STATISTICS OF GRAVITATIONAL MICROLENSING MAGNIFICATION.
I. TWO-DIMENSIONAL LENS DISTRIBUTION

LEV KOFLER
Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822; kofman@ifa.hawaii.edu

NICK KAISER
CIAR Cosmology Program, Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 1A7, Canada; kaiser@ifa.hawaii.edu

MAN HOI LEE
Department of Physics, Queen’s University, Kingston, ON K7L 3N6, Canada; mhlee@astro.queensu.ca

AND

ARIF BABUL
Department of Physics, New York University, 4 Washington Place, New York, NY 10003; babul@almuhit.physics.nyu.edu

Received 1996 August 21; accepted 1997 June 24

ABSTRACT

The propagation of light from distant sources through a distribution of clumpy matter, acting as point-mass lenses, produces multiple images that contribute to the total brightness of the observed macroimages. In this paper, we refine the theory of gravitational microlensing for a planar distribution of point masses. In an accompanying paper, we extend the analysis to a three-dimensional lens distribution.

In the two-dimensional case, we derive the probability distribution of macroimage magnification, \( P(A) \), at \( A - 1 \gg \tau^2 \) for a low optical depth lens distribution by modeling the illumination pattern as a superposition of the patterns due to individual “point-mass plus weak-shear” lenses. A point-mass lens perturbed by weak shear \( S \) produces an astroid-shaped caustic. We show that the magnification cross section \( \sigma(A|S) \) of the point-mass plus weak-shear lens obeys a simple scaling property, and we provide a useful analytic approximation for the cross section. By convolving this cross section with the probability distribution of the shear due to the neighboring point masses, we obtain a caustic-induced feature in \( P(A) \) that also exhibits a simple scaling property. This feature results in a 20% enhancement in \( P(A) \) at \( A \approx 2/\tau \).

In the low-magnification \( (A - 1 \lesssim 1) \) limit, the macroimage consists of a single bright primary image and a large number of faint secondary images formed close to each of the point masses. The magnifications of the primary and the secondary images can be strongly correlated. Taking into account the correlations, we derive \( P(A) \) for low magnification and find that \( P(A) \) has a peak of amplitude \( \approx 0.16/\tau^2 \) at \( A - 1 \approx 0.84\tau^2 \). The low-magnification distribution matches smoothly the distribution for \( A - 1 \gg \tau^2 \) in the overlapping regimes \( A - 1 \gg \tau^2 \) and \( A \ll 1/\tau \).

Finally, after a discussion of the correct normalization for \( P(A) \), we combine the results and obtain a practical semianalytic expression for the macroimage magnification distribution \( P(A) \). This semianalytic distribution is in qualitative agreement with the results of previous numerical simulations, but the latter show stronger caustic-induced features at moderate \( A \) \( (1.5 \lesssim A \lesssim 10) \) for \( \tau \) as small as 0.1. We resolve this discrepancy by reexamining the criterion for low optical depth. A simple argument shows that the fraction of caustics of individual lenses that merge with those of their neighbors is approximately \( 1 - \exp(-8\tau) \). For \( \tau = 0.1 \), the fraction is surprisingly high: \( \approx 55\% \). A simple criterion for the low optical depth analysis to be valid is \( \tau \ll 1/\tau \), though the comparison with numerical simulations indicates that the semianalytic distribution is a reasonable fit to \( P(A) \) for \( \tau \) up to 0.05.

Subject headings: gravitational lensing — methods: statistical

1. INTRODUCTION

Gravitational lensing provides a powerful independent tool to probe the distribution of matter in the universe and in individual astronomical objects. There are several situations in which the lensing objects can be modeled as an ensemble of point masses that produce multiple images (microimages) of distant sources. If the angular separations of the microimages are too small to be resolved obser-

1 Present address: Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822.
2 Present address: Department of Physics and Astronomy, University of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada.
that a substantial fraction of the dark matter in the universe is in the form of compact objects. Then the universe as a whole has a significant optical depth to microlensing (Press & Gunn 1973). Finally, microlensing can arise in a globular cluster if brown dwarfs are present (Paczynski 1994).

The main motivation of this study is to provide a systematic theory of gravitational microlensing by a random distribution of point masses. There are different regimes of microlensing and they often require different types of modeling. First, the spatial distribution of the lenses along the line of sight is usually either “compact” (e.g., stars in a distant galaxy) or “extended” (e.g., cosmologically distributed compact objects). A lens distribution is compact if the lenses are distributed in a region of scale R that is much smaller than the other distances (those between the source and the lenses and between the observer and the lenses) in the problem; it is extended if R is comparable to the other distances in the problem. Second, one has to distinguish between the cases of low, moderate (or nearly critical, \(\tau \sim 1\)), and high (or supercritical, \(\tau \gg 1\)) optical depth, where optical depth \(\tau\) is the fraction of the sky covered (locally) by the Einstein circles of all contributing lenses: \(\tau = \pi \sum n_i x_{E,i}^2\), where \(n_i\) is the projected (onto the observer plane) surface number density of lenses with Einstein radius \(x_{E,i}\) (see § 2). In the low optical depth limit, the Einstein circles of the individual lenses rarely overlap, and lensing by close groups of two (or more) point masses (in projection) is negligible. Although it is often assumed that \(\tau < 0.1\) is low optical depth, we are not aware of any analysis that has quantified how small \(\tau\) must be before the low optical depth condition is reached. In fact, as we shall see, \(\tau = 0.1\) is not sufficiently small in the case of compact lens distributions.

For a compact lens distribution, the deflection of light rays is generally small and essentially occurs in the vicinity of the lens distribution. Consequently, a compact lens distribution is usually approximated as a planar (two-dimensional) structure corresponding to its projection onto a lens plane that is oriented perpendicular to the line of sight between the observer and the source. This “single lens plane” approximation allows us to pose and explore well-defined questions concerning, for example, the caustic structure of the lens configuration, the resulting illumination pattern, the probability distribution of macroimage magnification, and the temporal variation of the observed brightness of a source (which is the magnification profile along a track through the illumination pattern). While these issues have been explored in great detail both analytically and through computer simulations (e.g., Chang & Refsdal 1979, 1984; Young 1981; Nityananda & Ostriker 1984; Kayser, Refsdal, & Stabell 1986; Paczyński 1986a; Schneider 1987a, 1987b; Lee & Spergel 1990; Wambsganss 1990, 1992; Witt 1990, 1993; Kaiser 1992; Mao 1992; Rauch et al. 1992; Lewis et al. 1993; Lewis & Irwin 1995; see also Schneider, Ehlers, & Falco 1992 and references therein), they are by no means completely resolved.

The greatest progress has been made in understanding the properties of single-plane lensing under low optical depth conditions. One of the earliest approaches (e.g., Turner, Ostriker, & Gott 1984) advocated modeling the low-\(\tau\) illumination pattern as a superposition of what one would obtain for a collection of point masses acting independently. An isolated point-mass lens produces a circularly symmetric illumination pattern with a divergent spike at the origin. Its differential cross section for magnification \(A\), \(\sigma(A]\), is proportional to the mass of the lens and scales as \(A^{-3}\) for \(A \gg 1\). A simple superposition of the differential cross sections implies that the probability distribution of macroimage magnification \(P(A)\) is

\[
P(A)dA = 2\tau A^{-3}dA \quad \text{for} \ A \gg 1
\]

(Paczynski 1986b) or, more generally,

\[
P(A)dA = 2\tau(\tau^2 - 1)^{-3/2}d\tau \quad \text{for} \ A - 1 \gg \tau^2.
\]

However, it is well known that neither the \(\tau\)-nor the \(A\)-dependence is correct in the above distributions, even in the limit \(A \gg 1\). As shown by Nityananda & Ostriker (1984; also Chang & Refsdal 1979, 1984), each lens is subject to an external shear from the macroscopic mass distribution, as well as from its nearest neighbors, and this external shear breaks the degeneracy of the pointlike caustic of the isolated point-mass lens. The resulting caustic has the shape of an “astroid,” with its width proportional to the dimensionless magnitude \(S\) of the shear perturbation, and the corresponding magnification cross section \(\sigma(A|S)\) shows a strong feature at \(A \sim 1/S\) (see § 3 for a more detailed analysis). This suggests that there ought to be a feature in \(P(A)\) at \(A \sim 1/\tau\), since it can readily be shown that both the typical magnitude of the macroscopic shear and its random component are of order \(\tau\). The natural extension of the superposition approach is to model the illumination pattern as a superposition of the patterns due to individual “point-mass plus weak-shear” lenses. Briefly, the resulting probability distribution of magnification is a convolution of the magnification cross section \(\sigma(A|S)\) for a single lens with the probability distribution function for the shear \(p(S)\). In a subsequent section, we pursue this approach and discuss the properties of the \(P(A)\) thus acquired.

Schneider (1987a) adopted a very different approach in attempting to determine the asymptotic behavior of \(P(A)\) at the limit of high magnification. He argued that the high-magnification events are dominated by observers close to the fold caustics and that the illumination pattern is hence dominated by two images of nearly equal brightness. This greatly simplifies the problem. Instead of having to identify all the microimages that make up a macroimage and then calculate the probability distribution function for their total magnification, one can simply use the probability distribution function for the individual image magnification. The latter can be calculated regardless of the optical depth. However, this approach has one drawback—the difficulty in quantifying how large \(A\) must be before \(P(A)\) is well approximated by its asymptotic form. In the low optical depth derivation discussed above, the asymptotic form is a good approximation only for very high magnification (\(A \gg \tau^{-1}\)). Nonetheless, it should be noted that Schneider’s derivation is the only rigorous analytic result for finite \(\tau\).

Understanding the behavior of \(P(A)\) at low magnification (i.e., at \(\delta A \equiv A - 1 \ll 1\)) is also problematic. The distribution quoted above (eq. [2]) is not properly normalized, as the integral \(\int dA \ P(A)\) diverges as the lower limit of the integral approaches \(A = 1\). One rather cavalier approach is simply to impose a cutoff at some \(A_{\min}\) such that the total probability is unity. Schneider (1987b) adopted a more sophisticated approach. He argued that a typical observer

\[ \sim \] denotes order-of-magnitude relationships.

\footnote{Throughout this paper, we frequently state that an expression or a derivation is valid if a variable \(x \gg x_E\) (or \(x \ll x_E\)). This usually means that \(x\) should be an order of magnitude larger (or smaller) than \(x_E\). Similarly, \(\sim\) denotes order-of-magnitude relationships.}
lensing, discussing the relevant equations and defining our total macroimage magnification associated with the lensing deflected rays—the so-called diffuse component. Then the position of the source and a large number of faint secondary lens) will see one primary image close to the unperturbed 510 KOFMAN ET AL. Vol. 489

and are statistically independent, Schneider magnification of the secondary images. Assuming that $P$ is normalized, and has the same amplitude $A_i$ formula that is valid for the primary and secondary images This yields a($magnification, we resolve the discrepancy in Schneider’s caustic-induced ” bump Ï” for low optical depth. For low 20% level. We also find a simple scaling property for the approach is two-pronged. For and that is useful, as well as physically motivated. Our position of the point-mass plus weak-shear lenses. We find that the random shear tends to smear out the strong caustic features of the individual lenses and that is small as 0.1. The additional $\phi(x)$ at distance $\chi$ from the source is given by

$\chi = \chi_0 \theta$. This defines our planar Lagrangian coordinates $x$. In the presence of density inhomogeneities, the light ray would suffer deflections and pierce the observer plane at the Eulerian coordinates $\rho$. The mapping from Lagrangian to Eulerian coordinates in the observer plane is

$$r(x) = x + s(x, \chi_0) ,$$

where the comoving displacement vector $s(x, \chi)$ at distance $\chi$ from the source is given by

$$s(x, \chi) = -2 \int_0^x d\chi (\chi - \chi') V\phi[x' + s(x, \chi'), \chi'] .$$

For most compact lens distributions of interest, the deflection of a light ray is small and occurs essentially in the vicinity of the lens. In effect, the deflection can be thought of as occurring at a plane that is situated close to the location of the lens(es). In this (single) thin-screen approximation, the lens mapping (eq. [5]) is a gradient mapping with the displacement vector $s = \Psi'$, where $\Psi$ is an effective surface gravitational potential. For the particular case of single-plane lensing by point masses (for the sake of simplicity, we assume that all point masses have the same mass $m$,)

$$s(x) = -4Gm \frac{\chi_{Lo} \chi_{so}}{d_L \chi_{Li}} \sum_k \frac{x - x_k}{|x - x_k|^2}$$

$$= -x_0^2 \sum_k \frac{x - x_k}{|x - x_k|^2} ,$$

where $x_k$ is the position of the $k$th lens projected onto the observer plane (i.e., $x_k = \chi_0 \theta_k$, where $\theta_k$ is the angular position of lens $k$ from the source), $a_L$ is the scale factor at the redshift of the lens plane, and $\chi_{Lo}$, $\chi_{so}$ and $\chi_{Li}$ are the comoving distances between the lens plane and the observer plane, between the source and the observer plane, and between the source and the lens plane, respectively (see Fig. 1). In addition,

$$x_E = \left(4Gm_0 \frac{\chi_{so}}{d_L \chi_{Li}}\right)^{1/2}$$

2. LENS EQUATIONS

We begin by outlining the basic concepts of gravitational lensing, discussing the relevant equations and defining our notation. Regarding the latter, we have adopted the notation of Kaiser (1992), some of which differs slightly from that in general use.

For the sake of simplicity, we shall only consider gravitational lensing in an Einstein–de Sitter cosmological background weakly perturbed by the gravitational field of the lenses. The line element for such a universe is

$$ds^2 = a^2(\eta)[(1 + 2\phi)d\eta^2 - (1 - 2\phi)(dx^2 + x^2 d\Omega^2)] .$$

Let the conformal time $\eta = 1$ at the present. The scale factor is $a = a_0 \eta^2$ and the Hubble parameter is $H = H_0 \eta^{-3}$ with $H_0 = 2/a_0$. In this convention, the unit comoving length is $a_0 = 6000 h^{-1}$ Mpc, where $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$. The Newtonian (peculiar) potential $\phi$ induced by the density inhomogeneities is given by Poisson’s equation (in comoving coordinates):

$$\nabla^2 \phi(\chi) = 6\eta^{-2} A(\chi) , \quad A(\chi) \equiv [\rho(\chi) - \bar{\rho}]\rho .$$

We consider an isotropic point source at the origin of our coordinates and an observer plane that is perpendicular to the $z$-axis and at a comoving distance $\chi_{so} \theta$ from the source (see Fig. 1). In the absence of perturbations, the observer plane would be uniformly illuminated, and a ray that leaves the source at angle $\theta = (\theta_1, \theta_2)$ would pierce the observer plane at $x = \chi_{so} \theta$. This defines our planar Lagrangian coordinates $x$. In the presence of density inhomogeneities, the light ray would suffer deflections and pierce the observer plane at the Eulerian coordinates $r$. The mapping from Lagrangian to Eulerian coordinates in the observer plane is

$$r(x) = x + s(x, \chi_0) ,$$

where the comoving displacement vector $s(x, \chi)$ at distance $\chi$ from the source is given by

$$s(x, \chi) = -2 \int_0^x d\chi (\chi - \chi') V\phi[x' + s(x, \chi'), \chi'] .$$

where $x' = \chi' \theta$. The mapping (eq. [5]) is a gradient mapping with the displacement vector $s = \Psi'$, where $\Psi$ is an effective surface gravitational potential. For the particular case of single-plane lensing by point masses (for the sake of simplicity, we assume that all point masses have the same mass $m$,)

$$s(x) = -4Gm \frac{\chi_{Lo} \chi_{so}}{d_L \chi_{Li}} \sum_k \frac{x - x_k}{|x - x_k|^2}$$

$$= -x_0^2 \sum_k \frac{x - x_k}{|x - x_k|^2} ,$$

where $x_k$ is the position of the $k$th lens projected onto the observer plane (i.e., $x_k = \chi_0 \theta_k$, where $\theta_k$ is the angular position of lens $k$ from the source), $a_L$ is the scale factor at the redshift of the lens plane, and $\chi_{Lo}$, $\chi_{so}$ and $\chi_{Li}$ are the comoving distances between the lens plane and the observer plane, between the source and the observer plane, and between the source and the lens plane, respectively (see Fig. 1). In addition,

$$x_E = \left(4Gm_0 \frac{\chi_{so}}{d_L \chi_{Li}}\right)^{1/2}$$
is the (Lagrangian) Einstein radius of a point mass on the lens plane. For an ensemble of randomly distributed point masses, the optical depth to microlensing is \( \tau = \pi n x_0^2 \), where \( n = (a_L x_A / x_M)^2 \Sigma/m \) is the projected (onto the observer plane) surface number density of point masses and \( \Sigma \) is the physical surface mass density of the lens.\(^4\)

The generic behavior and the properties of gradient mappings such as that defined by equation (5) are relatively well known. In fact, gradient mapping has been used to describe a variety of phenomena. The mapping ought to be familiar to aficionados of the Zeldovich approximation (Zel'dovich 1970), used to describe the growth of cosmological velocities due to gravitational instability. The evolution of the density probability distribution function in the Zeldovich approximation prior to significant orbit crossing (or multiple streaming\(^5\)) was studied by Kofman et al. (1994). The two-dimensional mapping has been used to describe the brightness distribution behind a phase screen (which is similar to the illumination pattern that appears at the bottom of a swimming pool; Longuet-Higgins 1960; Gurbatov, Malakhov, & Saichev 1991). However, in all cases the results yield the brightness distribution (or the density of particles) in the single-stream approximation. We are not aware, for instance, of any solution for the brightness distribution at the bottom of a swimming pool that accounts for multiple streaming.

The situation in microlensing is quite different (and much more complex), in that one must solve the gradient mapping equation to locate all the microimages, and then sum the image magnifications to obtain the macroimage magnification. In general, gravitational lensing maps a number of (micro)images with Lagrangian coordinates \( x^1, \ldots, x^N \) onto the same Eulerian coordinates \( r \). The observed flux of the image with Lagrangian coordinates \( x \) is magnified (or amplified) by \( A_i = 1/|\tilde{D}(x^i)| \), where

\[
\tilde{D} = \frac{\partial r}{\partial x} = \mathbf{I} + \frac{\partial s}{\partial x}
\]

is the deformation tensor (or Jacobian matrix) associated with the lens mapping and \( \mathbf{I} \) is the identity matrix. If the individual microimages are not resolved observationally, one works with the macroimage, which has a total magnification (or amplification) \( A = \sum_{i=1}^{N} A_i \). In Lagrangian space \( (x) \), the loci of points on which \( |\tilde{D}| = 0 \) (i.e., infinite magnification) trace out curves that are called critical curves, while the mapping of these curves to Eulerian space \( (r) \) defines the caustics. Because we consider a single source and an ensemble of observers, the caustics lie on the observer plane in our notation. This differs from the usual practice of considering a single observer and an ensemble of sources, in which the caustics lie on the source plane.

For a given lens configuration, one can define a differential cross section \( \sigma(A) \) such that \( \sigma(A)dA \) is the area in the observer plane, where the total magnification is between \( A \) and \( A + dA \). As we shall see, it is also useful to define a “normalized” differential cross section \( \varphi(A) = \sigma(A)/\sigma_{(0)}(A) \), where \( \sigma_{(0)}(A) \) is the differential cross section associated with the simple lens configuration of a single isolated point mass (see eq. [11] below). The function \( \varphi(A) \) is a measure of the deviations of the cross section of a more complex lens configuration from that of the single isolated point-mass lens. Finally, regarding lensing by an ensemble of randomly distributed lenses, the generalization of the cross section is the probability distribution of macroimage magnification \( P(A) \).

### 3. CAUSTIC-INDUCED FEATURE AT HIGH MAGNIFICATION

In § 3.1, we solve the lens equation for a point-mass lens perturbed by weak shear. This allows us to study analyti-
cally the caustic-induced features in the differential cross section (§ 3.2). Finally, in § 3.3 we consider the superposition of the point-mass plus weak-shear lenses and derive the caustic-induced feature in \(P(A)\) at high magnification \((A \sim 1/\varepsilon)\).

3.1. “Point-Mass Plus Weak-Shear” Lens: Solution of the Lens Equation

The lens equation for an isolated point-mass lens located at comoving distance \(x_{L0}\) from the observer and at the origin of the lens plane is

\[
r = x - x_L \frac{x}{|x|^2} .
\]

(10)

The corresponding differential cross section is

\[
s_0(A) = \frac{2 \pi x_L^2}{(2^2 - 1)^{3/2}} .
\]

(11)

As is well known, the caustic associated with this lens is a degenerate point located at \(r = 0\), the degeneracy being due to the spherical symmetry of the lens.

The degeneracy is lifted if the point mass is not strictly isolated and is perturbed by shear, as in the typical lensing configuration in single-plane, low optical depth microlensing. Under low optical depth conditions, lensing (other than low-magnification events) is dominated by the point mass closest to the location where the light ray pierces the lens plane, and the contribution of the other distant point masses can be treated as perturbations. If we choose a coordinate system such that the dominant lens is located at the origin of the lens plane and expand the perturbations to the first order in \(x\), the perturbations correspond to a constant deflection and a small constant shear. The lens mapping (eqs. [5] and [7]) then simplifies to

\[
r = x + s(x) , \quad s(x) \approx d_L(x) + \Delta_L + \hat{S}_L x ,
\]

(12)

where

\[
d_L(x) = -x_L \frac{x}{|x|^2} .
\]

(13a)

denotes the influence of the dominant lens, while \(\Delta_L\) (the constant deflection) and \(\hat{S}_L\) (the shear matrix) represent the perturbative influence of the other lenses on the plane:

\[
\Delta_L = x_L \sum_k \frac{x_k}{x_k^2} ,
\]

\[
\hat{S}_L = x_L^2 \sum_k \frac{1}{|x_k|^2} \left( \begin{array}{cc} x_k^2 - y_k^2 & 2 x_k y_k \\ 2 x_k y_k & y_k^2 - x_k^2 \end{array} \right) \equiv S \hat{T}(\phi_L) .
\]

(13b)

where

\[
s_0(A) \approx \frac{2 \pi x_L^2}{(2^2 - 1)^{3/2}} .
\]

(11)

As is well known, the caustic associated with this lens is a degenerate point located at \(r = 0\), the degeneracy being due to the spherical symmetry of the lens.

The degeneracy is lifted if the point mass is not strictly isolated and is perturbed by shear, as in the typical lensing configuration in single-plane, low optical depth microlensing. Under low optical depth conditions, lensing (other than low-magnification events) is dominated by the point mass closest to the location where the light ray pierces the lens plane, and the contribution of the other distant point masses can be treated as perturbations. If we choose a coordinate system such that the dominant lens is located at the origin of the lens plane and expand the perturbations to the first order in \(x\), the perturbations correspond to a constant deflection and a small constant shear. The lens mapping (eqs. [5] and [7]) then simplifies to

\[
r = x + s(x) , \quad s(x) \approx d_L(x) + \Delta_L + \hat{S}_L x ,
\]

(12)

where

\[
d_L(x) = -x_L \frac{x}{|x|^2} .
\]

(13a)

denotes the influence of the dominant lens, while \(\Delta_L\) (the constant deflection) and \(\hat{S}_L\) (the shear matrix) represent the perturbative influence of the other lenses on the plane:

\[
\Delta_L = x_L \sum_k \frac{x_k}{x_k^2} ,
\]

\[
\hat{S}_L = x_L^2 \sum_k \frac{1}{|x_k|^2} \left( \begin{array}{cc} x_k^2 - y_k^2 & 2 x_k y_k \\ 2 x_k y_k & y_k^2 - x_k^2 \end{array} \right) \equiv S \hat{T}(\phi_L) .
\]

(13b)

Finally, transforming the Eulerian variable according to \(r \to r' = r - \Delta_L\) and rotating the coordinate axes such that they coincide with the principal shear axes, i.e., \(\phi_L = 0\), we can cast equation (12) into a more conventional form:

\[
r' = x - x_L^2 \frac{x}{|x|^2} + S \hat{T}(0) x ,
\]

(15)

where

\[
\hat{T}(0) = \left( \begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right) .
\]

(16)

Equation (15) yields a fourth-order equation for the image positions (see Appendix A), which has four solutions inside the caustic and two solutions outside.

For an arbitrary value of \(S\), it is difficult to proceed any further, especially with regard to determining the cross section \(s(A)\) analytically. Fortunately, the magnitude of the shear is typically \(S \sim \tau\) and, in the limit of low optical depth, is small (see § 3.3). In Figure 2, we show the caustic and the contours of constant magnification on the observer plane for a point mass perturbed by weak shear \((S \ll 1)\). (The results shown in Fig. 2 were obtained numerically for \(S = 0.02\).) For \(S \ll 1\), one can observe that the entire four-stream region (i.e., the region inside the caustic) has high magnification. The minimum (though high) magnification in this region occurs at \(r' = 0\); it can be found from equations (A1)–(A3) that

\[
A = (S - S^2)^{-1} \approx S^{-1} \gg 1
\]

(17)

at \(r' = 0\) (see also Mao 1992). Elsewhere within the caustic, \(A\) is even higher and diverges to infinity at the caustic. This means that the light rays that delineate the four-stream region pass very close to the critical curve in Lagrangian space. To first order in \(S\), the equation for the critical line (eq. [A4]) can be expressed as

\[
x^2 = x_L^2 (1 + S \cos 2\phi) ,
\]

(18)

where we have used the polar representation \((x, \phi)\) for the Lagrangian coordinates \(x\). Hence, the critical line is a slight-
ly flattened ellipse. We can study the behavior of the rays that pass near the critical line by considering rays with $(x/x_E)^2 = 1 + t$ and $t \ll 1$. If we now rescale the variables so that $t' = t/S$ and $q' = q/S = r'(x_E/S)$, equation (A1) reduces at the lowest order in $S$ to a simple quartic equation for $t'$:

$$t'^4 - (2 + q'^2)t'^2 + 2q'^2t' \cos 2\phi + (1 - q'^2) = 0. \quad (19)$$

In general, equation (19) has either two or four real solutions (corresponding to the positions of the two images outside the caustic or the four images inside). Once the images are identified, their magnifications can be summed to obtain the macroimage magnification factor:

$$A(r') = \frac{1}{2S} \sum |t' - (t'^2 + 1) \cos 2\phi - 2t'|^{-1}. \quad (20)$$

Therefore, the magnification inside and around the caustic depends on the parameter $S$ as $A \propto 1/S$. This means that for $S \ll 1$ the distribution of $A$ over the observers is a function of the combination $SA$ only, and any function of $A$ obeys a simple scaling property.

We can obtain a parametric expression for the caustic in Eulerian space by solving equations (19) and (20):

$$q'^2(\phi) = \frac{(t'^2 - 1)^2}{t'^2 - 2t' \cos 2\phi + 1},$$

$$t'(\phi) = \cos 2\phi - \left[\sin^2 2\phi(1 + \cos 2\phi)\right]^{1/3} + \left[\sin^2 2\phi(1 - \cos 2\phi)\right]^{1/3}. \quad (21)$$

After some algebra, equation (21) can be simplified to

$$q'^{2/3}(\cos^{2/3} \phi + \sin^{2/3} \phi) = 2^{2/3}. \quad (22)$$

Thus the caustic has the shape of an astroid, with the four cusp catastrophes at $q = 0$ and $q = 2$ (or $r = 2x_E$) from the center of the astroid, connected to each other by four fold catastrophes. The astroid shown in Figure 2, which was obtained numerically for $S = 0.02$, is described by equation (22).

3.2. Analytic Determination of the Caustic-induced Features in $\sigma(A)$

One consequence of the scaling relationship noted above is that the caustic-induced features in the differential cross section of a point-mass plus weak-shear lens scale as

$$\sigma(A) = \sigma_0(A) \varphi(SA), \quad (23)$$

where $\sigma_0(A)$ is the differential cross section of an isolated point mass (eq. [11]) and $\varphi(SA)$ is a function describing the scaling behavior. In Figure 3, we show the “normalized” differential cross section $\varphi(SA)$ obtained numerically for $S = 0.02$ (solid line). The modification of the cross section at $SA \sim 1$ is quite pronounced (a factor of a few higher or lower), and the deviations persist over a decade in $SA$. These features were first found numerically by Nityananda & Ostriker (1984; see also Mao 1992). In this subsection, we investigate the properties of $\varphi(SA)$ and find analytically the critical values of $A$ and $\varphi(SA)$. We will also provide an analytic fit to $\varphi(SA)$.

First, it can be seen that $\varphi(SA)$ tends toward unity asymptotically at $SA \ll 1$ and $SA \gg 1$. As we have noted previously, $\sigma(A) \propto dA$ is the area in the Eulerian plane between the $A$ and $A + dA$ contours. Essentially, this is equal to the length of the isomagnification contour corresponding to magnification $A$, multiplied by the incremental separation $dr$ between the $A$ and $A + dA$ contours. It is well known that in the vicinity of the fold caustic (i.e., in the limit of very high magnification $A \gg 1/S$), $A \propto 1/[r_c(\phi) - r_c(\phi)]^{1/2}$, where $r_c(\phi)$ is the distance to the caustic from the center of the astroid. On the other hand, the lengths of the isomagnification contours are nearly independent of magnification, because both in size and in shape the contours approach the astroid-shaped caustic asymptotically. Consequently, $\sigma(A) \propto 1/A^3$ in the high-magnification limit, $\varphi(SA)$ tends toward a constant value. In fact, to first order in $S$, the magnification cross sections of the isolated point mass and the point-mass plus weak-shear cases coincide at $A \gg 1/S$ and $\varphi \rightarrow 1$.

We can solve equations (A1) and (A2) at great distances from the astroid by only considering terms to first order in $S$ and using a perturbation series with respect to $1/r'$. We find that the resulting isomagnification contours are given by $A \approx 1 + 2(x_E/r')^2 + 2S(x_E/r')^2 \cos 2\phi$. In the limit $A < 1/S$ (but with $A - 1 > S^2$), the contours are similar to those associated with an isolated point-mass lens. Consequently, the magnification cross section converges to that of an isolated point mass and $\varphi \rightarrow 1$.

The interesting features in $\varphi(SA)$ occur at $SA \sim 1$. We can understand these features by studying the geometry of the isomagnification contours on the observer plane (see Fig. 2). As noted above, the isomagnification contours at great distances from the astroid are slightly deformed circles. In going from low ($A \ll 1/S$) to high ($A \sim 1/S$) magnification, the quadrupole moment of the contours continues to increase and the isomagnification contours are continuously deformed. At the first critical value, $A_1$, the contour oscillates the astroid and, because of symmetry, touches it at four Eulerian points with coordinates $r = Sx_E$, $\phi = (2n - 1)\pi/4$, and $n = 1, \ldots, 4$. There are four images associated with each of these points, and we can determine their Lagrangian coordinates as well as their magnifications, from equations (19) and (20). Two of these images have infinite magnification, while the other two, which are of interest here, have finite magnification. From equation (20), we find that the total magnifi-
The next contours of $A$ that are slightly larger than $A_1$ consist of four symmetric arcs around and outside the four cusps. Because of the loss of the area inside the astroid, the corresponding $\sigma(A)dA$ decreases rapidly. This explains the first break in $\sigma(SA)$ at $SA_1$ (see Fig. 3).

Next we consider the isomagnification contours inside the caustic, i.e., in the four-image region. The minimum magnification $A_2$ inside the caustic is located at $r = 0$, the origin of the Eulerian plane. Using equations (19) and (20), we find that

$$A_2 = 1/S.$$  

[As mentioned above, the exact solution is $A_2 = 1/S(1 - S^2)$ for $0 < S < 1$, which reduces to $1/S$ for $S \ll 1$; see eq. (17)]. We can also calculate the magnification in the vicinity of the origin by evaluating equations (19) and (20) for $r' < SX_e$ or $\theta' < 1$. We find that the Eulerian position $(r', \phi)$ is associated with four images with $t_{1,2} = 1 \pm (r'/SX_e) \sin \phi$ and $t_{3,4} = -1 \pm (r'/SX_e) \cos \phi$. We also find that the total magnification is

$$A(r') = 1 + 1/S \left( \frac{r'}{SX_e} \right)^2.$$  

Therefore, the differential cross section is

$$\sigma(A | S) = 2\pi r' \frac{d\sigma}{dA} = 4\pi A^2 \Theta(SA - 1) + \delta\sigma(A | S),$$

where $\Theta$ is the step function and the term $\delta\sigma(A | S) \propto A^{-7/2}$ (Mao 1992) takes into account the contribution to the differential cross section from contours with $A > A_2$ that are outside the caustic and in the vicinity of the cusps (see Fig. 2). The discontinuous jump at $A_2$ described by equation (27) is in good agreement with the jump of height 2 in $\sigma(SA)$ shown in Figure 3.

Finally, we provide a simple analytic fit (with three segments for the different $A$ intervals) to the numerically obtained differential cross section of the point-mass plus weak-shear lens shown in Figure 3 (dotted lines): \[ \begin{align*}
\phi(SA) &= \begin{cases}
1 + 7.7(SA)^{3.5}, & \text{for } A \leq A_1, \\
0.17(SA - 0.33)^{1.72} + 0.023, & \text{for } A_1 \leq A \leq A_2, \\
1 + 0.85(\frac{SA}{SA})^{3.3} + 0.37, & \text{for } A \geq A_2.
\end{cases}
\end{align*} \]

3.3. Superposition of Cross Sections with Distributed Shear

Let us now consider a collection of point masses (all with the same mass $m$) randomly distributed on a single lens plane in the low optical depth limit. The combination of individual lenses produces several important effects. First, as we discussed above, the lens distribution as a whole induces shear in the neighborhoods of the individual point masses. We found in §§ 3.1–3.2, for a point mass perturbed by weak shear, that the shape of the isomagnification contours has a uniform form—every lens produces a caustic with the same astroid shape (Fig. 2)—and that the normalized differential cross section $\phi$ is a function only of the combination $SA$ (Fig. 3). Since the magnitude of the shear $S$ varies from lens to lens, there is a distribution of $S$ for the cross section $\sigma(A | S) = \sigma_0(A)\phi(SA)$ near each point mass. In this subsection, we consider the superposition of the cross sections, with the caustic-induced features at different $A \sim 1/S$. Other effects of the combination of individual lenses will be considered systematically in subsequent sections.

Under the condition of low optical depth, the point masses are well separated, and the caustics and the high-magnification contours on the observer plane are relatively isolated. Consequently, the macroimage magnification distribution $P(A)$ at high $A$ can be approximated as a superposition of the cross sections of the individual point-mass plus weak-shear lenses:

$$P(A) = \int_0^\infty dS \sigma(SA) \sigma(A | S),$$

where $\sigma(S)$ is the probability distribution of the shear due to the other lenses on the plane. For a random lens distribution with optical depth $\tau = \pi A^2$, $\sigma(S)$ is given by

$$\sigma(S) = \frac{\tau S}{(\tau^2 + S^2)^{3/2}}$$

(Nityananda & Ostriker 1984; Schneider 1987b; Lee & Spergel 1990). Note that the shear distribution tends toward $\sigma(S) = \tau/S^2$ for $S \gg \tau$, which is just the shear from the nearest neighbor. Also note that $\sigma(S)$ has a prominent peak at $S = \tau/\sqrt{2}$ for $\tau \ll 1$.

Substituting equation (30) into equation (29), we obtain

$$P(A) = n \sigma_0(A)f_1(t_A) = \frac{2\tau}{(A^2 - 1)^{3/2}} f_1(t_A),$$

where we have introduced the function

$$f_1(t_A) = \int_0^\infty dy \frac{\phi(\tau Ay)}{(1 + y^2)^{3/2}}$$

$$= 1 + \int_0^\infty dy \frac{\phi(\tau Ay) - 1}{(1 + y^2)^{3/2}}$$

(32)

to describe the caustic-induced feature in the macroimage magnification distribution. Since $f_1(t_A)$ depends on $\tau$ and $A$ only through the combination $t_A$, the caustic-induced feature in $P(A)$ also has a simple scaling property. The function $f_1(t_A)$, obtained numerically with $\phi(SA)$ as shown in Figure 3, is plotted in Figure 4 (solid line). There is a mild "bump" located at $A \approx 2/\tau$, which is a 20% enhancement at its peak. The bump is similar in shape to the features in $\phi$, but it is much weaker and smoother because of the convolution over the shear distribution. The function $f_1(t_A)$ provides a semianalytic description for the caustic-induced bump found numerically by Rauch et al. (1992), but only in the low optical depth limit (see § 5).

In Figure 4 we also show an analytic fit (dotted line) to the numerically evaluated $f_1(t_A)$:

$$f_1(t_A) = 1 - 0.81 \frac{(t_A)^2(1 - 3t_A)}{[1 + 1.5(t_A)]^{8/3}}.$$  

(33)

The maximum deviation of this fit from the numerical result is less than 0.01.
expression for $P$ of induced feature in the macroimage magnification distribution, lenses that results in the caustic-induced feature is No. 2, 1997 STATISTICS OF MICROLENSING MAGNIFICATION. I. 515 that yields a turnover at $\tau$. While terms in equation (35) are statistically independent. While that results by no more than 0.6%, and it is indistinguishable from the numerical integration results by no more than 0.6%, and it is indistinguishable from the numerical integration results on the scales of Figure 5.

The distribution given by equation (37) has the following properties: it is a normalized distribution, i.e., $\int dA P(A) = 1$. It has a sharp peak $P(A) \approx 0.16/\tau^2$ at $\delta A \approx 0.84\tau^2$ (see Fig. 5). For $A - 1 \gg \tau^2$, it has the asymptote

$$P(A) \approx \frac{\tau}{\sqrt{2(A - 1)^{3/2}}} ,$$

which matches the low-magnification ($\delta A \ll 1$) asymptote (eq. [34]) of the distribution in equation (2). Since both equations (31) and (37) yield the same result in two overlapping regimes, $A \ll 1/\tau$ and $A - 1 \gg \tau^2$, it is easy to construct a macroimage magnification distribution that takes into account the effects of both the caustics and the diffuse component:

$$P(A) = \frac{2\tau}{(A^2 - 1)^{3/2}} f_1(\tau^2) g_1 \left( \frac{A - 1}{\tau^2} \right) .$$

In the appropriate limits, this distribution becomes either equation (31) or (37).

5. FINAL FORM OF $P(A)$ AND COMPARISON WITH NUMERICAL SIMULATIONS

5.1. Renormalization and the Final Form of $P(A)$

In § 3.3, we assumed that the surface density of the astroid-shaped caustics on the observer plane is the same as
the projected surface density of point masses (see eq. [29]). There is, however, a net convergence of the light rays by the overall lens distribution. This increases the density of asteroids by a factor where is the average overall lens distribution. This increases the density of the projected surface density of point masses (see eq. [29]).

Fig. 5.—(a) The function \( q(y = (A - 1)/\tau^2) \) (eqs. [37] and [B14]) that describes the modification of \( P(A) \) at low magnification \((A - 1 < 1)\) in the low optical depth limit. (b) The macroimage magnification distribution \( P(A) = g(y)/(\sqrt{2}y^{3/2}\tau^2) \) at low magnification in the low optical depth limit (solid line). The dashed line shows the asymptote \( P(A) = 1/(\sqrt{2}y^{3/2}\tau^2) \) for \( A - 1 \gg \tau^2 \).

To obtain the probability distribution function for the magnification in Eulerian space, \( P(A) \), we have to first multiply \( P_1(A) \) by the factor \( \bar{A}/\bar{A} \), which takes into account the focusing of the light rays by the lenses (Schneider 1987b; Lee & Spergel 1990). We then double the magnification \((A \rightarrow 2A)\), which takes into account the fact that the high-magnification events are dominated by a pair of very bright microimages that form close to a fold catastrophe. The final result is

\[
P(A) = \frac{2\tau\bar{A}}{(1 + \tau^2)^{3/2}} A^{-3} = \frac{2\tau}{(1 - \tau^2)(1 + \tau^2)^{3/2}} A^{-3}, \tag{42}
\]

which has the usual \( A^{-3} \) form. Although this is an exact analytic result for finite \( \tau \), it is only applicable to sufficiently high magnification so that the macroimage is dominated by two microimages. As we saw in § 3, for low \( \tau \) one actually has to go to very high magnification \((A \gg 1/\tau)\) for this to be a good approximation. The normalization in equation (42) differs from the normalization \( 2\tau \) in equation (40) by the factors \( \bar{A} \) and \((1 + \tau^2)^{-3/2} \). For \( \tau \leq 0.1 \), we have \( \bar{A} \approx 1 + 2\tau \) and \((1 + \tau^2)^{-3/2} \approx 1 - 3\tau^2/2 \), and the correction to the normalization \( 2\tau \) is mainly due to the factor \( \bar{A} \). Note that we have already derived the factor \( \bar{A} \) in the previous paragraph.

We can now collect the three effects of the combination of individual lenses discussed above—the caustic-induced feature from equation (31), the low-magnification modification due to the diffuse component (eq. [37]), and the renormalization due to the net convergence of light rays and multiple streaming (eq. [42])—and find that, for low optical depth, the final form of the macroimage magnification distribution due to a two-dimensional distribution of point masses is

\[
P(A) = \frac{2\tau}{(1 - \tau^2)(1 + \tau^2)^{3/2}} (A^2 - 1)^{-3/2} f_1(\tau A) g_1(A - 1) \left( \frac{\tau}{\tau^2} \right). \tag{43}
\]

5.2. Comparison with Numerical Simulations

Several authors (e.g., Paczyński 1986a; Wambsganss 1990, 1992; Rauch et al. 1992; Lewis & Irwin 1995) have used numerical simulations to calculate the macroimage magnification distribution, \( P(A) \), produced by a two-dimensional distribution of point masses. In particular, distributions with high resolution in \( A \) have been obtained from Monte Carlo simulations by Rauch et al. for \( \tau = 0.1, 0.2, \) and 0.3. We now compare the semianalytic \( P(A) \) derived in this paper (eq. [43]) with these numerical results. As noted, our derivation assumes low optical depth. It is not immediately obvious whether optical depths in the range \( 0.1 \leq \tau \leq 0.3 \) are sufficiently small and whether we should expect a good agreement between the semianalytic and numerical results. In fact, as we shall see, \( \tau = 0.1 \) is not sufficiently small. There are differences between the semianalytic and the numerical results that indicate finite optical depth effects. Ideally, comparison should also be made for smaller \( \tau \), but numerical \( P(A) \) with the required accuracy is not available [the computational requirement increases rapidly with decreasing \( \tau \) because the caustic-induced feature shifts to higher \( A (\propto 1/\tau) \) while the amplitude of \( P(A) \) at large \( A (\propto \tau) \) decreases].

In Figure 6, we show the high-resolution \( P(A) \) obtained by Rauch et al. for \( \tau = 0.1 \) and 0.2 (histograms). The data are
is due to that point mass and the shear perturbation from the other lenses, and an astroid-shaped caustic is associated with each point mass. However, since the lenses are randomly distributed on the lens plane, some of the lenses have close neighbor(s), separated by less than a few Einstein radii \( x_E \). In these cases, the caustics are not isolated astroids but rather are more complicated structures produced by the collective effect of two (or more) point masses. If the surface density (or optical depth) of the lens distribution is low enough so that the fraction of lenses with close neighbor(s) is small, the contribution to \( P(A) \) by these configurations is negligible and the analysis in § 3 is valid. For larger \( \tau \), however, we have to take into account the more complex configurations, which we can do so by evaluating the macroimage magnification distribution as a series: \( P(A) = P_1(A) + P_2(A) + \cdots \), where \( P_1(A) \) is the contribution by the point masses perturbed by shear, \( P_2(A) \) is the contribution by close pairs of point masses, etc. The first term, \( P_1(A) \), is the distribution derived in § 3 (but with a maximum cutoff in the convolution over shear). The second term, \( P_2(A) \), can be evaluated (approximately) as a convolution of the cross section \( \sigma |A| |d| \) for two point masses separated by distance \( d \) with the probability distribution for \( d \) if we consider a point mass and its nearest neighbor as a two–point-mass lens and ignore the perturbation from the other lenses.

To understand the contribution to \( P(A) \) from close pairs of point masses, we must first look at some of the properties of lensing by two equal point masses on a single lens plane (see § 3 of Paper II and Schneider & Weiss 1986 for details). For consistency with the notation in Paper II, we shall express the Lagrangian separation \( d \) between the lenses in units of \( \sqrt{2} x_E \). In the limit \( d \gg 2 \), the region near each of the point masses is perturbed by the weak shear [\( S = 1/(2d^2) \)] from the other point mass, and there are two astroid-shaped caustics. The caustics move toward each other (and become asymmetric) as \( d \) decreases, and they touch when \( d = 2 \). As \( d \) decreases below 2, the number of caustics changes from two to one to, finally, three (see Fig. 2 of Paper II). In Paper II, this sequence of caustic topologies is denoted as topology types A’ (for \( d > 2 \)), B’ (for \( 1/\sqrt{2} < d < 2 \)), and C’ (for \( d < 1/\sqrt{2} \)). In all cases, the normalized differential cross sections \( \sigma(A|d) \) are qualitatively similar to that for the point-mass plus weak-shear lens, but the caustic-induced features can be significantly stronger (compare Fig. 4 of Paper II to Fig. 3 of this paper). As in the point-mass plus weak-shear case, there is a discontinuous jump in \( \phi(A) \) at the minimum magnification \( A_{\text{min}} \) inside the caustic(s). In the limit \( d \gg 2 \), the minimum magnification inside the two astroid-shaped caustics is \( A_{\text{min}}(d) \approx A_{\text{min}}[S = 1/(2d^2)] = 2d^2 \) (eq. [25]). However, as Witt & Mao (1995) have shown, \( A_{\text{min}} \) is a nonmonotonic function of \( d \) and has a global minimum of 3 when \( d = \sqrt{2} \). If we now consider the convolution of these cross sections with the probability distribution for \( d \), it is clear that the resulting distribution \( P_2(A) \) should show an enhancement in the caustic-induced feature near \( A = 3 \).

In their analysis of the caustic-induced feature in \( P(A) \), Rauch et al. (1992) compared the results from the full-scale Monte Carlo simulations with the results from a simpler two–point-mass model (see their Fig. 6 for the \( \tau = 0.2 \) case). Their two–point-mass calculation is in fact an approximate numerical evaluation of \( P_1(A) + P_2(A) \). They found that the two–point-mass model is able to reproduce in part the caustic-induced feature at intermediate \( A \). In particular, the distribution \( P(A) \) from the two–point-mass model shows a

5.3. Reexamining the Criterion for Low Optical Depth

In our analysis of the caustic-induced feature in \( P(A) \) for low optical depth (§ 3), we assumed that all the point masses are well separated. Then the deflection near each point mass was simply horizontal lines of amplitude \( 2 \tau(1 - \tau)^{-2}(1 + \tau^2)^{-3/2} \) (dashed lines).

At \( \log A < 0.1 \) (or \( \delta A < 0.3 \)), the semianalytic and numerical results for \( \tau = 0.1 \) are in good agreement, but the results for \( \tau = 0.2 \) are slightly different in shape and amplitude. Since this range of \( A \) is dominated by the low-magnification modification \( g_1 \), we conclude that the function \( g_1 \) derived in § 4 is valid for \( \tau < 0.2 \). [Recall that the peak in \( P(A) \) produced by \( g_1 \) is located at \( \delta A \approx 0.84 \tau^2 \), which is much less than 0.3 in both cases.] The semianalytic caustic-induced feature also provides a reasonably good fit to the numerical results at \( \log A > 1 \) and 1.2 for \( \tau = 0.1 \) and 0.2, respectively. At intermediate \( A \), there are significant differences between the numerical and the semianalytic results. The numerical distribution is lower than the semianalytic distribution in the neighborhood of the minimum at \( A \approx 2.5 \), and it is higher than the semianalytic distribution at smaller and larger \( A \). This pattern of deviations is similar to the cross section shown in Figure 3 and suggests that the function \( f_1 \) derived in § 3 underestimates the strength of the caustic-induced feature at intermediate \( A \). It is, however, important to note that the differences between the semianalytic and numerical results decrease in strength with decreasing \( \tau \) and should be relatively small for \( \tau < 0.05 \). In the next subsection, we analyze the cause of the differences at intermediate \( A \).
dip at $A \approx 3$. Rauch et al. suggested that the slightly lower value ($A \approx 2.5$) of the location of the dip found in the full Monte Carlo simulations is due to nonnegligible contributions from configurations of three or more point masses.

Is there a simple explanation for the relatively strong contribution to $P(A)$ by the collective effect of two (or more) point masses for $\tau$ as small as 0.1? For a Poisson distribution of point masses, the probability that the nearest neighbor of a point mass is at a distance less than $d$ (again in units of $\sqrt{2}\,x_f$) is simply $1 - \exp (-2\pi n x_f^2 d^2)$. If we ignore the deflection due to the other lenses, the point mass and its nearest neighbor make up a two-point-mass lens. As we mentioned above, a two-point-mass lens produces two astroid-shaped caustics only if $d > d_{AB}$, where $d_{AB} = 2$ is the separation at which the caustic topology changes from type $A'$ to $B'$. Therefore, a simple estimate for the fraction of point masses whose caustics are not isolated astroids is

$$P_{na}(\tau) = 1 - \exp (-2\pi n x_f^2 d_{AB}^2) = 1 - \exp (-8\tau). \quad (44)$$

For the contribution to $P(A)$ from $P_{na}(A)$ (and higher order terms) to be negligible, $P_{na}(\tau) \ll 1$ or $\tau \ll \frac{1}{8}$. It is immediately clear that $\tau = 0.1$ is not sufficiently small for $P_{na}(A)$ to be negligible. For $\tau = 0.1$, we estimate that about half the point masses produce caustic structures that are more complex than the astroid shape: $P_{na} = 0.55$. This is consistent with the illumination pattern shown in Figure 1 of Rauch et al. (1992). From the comparison in § 5.2, we estimated that the contribution from $P_{na}(A)$ should be reasonably small for $\tau < 0.05$; this corresponds to $P_{na} < 0.33$.

6. CONCLUSIONS

In this paper, we have attempted to build upon the various approaches developed in previous studies in order to develop a systematic theory of gravitational microlensing by a planar distribution of point masses in the limit of low optical depth. In particular, we have derived a practical semianalytic expression for the probability distribution of macroimage magnification, $P(A)$.

At $\delta A \gg \tau^2$, we model the illumination pattern as a superposition of the patterns due to individual “point-mass plus weak-shear” lenses. The shear perturbation near each point mass is induced by the neighboring point masses. It breaks the degeneracy of the caustic of an isolated point-mass lens and produces an astroid-shaped caustic. The convolution of the magnification cross section of the point-mass plus weak-shear lens with the probability distribution of shear yields $P(A) = 2\sigma(A^2 - 1)^{-3/2}f_1(\sigma A)$, where the function $f_1(\sigma A)$ (eq. [32]; Fig. 4) describes the caustic-induced feature in the macroimage magnification distribution. In effect, $f_1(\sigma A)$ introduces a mild “bump,” a 20% enhancement, at $A \approx 2/\sigma$. Since $f_1(\sigma A)$ depends on $\sigma$ and $A$ only through the combination $\sigma A$, the caustic-induced feature in $P(A)$ exhibits a simple scaling property. To facilitate future computations, we have provided a useful analytic fit to $f_1(\sigma A)$ (eq. [33]).

We should point out that the results derived in §§ 3.1–3.2 for the point-mass plus weak-shear lens may also have applications in the analysis of gravitational microlensing by physical binary systems. Microlensing searches toward the Galactic bulge and the Large Magellanic Cloud have already discovered microlensing events by close binaries with a single merged caustic (Udalski et al. 1994; Alard, Mao, & Guibert 1995). Since wide binaries are more common than close ones, there should be a significant number of events due to wide binaries (Di Stefano & Mao 1996). If the components of the binary are sufficiently far apart, the region near each point mass is perturbed by the weak shear from the other point mass. In these cases, there are two separate astroid-shaped caustics, and the results derived in §§ 3.1–3.2 are applicable.

At low magnification ($A - 1 \ll 1$), the macroimage consists of a bright primary image barely deflected from the unperturbed position of the source and a large number of faint secondary images (the diffuse component) formed close to each of the lenses. The magnifications of the primary and the secondary images can be strongly correlated. Taking into account the correlations, we find that $P(A) = 2(\sigma A - 1)^{-3/2}g_1(A - 1/\tau^2)$, where $g_1$ (eqs. [37] and [B14]; Fig. 5) represents the low-magnification correction. An excellent analytic fit to $g_1$ is given by equation (38). The function $g_1$ prevents $P(A)$ from diverging at $A = 1$, and it introduces a sharp peak of amplitude $P(A) \approx 0.16/\tau^2$ at $A - 1 \approx 0.84\tau^2$.

We also discussed the renormalization of $P(A)$ due to the net convergence of light rays and multiple streaming. Finally, collecting the above results, we find that in the low optical depth limit

$$P(A) = \frac{2\sigma}{(1 - \tau^2)(1 + \tau^2)^{3/2}} \left( A^2 - 1 - \frac{3/2}{f_1(\tau A)}g_1 \left( \frac{A - 1}{\tau^2} \right) \right)$$

(eq. [43]). In order to determine the realm of validity of the above semianalytic expression, we compared it against $P(A)$ obtained from Monte Carlo simulations by Rauch et al. (1992) for $\tau = 0.1$ and 0.2. At low magnifications ($\log A < 0.1$), we find that the $\tau = 0.1$ semianalytic and numerical results are in good agreement with each other. At greater optical depths, differences arise both in shape and in amplitude. We therefore conclude that the low-magnification modification $g_1$ is valid for $\tau < 0.2$. The numerical and the semianalytic results are also in good agreement in the high-magnification regime ($\log A > 1$ and 1.2 for $\tau = 0.1$ and 0.2, respectively). At intermediate $A$, however, the semianalytic expression does not match the numerical result even for optical depth $\tau = 0.1$. The deviations occur because the function $f_1$ derived in § 3 underestimate the strength of the caustic-induced feature at intermediate $A$. The deviations tend to diminish with decreasing optical depth, and we expect them to be relatively small for $\tau < 0.05$.

In order to understand the discrepancy between the numerical and semianalytic results for $\tau$ as small as 0.1, we reexamined our derivation of the function $f_1$ that describes the caustic-induced feature in $P(A)$. In our derivation, we assumed that a unique astroid-shaped caustic is associated with each point mass. In a random distribution of point masses, there may arise groups of two (or more) lenses that lie sufficiently close to each other and give rise to caustic configurations more complicated than isolated astroids. If the surface density of point masses is small, such lens configurations will be rare. If, however, the surface density is large, they will be more common and their contributions cannot be neglected. For a distribution of optical depth $\tau$, we estimate that the fraction of point masses whose caustics are not simple astroids is $P_{na}(\tau) = 1 - \exp (-8\tau)$. For $\tau = 0.1$, about half the point masses fall into this category. A simple criterion for the low optical depth analysis to be valid is $P_{na} \ll 1$ or $\tau \ll 0.1$, though the comparison with numerical simulations indicates that the semianalytic distribution is a reasonable fit to $P(A)$ for $\tau$ up to 0.05.
In the accompanying Paper II, we extend our analysis to the more complicated situation of a three-dimensional distribution of point-mass lenses.

We thank K. Rauch for providing us with the results of Rauch et al. (1992). We also thank S. Mao and K. Rauch for helpful discussions. A. B. and L. K. acknowledge the hospitality shown to them at CITA during their visits in 1995. This work was supported in part by NSERC (Canada), the CIAR cosmology program, CITA, the Institute for Astronomy (L. K.), the Dudley Observatory (A. B.), and a CITA National Fellowship (M. H. L).

APPENDIX A

LENS EQUATIONS FOR POINT MASS PLUS ARBITRARY SHEAR

Expressing $r'$ and $x$ in polar coordinates $(r', \phi)$ and $(x, y)$, respectively, equation (15) can be written as a fourth-order equation for the Lagrangian coordinates $x$ in terms of the Eulerian coordinates $r'$:

\[
(\zeta - 1)^2(\zeta^2 - 2 + \varphi^2)\zeta + 1) = -2S\varphi^2 \cos 2\psi\zeta^2(\zeta - 1) + S^2\varphi^2\zeta^3 + 2S^2\zeta^2(\zeta - 1)^2 - S^4\zeta^4,
\]

(A1)

\[
\tan \theta = \frac{1 - \zeta - S\zeta}{1 + \zeta + S\zeta} \tan \phi,
\]

(A2)

where we have defined $\zeta \equiv (x/y)^2$ and $\varphi \equiv r'/x_P$. In this notation, the Jacobian of the mapping is

\[
|\mathbf{D}| = \left| \frac{\partial r'}{\partial x} \right| = 1 - \frac{2S \cos 2\phi}{\zeta} - S^2.
\]

(A3)

The critical line (i.e., the loci of Lagrangian points where the Jacobian $|\mathbf{D}|$ vanishes) is given by

\[
\zeta = \frac{S}{1 - S^2} \left( \cos 2\phi + \sqrt{\cos^2 2\phi + \frac{1 - S^2}{S^2}} \right),
\]

(A4)

which corresponds to the so-called Cassini oval. The associated caustic is nontrivial and it is not a point singularity. Based on equations (A1) and (A2), we expect four images in the interior of the caustic and two outside.

APPENDIX B

LOW-MAGNIFICATION LIMIT OF $P(A)$

Consider a random distribution of $N = \pi n R^2$ point-mass lenses within a circle of Lagrangian radius $R$, where $n$ is the projected surface number density of lenses. The total shear at the origin due to the $N$ point masses is given in matrix form by equation (13b). We will find it convenient to write the two distinct components of the shear matrix $\mathbf{S}_L$ as a vector: $\mathbf{S} = \sum_{k=1}^{N} s_k$, where $s_k = (x_k/|x_k|)^2(\cos 2\phi_k, \sin 2\phi_k)$ is the shear perturbation from the $k$th point mass.

To determine the low-magnification $(A - 1 \ll 1)$ limit of $P(A)$, we need to evaluate the probability distribution function for the following combination of shears (eq. [35]):

\[
\delta A = A - 1 = \left| \sum_{k=1}^{N} s_k \right|^2 + \sum_{k=1}^{N} |s_k|^2.
\]

(B1)

The first term in equation (B1) is just $S^2 = |\mathbf{S}|^2$, and the probability distribution of this term is found easily from the distribution (30). The second term in equation (B1) is the so-called diffuse component, and its probability distribution has previously been considered by Schneider (1987b). The evaluation of the probability distribution of the sum of these two terms, which is $\delta A$, is more complicated.

Let us start with the formal expression for the probability distribution of $\delta A$ as a function of the $N$ statistically independent shears (eq. [36]):

\[
P(\delta A) = \int \cdots \int |p(s_k)p(s_\delta)(A - 1 - \left| \sum_{k=1}^{N} s_k \right|^2 - \sum_{k=1}^{N} |s_k|^2)|,
\]

(B2)

where $p(s_k)$ is the probability distribution of the shear from the $k$th point mass and $\delta$ is the Dirac delta function. Equation (B2) can be simplified if we express each term in the integrand as a Fourier transform. For $p(s_k)$, we have

\[
p(s_k) = \frac{1}{(2\pi)^2} \int d^2t_k q(t_k)e^{-i\mathbf{t_k}\mathbf{s_k}},
\]

(B3)

where the characteristic function $q(t_k)$ is

\[
q(t_k) = 1 - \frac{t}{N} t_k
\]

(B4)
Finally, after performing the integration with respect to $d$ and then integrate with respect to $v$ and we are left with two integrations, with respect to $v$ using variables $w$ and $A$. In equation (B8), $w = A - 1$. Substituting equations (B3) and (B5) into equation (B2), we can perform the integration with respect to $s_k$,

$$
\int d^2 s_k \exp \left[ i w s_k^2 - i (t_k - u) \cdot s_k \right] = \frac{i \pi}{v} \exp \left( -i \frac{|t_k - u|^2}{4v} \right),
$$

and then integrate with respect to $t_k$:

$$
\frac{1}{(2\pi)^2} \int d^2 t_k q(t_k) \left( \frac{i \pi}{v} \right) \exp \left( -i \frac{|t_k - u|^2}{4v} \right) = 1 - \frac{f(v, u)}{N},
$$

where

$$
f(v, u) = \frac{i \pi}{2v} e^{-i u^2/4v} \int_0^\infty dt J_0 \left( \frac{tu}{2v} \right) t^2 e^{-it^2/4v} = \frac{\pi |v|}{2} \left( 1 - i \frac{v}{|v|} \right) g(z)
$$

and $J_0$ is the zeroth-order Bessel function. In equation (B8), $z = iv^2/8v$,

$$
g(z) = [(1 + 2z)I_0(z) + 2zI_1(z)]e^{-z},
$$

and $I_0$ and $I_1$ are the modified Bessel functions.

We can now group together the $N$ integrations over $t_k$ (eq. [B7]) to obtain

$$
\left[ 1 - \frac{f(v, u)}{N} \right]^N = e^{-f(v, u)}
$$

in the limit $N \to \infty$. Then

$$
P(\delta A) = \frac{1}{(2\pi)^2} \int d^2 X \int dv \int d^2 u \exp \left[ -i v (w - X^2) - i u \cdot X - f(v, u) \right].
$$

In equation (B11), the integration with respect to $X$ is

$$
\int d^2 X e^{i (w X^2 - u \cdot X)} = \frac{i \pi}{v} e^{-i u^2/4v},
$$

and we are left with two integrations, with respect to $v$ and to $u$. Substituting equations (B8) and (B12) into equation (B11), and using variables $v$ and $z$ instead of $v$ and $u$, we obtain

$$
P(\delta A) = \frac{2}{\pi} \int_0^\infty dz e^{-2z} \int_0^\infty dv \cos \left( wz - \frac{\pi v}{2} g(z) \right) \exp \left[ -\tau \frac{\pi v}{2} g(z) \right].
$$

Finally, after performing the integration with respect to $v$, we have $P(\delta A)$ in the form

$$
P(\delta A) = \frac{\tau}{w^{3/2}} \int_0^\infty dz g(z) \exp \left[ -2z - \frac{\pi^2 z^2}{4w} g^2(z) \right],
$$

where $w = \delta A$ and where $g(z)$ is defined in equation (B9). Equation (B14) must be integrated numerically in general, but it has the following analytic asymptotes:

$$
P(\delta A) =
\begin{cases}
\frac{2}{\pi \tau^2} y^{-1/2} e^{-\pi/4y}, & \text{for } y = \delta A/\tau^2 \ll 1, \\
\frac{1}{\sqrt{2 \pi^2}} y^{-3/2}, & \text{for } y = \delta A/\tau^2 \gg 1.
\end{cases}
$$

REFERENCES

Alard, C., Mao, S., & Guibert, J. 1995, A&A, 300, L17
Alcock, C., et al. 1993, Nature, 365, 621
Ansari, R., et al. 1996, A&A, 314, 94
Aubourg, E., et al. 1993, Nature, 365, 623
Chang, K., & Refsdal, S. 1979, Nature, 282, 561
Chang, K., & Refsdal, S. 1984, A&A, 132, 168
Di Stefano, R., & Mao, S. 1996, ApJ, 457, 93
Gott, J. R. 1981, ApJ, 243, 140
Gurbatov, S. N., Malakhov, A. N., & Saichev, A. I. 1991, Nonlinear Random Waves and Turbulence in Nondispersive Media: Waves, Rays, Particles (Manchester: Manchester Univ. Press)
Kaiser, N. 1992, in New Insights into the Universe, ed. V. J. Martínez, M. Portilla, & D. Sáez (Berlin: Springer), 247
Kayser, R., Refsdal, S., & Stabell, R. 1986, A&A, 166, 36
Kofman, L., Bertchinger, E., Gelb, J. M., Nusser, A., & Dekel, A. 1994, ApJ, 420, 44
Lee, M. H., Babul, A., Kofman, L., & Kaiser, N. 1997, ApJ, 489, 522 (Paper II)
Lee, M. H., & Spergel, D. N. 1990, ApJ, 357, 23
Lewis, G. F., & Irwin, M. J. 1995, MNRAS, 276, 103
Lewis, G. F., Miralda-Escudé, J., Richardson, D. C., & Wambsganss, J. 1993, MNRAS, 261, 647
Longueville-Higgins, M. S. 1960, Proc. Cambridge Philos. Soc., 56, 27
Mao, S. 1992, ApJ, 389, 63
Nityananda, R., & Ostriker, J. P. 1984, J. Astrophys. Astron., 5, 235
Paczynski, B. 1986a, ApJ, 301, 503
———. 1986b, ApJ, 304, 1
———. 1994, Acta Astron., 44, 235
Press, W. H., & Gunn, J. E. 1973, ApJ, 185, 397
Rauch, K. P., Mao, S., Wambsganss, J., & Paczynski, B. 1992, ApJ, 386, 30
Schneider, P. 1987a, ApJ, 319, 9
———. 1987b, A&A, 179, 80
Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (New York: Springer)
Schneider, P., & Weiss, A. 1986, A&A, 164, 237
Turner, E. L., Ostriker, J. P., & Gott, J. R. 1984, ApJ, 284, 1
Udalski, A., Szymański, M., Mao, S., Di Stefano, R., Kubiak, M., Mateo, M., & Krzeminski, W. 1994, ApJ, 436, L103
Wambsganss, J. 1990, Ph.D. thesis, Ludwig-Maximilians-Univ., Munich
———. 1992, ApJ, 386, 19
Witt, H. J. 1990, A&A, 236, 311
———. 1993, ApJ, 403, 530
Witt, H. J., & Mao, S. 1995, ApJ, 447, L105
Young, P. 1981, ApJ, 244, 756
Zeldovich, Ya. B. 1970, A&A, 5, 84