensing. The study and observation of gravitationl lenses is one of the major areas of current research in astronomy.

6 Thin Lens Approximation

Our derivation of the lens equations (15) made the following, well-justified approximations: the spacetime is a weakly perturbed Roberston-Walker model with small-amplitude curvature fluctuations ($\phi^2 \ll 1$), the perturbing mass distribution is slowly-

evolving ($\partial_t \phi$ neglected), and the angular deflections are small ($|\vec{\xi}_S - \vec{\xi}_0| \sim \phi < 10^{-3}$).

Nearly all calculations of lensing are made with an additional approximation, the thin-lens approximation. This approximation supposes that the image deflection occurs in a small range of distance δr about $r = r_L$. In this case, the first of equations (15) gives the thin lens equation

$$\vec{\xi}_S = \vec{\xi}_0 - \frac{R_{LS}}{R_S}\,\vec{\gamma}(\vec{\xi}_0, R_L)\ , \quad \vec{\gamma}(\xi, R) = 2\int \frac{\partial\phi}{\partial\vec{\xi}}(\vec{\xi}, r')\,\frac{dr'}{R}\ , \tag{17}$$

where $R_S \equiv R(r_S)$, $R_L \equiv R(r_L)$ and $R_{LS} \equiv R(r_S - r_L)$. The deflection angle $\vec{\gamma} = -2\int \vec{g}\,dr$ where $\vec{g} = -\vec{\nabla}_\perp \phi = -(1/R)\partial\phi/\partial\vec{\xi}$ is the Newtonian gravity vector (up to factors of a from the cosmology).

Let us estimate the deflection angle γ for a source directly behind a Newtonian point mass with $g = GM/r^2$ (here r is the proper distance from the point mass to a point on the light ray). The impact parameter in the thin-lens approximation is $b = \xi_0 R_L$. Because the deflection is small, the path is nearly a straight line past the lens, and the integral of g along the path gives, crudely, $2bg(b) = 2GM/b = 2GM/(\xi_0 R_L)$. (The factor of two is chosen so that this is, in fact, the exact result of a careful calculation.) With the source lying directly behind the lens, $\xi_S = 0$.

Substituting this deflection into the thin lens equation (17) gives

$$0 = \xi_0 - \frac{R_{LS}}{R_L R_S}\,\frac{4GM}{\xi_0}\ . \tag{18}$$

Vectors are suppressed because this lens equation holds at all positions around a ring of radius $\xi_0 = |\vec{\xi}_0|$ in the image plane. An image directly behind a point mass produces an Einstein ring. Solving for ξ_0 gives the Einstein ring radius:

$$\xi_0 = \left(\frac{4GMR_{LS}}{R_L R_S c^2}\right)^{1/2}\ . \tag{19}$$

References

[1] Barkana, R. 1997, MIT PhD thesis.

[2] Bertschinger, E. 1998, 8.962 notes *Hamiltonian Dynamics of Particle Motion.*

[3] Blandford, R. D. and Narayan, R. 1992, *Ann. Rev. Astr. Ap.*, 30, 311.

[4] Schneider, P., Ehlers, J., and Falco, E. E. 1992, *Gravitational Lenses* (Berlin: Springer-Verlag).