PLANET PARAMETERS IN MICROLENSING EVENTS

B. SCOTT GAUDI AND ANDREW GOULD[1]
Department of Astronomy, Ohio State University, Columbus, OH 43210;
gould@payne.mps.ohio-state.edu, gaudi@payne.mps.ohio-state.edu

ABSTRACT

A planetary microlensing event occurs when a planet perturbs one of the two images created in a point-mass microlensing event, causing a deviation from the standard Paczyński curve. Determination of the two physical parameters that can be extracted from a planetary microlensing event, the planet/star mass ratio q, and the planet/star separation in units of the stellar Einstein ring, y_p, is hampered by several types of degeneracies. There are two distinct and qualitatively different classes of planetary events: major and minor image perturbations. For major image perturbations, there is a potentially crippling continuous degeneracy in q which is of order δ_d^{-1}, where δ_d is the maximum fractional deviation of the planetary perturbation. Since the threshold of detection is expected to be $\delta_d \sim 5\%$, this degeneracy in q can be a factor of ~ 20. For minor image perturbations, the continuous degeneracy in q is considerably less severe, and is typically less than a factor of 4. We show that these degeneracies can be resolved by observations from dedicated telescopes on several continents together with optical/infrared photometry from one of these sites. There also exists a class of discrete degeneracies. These are typically easy to resolve given good temporal coverage of the planetary event. Unambiguous interpretation of planetary microlensing events requires the resolution of both types of degeneracy. We describe the degeneracies in detail and specify the situations in which they are problematic. We also describe how individual planet masses and physical projected separations can be measured.

Subject headings: gravitational lensing — planetary systems

1. INTRODUCTION

Two worldwide networks are currently searching for extra-solar planetary systems by making densely sampled observations of ongoing microlensing events toward the Galactic bulge (PLANET, Albrow et al. 1996; GMAN, Pratt et al. 1996). Several other groups will join the search shortly, and there is serious discussion of new initiatives that would intensify the search by an order of magnitude. More than 100 microlensing events have been detected to date by four groups, MACHO (Alcock et al. 1997), EROS (Ansari et al. 1996), OGLE (Udalski et al. 1994), and DUO (Alard 1996) based on observations made once or twice per night. The events typically last one week to a few months. MACHO and OGLE have reported "alerts," events detected before peak. This alert capability is what has allowed PLANET and GMAN to make intensive, sometimes round-the-clock, follow-up observations in hopes of finding the planetary perturbations which are expected to last a day or less.

In sharp contrast to this explosion of observational activity, theoretical work on planet detection has been rather sparse, amounting to only five papers in as many years. Mao & Paczyński (1991) originally suggested that planets might be detected in microlensing events. Gould & Loeb (1992) developed a formalism for understanding the character of planetary perturbations and made systematic estimates of the rate of detection for various planetary-system parameters. Bolatto & Falco (1994) studied the detection rate in the more general context of binary systems. These early works assumed that the lensed star could be treated as a point source. The usefulness of this approximation depends primarily on the angular size of the source θ_*,

relative to the planetary Einstein ring, θ_p,

$$\theta_p = \left(\frac{m}{M}\right)^{1/2} \theta_e, \quad \theta_e = \left(\frac{4GMD_{ls}}{c^2 D_{ol} D_{os}}\right)^{1/2}. \quad (1.1)$$

Here θ_e is the Einstein ring of the lensing star, m and M are the masses of the planet and its parent star, and D_{ol}, D_{ls}, and D_{os} are the distances between the observer, lens, and source. For Jupiter-mass planets at typical distances ($D_{ls} \sim 2$ kpc) from bulge giant sources, $\theta_p \sim 3\theta_*$ so the approximation is a reasonable one. However, for Saturn-mass, Neptune-mass and especially Earth-mass planets, the finite size of the source becomes quite important, and even for Jupiter-mass planets it is not completely negligible. Moreover, as we will stress below, it is quite possible to mistake a "Jupiter event" in which the source size is negligible for a "Neptune event" with $\theta_* > \theta_p$. Hence it is essential to understand finite-source effects even to interpret events where the source is in fact small.

Progress on finite-source effects was substantially delayed by problems of computation. Like all binary lenses, planetary systems have caustics, curves in the source plane where a point source is infinitely magnified as two images either appear or disappear. If one attempts to integrate the magnification of a finite source that crosses a caustic, one is plagued with numerical instabilities near the caustic. While it is straightforward to solve these problems for any given geometry, the broad range of possible geometries makes it difficult to develop an algorithm sufficiently robust for a statistical study of lensing events. Bennett & Rhie (1996) solves this problem by integrating in the image plane (where the variation of the magnification is smooth) rather than the source plane (where it is discontinuous). They were thereby able to investigate for the first time the detectability of Earth to Neptune-mass planets. Gould & Gaucherel (1996) showed that this approach could be simplified from a two-

[1] Alfred P. Sloan Foundation Fellow.

dimensional integral over the image of the source to a one-dimensional integral over its boundary. The implementation of this method requires some care. We describe the practical procedures elsewhere (Gaudi 1997). The difficult computational problems originally posed by finite-source effects are now completely solved.

To date, the analysis of planetary-system lensing events has focused on the question of "detectability" which was quantified by Gould & Loeb (1992) as a certain minimal fractional deviation from a standard Paczyński (1986) light curve having magnification

$$A(x) = \frac{x^2 + 2}{x(x^2 + 4)^{1/2}}, \quad x(t) = \left[\frac{(t - t_0)^2}{t_e^2} + \beta^2\right]^{1/2}, \quad (1.2)$$

where x is the projected lens-source separation in units of θ_e. Note that this curve is characterized by just three parameters: t_0 the time of closest approach, β, the impact parameter in units of θ_e, and t_e, the Einstein radius crossing time. Bennett & Rhie (1996) adopted a similar approach but added the qualification that the deviation persist for a certain minimum time.

Here we investigate a different question: How well can the parameters of the planetary-system be measured? As discussed by Gould & Loeb (1992), if there are light-curve data of "sufficient quality," two planetary-system parameters can generically be extracted from a microlensing event that displays a planetary perturbation. These are the planet/star mass ratio, q, and the planet/star projected separation in units of the stellar Einstein ring, y_p,

$$q \equiv \frac{m}{M}, \quad y_p \equiv \frac{a_p}{r_e}. \quad (1.3)$$

Here a_p is the physical projected separation, and $r_e = D_{ol} \theta_e$. As we discuss in § 8, it will often be possible to make additional observations that specify that the mass and distance of the lensing star, or equivalently M and r_e. For these cases, the measurements of q and r_e yield the mass $m = qM$ and projected separation $a_p = y_p r_e$.

If a planet were detected by observing a deviation from the standard curve, but its mass ratio remained uncertain by a factor of 10, the scientific value of the detection would be severely degraded. Indeed, such "detections" would probably not receive general acceptance. Thus, the problems of planet detection and parameter measurement are intimately connected. Microlensing planet-detection programs must monitor a total of at least several hundred events in order to obtain representative statistics on the frequency of planets. These observations require large blocks of 1–2 m class telescope time coordinated over several continents. For funding agencies and time allocation committees to make rational decisions about the allocation of scarce resources, and for observers to make rational choices among prospective targets, it is essential to determine what are the minimum observational requirements for detecting planetary systems *and* measuring the characteristics of the detected systems.

As we discuss below, there are two distinct classes of degeneracies which can hamper the determination of the planetary parameters. Discrete degeneracies are typically less severe and can usually be broken by the usual technique of obtaining accurate and densely sampled light curves. Continuous degeneracies are more problematic and often require additional information. While, in general, micro-lensing events are achromatic, if the magnification gradient is locally very large (i.e., near a caustic), then the lens will resolve the source as the source passes through this region. Any difference in the surface brightness profile of the source in two different flux bands will produce a color change during the event. As we demonstrate below, measurement of this color change can often be used to break the continuous degeneracy.

2. TYPES OF DEGENERACY

2.1. *Discrete*

Planetary-system lensing events are subject to two different discrete degeneracies. The first ambiguity relates to which image the planet is perturbing: the major image outside the Einstein ring or the minor image inside the Einstein ring. For almost all cases, this degeneracy is easily broken provided there is good temporal coverage of the light curve. However, if it is not broken the uncertainty in q and y_p can be a factor of a few. The magnitudes of these uncertainties depend only on the overall geometry of the event and not on the mass of the planet. The second ambiguity relates to whether the planet lies closer to or farther from the star than does the position of the image that it is perturbing. This degeneracy is more difficult to break, but it does not seriously affect the determination of q, and the uncertainty induced in y_p is proportional to $q^{1/2}$ and is therefore often much smaller than one induced by first degeneracy. These two discrete degeneracies are illustrated in Figure 1. The values of q and y_p for each of the four possible solutions are displayed in Table 1. Note that the example illustrated in Figure 1 and tabulated in Table 1 is for a maximum fractional deviation from the standard Paczyński curve of $\delta_d = 0.15$. This value was chosen for reasons of clarity. The majority of detectable planetary perturbations will have values of δ_d smaller than this (i.e., $\delta_d \sim 0.05$), simply because low-amplitude perturbations are a priori more likely because the cross section for lensing is higher. For perturbations with $\delta_d \sim 0.05$, the uncertainties in q and y_p will be significantly higher than those illustrated in Table 1.

2.2. *Continuous*

In addition, there is a continuous degeneracy arising from finite-source effects being misinterpreted as a larger value of q. This is because q is determined from the (square of the) duration of the planetary perturbation relative to the total duration of the event. If the size of the source is larger than the Einstein ring of the planet, then the duration of the planetary perturbation will be the crossing time of the source, not of the planet Einstein ring. Figure 2 shows 10 light curves all with the same maximum fractional deviation, δ_d, and same full width half maximum (FWHM) of perturbation, t_d. The parameter that differs in each of these

TABLE 1
DEGENERATE PARAMETER VALUES: DISCRETE

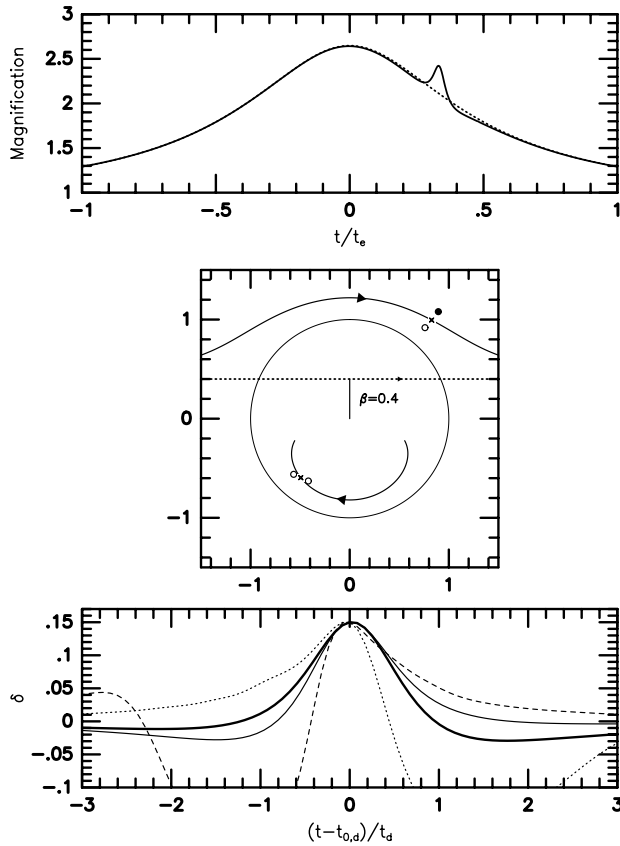|	Planet/Star Separation y_p	Planet/Star Mass Ratio q/q_0
Major image	1.40	1.00
	1.19	0.91
Minor image......	0.75	1.08
	0.80	0.88

FIG. 1.—Discrete degeneracies. *Top*: lensing light curve with (*solid curve*) and without (*dashed curve*) taking account of the presence of a planet with mass ratio $q = 10^{-3}$. *Middle*: associated lensing geometry. The two solid curves represent the path of the images relative to the lens. The crosses represent the image positions at the time of the perturbation. The circles are the four planet positions for which the light curves reproduce the measured parameters δ_d (maximum fraction deviation) and t_d (FWHM of deviation) at the peak of the disturbance when the source-lens separation is x_d. The filled circle is the "actual" planet position. *Bottom*: four associated light curves for times near the peak of the perturbation, $t_{0,d}$. Note that time is expressed in units of the perturbation time scale, t_d, not t_e. The bold curve corresponds to the "actual" planet position. Clearly, if the light curve is well sampled, the two dashed curves corresponding to the image position inside the Einstein ring in the middle panel could be ruled out immediately. However, the two solid curves are less easily distinguished. These differ by $\sim 15\%$ in planet/star separation and 10% in mass. See middle panel and Table 1.

curves is the ratio of source radius, θ_*, to planet Einstein radius, $\theta_p = q^{1/2}\theta_e$,

$$\rho = \frac{\theta_*}{\theta_p}. \tag{2.1}$$

Table 2 gives the inferred values of q and of the proper motion μ (of the planetary system relative to the observer-source line of sight) associated with each curve in units of the arbitrary chosen "fiducial" values associated with $\rho = 0.3$. In so far as one could not distinguish among these curves, any of these parameter combinations would be acceptable. The fiducial parameters q_0 and μ_0 would then be measurable, i.e., by fitting the observed lightcurve for $\rho = 0.3$, but the actual values of μ and q would not. The proper motion of both bulge and disk lenses is typically $\mu \sim O(V_{LSR}/R_0) \sim 30$ km s^{-1} kpc^{-1}, where $V_{LSR} \sim 220$ km s^{-1} is the rotation speed of the local standard of rest, and $R_0 \sim 8$ kpc is the Galactocentric distance. If, for the
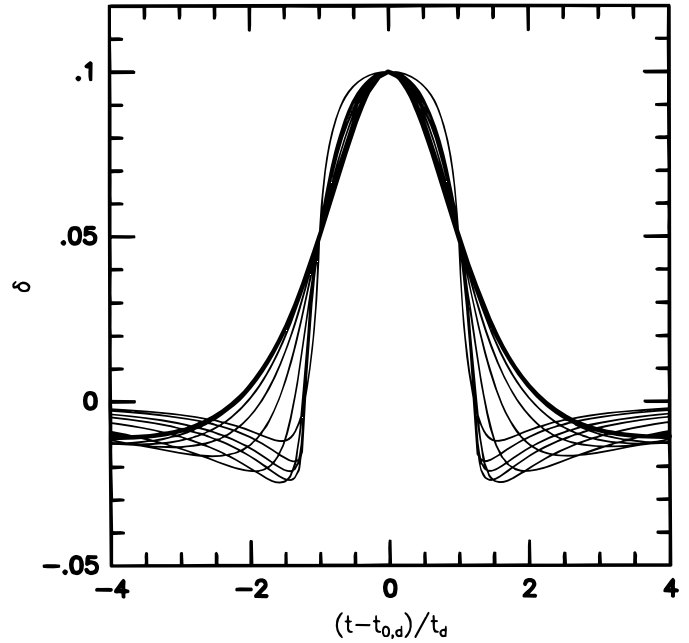
FIG. 2.—Ten curves showing the fractional deviation, δ, as a function of time in units of the perturbation timescale, t_d, for a geometry in which the perturbation occurs near the peak of the unperturbed lightcurve, and the planet/star projected separation is $y_p = 1.29$. In terms of the formalism of § 3, the geometry is $\gamma = 0.6$ and $\phi = 90°$ (see eqs. [3.1] and [3.2]). All 10 curves have maximum deviation $\delta_d = 10\%$ and FWHM $t_d = 0.06t_e$. The ratios of source radius to planet Einstein ring range from $\rho = 0.1$ to $\rho = 2.87$, the largest source radius consistent with this maximum deviation. Table 2 gives the corresponding values of $q = m/M$, and proper motion, μ, relative to the fiducial values q_0 and μ_0 at the arbitrarily chosen value $\rho = 0.3$.

example shown in Table 2, the fiducial value of μ_0 obtained by fitting the lightcurve with $\rho = 0.3$ was determined to be $\mu_0 \sim V_{LSR}/R_0$, one might then choose to argue that although one could not discard low-mass solutions based simply on the observed light curve, the proper motions associated with the low-mass solutions (i.e., $\mu \sim \mu_0/3$) would be so low as to be a priori unlikely, thus making these solutions improbable. However, these solutions could not actually be conclusively ruled out by such an argument, since the distribution of μ is rather broad (see Han & Gould 1995). Thus, there would remain a factor ~ 15 uncertainty in the planet/star mass ratio.

TABLE 2

DEGENERATE PARAMETER VALUES: CONTINUOUS MAJOR IMAGE

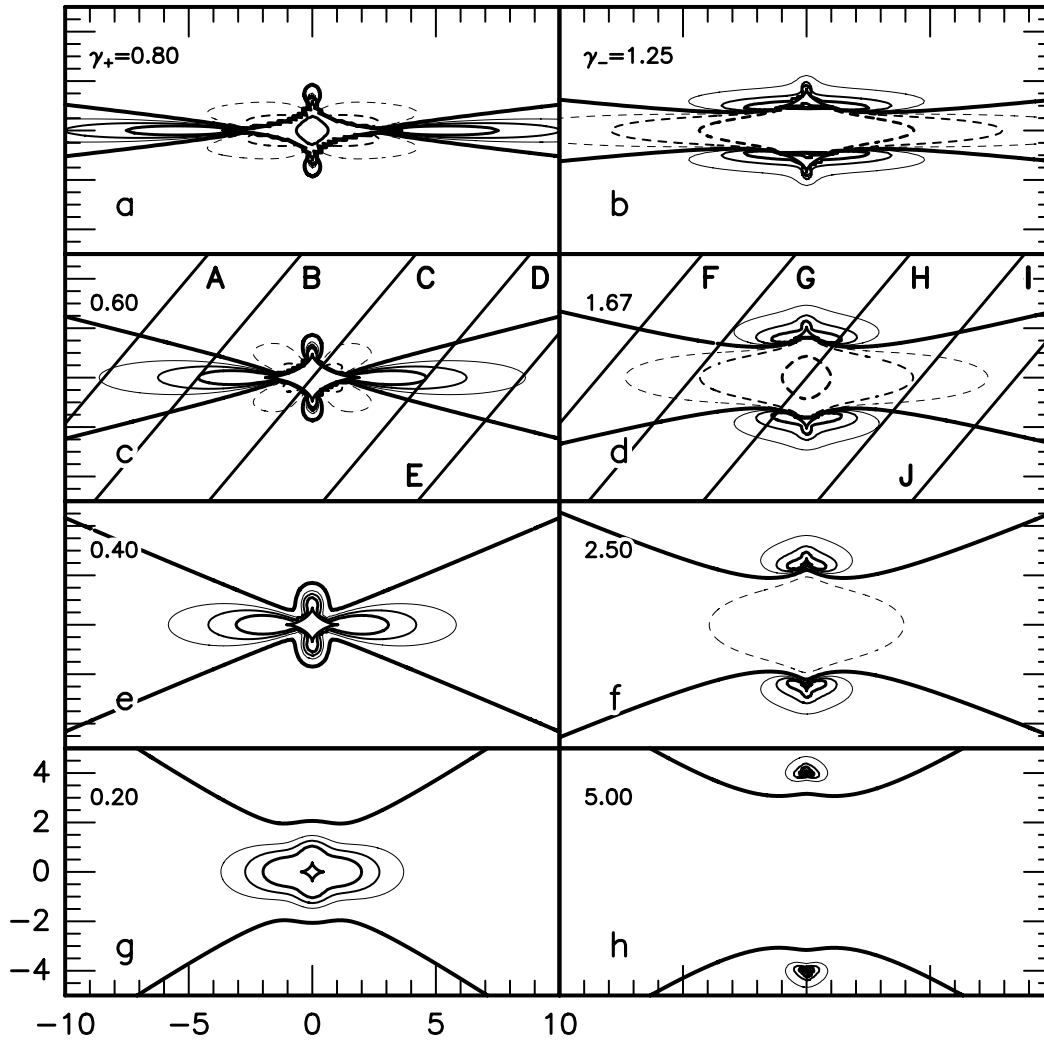| Dimensionless Source Radius ρ	Planet/Star Mass Ratio q/q_0	Proper Motion μ/μ_0
0.10	1.095	2.867
0.20	1.041	1.470
0.30	1.000	1.000
0.60	0.957	0.511
0.90	0.767	0.381
1.20	0.566	0.332
1.50	0.373	0.327
1.80	0.236	0.343
2.10	0.163	0.354
2.40	0.127	0.351
2.70	0.093	0.364
2.87	0.074	0.383

FIG. 3.—Chang-Refsdal magnification contours of a point source as a function of source position in units of the planet Einstein ring, $\theta_p = q^{1/2}\theta_e$, for various pairs of shears (γ_+ and $\gamma_- = \gamma_+^{-1}$) corresponding to planetary perturbations of the major and minor images, respectively. Magnification contours are calculated including the contribution of the unperturbed image. Contour pairs are for $\gamma_+ = 0.8$, 0.6, 0.4, and 0.2 in panels (a)–(b), (c)–(d), (e)–(f), and (g)–(h), and correspond to source positions at perturbation of $x_d = 0.22$, 0.52, 0.95, and 1.79. Solid contours are $\delta = 5\%$ (lightest), 10%, 20%, and ∞ (boldest). Super-bold solid contour is no deviation. Dashed contours are -5% (lightest), -10%, and -20% (boldest). Diagonal lines in panels (c) and (d) represent possible trajectories assuming that the overall light curve shows $x_d = 0.52$ (i.e., $\gamma_+ = 0.6$) and $\beta = 0.4$ (i.e, $\phi = \sin^{-1}\beta/x_d \sim 50°$). If the maximum deviation were observed to be $\delta_d = 20\%$ (and the point-source approximation were known to be valid), then the trajectory must be either B or D.

2.3. Relation between Degeneracies in q and μ

From the relation $\mu = \theta_e/t_e$, we obtain the identity $\mu = (\theta_e/\theta_p)(\theta_p/\theta_*)(\theta_*/t_e)$ or

$$\mu\rho p^{1/2} = \frac{\theta_*}{t_e}. \qquad (2.2)$$

Since the quantities on the right-hand side of this equation are observables, the product on the left-hand side must be constant for all allowed parameter combinations in any given planetary event: $\mu\rho q^{1/2} = $ constant. This equation then establishes a relationship between degeneracies in q and degeneracies in μ. If a range of solutions are permitted that have different values of ρ but very similar values of q, then we say that the mass ratio is not degenerate. However, it follows from equation (2.2) that the proper motion μ then varies inversely as ρ and therefore that it is degenerate. Similarly, if the range of allowed solutions all have the same value of μ, then the proper motion is not degenerate but then $q \propto \rho^{-2}$ and so the mass ratio is degenerate. This

relationship is illustrated by Table 2. The region $\rho \leq 0.3$ has well-determined q but degenerate μ, while the region $\rho \gtrsim 1$ has well-determined μ but degenerate q.

3. THE CHANG-REFSDAL LENS APPROXIMATION

In order to systematically investigate the role of these degeneracies and to determine the data that are required to break them, we follow Gould & Loeb (1992) and approximate the planetary perturbation as a Chang-Refsdal lens (Chang & Refsdal 1979; Schneider, Ehlers, & Falco 1992). A Chang-Refsdal lens is a point mass (in this case the planet) superimposed on a uniform background shear γ. For any given lensing event, the value of γ is simply the shear due to the lensing star at the *unperturbed* position of the image that is perturbed by the planet. The evaluation of γ is made at the midpoint of the perturbation. The projected source-lens separation in units of θ_e at the mid-point of the perturbation, x_d, is known from the light curve (see Figs. 1a, 1b). The associated projected image-lens separations in units of
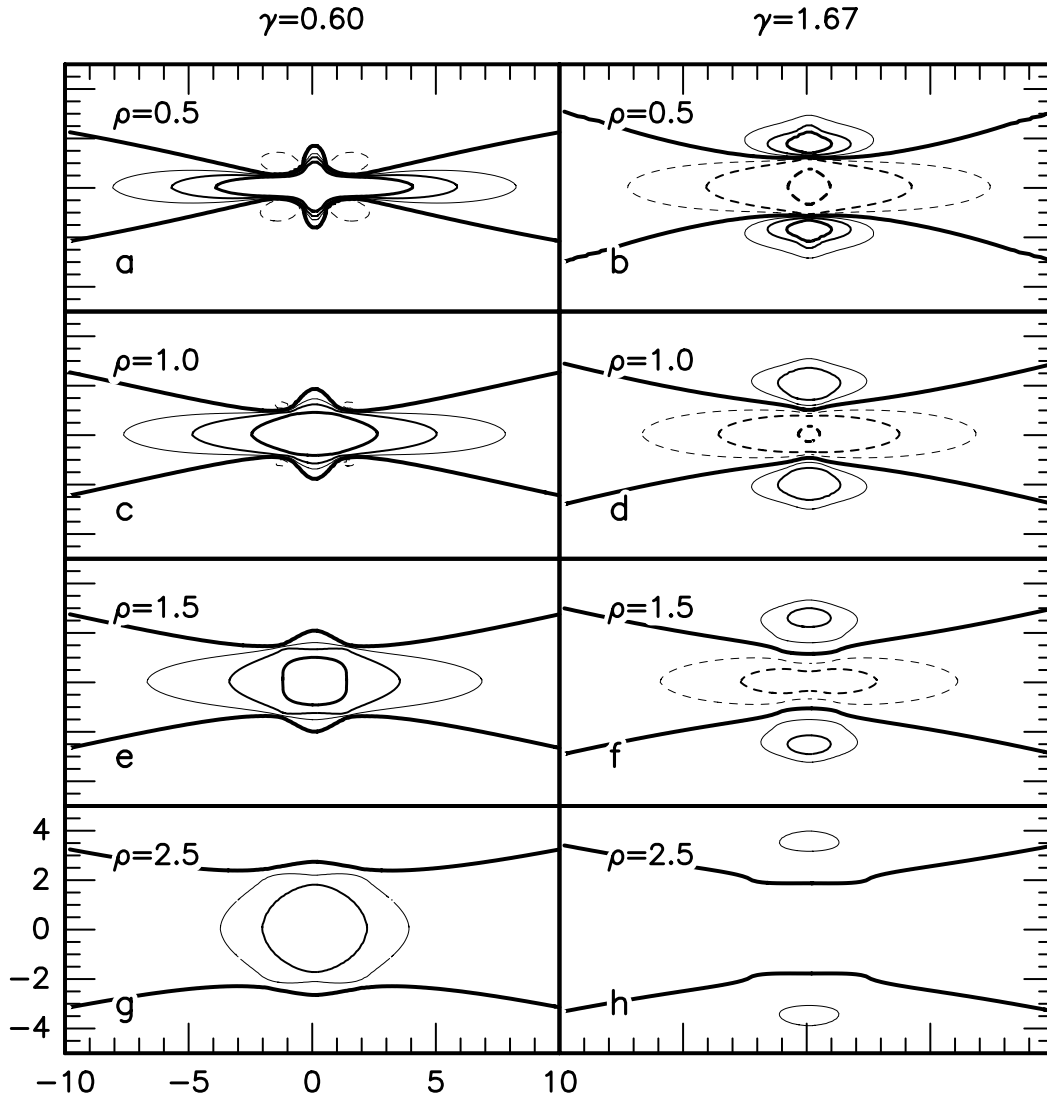
FIG. 4.—Chang-Refsdal magnification contours of a nonpoint source as a function of source position in units of θ_p, for the pair of shears $(\gamma_+, \gamma_-) = (0.6,$ 1.67) corresponding to planetary perturbations at source position $x_d = 0.52$. The ratios of planet Einstein radius to source radius are $\rho = 0.5, 1.0, 1.5,$ and 2.5. Compare with point-source case shown in Figs. 3c and 3d. Contours levels are the same as for Fig. 3.

θ_e are $y_{d,\pm}$ with shears $\gamma_\pm$ given by

$$\gamma_\pm = y_{d,\pm}^{-2}, \quad y_{d,\pm} = \frac{(x_d^2 + 4)^{1/2} \pm x_d}{2} \tag{3.1}$$

Thus the shear is known (up to a two-fold ambiguity) simply from the position of the planetary perturbation on the overall light curve.

When computing the fractional deviation, δ, of the Chang-Refsdal lens from the standard Paczyński curve, one always normalizes relative to the total unperturbed magnification (eq. [1.2]) which includes both the image perturbed by the planet and the image that remains unperturbed (Gould & Loeb 1992).

The Chang-Refsdal approximation permits an immense conceptual simplification of the problem. For a point source, all possible light curves of an event with a given x_d can be represented on a pair of diagrams, one for γ_+ and one for γ_-. All possible planetary perturbations can therefore be represented by a single-parameter family of such diagrams (see Fig. 3). For a given event, one knows β and x_d

from the overall light curve. One can therefore compute $\gamma_\pm$ using equation (3.1) and thereby pick out which two diagrams are relevant. One also knows the angle ϕ at which the source cuts through the diagram,

$$\sin \phi = \frac{\beta}{x_d}. \tag{3.2}$$

If for example, $x_d = 0.516$ and $\beta = 0.4$, then all possible light curves are represented by the parallel lines indicated in Figures 3c. If the light curve is well sampled, it is easy to distinguish between γ_+ and γ_-. Suppose that γ_+ (Fig. 3b) is correct, and say that the maximum fractional deviation is $\delta_d = 20\%$. Then one can immediately identify the correct curve as being either B or D. The observed duration of the perturbation relative to that of the whole event then sets the scale of diagram relative to θ_e and thus determines the mass ratio.

Of course, one does not know a priori that finite source effects can be ignored. However, for any x_d, all possible events can still be represented by a single-parameter family
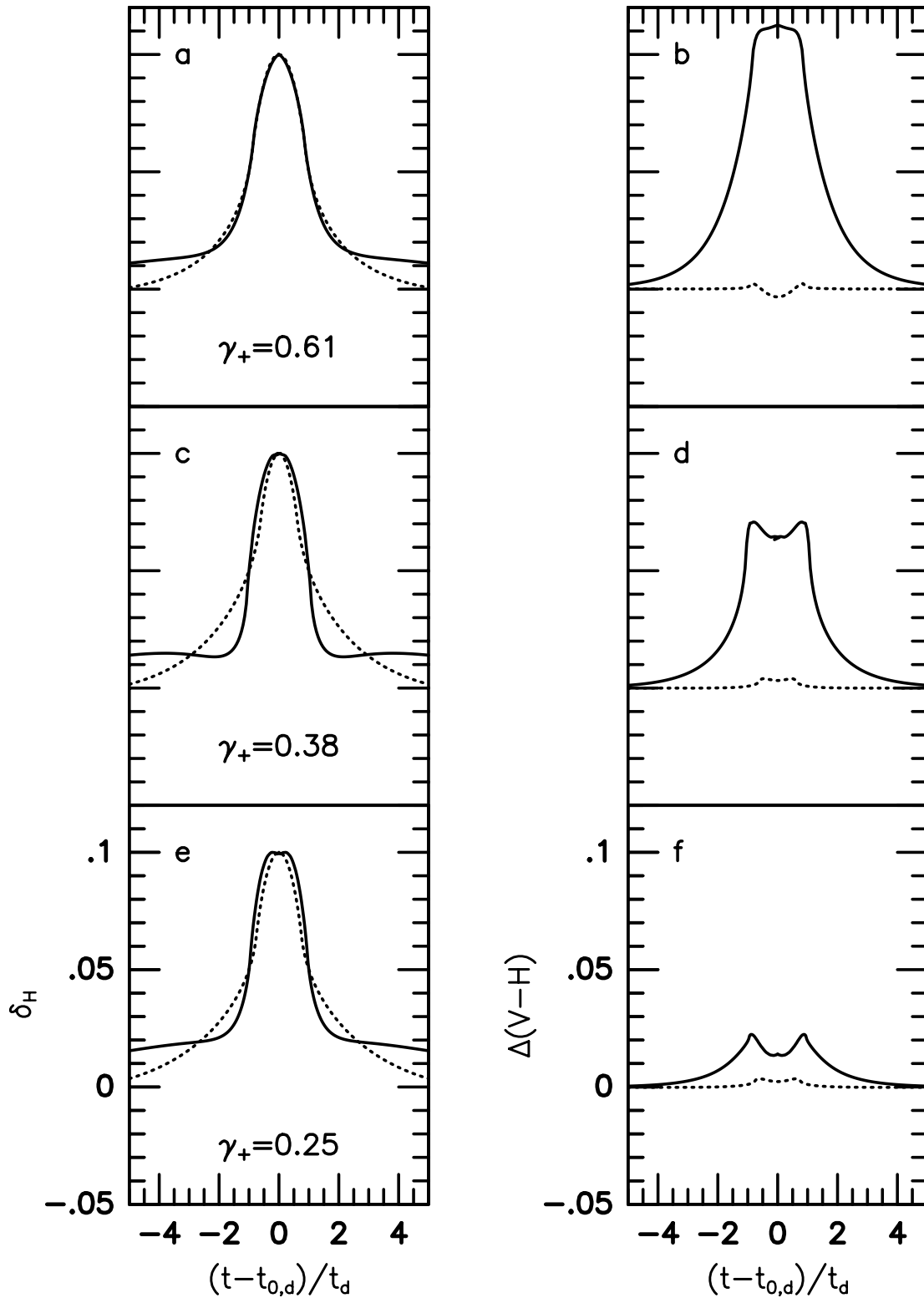
FIG. 5.—Fractional deviations δ for H-band light curves, demonstrating the major/minor image degeneracy for $\phi = 0$, with $\gamma_+ = 0.61, 0.38, 0.25$, in panels ($a$), ($c$), and ($e$). The solid curve corresponds to γ_+, the dashed curve to $\gamma_- = \gamma_+^{-1}$. Also shown in panels (b), (d), and (f) are the associated fractional color changes $\Delta(V - H)$.

of diagrams. The relevant parameter is, ρ, the ratio of the angular radius of the source to the Einstein radius of the planet. Hence, it is quite easy to study all possible degeneracies (see Fig. 4).

The drawback of using the Chang-Refsdal approximation is that it is not exact. Moreover, for any given lensing event, it is straight forward to construct models that are exact. As we argue below, however, the lack of exactness has no significant impact in the analysis of degeneracies. On the other hand, using the exact solution increases the dimensionality of parameter space and thereby the conceptual complexity of the problem, without any compensating

benefits. We therefore strongly advocate using the Chang-Refsdal framework. We present a more detailed analysis in the Appendix.

4. DEGENERACY BETWEEN MAJOR AND MINOR IMAGES

Figure 3 indicates that it should generally be quite easy to distinguish between perturbations of the major image have one major positive excursion, while perturbations of the minor image have two positive excursions separated by a large negative excursion. However, if the observations are made from only one site, then good temporal coverage is far from automatic. The timescale for these excursions is the minimum of the crossing time of the star, $\theta_*/\mu \sim 10$ hr for a giant, and the crossing time of the planet Einstein ring $\theta_p/\mu \sim 10(m/50\ M_\oplus)^{1/2}$ hr. Thus it would be quite possible to observe a positive excursion (or a significant fraction of it) one night and then miss any subsequent excursions due to one or two nights of bad weather. However, if there were three observing sites on different continents, such large data gaps would be rare.

There is nevertheless another possible source of degeneracy between the major and minor images. If ϕ is sufficiently small, then a source coming close to one of the caustics of the perturbations of the minor image could cross the star-planet axis at a point far enough from the planet that the negative excursion along this axis would be very small. In Figure 5, panels a, c, and d, we present examples of this degeneracy for three different values of γ_+ for $\phi = 0$. The parameter ρ was chosen in each case such that the curves for γ_+ and γ_- would be most similar. It is clear that the degenerate curves could be distinguished only if precise measurements could be made at the wings of the perturbation. We therefore investigate another method of breaking the degeneracy between the γ_+ and γ_- cases. Following the discussion in § 1, if the gradient of the magnification across the source is large, and if the surface brightness profile of the source is a function of wavelength, then we would expect a change in color during the perturbation. This color change depends on the magnitude of the average gradient across the source is large, i.e., a larger gradient results in a larger color change. In fact, the gradient of the magnification across the face of the source is much larger in the γ_+ case than the γ_- case. Measuring the color change would therefore enable one to distinguish between the two cases. Giant stars are more limb-darkened in the optical than in the infrared (Manduca, Bell & Gustafsson 1977; Manduca 1979), and thus lensing of a giant source will cause a color change. We define the fractional color change as

$$\Delta(V-H) = -2.5 \log \delta_V/\delta_H , \qquad (4.1)$$

where V and H are the usual broadband magnitudes, and δ_V and δ_H are the fractional deviations for V-band and H-band respectively. Thus $\Delta(V-H)$ is a measure of difference between the lensed and unlensed color (see § 6.3 for a more detailed discussion). In Figure 5, panels b, d, and e, we show the fractional color change $\Delta(V-H)$ corresponding to the perturbations in panels a, c, and d, respectively. There are relatively large (2%–10%) fractional color changes associated with the γ_+ curves throughout the event, while the fractional color changes associated with the γ_- curves are always negligible. Thus by measuring $\Delta(V-H)$, one can distinguish between the two degenerate cases even during

the peak of the perturbation. The larger the value of ϕ, the larger the negative excursion in the minor-image perturbation, and thus curves with $\phi > 0$ will be less degenerate than the examples shown in Figure 5.

5. DEGENERACY OF PLANET POSITION RELATION TO UNPERTURBED IMAGE

Henceforth we will assume that the major/image degeneracy has been broken and the true value of γ is known. Thus we will omit the subscripts $\pm$ from γ. Note that values of $\gamma < 1$ refer to major image perturbations and $\gamma > 1$ refer to minor image perturbations.

In general, the degeneracy of the planet position relative to the unperturbed image introduces an uncertainty in y_p which is $\Delta y_p \sim 2\alpha\theta_p/\theta_e$ where $\alpha\theta_p$ is the separation between the planet and the unperturbed image at the midpoint of the perturbation. Assume that perturbations of $\delta_d \sim 0.05$ are detectable, and consider only perturbations of the major image ($\gamma < 1$). As we will show below, this degeneracy is not problematic for perturbations of the minor image ($\gamma > 1$). From Figure 3 one sees that the majority of these perturbations will be single-peaked, and the peak, δ_d, will occur near the planet-star axis (the x-axes in Fig. 3; see, i.e., trajectory E). Thus α can be estimated by the point at which the $\delta = 0.05$ contour crosses the planet-star axis. For $\gamma = 0.20$, $\alpha \sim 4$, whereas for $\gamma = 0.80$, $\alpha \sim 14$. The range of α will be smaller for any perturbations with $\delta_d > 0.05$. The fractional uncertainty in $y_p = (\gamma)^{-1/2}$ is $\Delta y_p/y_p \sim 2\alpha(\gamma q)^{1/2}$. Consider events with $\gamma \sim 0.60$. For these events, $\alpha \sim 8$ and hence if the degeneracy remains unbroken, the fractional error is $\Delta y_p/y_p \sim (150m/M)^{1/2}$. For Jupiter-mass planets in orbit around M dwarfs, this error is of order unity, while for Neptune-mass planets it is $\sim 20\%$. On the other hand, the degeneracy in q is small (see Table 1 and also Appendix).

In this section we consider only point sources. If the planet has a low mass, so that finite source effects are important (see § 1 and eq. [1.1]), then (as mentioned above) the difference between the two degenerate solutions is small and distinguishing between them is relatively less important. In addition, finite-source effects are more properly addressed in the context of the mass/finite-source degeneracy discussed in § 6.

5.1. *Perturbations of the Major Image*

It is clear from Figure 3 that it is impossible to break the degeneracy if $\phi = 90°$, i.e., if the planetary perturbation takes place at the peak of the light curve ($x_d = \beta$). If the source crosses the perturbation structure in the region of the caustic ($\alpha \lesssim 1$) then degeneracy is relatively unimportant and, in any event, is easily broken (provided $\phi < 90°$) due to the richness of the structure in this region. We therefore focus on the case $\alpha > 1$. From Figure 3, one sees that there is an asymmetry in the light curve which has opposite senses depending on whether the source crosses to the left or the right of the planet. If it passes to the right, then the deviation is more pronounced at the beginning of the perturbation than at the end, and if it passes to the left, the deviation is more pronounced at the end. We define the asymmetry factor P_ϕ as the maximum over all times t of the fractional difference,

$$P_\phi(\gamma, \delta_d) \equiv \max \left[| \delta(t_{0,d} + t) - \delta(t_{0,d} - t)| \right] , \qquad (5.1)$$

where $t_{0,d}$ is the midpoint of the perturbation and $\delta(t)$ is the fractional deviation as a function of time. To lowest order,
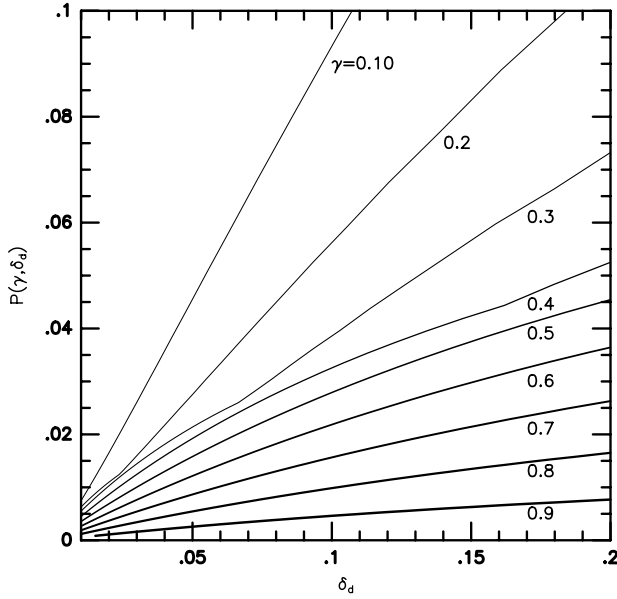
FIG. 6.—Asymmetry factor P (see eqs. [5.1] and [5.2]) for nine values of γ, as a function of maximum deviation, δ_d. The actual asymmetry of a given light curve, $P_\phi(\gamma, \delta_d)$, is given by $P_\phi(\gamma, \delta_d) \sim P(\gamma, \delta_d) \cot \phi$. Using this figure and formula, one can therefore determine whether the degeneracy can be broken for any given sensitivity threshold.

one may approximate

$$P_\phi(\gamma, \delta_d) = P(\gamma, \delta_d) \cot \phi . \tag{5.2}$$

From Figure 3, one can see that as γ increases, the positive contours of δ become more stretched along the planet-star axis, and thus low-peak perturbations occur farther from the areas of negative excursion. One would therefore expect smaller values of P for larger values of γ. Figure 6 shows P as a function of δ_d for several values of γ. As expected, P generally decreases with increasing γ. Analytically, $P(\gamma, \delta_d)$ is roughly given by,

$$P(\gamma, \delta_d) \simeq \begin{cases} \frac{1}{2}(1 - \gamma)\delta_d , & \gamma \gtrsim \frac{1}{3} ; \\ \frac{2}{9}(5 - 12\gamma)\delta_d , & \gamma \lesssim \frac{1}{3} . \end{cases} \tag{5.3}$$

The break at $\gamma \sim \frac{1}{3}$ occurs because the basic structure of the contours of δ changes between $\gamma = 0.40$ and $\gamma = 0.20$ (see Fig. 3). For $\gamma \gtrsim \frac{1}{3}$, the $\delta > 0$ contours converge to the caustic, whereas for $\gamma \lesssim 0.3$, these contours encircle the caustic. This effect causes the analytical form of P to exhibit a break, and is also reflected in Figure 6. From Figure 6, or using equation (5.3), one can determine, for a given sensitivity, whether the degeneracy can be broken for any perturbation. For example, consider trajectories such that $\phi \gtrsim 75°$. If one were sensitive to asymmetries of $P_\phi \sim 1\%$, i.e., $P \sim 0.04$, then the degeneracy could be broken only for $\gamma \lesssim 0.3$, and then only if $\delta_d > 0.1$. On the other hand, if $\phi \sim 30° (P \sim 0.004)$, then the degeneracy could be broken for essentially all values of γ.

5.2. Perturbations of the Minor Image

As we discussed in § 4, to distinguish minor-image from major-image perturbations, it is necessary to observe the negative excursion (centered on the x-axes of the right-hand side of Figure 3. If these are observed, then one can easily distinguish the case where the source transits the right side from the case where it transits the left side of the x-axis

(provided $\phi < 90°$). Hence, there is no degeneracy of minor-image perturbations unless the more severe major/minor-image degeneracy remains unbroken.

6. CONTINUOUS MASS DEGENERACY OF MAJOR IMAGE PERTURBATIONS

By far the potentially most crippling form of parameter degeneracy is the one that is illustrated in Figure 2 and is tabulated in Table 2. The basic character of this degeneracy can be understood analytically using the following theorem (Gould & Gaucherel 1996): if the unperturbed major image crosses the position of the planet and the source is larger than the major-image caustic structure, then

$$\delta_d \simeq \frac{2}{\rho^2 A(\gamma)} , \quad A(\gamma) = \frac{1 + \gamma^2}{1 - \gamma^2} . \tag{6.1}$$

The FWHM of such an event is $t_d \sim 2(\csc \phi)\rho q^{1/2} t_e$. On the other hand, the FWHM of a low-peak perturbation of a point-source is $t_d \sim 2(\csc \phi)q^{1/2}t_e$. Suppose that a low-peak perturbation has observables t_e, β, x_d, t_d, and δ_d. Now consider the combination of observables $Q \equiv [(\beta/x_d)(t_d/t_e)/2]^2$. Then the solution

$$q \sim Q , \quad \rho \lesssim 1 , \tag{6.2}$$

reproduces the observed values of δ_d and t_d. However, the solution

$$q \sim \frac{Q}{\rho_{max}^2} , \quad \rho \sim \rho_{max} , \quad \rho_{max} \equiv \left(\frac{2}{\delta_d} \frac{1 - \gamma^2}{1 + \gamma^2} \right)^{1/2} , \tag{6.3}$$

also reproduces the values of δ_d and t_d. Note that the ratio of masses for these two solutions is ρ_{max}^2. For $\delta_d \sim 5\%$, this ratio is typically $\lesssim 20$ and can be as high as 40. Thus, unless this degeneracy is broken, any low-peak perturbation of a point source by a Jupiter-mass planet can masquerade as a Neptune-mass event, and vice versa. All intervening masses are permitted as well. Clearly, unless this degeneracy is broken, low-peak perturbations will contain very little information about mass, and unambiguous detection of low-mass planets will be impossible. There are three possible paths to breaking this degeneracy.

6.1. Proper-Motion Measurement

If the proper motion μ of the lens relative to the source were measured, then one could partially break the degeneracy between equations (6.2) and (6.3). The angular size of the source θ_* would be known from its dereddened color and magnitude and Stefan's law. The time for the source to cross the perturbation x-axis, $t_c \sim 2\theta_* \csc \phi/\mu$, would then also be known. If one found $t_c < t_d$, this would imply that t_d was dominated by the size of the planet Einstein ring, not the source. Hence, the solution equation (6.2), would be indicated. On the other hand, if $t_c \sim t_d$, one would know only that the solution (6.2) was not correct, and that the mass ratio lay somewhere in the interval $\rho_{max}^{-2} Q \leq qQ$ (see Table 2).

Proper motions can be measured if the lensing star transits or nearly transits the source or by imaging the split image of the source using infrared interferometry. See Gould (1996) for a review. Another approach would be simply to wait a few decades and measure the angular separation of the lens and source. Since typical proper motions are 5 mas yr^{-1}, the separation should be 0″.1 after a few decades. Unfortunately, most lenses are probably
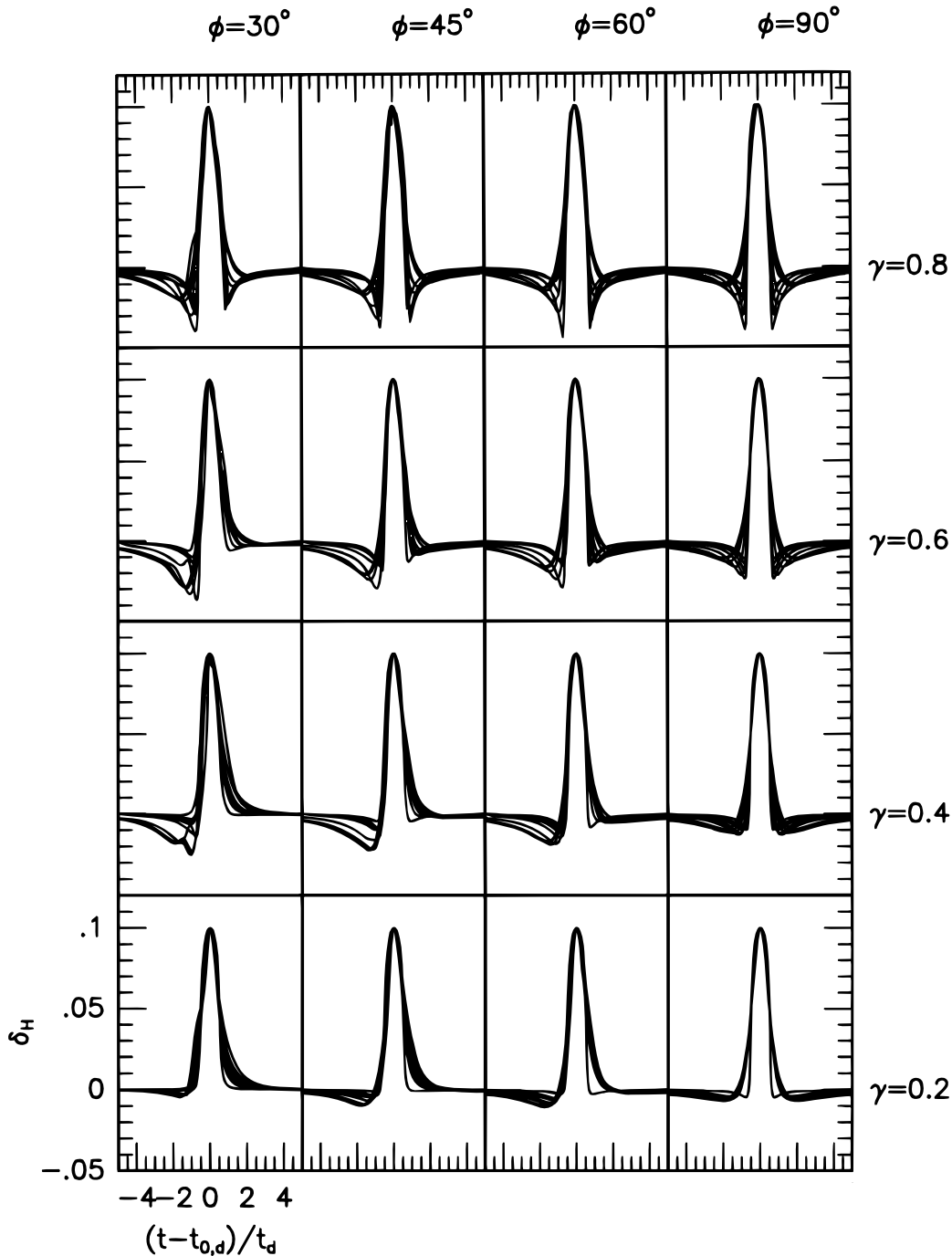
Fig. 7.—Fractional deviations δ for H-band light curves. Similar to Fig. 2, except now shown for shears $\gamma = 0.2$, 0.4, 0.6, and 0.8, and for trajectory of source motion $\phi = 30°$, 45°, 60°, and 90°. (Corresponding curves for $\phi \to -\phi$ can be found by reversing the x-axes). In each curve, the maximum deviation is $\delta_d = 10\%$ and the FWHM is $t_d = 0.06t_e$.

fainter than $M_I = 10$, while typical giant sources are $M_I \sim 0$, and even turnoff stars are $M_I \sim 0$, and even turnoff stars are $M_I \sim 3$. Thus it would be difficult to image the lens until it was quite well separated from the source at which point it might be hard to distinguish it from random field stars. We explore this possibility further in § 8.

6.2. Detailed Light Curves

Although the parameter combinations, equations (6.2) and (6.3), reproduce the gross features of the perturbation (peak and FWHM) equally well, the detailed structures of the light curves are different. Figure 2 illustrates the prin-

cipal difference for elongated perturbation structures, in this case $\gamma = 0.6$. When $\rho \sim \rho_{max}$, the wings show a dip because the source passes over the caustic which is surrounded by regions of negative perturbation (see Fig. 3c and 4). On the other hand, when $\rho \lesssim 1$ the approach to the peak is smooth because the source is passing over the smooth outer portion of the ridge seen in Figure 3c. From Figure 2 it is clear that if these wing structures could be resolved at the $\sim 1\%$ level, then the degeneracy in mass could be reduced from the factor ~ 15 seen in Table 2, to a factor ~ 1.5.

Figure 7 is an array of 16 diagrams each similar to Figure 2, but with different values of γ (0.2, 0.4, 0.6, and 0.8) and
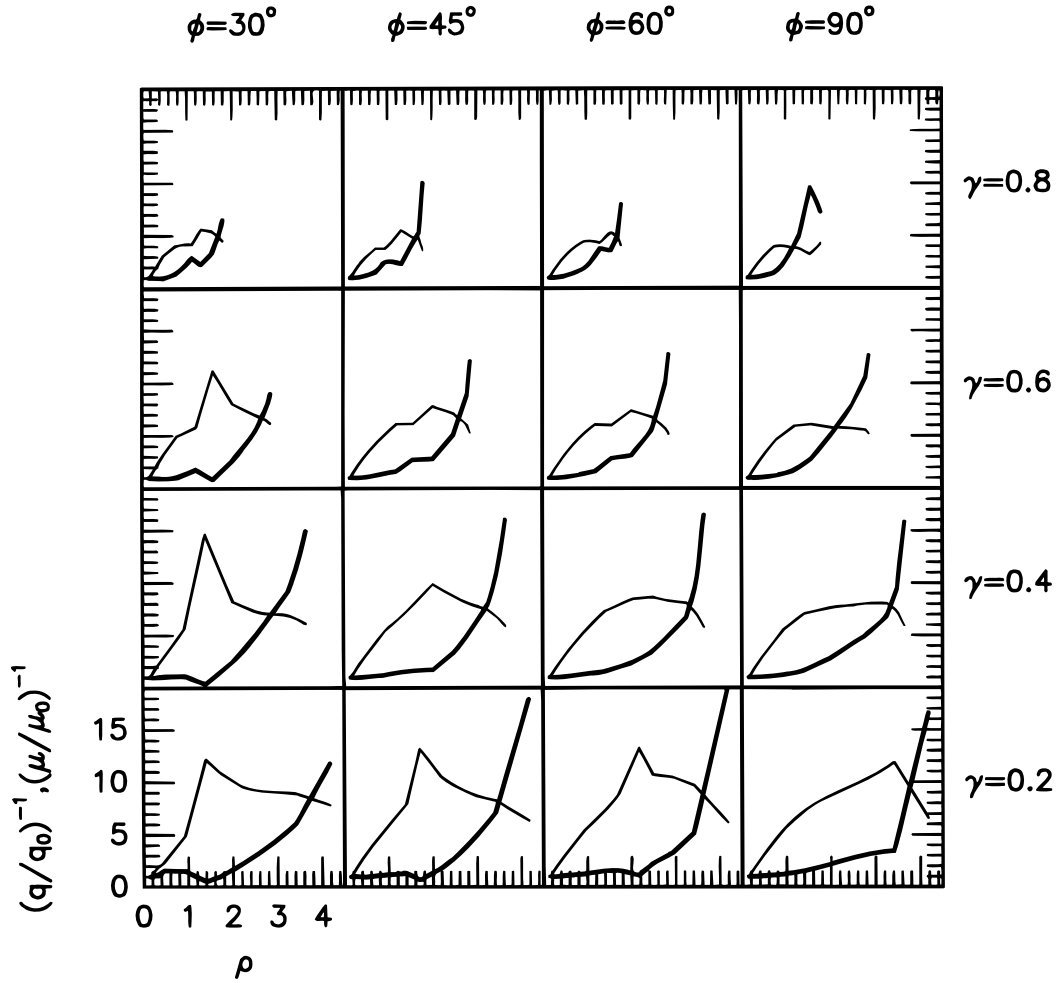
FIG. 8.—Values of $(q/q_0)^{-1}$ (*bold lines*) and $(\mu/\mu_0)^{-1}$ as functions of ρ for each set of degenerate curves in Fig. 7. The fiducial values q_0 and μ_0 are associated with the curve with $\rho = 0.1$.

different angles of source motion ϕ (30°, 45°, 60°, and 90°). It is clear that it is easier to break the degeneracy as γ increases and as ϕ decreases. For $\gamma \gtrsim 0.4$ the uncertainty in q could be significantly reduced if the wing structures could be resolved at the ~1% level. For $\gamma \lesssim 0.2$, however, distinguishing between the degenerate curves would require an accuracy $\ll 1\%$. Figure 8 shows the values of q/q_0 and μ/μ_0 as a function of ρ for each of the combinations of ϕ and γ in Figure 7. Note that larger values of ρ_{max} are allowed for smaller values of γ, and thus the range of acceptable values of q is largest for small γ. This is especially disturbing in light of the fact that the degenerate curves are most similar for small γ.

6.3. Optical/Infrared Colors

A major shortcoming of the details light curve method for breaking the degeneracy is that it depends critically on obtaining accurate observations during two brief intervals covering the wings of the light curve. As a practical matter, it may be difficult to obtain such coverage for a variety of reasons. Once the event is noticed, observatories that are dedicated to the planet search can engage in frequent monitoring and thereby obtain very accurate light curves. However, it is quite possible, indeed likely, that the planetary perturbation will not be recognized in time for intensive monitoring of the first wing. Sometimes observation of the

first wing is crucial to breaking the degeneracy. Moreover, the second wing will likely be observable from at most one observatory which could be affected by bad weather.

Optical/infrared color measurements by contrast yield degeneracy-breaking information throughout the event. The reason is that by the principle of equivalence, lensing of a point source is achromatic. If lensing introduces color changes, the lens must be resolving the source (Witt 1995; Loeb & Sasselov 1995). The best opportunity to observe this effect is by looking for optical/infrared color differences (Gould & Welch 1996) because giant stars are more limb-darkened in the optical than in the infrared (Manduca et al. 1977). Thus, if the planet Einstein ring is larger than the source (and the low peak is due to the source passing over regions of small perturbation), the color changes will be very small. On the other hand, if $\theta_* > \theta_p$ (and the low peak occurs when the large source passes over the caustic), the caustic structure will resolve the differential limb darkening of the star and the color changes will be more pronounced. We determine the fractional color change using equation (4.1), using the limb-darkening model parameterized by the surface brightness as a function of angular distance from the center of the source,

$$\frac{S(\theta)}{S(0)} = 1 - \kappa_1 Y - \kappa_2 Y^2 , \quad Y \equiv 1 - \sqrt{1 - \frac{\theta^2}{\theta_*^2}} . \quad (6.4)$$
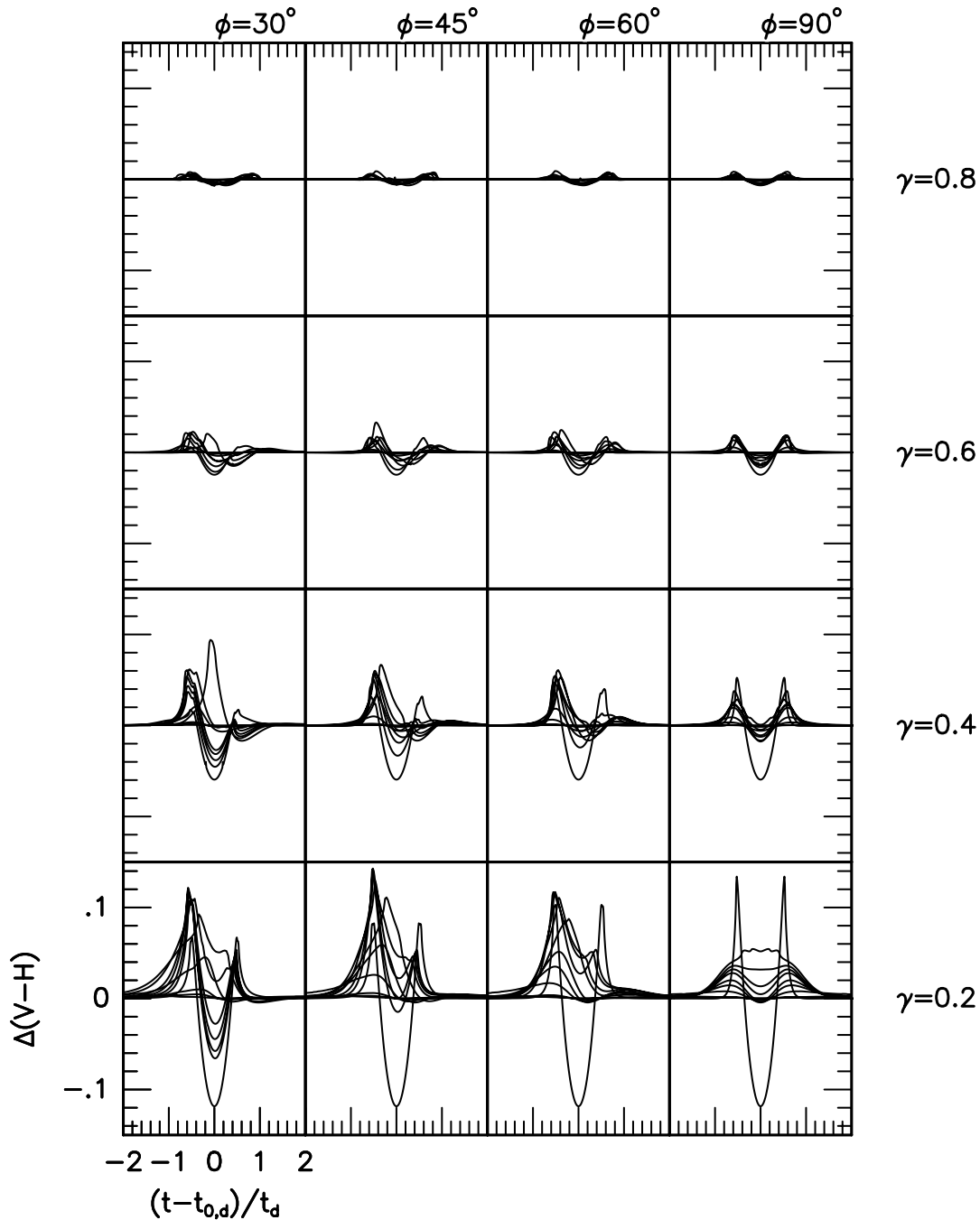
FIG. 9.—Fractional color change $\Delta(V-H)$ for light curves shown in Fig. 7

The coefficients for a cool (4500 K) giant (log $g = 1.5$) of solar metallicity in V and H are $\kappa_1^V = 0.798$, $\kappa_2^V = -0.007$, $\kappa_1^H = 0.206$, and $\kappa_2^H = 0.331$ (Manduca et al. 1997; Manduca 1979). See Gould & Welch (1996) for further discussion and their Figure 3 for a graphical representation of $-2.5 \log S_V(\theta)/S_H(\theta)$. Figure 9 shows the $V-H$ colors for the same parameters as are used for the H-band curves as in Figure 7. The magnitude of the fractional color change is largest for smallest γ. This is fortunate, since, as discussed in § 6.2, the degeneracy is most severe for small γ, both in terms of the similarity in the light curves, and in the range of allowed values of q. It is therefore essential to have optical/

infrared color measurements to ensure that the continuous degeneracy can be broken for all possible values of γ.

7. CONTINUOUS DEGENERACY OF MINOR IMAGE PERTURBATIONS

There is also a continuous degeneracy for minor-image perturbations, but the degeneracy is considerably less severe than for major-image perturbations because the caustic structure is qualitatively different. As with major image perturbations, the basic character of the minor image degeneracy can be understood analytically. Consider the following theorem (Gould & Gaucherel 1996): if the unperturbed
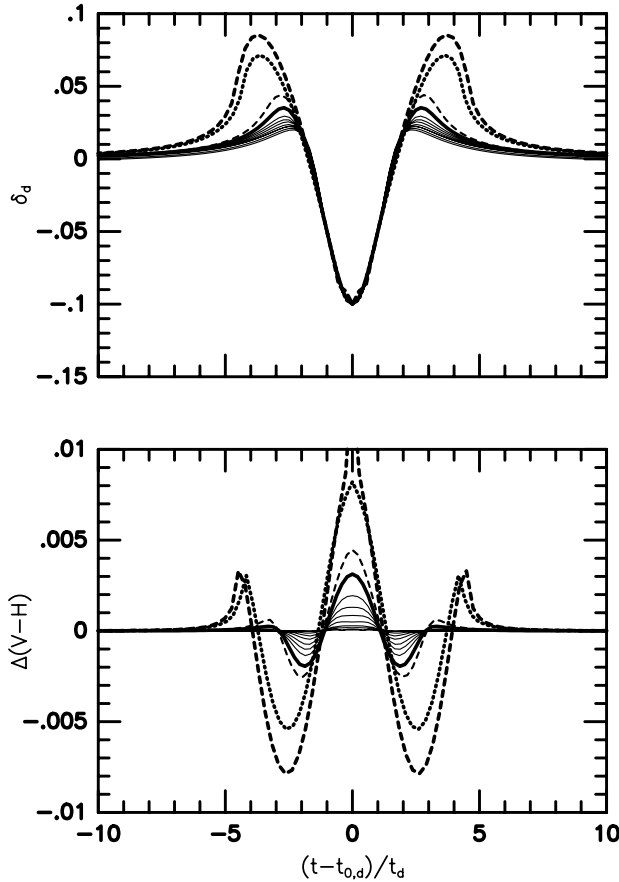
FIG. 10.—*Top*: ten light curves with $\gamma = 1.67$, $\phi = 90°$, all with maximum deviation $\delta_d = -10\%$ and FWHM $t_d = 0.06t_e$. The ratios of source radius to planet Einstein ring range from $\rho = 0.1$ to $\rho = 1.83$, the largest source radius consistent with this maximum deviation. The corresponding relative values of $q = m/M$, and the relative proper motion, μ/μ_0, are given in Table 4. The parameter values given in last four rows of Table 4 (for source radii of $\rho = 1.6$, 1.7, 1.83, and 1.80), correspond to those derived from the bold, dashed, bold dotted, and bold dashed curves, respectively. *Bottom*: fractional color change $\Delta(V - H)$ for the ten curves in the top panel.

minor image crosses the position of the planet and the source encloses both minor-image caustics, then

$$\delta_d \simeq -\frac{f(\rho, \gamma)}{\rho^4 A(\gamma)} \rightarrow -\frac{2}{\rho^4 A(\gamma)}, \quad A(\gamma) = \frac{\gamma^2 + 1}{\gamma^2 - 1}, \quad (7.1)$$

where $f(\rho, \gamma) = [(\frac{1}{2} + \rho^{-2})^2 - \gamma^2\rho^{-4}]^{-1/2}$, and the limit applies for $\rho \gg \gamma^{1/2}$. That is, in contrast to the major-image perturbation (cf. eq. [6.1]), the minor-image perturbation goes to zero rapidly for large sources. For minor image perturbations, the caustics are located at (Schneider et al. 1992)

$$d_{\text{caus}} \sim 2(\gamma - 1)^{1/2}. \quad (7.2)$$

Thus the source must have $\rho \gtrsim 2(\gamma - 1)^{1/2}$ to enclose both caustics. If the source is significantly larger than this, the perturbation will be negligibly small. Hence we can restrict attention to sources $\rho < d_{\text{caus}}$.

For minor-image perturbations, t_d is the FWHM of the negative deviation. For a low-peak point source perturbation, ($\alpha \gtrsim 2$), this duration scales as the distance between the contours of $\delta = 0$ at α (see Fig. 3). For finite sources of

increasing ρ, the trajectories must move close to the center to maintain the observed value of δ_d (see Fig. 4). However, until ρ becomes so large as to cover the caustics, the positions of the $\delta = 0$ contours basically do not change (compare Fig. 3d with Fig. 4). Since these contours are approximately horizontal, t_d is not greatly influenced by changes in ρ. Finally, since the largest permitted source has $\rho \sim d_{\text{caus}}$, which is also approximately equal to the separation of the $\delta = 0$ contours, t_d remains roughly the same even for this extreme case (see Fig. 3). The small degeneracy that does exist arises from the difference between this extreme case on the one hand and the smaller sources and point sources on the other. Examining Figure 3, we can expect that the degeneracy will be somewhat larger for larger values of γ, since the contours of $\delta = 0$ become less horizontal as γ increases. In Table 3 we give the degeneracy in the inferred values of q for $\delta_d = -10\%$ and -5%, and for several values of γ. Note that the largest degeneracy in q is only a factor of ~ 4.

Figure 10a shows 12 light curves for $\gamma = 1.67$ and $\phi = 90°$, all with maximum negative perturbation $\delta_d = -10\%$, and all with the same FWHM. Table 4 gives the inferred values of q and μ for each curve, relative to the fiducial values q_0 and μ_0 associated with $\rho = 0.3$. Note that the degeneracy in the derived mass ratios is only a factor of ~ 1.5. Also note that the inferred mass ratios of the first nine curves agree to $\sim 4\%$. Thus to resolve the small degeneracy

TABLE 3

CONTINUOUS MINOR IMAGE DEGENERACY

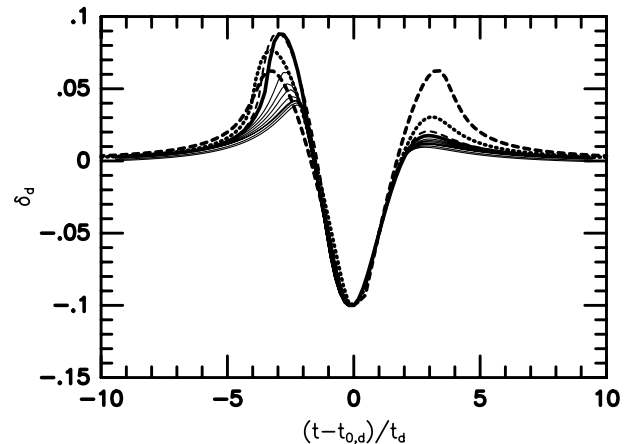| δ_d	γ	Mass Ratio Degeneracy $q_{\text{max}}/q_{\text{min}}$
-10%	1.25	1.45
	1.43	1.09
	1.67	1.68
	2.00	3.82
-5%	1.25	1.66
	1.43	1.16
	1.67	1.22
	2.00	1.67
	2.50	2.34

FIG. 11.—Ten light curves with $\gamma = 0.6^{-1}$, $\phi = 60°$, all with maximum deviation $\delta_d = -10\%$ and FWHM $t_d = 0.06t_e$. All other parameters are the same as Fig. 10, and are given in Table 4. Curves are as in Fig. 10.

TABLE 4

DEGENERATE PARAMETER VALUES: CONTINUOUS MINOR IMAGE

Impact Parameter α	Dimensionless Source Radius ρ	Planet/Star Mass Ratio q/q_0	Proper Motion μ/μ_0
4.32	0.10	0.993	4.014
4.30	0.20	1.016	1.984
4.25	0.40	1.000	1.000
4.21	0.60	1.006	0.665
4.00	0.80	1.020	0.495
3.70	1.00	1.033	0.394
3.50	1.20	1.017	0.331
3.20	1.40	0.994	0.287
2.60	1.60	1.004	0.249
2.00	1.70	1.093	0.225
1.00	1.83	1.591	0.173
0.00	1.80	1.628	0.174

in q, one only needs to distinguish between the last four curves. From Figure 10a it is clear that this would be possible if one could resolve the positive perturbation structures at the $\sim 1\%$ level. Furthermore, the situation presented in Figure 10, for which $\phi = 90°$, is the worst case scenario. Due to the structure of the caustics of minor-image perturbations, trajectories with $\phi < 90°$ display marked asymmetry about $t_{0,d}$, excepting trajectories with $\alpha \sim 0$, which are nearly symmetric. This enables one to distinguish between curves with $\alpha \gtrsim 1$ and $\alpha \sim 0$ more easily when $\phi < 90°$. This is demonstrated in Figure 11, which shows 12 light curves with the same parameters as Figure 10, except that now $\phi = 60°$. Comparing Figures 10 and 11, it is clear that the curves are appreciably less degenerate for $\phi = 60°$ than for $\phi = 90°$. From Figure 10b, we see that the magnitude of the fractional color change for perturbations with $\gamma = 1.67$ is always small, $\Delta(V-H) \lesssim 1\%$. From numerical calculations, we find that $\Delta(V-H) \lesssim 1\%$ regardless of the value of γ. Thus, in contrast to major image perturbations, optical/infrared colors are not useful in resolving the degeneracy in minor image perturbations since the magnitude of the fractional color change is always small.

8. FROM MASS RATIOS TO PLANET MASSES

If the various degeneracies described in this paper are broken, one generally recovers two planetary-system parameters from a planetary microlensing event: q and y_p. While q is of some interest in its own right, y_p is not. The quantities one would most like to know are the planet mass $m = qM$ and the physical projected separation $a_p = r_e y_p$. One could take a purely statistical approach to estimating these quantities: given the measured timescale t_e of the event and a plausible model of the distribution and velocities of lenses and sources along the line of sight, r_e and M can be estimated to a factor of 3. In this section, we discuss what further constraints might be obtained on M and r_e in order to determine m and a_p.

The single most powerful method of acquiring additional information would be to launch a parallax satellite (Refsdal 1966; Gould 1995a; Gaudi & Gould 1997) which would routinely measure $\tilde{r}_e \equiv (D_{os}/D_{ls})r_e$ and often measure the

direction of motion as well. This information would, by itself, narrow the uncertainty in the mass to a factor ~ 1.7 (see Han & Gould 1995, especially Fig. 7). However, if the proper motion μ were also measured, this would yield a complete solution of the lensing geometry including both M and r_e (e.g., Gould 1996). In general, one expects to measure μ only in $\sim 20\%$ of giant events even with relatively aggressive observations (Gould 1996). However, for events with planetary perturbations, μ can be measured much more frequently. Recall from § 6 that for giant sources, major image perturbations, and planetary masses $m \lesssim 100$ $M_{\oplus}$, the planet usually resolves the source (if it is detected at all) and that in the process of resolving the resulting degeneracy, one measures μ. Even when $\theta_* < \theta_p$, the source will sometimes cross a caustic in which case μ can be measured. Finally, for $\theta_* < \theta_p$ one can obtain a lower limit $\mu > \mu_{\min}$ based on the lack of detection of finite-source effects. Since the mass is given by $M = (c^2/4G)\tilde{r}_e t_e \mu$, and since $\tilde{r}_e$ and t_e are measured, this gives a lower limit on the mass (Gould 1995b).

However, it is much more difficult to resolve the finite source degeneracy for minor-image perturbations. Even though (or rather, because) the measurement of the mass ratio, q, is not seriously hampered by this degeneracy, the proper motion μ is poorly determined (see § 2.3. Thus, it would be necessary to measure μ using other methods (see § 6.1).

We now address several questions related to one of those methods: direct imaging of the source and lensing star several decades after the event. For definiteness, we suppose that the measurement is made after 20 years. The expected separation is $\sim 0\rlap{.}''1$, but could plausibly be $\sim 0\rlap{.}''3$. At Baade's Window, the expected number of stars $M_I < 10$ inside this radius is ~ 0.5 (Light et al. 1996). Thus one would not be overwhelmed with candidates. On the other hand, the great majority of lensing events are almost certainly due to objects that are fainter than $M_I = 10$ simply because one does not come close to accounting for the observed events from the observed ($M_I < 10$) stars alone (Han 1996). Thus, to positively identify a candidate star as the lens, one needs additional information. A parallax satellite could provide two pieces of corroborating data. First, the measured $\tilde{r}_e$ together with the proper motion inferred from the candidate-source angular separation would give a mass and distance to the lens (Gould 1995b). One could then predict an apparent magnitude and see if it agreed with that of the candidate. Second, if the parallax measurement gave the angle of motion, one could check this against the direction of the source-candidate separation vector. In addition, the candidate's inferred proper motion must satisfy the lower limit derived from lack of finite-source effects as discussed above. Finally, one could wait another decade or so to see if the direction of the candidate's proper motion was indeed away from the source. These methods would allow one to unambiguously identify the lens responsible for the event from the available candidates, thereby enabling one to measure μ directly.

APPENDIX

JUSTIFICATION FOR THE CHANG-REFSDAL APPROXIMATION

What errors are introduced by the Chang-Refsdal approximation? The unperturbed image structure consists of two images separated by greater than $2\theta_e$. The planet, with an effective sphere of influence $\sim \theta_p \ll \theta_e$ can have a major effect on at most one of these. For definiteness, say this is the major image. In the Chang-Refsdal approximation, the minor image is then treated as being completely unaffected by the presence of the planet. In fact, the planet will change the shear at the minor image by $O(\theta_p/\theta_e)$ and therefore change the magnification by a similar amount. However, what is directly of interest for analyzing the planetary perturbation is not the absolute difference in magnification with and without the planet. Rather, it is the *change* in this difference over the lifetime of the planetary perturbation. Hence, the net effect is $O[(\theta_p/\theta_e)^2]$, i.e., of higher order than the effects being analyzed.

We now turn to the errors in the Chang-Refsdal estimate of the magnification of the perturbed image. In general, the perturbed image is split by the planet into two or four images. For each such image, i, the shear due to the parent star is γ_i. If this value were exactly equal to γ, the shear at the position of the unperturbed image, then the Chang-Refsdal approximation would be exact. Typically, $\Delta\gamma_i \equiv \gamma_i - \gamma$ is small, $\Delta\gamma_i/\gamma \sim O(\theta_p/\theta_e)$, so one expects that the errors induced by the approximation are small.

We focus first on perturbations of the major image. Let $\Delta\theta$ be the separation between the planet and the unperturbed image and define $\alpha = \Delta\theta/\theta_p$. Consider first the case $\alpha \gg 1$ which is important when $\gamma \gtrsim 0.5$ because the magnification contours then become significantly elongated (see Fig. 3). The image is then split into two images, one very close to the planet and the other very close to the unperturbed image. For the image close to the planet, the shear due to the parent star may be significantly misestimated, $\Delta\gamma_i/\gamma \sim \alpha\theta_p/\theta_e$. However, for this image, the total shear is dominated by the planet and is $O(\alpha^2)$, so true fractional error is only $\sim \alpha^{-1}\theta_p/\theta_e$. Moreover, the magnification of this image is small, $O(\alpha^{-4})$, so the total error induced by the approximation is $\sim \alpha^{-5}\theta_p/\theta_e$ and is completely negligible. The other image is displaced by $\sim \alpha^{-1}\theta_p$ from the unperturbed image, so $\Delta\gamma_i/\gamma \sim \alpha^{-1}\theta_p/\theta_e$ which induces a similar small change in magnification. Recall from Figure 3c, that the source trajectory is determined up to a two-fold degeneracy from the maximum magnification. Since the sign of the image displacement is different for the two allowed solutions, the error in estimating the magnification structure could result in two types of errors. First, there is an error in the planet star separation, but this is only $\sim \alpha^{-1}\theta_p$ and is therefore lower by α^{-1} than the basic degeneracy indicated in Figure 3c. Second, there is an error in the estimate of q and, in fact a degeneracy because the error has opposite sign for the two allowed solutions. This could in principal be significant because, within the Chang-Refsdal framework, the two allowed solutions indicated in Figure 3c have identical values of q, and this effect is therefore the lowest order degeneracy. However, the mass ratio is estimated from the FWHM of the light curve which is only a weak function of position along the elongated magnification contours. Moreover, the misestimate of that position is small. We therefore estimate a fractional mass degeneracy of $\Delta q/q \sim \alpha^{-2}q^{1/2}$.

For $\alpha \lesssim 1$ and sources that are small compared to the caustic structure (seen e.g., in Fig. 3), the situation is similar to that of caustic-crossing binary-lens events. The light curves are highly nondegenerate, and one determines not only q and x_p, but also ρ. From the standpoint of understanding degeneracies, the important case is when the source is of order or larger than the caustic. Here, there are roughly equally magnified images displace roughly by θ_p on either side of the planet. Hence, the lowest order errors cancel and the next order errors are $\sim O[\theta_p/\theta_e)^2]$, and can therefore be ignored.

There is one exception to this conclusion. In the argument given above, we implicitly assumed that the planetary perturbation would be significant only over an interval of source motion $\sim \theta_p$. This assumption fails when the perturbation structure is elongated ($\gamma \gtrsim 0.5$) *and* when the angle of source motion is low ($\sin \phi = B/x_d \ll 1$). In this case, the local shear is no longer well approximated by the shear at the center of the perturbation. A proper calculation would then require that the shear be recalculated at every point along the source trajectory, holding the planet fixed. This was the approach of Gould & Loeb (1992) and the resulting magnification for fixed planet position can be seen in their Figure 3. (In the present work, by contrast, what is held fixed in constructing Figs. 3 and 4 is the observable: the shear at midpoint of the perturbation). As can be seen by comparing Figure 3c of Gould & Loeb (1992) and Figure 3c of the present work, for Jupiter mass planets the difference in contours can be significant. However, there are three points to note. First, such events are rare both because the conditions ($\gamma \gtrsim 0.5$, $\beta \ll x_d$) together imply $\beta \lesssim 0.2$ and because the elongated contours are encountered "edge on," so the cross section is only $\sim \theta_p/\theta_e$. Second, the effect is proportional to $q^{1/2}$ and so would not be significant for, e.g., Earth-mass planets. Third, the nature of the effect is to provide information to break degeneracies in cases when the Chang-Refsdal approximation would lead one to believe that there is no information. In brief, in certain rare cases, the Chang-Refsdal approximation leads one to underestimate the amount of information available.

For perturbations of the minor image, the two principle sources of degeneracy are first, confusion of the two caustic peaks with each other and second, confusion of one of these peaks with a perturbation of the major image. Because these peaks are offset in the direction perpendicular to the star-planet axis, the error in their location is $O[(\theta_p/\theta_e)^2]$ and hence of higher order than their separation. As in the case of the major image, there are certain rare events with $\sin \phi \ll 1$ for which the Chang-Refsdal approximation makes the degeneracy seem somewhat worse than it is.

REFERENCES

Alard, C. 1996, in IAU Symp. 173, Astrophysical Applications of Gravitational Lensing, C. S. Kochanek & J. N. Hewitt (Dordrecht: Kluwer), 214
Albrow, M., et al. 1996, in IAU Symp. 173, Astrophysical Applications of Gravitational Lensing, ed. C. S. Kochanek & J. N. Hewitt (Dordrecht: Kluwer), 227
Alcock, C., et al. 1997, ApJ, 479, 119
Ansari, et al. 1996, A&A, 314, 94
Bennett, D., & Rhie, H. 1996, ApJ, 472, 660
Bolatto, A. D., Falco, E. E. 1994, ApJ, 436, 112
Chang, K., & Refsdal, S. 1979, Nature, 282, 561

Gaudi, B. 1997, in preparation
Gaudi, B., & Gould, A. 1997, ApJ, 477, 152
Gould, A. 1995a, ApJ, 441, L21
———. 1995b, ApJ, 447, 491
———. 1996, PAPS, 108, 465
Gould, A., & Gaucherel, C. 1996, ApJ, 477, 580
Gould, A., & Loeb, A. 1992, ApJ, 396, 104
Gould, A., & Welch, D. 1996, ApJ, 464, 212
Han, C. 1997, ApJ, 484, 555
Han, C. Gould, A. 1995, ApJ, 447, 53
Light, R. M., Baum, W. A., & Holtzman, J. A., 1997, in preparation
Loeb, A., & Sasselov, D. 1995, ApJ, 449, 33L

Manduca, A. 1979, A&AS, 36, 411
Manduca, A., Bell, R. A., & Gustafsson, B. 1977, A&A, 61, 809
Mao, S., & Paczyński, B. 1991, ApJ, 374, 37
Paczyński, B. 1986, ApJ, 304, 1
Pratt, M. R. et al. 1996, in IAU Symp. 173, Astrophysical Applications of Gravitational Lensing, ed. C. S., Kochanek & J. N. Hewitt (Dordrecht: Kluwer), 221
Refsdal, S. 1966, MNRAS, 134, 315
Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Berlin: Springer)
Udalski, A., et al. 1994, Acta Astron., 44, 165
Witt, H. 1995, ApJ, 449, 42