SIAM REVIEW
Vol. 44, No. 1, pp. 111–130

© 2002 Society for Industrial and Applied Mathematics

Inverting Gravitational Lenses*

Peter R. Newbury†
Raymond J. Spiteri‡

Abstract. Gravitational lensing provides a powerful tool to study a number of fundamental questions in astrophysics. Fortuitously, one can begin to explore some nontrivial issues associated with this phenomenon without a lot of very sophisticated mathematics, making an elementary treatment of this topic tractable even to senior undergraduates. In this paper, we give a relatively self-contained outline of the basic concepts and mathematics behind gravitational lensing as a recent and exciting topic for courses in mathematical modeling or scientific computing. To this end, we have designed and made available some interactive software to aid in the simulation and inversion of gravitational lenses in a classroom setting.

Key words. gravitational lensing, inverse problems, constrained optimization

AMS subject classifications. 85-08, 65K10, 49M30, 83B05

PII. S0036144500380934

1. Overview. Gravitational lensing has evolved into a powerful tool for studying a number of important questions in astrophysics. For example, gravitational lenses have been described as *cosmic telescopes*, natural magnifying glasses that allow astronomers to view objects that would otherwise be too distant or too faint to be picked up by observational instrumentation. Moreover, gravitational lenses provide a great way to detect and analyze dark matter. This is because gravitational lensing effects depend only on the total mass distribution, and not on the luminosity and composition of the mass distribution. Also, observations from gravitational lenses can provide bounds for universal constants such as the Hubble constant and the density parameter of the universe. Consequently, gravitational lenses have received a lot of attention from researchers in applied mathematics, astronomy, and physics (see, e.g., [1], [19], [31], [34], [35]).

In this article, we present two aspects of gravitational lensing that are nontrivial yet interesting and tractable in a classroom setting. First, we investigate arcs of light produced by a gravitational lens by studying the optics of the lens and describing a ray-tracing procedure for the *forward modeling* of gravitational lenses. Second, we give a procedure for *inverting* the gravitational lens in order to estimate the distribution of mass required to produce the arcs. We also describe software to aid in the classroom presentation of these concepts.

The material presented here is most suitable as a module in an introductory course on mathematical modeling or inverse problems. It could also serve as the basis

*Received by the editors November 9, 2000; accepted for publication (in revised form) August 16, 2001; published electronically February 1, 2002.
http://www.siam.org/journals/sirev/44-1/38093.html

†Department of Mathematics and Statistics, Langara College, 100 West 49th Avenue, Vancouver, BC, Canada, V5Y 2Z6 (pnewbury@langara.bc.ca).

‡Faculty of Computer Science, Dalhousie University, 6050 University Avenue, Halifax, NS, Canada, B3H 1W5 (spiteri@cs.dal.ca).

for a student project. Some of the techniques described also make it suitable as an important and novel application for topics in a course on scientific computing.

2. Introduction. Cosmologists concern themselves with questions about the origin and fate of the universe. In particular, one of the fundamental questions of cosmology is, "What is the mass of the universe?" The reason that cosmologists are concerned with such a question is that the answer could help them predict the ultimate fate of the universe. Since its explosive start at "the Big Bang," the universe has been expanding and cooling. So the question cosmologists ask is, "Will this present state of expansion and cooling persist forever, or will the universe someday stop expanding and eventually collapse in what is affectionately called 'the Big Crunch'?"

Observations suggest that the luminous mass in the universe, mass which emits "visible" radiation, accounts for less than 10% of what is needed to precipitate a Big Crunch (see, e.g., [28, Chapter 18]). There is strong evidence, however, for the presence of a very large amount of *dark matter*, perhaps even large enough to "close" the universe. Although this dark matter emits no detectable radiation (hence the terminology), it still interacts gravitationally with its surroundings. Astronomers searching for dark matter look for signs of gravitational interaction, much like the way planet hunters look for the periodic wobbling of stars rather than light coming from the planets themselves.

One of the tools used in the search for dark matter is the gravitational lens (see, e.g., [33], [41] for a full description of gravitational lenses). According to Newton's theory of gravitation, light (which is considered to have no mass) feels no gravitational interaction with its surroundings and travels in straight lines throughout the universe. However, according to Einstein's theory of general relativity [12], both massive particles and massless light rays move in geodesic paths in four-dimensional spacetime. If the spacetime is curved, as it is around any massive object, the geodesic paths are themselves curved. So light rays as well as particles are deflected as they pass by any massive object. If the object is massive enough, the deflection is even observable. This prediction of light-ray deflection was validated in 1919 with Eddington's observations of the deflection of starlight past the eclipsing Sun [10]. From that day on, general relativity has been the preeminent theory of gravitation.

If the interaction is strong enough, two rays of light diverging from a distant source may be deflected by an intermediate mass and focused onto an observer. This process is known as *gravitational lensing*, and the intermediate mass is said to create a *gravitational lens*. As shown schematically in Figure 2.1, this observer may in fact see two images of the source.

Forward simulations of gravitational lenses are extremely useful when studying gravitational lensing. Through simulations of lensing and before actually analyzing real observations of lenses, we can begin to build a catalogue of lensing phenomena. Later, observed lenses can be "looked up" in the catalogue, much the way a police sketch artist has a catalogue of facial features for a witness to consult. Also, because forward modeling involves ray-tracing and visualization of two-dimensional data, we can naturally exploit modern computer processor and graphics capabilities. To quickly produce simulations, we must also understand how to efficiently manage and access computer memory (because simulations involve several two-dimensional arrays), how to effectively visualize data, and how to deal with numerical issues (e.g., how to treat potentially small divisors). This makes gravitational lensing an ideal problem from which to learn not only about mathematical modeling, but also about programming, graphics, and numerical analysis.

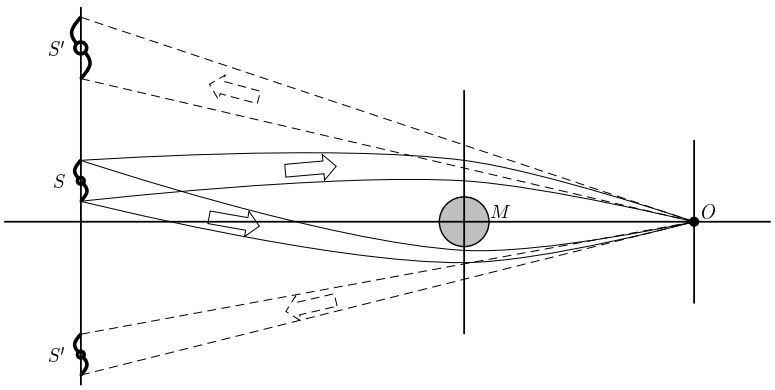

Fig. 2.1 *In this schematic diagram, light from the source S is deflected around the intermediate mass M and focused onto the observer O, who sees* two *distorted images S' of the source.*

As described below, the forward simulation of gravitational lenses is quite manageable. Moreover, the simple mathematics can still produce fascinating and intricate results. However, the ultimate goal of gravitational lens opticians and this article is to find a way to "invert" the lens: Given a collection of observed, gravitationally lensed objects and knowledge of the mathematics of lensing, determine what distribution of intermediate mass (the *deflector*) and what distribution of background light (the *source*) are needed to reproduce the observations. By comparing the gravitational mass of the deflector to the luminous mass of the deflector (i.e., the total mass associated with stars and galaxies in the deflector), astronomers can detect the presence of dark matter.

The remainder of this article is organized as follows. In section 3, we describe the forward problem of how lensed images are formed. In section 4, we outline a particular lens inversion scheme that is based on reconstructing the "best" background source of light using techniques from constrained optimization. Solutions to both of these problems are implemented as MATLAB functions that are available to interested readers via the Internet.[1] In section 5, we mention a few extensions from the idealizations in sections 3 and 4 and how they can be incorporated into the forward and inverse models. We end with some concluding remarks in section 6.

3. Gravitational Lens Optics. It is a surprisingly easy exercise to numerically simulate a gravitational lens; the mathematics describing the phenomenon turns out to be quite simple. From Einstein's theory of general relativity, we have that if the closest distance that a ray of light would have passed from a point mass M is ξ, then that ray will be deflected from a straight path towards the mass by an angle

$$\widetilde{\alpha} = \frac{4G}{c^2}\frac{M}{\xi},$$

where G is the gravitational constant (see, e.g., [32], [25] for derivations of this result). See also Figure 3.1. Under the *thin lens approximation*, which assumes that all the deflecting mass lies entirely in a two-dimensional intermediate deflector plane and

[1]The MATLAB functions `lens` and `invert` that are described in this article are available at http://www.cs.dal.ca/~spiteri/siam.

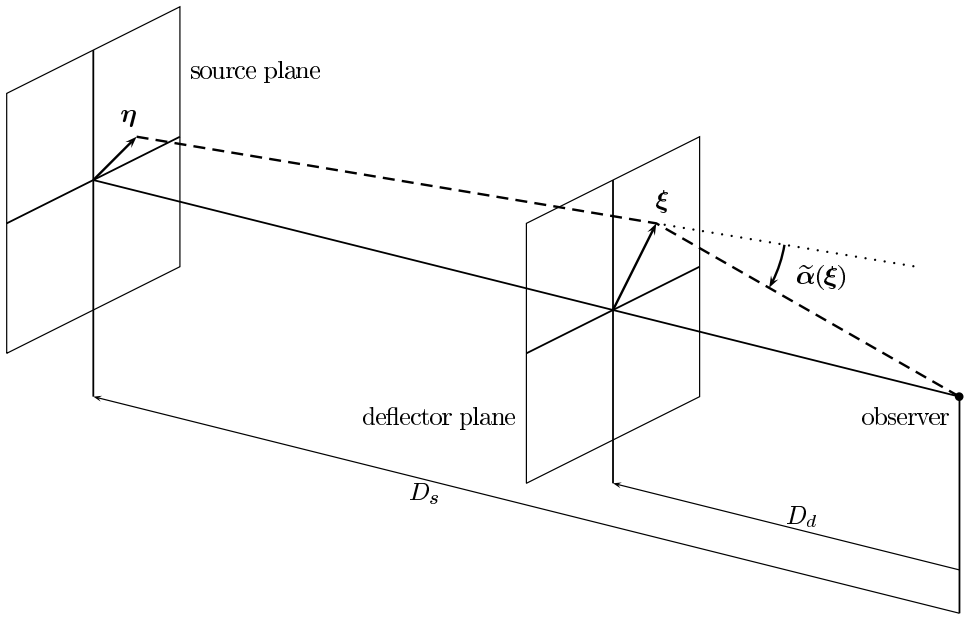

Fig. 3.1 *A ray emitted at $\boldsymbol{\eta}$ on the source plane pierces the deflector plane at $\boldsymbol{\xi}$. The ray is deflected by $\widetilde{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ onto the observer at O.*

that the ray is abruptly deflected as it passes through the deflector plane, this result can be extended to arbitrary (two-dimensional) mass distributions $M(\boldsymbol{\xi})$.[2] For every element of mass $dM = M(\boldsymbol{\xi}')d^2\xi'$ at position $\boldsymbol{\xi}'$ on the deflector plane, the ray of light piercing the deflector plane at position $\boldsymbol{\xi}$ is deflected by an angle with magnitude

$$d\widetilde{\alpha}(\boldsymbol{\xi}, \boldsymbol{\xi}') = \frac{4G}{c^2}\frac{dM}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|}$$

in the direction $-(\boldsymbol{\xi}-\boldsymbol{\xi}')/|\boldsymbol{\xi}-\boldsymbol{\xi}'|$ (a unit vector toward the mass element dM); i.e., we characterize the deflection due to the mass element by a two-dimensional vector. The total deflection $\widetilde{\boldsymbol{\alpha}}$ is therefore the sum of these deflection vectors over each element of mass in the deflector:

$$\widetilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \int\!\!\int \frac{4GM(\boldsymbol{\xi}')}{c^2}\frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}\, d^2\xi'. \tag{3.1}$$

The simple phenomenon of bending of light rays does not guarantee the existence of a lens, however. The relevant question is, Of all the rays emitted by a background source, which enter the observer's telescope? In other words, how does this lens *focus*? As described below, the *lens equation* selects the rays that leave the source and are deflected through exactly the correct angle to strike the observer's telescope.

We now derive the lens equation based on geometric arguments. The deflections involved in gravitational lensing are very small, only tens of arcseconds at most,[3] so

[2] We employ boldface symbols to represent two-component vectors; all other symbols are scalars.

[3] For comparison, a penny held at a distance of approximately a quarter of a mile appears to be about 10 arcseconds in diameter.

we assume certain angles are small where appropriate. Consider a mass distribution that produces a deflection $\widetilde{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ at point $\boldsymbol{\xi}$ on the deflector plane (see Figure 3.1). Light emitted at position $\boldsymbol{\eta}$ on the source plane strikes the deflector plane at $\boldsymbol{\xi}$ and is deflected to the observer at O if

$$\boldsymbol{\eta} = \frac{D_s}{D_d}\boldsymbol{\xi} - D_{ds}\widetilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{3.2}$$

where D_d, D_s, and D_{ds} are the distances from the observer to the deflector plane, from the observer to the source plane, and from the deflector plane to the source plane, respectively.

In the lens equation (3.2), $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are vectors that represent physical distances, usually measured in megaparsecs (Mpc), on the deflector and source planes. These vectors are not very useful for observers on Earth, however, as there is no way to directly measure them. As is the case with most mathematical models, the lens equation becomes much simpler to analyze and interpret if we nondimensionalize the coordinates. We define $\boldsymbol{x} = \boldsymbol{\xi}/\xi_o$, where

$$\xi_o = \frac{2\pi}{360 \cdot 60 \cdot 60} D_d = \frac{2\pi}{1\,296\,000} D_d;$$

the quantity $\boldsymbol{\xi}/D_d$ is an angle measured in radians (via the small angle assumption), and the leading factor of $2\pi/1\,296\,000$ converts $\boldsymbol{x}$ into an angle measured in arcseconds on the sky. Similarly, we define

$$\boldsymbol{y} = \frac{\boldsymbol{\eta}}{\frac{D_s}{D_d}\xi_o},$$

which locates objects on the source plane and which is also measured in arcseconds on the sky. The quantities $\boldsymbol{x}$ and $\boldsymbol{y}$ are still treated as vectors, with components taken in the north/south and east/west directions. With this scaling, the lens equation (3.2) simplifies to

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{\alpha}(\boldsymbol{x}), \tag{3.3}$$

where

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \frac{D_d D_{ds}}{\xi_o D_s}\widetilde{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

is the scaled deflection angle, also measured in arcseconds on the sky.

Simulations of gravitational lenses are generally numerical rather than analytical because *beams* (not simply rays) of light pass through the lens. During this journey, the beams are cross-sectionally distorted. This process can squeeze the beam of light, effectively forcing more rays through a smaller area and hence brightening the appearance of the background source. Astronomers refer to this as *magnification* of the image. Because of the vast distance to the source, often *only* these magnified images are visible; i.e., the source may be so faint it would be undetectable in the absence of gravitational lensing.

We can use the lens equation to determine analytically the amount by which light is magnified after passing through the lens. Suppose the observer detects light with flux I coming from a window of size $(dx)^2$ on the deflector plane. The *surface brightness* is the ratio $I/(dx)^2$, which is essentially a measure of the density of the observed signal. The source of this light is a window of size $(dy)^2$ on the source plane

emitting light with flux S. Gravitational lensing preserves the surface brightness of the light:

$$\frac{I}{(dx)^2} = \frac{S}{(dy)^2}.$$

This means the observed flux I is related to the emitted flux S by the ratio of the areas of the two windows,

$$I = \frac{(dx)^2}{(dy)^2} S.$$

The Jacobian $A_{ij} = \partial y_i / \partial x_j$ of (3.3) encodes the reciprocal ratio $(dy)^2/(dx)^2$, from which we can determine the magnification factor, $\mu(\boldsymbol{x}) = 1/|\det(A(\boldsymbol{x}))|$, where $\det(\cdot)$ denotes determinant. The points $\boldsymbol{x}$ where $\det(A(\boldsymbol{x})) = 0$ trace the *critical lines* of the lens. These are curves along which the beam of light is "squeezed shut" by the distortion, producing high (in theory, infinite) magnification.

For lenses with spherical symmetry, we can derive an analytical expression for the magnification. If the deflector mass is spherically symmetric about $\boldsymbol{x} = \boldsymbol{0}$, the lens equation can be written as

$$\boldsymbol{y} = \boldsymbol{x} - \alpha(r)\widehat{\boldsymbol{r}}, \tag{3.4}$$

where $\widehat{\boldsymbol{r}}$ is a radial unit vector and α depends only on the distance $r = |\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2}$ from the center of the mass distribution. A straightforward exercise in differentiation and algebraic manipulation[4] leads to the result

$$\det(A) = \left(1 - \frac{\alpha}{r}\right)\left(1 - \frac{d\alpha}{dr}\right). \tag{3.5}$$

As the magnification plays a critical role in the appearance and detection of gravitationally lensed images, it is useful to understand the network of critical lines throughout the lens.

3.1. Parametric Models: Singular Isothermal Spheres. One of the simplest ways to produce simulations of gravitational lenses is to use a *parametric model*, a procedure that turns out to be particularly well suited for computer simulations. In a parametric model, the mathematics and physics of a phenomenon are reasonably well understood, and the results of the simulations depend on the values of a number of variables, or parameters.

The toaster in your kitchen is an example of a parametric system with one parameter: the position of the lever that controls how dark you like your toast (e.g., light, brown, or burnt). Thanks to this parameter, you will not need different toasters to make different kinds of toast. The physics, mathematics, and engineering behind how toasters work (electricity coming on, elements heating up, etc.) do not change. The predictable result, your toast, depends only on the particular value of the model parameter, the position of the lever. With some trial and error, you can usually find an optimal parameter value that suits your taste.

To produce simulations of gravitational lenses, we choose a deflecting mass distribution $M(\boldsymbol{\xi})$ and a distribution of light on the background source plane to "observe" through the lens. As a first approximation to the dark matter that is thought

[4]A worksheet to derive (3.5) with the computer algebra software package Maple is available at http://www.cs.dal.ca/~spiteri/siam.

to inhabit the universe, modelers often turn to the spherically symmetric, *singular isothermal sphere* (SIS), which is a mass distribution described by

$$M(\xi) = \frac{1}{2G}\frac{\sigma^2}{\xi}. \tag{3.6}$$

A simple physical motivation for this distribution is as follows. If we imagine the dark matter to consist of an ideal gas of particles, the gas will reach hydrostatic equilibrium when the inward force of self-gravitation and the outward force of its own pressure are balanced. The density of the gas will be a function of the distance ξ from the center of the cloud, with the core of the cloud more densely packed than the outer reaches. In this equilibrium state, the gas will be isothermal, characterized by temperature T. Particles of dark matter are believed to behave in much the same way. When a cloud of dark matter relaxes to an equilibrium state, its density is a function of the distance from the center of the cloud. The "temperature" of the dark matter cloud is characterized by the cloud's line-of-sight velocity dispersion σ: distant galaxies are approaching or receding from our vantage point in the universe, so the *average* velocity of objects in these galaxies may be positive or negative. Astronomers measure the variations about the average velocity, the velocity *dispersion*, to gauge the mass of the cloud because, just like particles in a hot gas are moving relatively rapidly and particles in a cool gas are moving relatively slowly, more massive galaxies have relatively high velocity dispersions and less massive galaxies have relatively low velocity dispersions; hence, we can effectively treat σ as a *mass parameter*. The dark-matter SISs that produce gravitational lenses typically have velocity dispersions in the range 100–1000 km/s. Readers interested in seeing how this distribution solves the differential equations describing the hydrostatic equilibrium reached by a self-gravitating, isothermal gas are encouraged to consult [4, section 4.4].

The SIS is an ideal parametric model for studying lensing. Not only does it have a well-founded origin in galactic dynamics, but the mathematics of lensing by a SIS can be made very simple by making the following two observations. Imagine a ray of light passing a distance ξ from the center of a SIS. First we note that, as with any gravitating system, the gravitational interaction at radius ξ is completely determined by the interaction with annuli of mass having radii $\xi' < \xi$ because the contributions due to mass outside the annulus at ξ conspire to balance perfectly. Second, the gravitational interaction due to an elemental annular mass with radius ξ' is equal to the interaction with an equivalent point mass sitting at $\xi' = 0$. Thus, the deflection angle $\widetilde{\boldsymbol{\alpha}}$ points towards the center of the SIS, and the magnitude, given by a version of (3.1) that is simplified using these two observations, is *constant*:

$$\begin{aligned}\widetilde{\alpha}(\xi) &= \int_0^{\xi} \frac{4G}{c^2}\,\frac{\sigma^2}{2G\xi\xi'}\,2\pi\xi' d\xi' \\ &= 4\pi\left(\frac{\sigma}{c}\right)^2. \end{aligned} \tag{3.7}$$

Direct measurement of the distances D_d to the deflector plane and D_s to the source plane are clearly impossible. Instead, astronomers rely on the Hubble law [28], which states that the speed at which an object is receding from the Earth is directly proportional to the distance to that object. The redshift z of an object measures the speed of that object relative to observers on Earth (light from receding objects is said to be red-shifted; light from approaching objects is blue-shifted) so that distance is

directly proportional to redshift. The constant of proportionality is c/H_o, where H_o is the Hubble constant, giving

$$D = z\frac{c}{H_o} \text{ Mpc.}$$

The value of the Hubble constant is one of the "holy grails" of astronomy, for it is a direct measure of the expansion rate of the universe. Its value is hotly disputed, although most astronomers agree that H_o lies in the range 50–100 km/s/Mpc. The Hubble law holds well for "nearby" objects, where the expansion of the universe is fairly uniform, but breaks down over very large distances, where the large-scale curvature of the universe becomes important. If we make the assumption that distance is proportional to redshift for all redshifts, then the ratio D_{ds}/D_s in the lens equation is given by $(z_s - z_d)/z_s$, where z_s and z_d are the redshifts of the source and deflector, respectively. The lens equation (3.4) becomes

$$\boldsymbol{y} = \boldsymbol{x} - 28.8\,\frac{z_s - z_d}{z_s}\left(\frac{\sigma}{1000}\right)^2 \frac{\boldsymbol{x}}{|\boldsymbol{x}|}.$$

(Notice that the magnitude of the deflection angle is constant, and only the direction of the deflection changes.) Because $\widetilde{\alpha}$ and α are constant, (3.5) locates the critical lines of this lens at all points where $\alpha = |\boldsymbol{x}|$, namely the circular *Einstein ring*[5] that surrounds the center of the SIS at *Einstein radius*

$$R_E = 28.8\,\frac{z_s - z_d}{z_s}\left(\frac{\sigma}{1000}\right)^2.$$

The term $(\sigma/1000)^2$ in these last two equations reveals a practice that many scientists use when mathematical formulas arise in their research. Clearly, the factor of $(1000)^2$ in the denominator could be incorporated into the leading coefficient 28.8 without changing the numerical results. However, numbers like 0.0000288 or 2.88×10^{-5} are more difficult to work with in your head. This leading coefficient is then to be multiplied by σ^2, which is typically quoted in the range $10^4 - 10^6$ km^2/s^2. By writing this as $(\sigma/1000)^2$, the factor of 1000 normalizes the units of the velocity dispersion σ, making mental computations more tractable for typical parameter values and helping to build an intuitive feeling for the calculations. For example, if an Einstein ring is observed with a radius of about 4 arcseconds, and the background source is about twice as far as the lensing mass (so that $(z_s - z_d)/z_s \approx \frac{1}{2}$), then we know almost immediately that $(\sigma/1000)^2 \approx \frac{1}{4}$ and the SIS has $\sigma \approx 500$ km/s.

Changing the mass of the SIS or its location in the sky does not change the physics of lensing or the mathematics used to describe it, only the particular realization of the phenomenon. A parametric model can therefore be formulated (and programmed) once, and subsequently sets of model parameters can be read in for each simulation. There are several typical steps to produce a simulation.

(1) Choose the *model parameters*, which can be divided into three sets:
   (a) *Observation parameters*:
      - height and width of the observation, in pixels, N
      - angular size of each pixel, Δ

[5]It is interesting to note that Chwolson [6] was the first to indicate the possibility of a ring-shaped image formed by a gravitational lens. Accordingly, it has been suggested [3] that these rings would more correctly be called *Chwolson rings*.

Table 3.1 *Model parameters for the simulation shown in Figure* 3.2.

Parameters		
Observation	Deflector mass	Source
$N = 128$ $\Delta = 0.2$	$\boldsymbol{x}_{SIS} = (0, 0)$ $z_d = 0.3$ $\sigma = 600$	$\boldsymbol{y}_{ELL} = (1.0, 1.5)$ $z_s = 0.7$ $a = 1.5$ $e = 0.4$ $\theta = 40°$ $S = 1$

For example, the wide-field/planetary camera on the *Hubble Space Telescope* produces 800×800 pixel images with $\Delta = 0.0455$ arcseconds.[6] (By comparison, the full moon is about 0.5 degrees or 1800 arcseconds across, nearly 50 times wider than the *Hubble*'s field-of-view!)

(b) *Deflector mass parameters*:
- coordinates of the center of the SIS, $\boldsymbol{x}_{SIS}$
- redshift to the deflector plane, z_d
- mass parameter, σ

(c) *Source parameters*: A uniform, elliptical disk of light on the source plane is described by
- coordinates of the center of the elliptical disk, $\boldsymbol{y}_{ELL}$
- redshift to the source plane, z_s
- semimajor axis of the ellipse, in arcseconds, a
- ellipticity[7] of the ellipse, $e = 1 - b/a$
- orientation of the major axis, θ, measured counterclockwise from the positive x-axis
- flux of the source, S

(2) Shoot a ray through the center $\boldsymbol{x}$ of each of the N^2 pixels, calculate the deflection angle $\boldsymbol{\alpha}(\boldsymbol{x})$, and solve the lens equation for the point $\boldsymbol{y}$ where the ray pierces the source plane. This is opposite to the direction the light travels through the lens, but this is valid because gravitational lenses show *optical reciprocity*: light travels the same way in both directions through the lens.

(3) Determine whether the signal S was emitted from the source: The point $\boldsymbol{y} = (y_1, y_2)$ is inside the background ellipse (and hence light is emitted from point $\boldsymbol{y}$) if

$$\frac{y_1^2}{a^2} + \frac{y_2^2}{b^2} \leq 1,$$

where the coordinates (y_1, y_2) are relative to the principal axes of the ellipse.

(4) Modify the signal by the magnification factor $\mu(\boldsymbol{x})$ and fill the pixel at location $\boldsymbol{x}$ in the observation with the value μS.

These steps are coded in the MATLAB function `lens`, which efficiently generates simulations using MATLAB built-in `.*` and `./` array operations. The parameters listed in Table 3.1 are stored in the text file `SIS1`, and the simulation shown in Figure 3.2 is produced by executing the command `lens('SIS1')` in a MATLAB session.

[6] Full documentation on the *Hubble Space Telescope* is available on-line through the Space Telescope Science Institute at http://www.stsci.edu.

[7] Astronomers seem to prefer this measure of the shape of the ellipse, rather than the eccentricity $\sqrt{1 - (b/a)^2}$, again perhaps because it is easier to compute mentally.

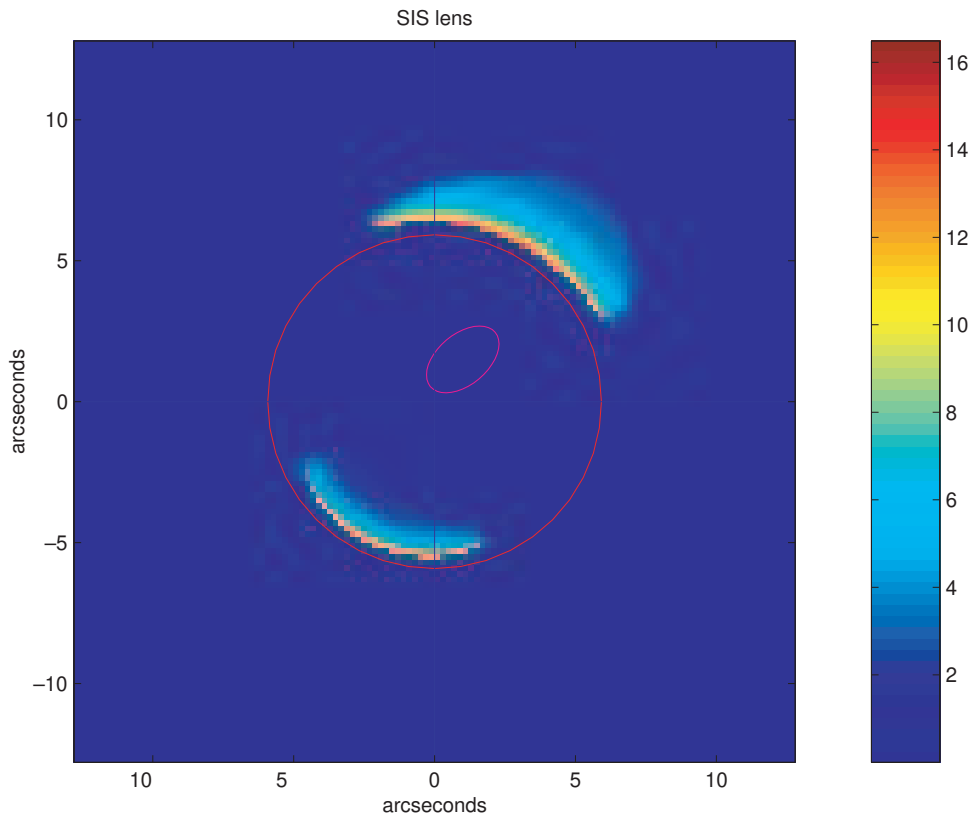

Fig. 3.2 *A SIS lens produces two distorted images of an elliptical background disk. The red circle is the Einstein ring of the lens, along which the magnification diverges. The position of the background source, as it would appear in the absence of lensing, is marked by the magenta ellipse. The background ellipse has uniform flux $S = 1$, so the color scale encodes the magnification factor.*

Interested readers are encouraged to modify the parameter file `SIS1` to explore the intricate nature of gravitational lenses. Minor modifications of the `lens` function allow the user to interactively move the background source, change the mass parameter σ, etc. Through this experimentation it becomes clear that gravitational lensing is an *ill-posed problem*: small changes in the model parameters can produce large changes in the results. For example, moving the background source by only one Einstein radius can transform the lensed arc from a spectacular Einstein ring into a single, barely distorted copy of the background source.

4. Inverting Gravitational Lenses. As we have seen in the previous section, parametric models are associated with the solution to the *forward problem*: determine the data produced by a completely specified parametric model. Closely associated with the forward problems are the *inverse problems* that try to recover the values of the model parameters necessary to reproduce a given observation. Unlike forward problems, inverse problems are often *not well posed* (or *ill posed*). Formally, we say a problem is well posed if it has a unique solution that depends continuously on the "data" (i.e., the initial and/or boundary conditions or any unknown parameters in the model). This means that small perturbations to the data of a well-posed problem re-

sult in small perturbations to the (unique) solution of the perturbed problem. Inverse problems are said to be ill posed because, for example, it may take wildly different models to account for sets of observations that look very similar, or possibly a number of models describe a given set of observations equally well. This is problematic because observations necessarily contain errors (or *noise*), making small perturbations a fact of life. Researchers often attempt to *regularize* the problem (see, e.g., [37]) so that the solution to the regularized problem behaves much like the solution to a well-posed problem that is in some sense "close to" the original problem. Fortunately, in practice, some solutions can also be discarded based on other considerations, like physics or common sense. Unfortunately, deciding the criteria that define the "solution" may be an ill-posed problem in itself; slightly different criteria may lead to drastically differing models.

Continuing with our earlier analogy, if someone hands you a burnt piece of toast, you instinctively solve the inverse-toaster problem and conclude with some certainty that the lever on the toaster was pushed all the way to burnt. Of course, there might be other solutions: maybe the toast was broiled in the oven for too long or maybe it was already burnt before it went into the toaster. However, we may have reason to suspect that these other solutions are improbable.

As we have seen in section 3, it is relatively simple to produce forward simulations of gravitational lensing. Students interested in gravitational lensing can gain a greater understanding of the nature of the lensing and its unique characteristics through experimentation with the forward problem. However, to use gravitational lensing as a tool to probe for dark matter, we must be able to invert the lens.

Suppose a set of observations reveals a collection of lensed images at positions $\boldsymbol{x}$. Because we know neither the deflection angle $\boldsymbol{\alpha}(\boldsymbol{x})$ (we do not know the mass distribution $M(\boldsymbol{x})$) nor the location $\boldsymbol{y}$ of the source of light on the background plane, we cannot solve the lens equation (3.3). The goal of lens inversion is to find a *physically meaningful* mass distribution and background source.

There are some natural constraints on this inverse problem. For example, a trivial solution for the lens shown in Figure 3.2 would be to conclude that the image was obtained from two crescent-shaped galaxies on the background source plane without any lensing whatsoever. This is *not* a probable solution, however, for galaxies are not generally crescent-shaped, in a pair, and with one very bright edge. As with many modeling problems in applied mathematics, we try to find the simplest solution that reproduces the observations because the simplest solution is likely to be the one that is least biased and most probable. With this philosophy, we arrive at a solution strategy based on two facts:

(1) Most galaxies are elliptical in appearance.
(2) If the lensed images seen in Figure 3.2 really are images of the same background (elliptical) source, then when we "push" the images back to the source plane (guided by the lens equation based on our current model parameters), the two bundles of rays must coincide and recreate the source.

Another way to view this second guideline is as a consistency check of the answer, offering another important tool in the modeling process. Under these guidelines, we adjust the model parameters until the "best" background source is reconstructed from the data.

Unfortunately, as is often the case, the definition of "best" varies from researcher to researcher. Let us choose a simple definition: the best background source is the smallest ellipse that encloses all of the points where the backward-traced rays pierce the source plane. This allows us to formulate the lens inversion problem as an op-

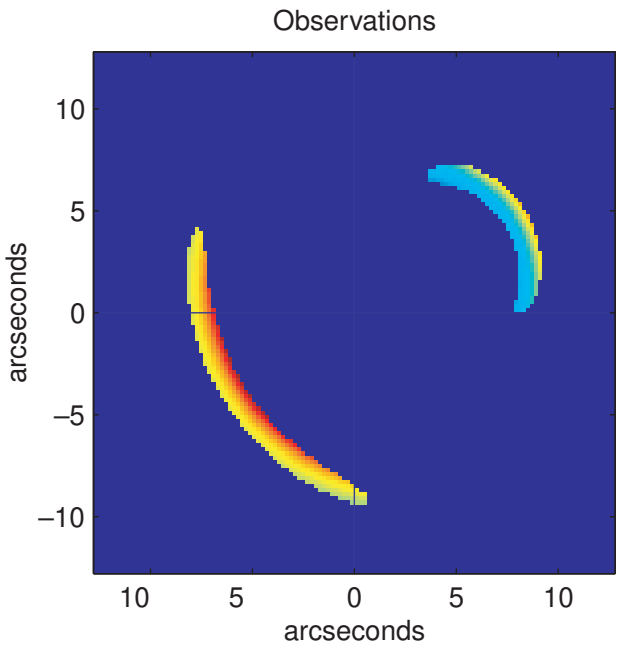

Fig. 4.1 *Observations of an unknown SIS lens.*

timization problem. The solution will be an optimal set of model parameters that describe the natural appearance of the background source and size and location of the SIS. This prediction can then be compared to the luminous mass of the deflector in the search for dark matter. Moreover, this also makes lensing useful for studying the appearance of very distant—and hence very old—elliptical galaxies.

More specifically, the inversion process goes as follows. We start with a set of observations:

- A description of the data structure: the height and width N of the observation (in pixels) and the size Δ of each pixel (in arcseconds). This is generally a description of the astronomical instrument used to collect the data.
- The data: an $N \times N$ array of values, which may be integers or floating-point numbers, again depending on the instrument used.

As in many applications of lens inversion routines like this one, the redshifts of the deflector and the source are observed and therefore can be treated as known in the model. For concreteness, we consider the observations shown in Figure 4.1, which is a 128×128 array of pixels of size 0.2 arcseconds. These observations were created using the '`w - write image to file`' option while running the forward modeling routine, `lens('SIS2')`. We assume that we know the deflector sits at redshift $z_d = 0.3$ and the source sits at redshift $z_s = 0.7$.

To begin the inversion, the user invokes the MATLAB program `invert('SIS2')` and uses the mouse to make an initial guess at the center $\boldsymbol{x}_{SIS}$ of the SIS. By default, the SIS mass parameter σ is initially set to a nominal value of 500 km/s. Using the numeric keypad, the user then interactively alters the values of $\boldsymbol{x}_{SIS}$ and σ. For each set of model parameters, the observations are traced back to the source plane. As a measure of how well the bundles of light coincide, the points on the source plane are surrounded by the ellipse with the smallest area. To find this best ellipse,

we employ perhaps the most basic optimization scheme for inequality constrained problems, namely, a modification of Courant's classical penalty method ([8]; see also, e.g., [15]). This allows us to solve a sequence of *unconstrained* optimization problems that converges to the solution of the constrained problem under mild assumptions. Such a strategy is a simple yet effective enhancement to unconstrained optimization routines, which are more common in software libraries and easier for students to understand and program from scratch.

For each choice of $\boldsymbol{x}_{SIS}$ and σ, the built-in MATLAB function `fminsearch` finds the values of the center $\boldsymbol{y}_{ELL}$, semimajor axis a, ellipticity $e = 1-b/a$, and orientation θ of the surrounding ellipse, which minimizes the objective function

$$\phi(\boldsymbol{y}_{ELL}, a, e, \theta; \lambda, \epsilon_i) = \pi a^2(1-e) + \frac{\lambda}{2}\sum_i \epsilon_i \left(\frac{(y'_{i,1})^2}{a^2} + \frac{(y'_{i,2})^2}{a^2(1-e)^2} - 1 \right).$$

The summation runs over the subset of the N^2 pixels in the observations that contain light. The first term corresponds to the area πab of the ellipse. The second term is a penalty term. The coordinates $\boldsymbol{y}'_i$ are the solutions to the lens equation $\boldsymbol{y}_i$ that pass through the pixels at $\boldsymbol{x}_i$ in the observations, transformed into the principal axes of the ellipse via the transformation

$$\begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} y_1 - y_{ELL,1} \\ y_2 - y_{ELL,2} \end{pmatrix};$$

hence $\boldsymbol{y}'_i = \boldsymbol{y}'_i(\boldsymbol{y}_{ELL}, \theta)$. Notice that the center $\boldsymbol{y}_{ELL}$ of the ellipse is mapped to the origin $\boldsymbol{y}' = (0,0)$ of the transformed coordinate system. By defining

$$\epsilon_i = \begin{cases} 0 & \text{if } \boldsymbol{y}'_i \text{ is already inside the ellipse,} \\ 1 & \text{if } \boldsymbol{y}'_i \text{ is outside the ellipse,} \end{cases}$$

this objective function does not give any weighting to points that are already inside the ellipse. As is typical in a penalty method such as this, we increase λ (say, from 10^1 to 10^{10} by factors of 10) to iterate to the best ellipse. At the end of this iteration, we will have found the formula for the smallest ellipse that contains all the points $\boldsymbol{y}_i$.

An intermediate result is shown in Figure 4.2. The ellipse drawn in the right panel is the smallest ellipse that surrounds the data points. The model parameters are not yet optimal, as the two beams of light in the right panel are not yet coincident. However, there is a very clear correspondence between the beams of light in the two panels.

By altering the model parameters, we arrive at a solution shown in Figure 4.3. The parameters are listed in Table 4.1, along with the actual parameters used to create the observations. We note that although the individual minimization problems have unique solutions, the user must supply an initial guess when running the routine `invert`, so the final values of the estimated parameters may vary slightly from user to user. This is another manifestation of the ill-posed nature of the inversion problem: the lack of a unique solution. And as we will see later, other reasonable solution strategies are possible, and these could also lead to different possible solutions. We halt our search for a solution at this point because the background source satisfies the criteria of our solution strategy:

(1) The data almost perfectly recreate an elliptical background source of light.
(2) The data from the two lensed arcs coincide.

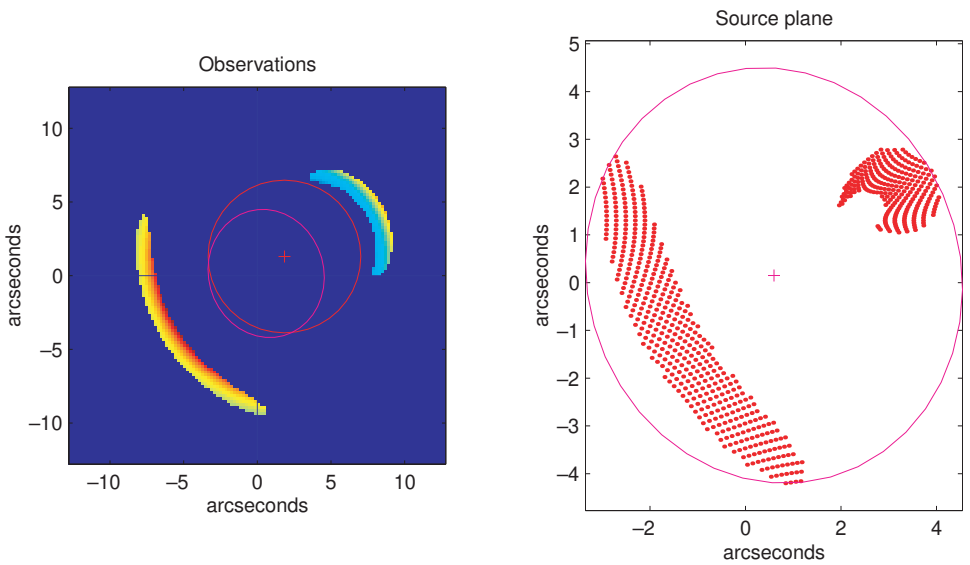

Fig. 4.2 *(left) The observations and the intermediate SIS mass, characterized by its center (marked with a red +) and its Einstein ring (red circle). (right) The center of each pixel in the observations containing light is traced back to the source plane and marked with a dot. The points are then surrounded by the ellipse with the smallest area, drawn in magenta in both panels. The model parameters are not yet optimal, as the two beams of light do not yet coincide.*

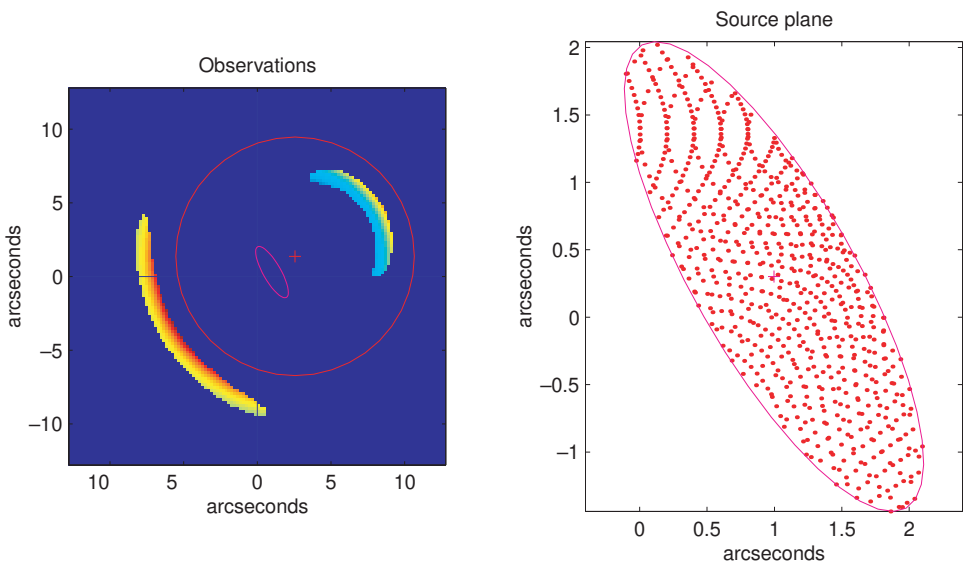

Fig. 4.3 *A solution to this SIS lens. (left) The two lensed arcs follow the Einstein ring of the lens. (right) When the observations are traced back to the source, an elliptical source is re-created. Notice that this source and ellipse occupy only a small fraction of the right panel of the intermediate result in Figure* 4.2*, emphasizing how closely the beams of light coincide.*

Table 4.1 *Model parameters estimated from the inversion routine and the actual values in the parameter file* `SIS2` *used to create the observations in Figure* 4.1.

Parameter	Estimated	Actual	Method of determination
Redshift z_d		0.3	Observed
Redshift z_s		0.7	Observed
Center $\boldsymbol{x}_{SIS}$	$(2.563, 1.373)$	$(2.600, 1.400)$	Chosen by user
Mass σ	750	750	Chosen by user
Center $\boldsymbol{y}_{ELL}$	$(0.997, 0.301)$	$(1.000, 0.300)$	Constrained minimization
Semimajor axis a	1.978	2.000	Constrained minimization
Ellipticity $e = 1 - b/a$	0.697	0.700	Constrained minimization
Orientation θ	119.7	120.0	Constrained minimization

We are quite satisfied with this solution qualitatively because the two arcs follow the Einstein ring and show magnifications that diverge towards the Einstein ring. Moreover, the background source is uniformly covered in data, mimicking the fact that these two arcs are observations of an elliptical disk with a constant, *uniform* signal $S = 1$. More importantly, this source requires no extraneous characteristics to reproduce the lensed arcs besides the postulated ellipticity and uniformity. To infer additional features of the background source would require additional data, which we do not (yet) have. In other words, we are extracting no more information from the result than is justified.

We note that there are well-known potential drawbacks to this penalty function method. In particular, the conditioning of the minimization problem with the augmented formulation of ϕ deteriorates with increasing penalty parameter λ. This means that it can be very difficult for the unconstrained minimization routine to converge to the minimum of ϕ for a given (large) value of λ. Despite this, we typically obtain satisfactory results for this problem. We comment on possible extensions to the computational aspects of the solution method in the next section.

5. Extensions. The forward and inverse modeling procedures described above are quite general. The particular examples we include are straightforward. In this section, we discuss some of the extensions to the procedures.

5.1. More Complex Lens Models. Although the singular isothermal spheres in the examples are useful as benchmark problems, they are not very realistic. The models become much more realistic when we add ellipticity to the mass distributions of the lens in order to produce singular isothermal ellipsoids described by

$$M(\boldsymbol{\xi}) = \frac{1}{2G} \frac{\sigma^2}{\sqrt{\xi_1^2 + (1-e)^2 \xi_2^2}},$$

where e is the ellipticity of the mass distribution and the coordinates $\boldsymbol{\xi}$ generally contain additional parameters because the ellipsoid can be rotated and translated in the coordinate frame of the observations. Even more realistic mass distributions include the pseudo-isothermal elliptical mass distribution [22], which removes the central singularity of the SIS by introducing a core radius, ξ_c. The spherical version is described by the mass distribution

$$M(\xi) = \frac{1}{2G} \frac{\sigma^2}{\sqrt{\xi^2 + \xi_c^2}}.$$

This distribution reduces to the SIS for large ξ, but it has no central singularity. Unfortunately, these small changes greatly increase the complexity of the computation of the deflection angle $\boldsymbol{\alpha}(\boldsymbol{\xi})$, although elegant methods do exist (see, e.g., [36], [24]).

The next step toward realistic models is to allow for multiple components in the mass distributions. Although single galaxies can act as gravitational lenses, there are many examples of lenses formed by clusters of galaxies that consist of a collection of galaxies surrounded by a large dark matter cloud (as an example, see the spectacular gravitational lens in the galaxy cluster Abell 2218 [2]). Fortunately, the deflection angle $\boldsymbol{\alpha}$ is the vector sum of the deflections due to each galaxy, so computations increase only linearly with the number of mass components. The intricacies of the observations are greatly increased, however, for the observations depend critically on the magnification factor $\mu(\boldsymbol{x})$ and hence on derivatives of the deflection angle $\boldsymbol{\alpha}$. Complex lenses often show a tangled network of critical lines along which $\mu(\boldsymbol{x})$ diverges.

Of course, as more and more mass components are added to make the models more realistic, the number of model parameters quickly increases. The number of data (the number of lensed arcs) remains the same, however. At the very least, the number of parameters should be comparable to the number of data. Ideally, there should be more data than parameters, so that the parameters can be checked for consistency. Adding more mass components without the addition of further data means the predictive power of the parametric models becomes weaker. Furthermore, as the number of parameters and the dimension of the parameter space increase, more and more local minima appear in the objective function. Finding the global minimum of a function is generally a more difficult problem, perhaps requiring optimization strategies that are more sophisticated than the simple interactive approach of our `invert` routine.

Some astronomers attack this problem with an exhaustive search of the many-dimensional parameter space, computing a statistic at each grid point of parameter space that measures the mis-fit between the observations and the simulation produced by the parametric model at that grid point. The precise form of the mis-fit varies from researcher to researcher. At the end of the search, the grid point that produces the smallest mis-fit is chosen to correspond to the optimal set of parameter values. Given sufficient computing power, this approach has led to the successful inversion of several gravitational lenses (see, e.g., [38], [23]).

Other astronomers (e.g., [40]) employ a more elegant approach using the so-called maximum entropy method (MEM). Optimal model parameters are selected by maximizing the amount of entropy, essentially minimizing the amount of information, stored in the collection of points where the backward-traced rays pierce the source plane. See [26] for an excellent review of the MEM, in which the authors alluded to the fact that there are essentially as many definitions of "entropy" as there are of "best."

5.2. The "Best" Solution. Readers who experiment with the interactive inversion routine `invert` will soon pose this problem for themselves: Why can't we get MATLAB to systematically search the $(\sigma, \boldsymbol{x}_{SIS})$-parameter space for the values that produce the optimal ellipse? Following up this problem reveals one of the frustrating yet intriguing aspects of gravitational lenses and lens inversion. When we implement an optimization scheme to find the values of σ and $\boldsymbol{x}_{SIS}$ corresponding to the ellipse around the data with the *globally* smallest area, we produce the result shown in Figure 5.1. The ellipse has an area $\pi ab = 3.321$, slightly smaller than the area $\pi ab = 3.722$ of the solution shown in Figure 4.3. However, one look at the solution shows us that something is wrong. For the same reasons that we are satisfied with

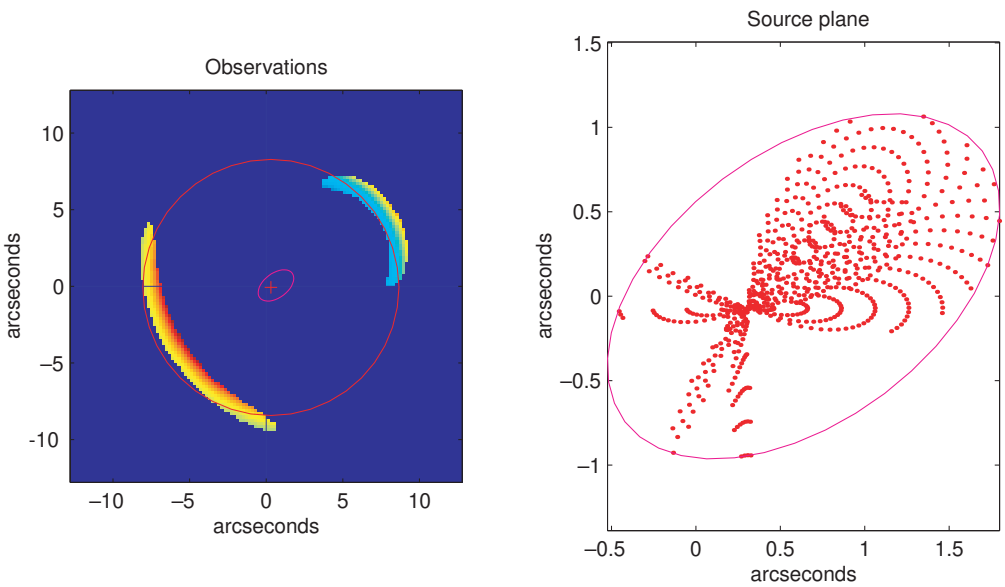

Fig. 5.1 *This lens produces a background ellipse whose area is even smaller than the one in Figure* 4.3. *However, it is not a reasonable solution to the problem.*

the solution shown in Figure 4.3, we reject this new solution. The lensed arcs do not follow the Einstein ring, nor do they exhibit the magnification associated with the ring. Furthermore, the reconstructed background source is not an elliptical cloud, but a highly distressed explosion of stars. This galaxy is not likely to exist, for its lobes and jets of matter are highly improbable. From the point-of-view of the scientific method, there is a great amount of additional structure (additional information) in the reconstructed background source, information we are not justified in extracting from the observations.

In our case, the problem can be interpreted as the lens in the unrealistic solution reducing the area of the background ellipse by "overdistorting" the beams of light. Evidently, our strategy of finding the smallest ellipse is insufficient to produce a reasonable solution. Fortunately, we have another set of data that can be used to avoid this situation: in addition to coinciding on the source plane, two backward-traced rays must carry the same demagnified signal. So far, our inversion scheme depends only on the coordinates of the pixels containing lensed light, not on the signals displayed in the pixels. By adding a third criterion to our inversion strategy, we may be able to avoid results like the one shown in Figure 5.1. Our strategy would thus be based on the following facts:

(1) Most galaxies are elliptical in appearance.
(2) If the lensed images are images of the same background (elliptical) source, then when we "push" the images back to the source plane (guided by the lens equation based on our current model parameters), the bundles of rays must coincide and recreate the source.
(3) When rays from different parts of the lensed images coincide on the source plane, the rays must carry the *same* demagnified signal.

We leave the implementation of this enhanced inversion scheme as a problem for further consideration.

A variation of this strategy that is used by astronomers (e.g., [23], [17]) is to first identify substructures with the lensed arcs, usually bright "knots" of light, and adjust the model parameters to make these knots coincident, both in position and signal, on the source plane. Then, with the parameters temporarily fixed, all of the data are traced back to the source plane and combined to check for consistency.

5.3. Other Computational Methods for Constrained Optimization. We have implemented a simple penalty function method that modifies the classical penalty method due to Courant into the computational method to solve the inequality constrained optimization problem. We seem to be fortunate enough to have an objective function for which more sophisticated techniques for nonlinear programming problems are not required. This makes the solution to this problem more tractable to students who may not have much training in nonlinear programming. At the same time, however, there is the opportunity for interested readers to implement their own techniques for constrained optimization. We briefly give an overview of some of the possibilities, in particular to give instructors some ideas as to other relevant topics.

Other forms for penalty terms can involve reciprocals [5] or logarithms [16] of the constraints. The latter form of penalty function has led to a popular class of constrained minimization techniques called *interior point methods* that preserve strict satisfaction of the constraints at all times during the minimization (see, e.g., [13], [42]). In particular, such methods have advantages for problems where the objective function is not defined when the constraints are violated. An *ideal* penalty function that avoids the ill-conditioning of the classical penalty method mentioned earlier is given in [14]. Existing MATLAB routines, in particular the suite of programs in the *Optimization Toolbox*, use a sequential quadratic programming algorithm to solve inequality constrained optimization problems.

We have not emphasized the use of techniques from global optimization for this particular application. This is partly due to the fact that our solution strategies may still be incomplete. In other words, there are no robust a priori criteria to determine whether the global minimum will yield the correct physical solution. Moreover, until we make the minimization problem much more difficult through the addition of more parameters in the model, we have found that the heavy machinery of global optimization routines is not worthwhile. However, this application affords an excellent stepping-off point for a discussion of global optimization strategies such as simulated annealing [20], genetic algorithms [9], or hybrid methods that combine both local and global minimization routines [30]. See also [27] for a useful survey on global optimization methods for nonconvex problems.

6. Conclusions. The study of gravitational lenses provides an interesting and relevant framework that incorporates a wide range of elements from applied mathematics and scientific computing. Although the routines described and the examples presented here are somewhat simplistic, they make a tractable (and perhaps even necessary) beginning to the study of gravitational lenses. Understanding even these simple lenses is the starting point for many researchers who use gravitational lenses as tools. The pattern of magnification of the arcs in Figure 3.2—both the height of the peaks and the rate at which the magnification rises and falls—is a signature of gravitational lensing. *Microlensing* studies search for missing mass in our own Milky Way galaxy in the form of massive compact halo objects (MACHOs) or weakly interacting massive particles (WIMPs) by searching for this pattern of magnification (see, e.g., [1], [35]). The arcs in Figure 3.2 are highly distorted images of the elliptical background source. Beams of light from elliptical background galaxies are only

weakly distorted and remain almost elliptical if they pass far from the center of the lens. *Weak lensing* studies (see, e.g., [21], [7]) determine the distribution of matter around galaxies and even clusters of galaxies by looking for small but statistically significant distortions of the background sky that are characteristic of gravitational lensing. Another effect of gravitational lensing is a possible *time delay* between the two (or more) lensed images, introduced because of the different path lengths and different parts of the gravitational potential well that the beams pass through on the way through the lens. This time delay makes it possible, in principle, to determine the Hubble constant H_o [29]; however, systematic errors have forced researchers, for now, to accept only ever-narrowing bounds on the value of this important constant [18].

As is typical with inverse problems, gravitational lens inversion is an ill-posed phenomenon. This makes the modeling a rather delicate proposition. As we have seen, it may be possible to invert a gravitational lens through an optimization scheme, but the choice of objective function to measure the quality of a solution is not obvious. Thus, the ensuing solutions may look very different despite the data or the objective functions looking very similar.

With the appropriate software, changing the values of the model parameters can be done interactively, with the new results displayed almost immediately on the computer screen. Based on our knowledge of the nature of lensing, the number of images formed, the characteristics of lens magnification and the Einstein ring, etc., students learning about lensing can adjust the parameters appropriately. We can get close to a "solution" in this way. Taking the final step of choosing the optimal set of model parameters is difficult, and consequently it is the focus of a great deal of lensing research. Fortunately, even these suboptimal results are useful to astronomers, who use lenses as tools to explore other phenomena such as the presence of dark matter, where often results accurate to within a factor of 2 are considered significant.

Acknowledgment. The authors would like to thank the referees for their helpful suggestions.

REFERENCES

[1] C. Alcock, *The dark halo of the Milky Way*, Science, 287 (2000), pp. 74–79.
[2] *Astronomy Picture of the Day*, NASA (2000 February 1). Retrieved from the World Wide Web: http://apod.gsfc.nasa.gov/apod/ap000201.html.
[3] J.M. Barnothy, *History of gravitational lenses and the phenomena they produce*, in Gravitational Lenses, J.M. Moran, J.N. Hewitt, and K.Y. Lo, eds., Springer-Verlag, Berlin, 1989, pp. 23–27.
[4] J. Binney and S. Tremaine, *Galactic Dynamics*, Princeton University Press, Princeton, 1987.
[5] C.W. Carroll, *The created response surface technique for optimizing nonlinear restrained systems*, Oper. Res., 9 (1961), pp. 169–184.
[6] O. Chwolson, *Über eine mögliche Form fiktiver Doppelsterne*, Astron. Nachr., 221 (1924), pp. 329–330.
[7] D. Clowe, G.A. Luppino, N. Kaiser, and I.M. Gioia, *Weak lensing by high-redshift clusters of galaxies.* I. *Cluster mass reconstruction*, Astrophys. J., 539 (2000), pp. 540–560.
[8] R. Courant *Variational methods for the solution of problems of equilibrium and vibration*, Bull. Amer. Math. Soc., 49 (1943), pp. 1–23.
[9] L. Davis, *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York, 1991.
[10] F.W. Dyson, A.S. Eddington, and C. Davidson, *A determination of the deflection of light by the Sun's gravitational field, from observations made at the total eclipse of May* 29, 1919, Monthly Notices of the Royal Astronomical Society, 62 (1920), pp. 291–333.
[11] A. Einstein, *Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes*, Ann. Physik, 35 (1911), pp. 898–908.
[12] A. Einstein, *Erklärung der Perihelbewegung des Merkur aus der allgemeinin Relativitätstheorie*, Sitzungsber. Preuß. Akad. Wiss. erster Halbband, (1915), pp. 831–839.

[13] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.
[14] R. Fletcher, *An ideal penalty function for constrained optimization*, J. Inst. Math. Appl., 15 (1975), pp. 319–342.
[15] R. Fletcher, *Practical Methods of Optimization*, Vol. 2, Wiley, Chichester, UK, 1981.
[16] K.R. Frisch, *The Logarithmic Potential Method of Convex Programming*, Memorandum, University Institute of Economics, Oslo, 1955.
[17] F. Hammer, I.M.Gioia, E.J. Shaya, P. Teyssandier, O. Le Fèvre, and G.A. Luppino, *Detailed lensing properties of the MS* 2137-2353 *core and reconstruction of the sources from* Hubble Space Telescope *imagery*, Astrophys. J. 491 (1997), pp. 477–482.
[18] D.B. Haarsma, J.N. Hewitt, J. Lehar, and B.F. Burke, *The Radio Wavelength Time Delay of Gravitational Lens* 0957+561, Astrophys. J., 510 (1999), pp. 64–70.
[19] P. Hewett and S. Warren, *Microlensing sheds light on dark matter*, Science, 275 (1997), pp. 626–627.
[20] L. Ingber, *Simulated annealing: Practice versus theory*, Math. Comput. Modelling, 18 (1993), pp. 29–57.
[21] N. Kaiser and G. Squires, *Mapping the dark matter with weak gravitational lensing*, Astrophys. J., 404 (1993), pp. 441–450.
[22] A. Kassiola and I. Kovner, *Elliptical mass distributions versus elliptic potentials in gravitational lenses*, Astrophys. J. 417 (1993), pp. 450–473.
[23] J.-P. Kneib, R.S. Ellis, I. Smail, W.J. Couch, and R.M. Sharples, *Hubble Space Telescope observations of the lensing cluster abell* 2218, Astrophys. J., 471 (1996), pp. 643–656.
[24] R. Kormann, P. Schneider, and M. Bartelmann, *Isothermal elliptical gravitational lens models*, Astronomy and Astrophysics, 284 (1994), pp. 285–299.
[25] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation*, Freeman, New York, 1973.
[26] R. Narayan and R. Nityananda, *Maximum entropy image restoration in astronomy*, Annual Review of Astronomy and Astrophysics, 24 (1986), pp. 127–170.
[27] P.M. Pardalos, D. Shalloway, and G.L. Xue, *Optimization methods for computing minima of nonconvex potential energy functions*, J. Global Optim., 4 (1994), pp. 117–136.
[28] P.J.E. Peebles, *Principles of Physical Cosmology*, Princeton University Press, Princeton, NJ, 1993.
[29] S. Refsdal, *On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect*, Monthly Notices of the Royal Astronomical Society, 128 (1964), pp. 307–310.
[30] J.M. Renders and S.P. Flasse, *Hybrid methods using genetic algorithms for global optimization*, IEEE Trans. Sys. Man Cyber. Part B, 26 (1996), pp. 243–258.
[31] J. Renn, T. Sauer, and J. Stachel, *The origin of gravitational lensing: A postscript to Einstein's* 1936 Science *paper*, Science, 275 (1997), pp. 184–186.
[32] W. Rindler, *Essential Relativity*, Springer-Verlag, New York, 1979.
[33] P. Schneider, J. Ehlers, and E.E. Falco, *Gravitational Lenses*, Springer-Verlag, Berlin, 1992.
[34] G. Schilling, *A magnifying glass for the Milky Way*, Science, 287 (2000), pp. 67–68.
[35] G. Schilling, *Signs of MACHOs in a far-off galaxy*, Science, 287 (2000), p. 779.
[36] T. Schramm, *Realistic elliptical potential wells for gravitational lens models*, Astronomy and Astrophysics, 231 (1990), pp. 19–24.
[37] A.N. Tikhonov and V. Ya. Arsenin, *Methods for Solving Ill-Posed Problems*, Wiley, New York, 1979.
[38] J.A. Tyson, G.P. Kochanski, and I.P. dell'Antonio, *Detailed Mass Map of CL* 0024+1654 *from Strong Lensing*, Astrophys. J., 498 (1998), pp. L107–L110.
[39] R.M. Wald, *General Relativity*, University of Chicago Press, Chicago, 1984.
[40] S. Wallington, C.S. Kochanek, and R. Narayan, *LensMEM: A gravitational lens inversion algorithm using the maximum entropy method*, Astrophys. J., 465 (1996), pp. 64–72.
[41] J. Wambsganss, *Gravitational Lensing in Astronomy*, *Living Reviews in Relativity*, Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Germany, 1998. Retrieved from the World Wide Web: http://www.livingreviews.org/Articles/Volume1/1998-12wamb/.
[42] S.J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.