Analytic Simplifications to Planetary Microlensing under the Generalized Perturbative Picture

Keming Zhang (张可名) 1

1Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA

ABSTRACT

The two-body gravitational lens equation underlying planetary microlensing is usually transformed into a quintic polynomial that can only be solved numerically. Here, I present methods to acquire approximate analytic and exact semi-analytic solutions. First, I propose the pure-shear approximation, which allows one to acquire closed-form magnification solutions that are accurate apart from a small region near the primary star. While previous works on the perturbative picture suggest that the uniform-shear Chang-Refsdal lens only describes the vicinity of planetary caustics and breaks down in the resonant regime, the pure-shear lens formalism allows for all three caustic topologies. I show that the recently proposed offset degeneracy is a direct consequence of the pure-shear approximation. Second, the sole recognition that there always exists one image that is largely unaffected by the presence of the planet allows one to easily factor out the corresponding root from the quintic polynomial, reducing it to an analytically solvable quartic polynomial. This allows one to acquire semi-analytic solutions that are exact. The two analytic simplifications proposed here not only can allow for substantially faster forward models, but also facilitates the use of gradient-based inference algorithms that provide additional factors of acceleration for the analysis of observed events.

Keywords: Binary lens microlensing (2136), Gravitational microlensing exoplanet detection (2147)

1. INTRODUCTION

In the simplest microlensing scenario, a foreground lens star splits a background source star into two images that are located inside and outside the Einstein radius of the lens star,

$$\theta_\text{E} = \sqrt{\frac{4GM}{D_\text{rel}c^2}},$$

where $G$ is the gravitational constant, $M$ is the lens-star mass, $c$ is the speed of light, and $D_\text{rel}^{-1} = D_\text{lens}^{-1} - D_\text{source}^{-1}$ is related to the relative distance between the lens and source. The image outside is Einstein ring is usually referred to as the major image and the inside image as the minor image. The locations of the major/minor images, along with their magnifications, can also be expressed as simple closed-form expressions of the projected source location.

A two-body lens, on the other hand, splits a source star into either three or five images, depending on whether the source is inside or outside of caustics. The locations of the images are found by solving the lens equation (Witt 1990)

$$\zeta = z - \frac{1}{z} - \frac{q}{z - s},$$

where $\zeta = \xi + i\eta$ is the true source location, $z = z_1 + iz_2$ is the image location, $q$ is the mass-ratio between the two lens components, and $s$ their projected separation in units of the Einstein ring radius of the more massive lens component. The above equation can be transformed into a quintic polynomial (Witt & Mao 1995) that can only be solved numerically. When finite source effects are considered, this quintic polynomial generally has to be solved repeatedly to account for the variance of magnification over the source area, thereby creating a computationally non-trivial problem for the modeling of observed events.

Partly for this reason, analytic simplifications to the two-body microlensing problem in the limit of planetary mass-ratios ($q \ll 1$) were studied soon after the utility of gravitational microlensing for exoplanet discovery was conceived by Mao & Paczynski (1991). In particular, Gould & Loeb (1992) (GL92) recognized that at most one of the major/minor images can be significantly affected by the presence of a planetary lens companion, and that “the perturbed images lie near the unperturbed
image.” By means of a Taylor expansion at the location of the image being perturbed, GL92 showed that the lens-equation could be reduced to a quartic polynomial with closed-form roots, which corresponds to either two or four images depending on whether the source is inside or outside of caustics. Nevertheless, the assumption that the perturbed images are close to the image being perturbed does not generally hold (Section 2), which results in qualitative differences to the exact calculations for both the planetary caustic shape and the magnification contours (see footnote 3 in GL92).

It should also be noted that GL92 only considered the situation where “the planet and the unperturbed image are separated by of order the planet’s own Einstein radius” and thus perturbations associated with planetary caustics, while central and resonant caustic perturbations are not studied until many years later (e.g., Griest & Safizadeh 1998; Dominik 1999; Bozza 1999). The GL92 work was further expanded in Gaudi & Gould (1997) (GG97), which studied the extent to which planetary parameters can be constrained under finite source effects, using a Chang-Refsdal approximation that can be considered as a further simplification to the GL92 formalism.

The approach of GL92 and GG97 was later named the “perturbative picture” by Dominik (1999) (D99), which formally identified the Chang-Refsdal approximation as a Taylor-expansion of the full lens equation, albeit with a slightly different definition of the shear. Nevertheless, none of the three shear-based formalism remains valid beyond approximating the behavior of planetary caustics. Because of this limitation, modeling of current observations still relies on numerically solving the full lens-equation, for which optimized quintic solvers have been developed that provide order-unity speed-up (Skowron & Gould 2012; Fatheddin & Sajadian 2022) compared to a baseline ZROOTs routine from Numerical Recipes.

Practical utilities aside, an important consequence of the Chang-Refsdal approximation of GG97 is the existence of degeneracies for planetary caustic perturbations, commonly referred to as the inner-outer degeneracy (Han et al. 2018). However, observed degeneracies that reference the inner-outer degeneracy rarely have well-isolated planetary caustics (Yee et al. 2021). Moreover, the equation underlying the inner-outer degeneracy was recently found (Zhang et al. 2022) and proved (Zhang & Gaudi 2022) to be valid also for perturbations associated with central and resonant caustics, exactly where the Chang-Refsdal approximation breaks down.

This new degeneracy is then proposed as the offset degeneracy, which also encompasses the close-wide degeneracy (e.g., Griest & Safizadeh 1998) for central caustics.

These results invite a reconsideration of the perturbative picture: if the offset degeneracy is viewed as a generalization of the GG97 degeneracy, then one may reasonably speculate the existence of a generalized perturbative picture that accommodates perturbations with central and resonant caustics as well. Indeed, as we argue in Section 2.1, the condition that one of the major/minor images is only weakly affected by the planet remains true even for central caustic perturbations. By removing the unaffected image outside the physical model, there is a maximum of four images, which may be sufficiently described by a quartic polynomial with closed-form roots.

In this paper, I present two approaches to reducing planetary microlensing to an analytically tractable problem. First, I present the pure-shear approximation, which generalizes the uniform-shear approximation in GG97 to describing magnifications under any caustic topology. I then present a second approach that utilizes the unperturbed location of weakly affected image to efficiently reduce the quintic polynomial to a quartic polynomial, thus allowing one to acquire exact semi-analytic solutions. This paper is organized as follows. In Section 2, the pure-shear lens formalism is proposed and compared to the previous formalism of GL92, GG97, and D99. The accuracy of the resulting magnification maps is examined in Section 3. In Section 4, we show that the pure-shear lens recovers known caustic properties of the exact planetary lens. Finally, the alternative approach of acquiring exact semi-analytic solutions is discussed in Section 5, along with a discussion of how the two analytic simplifications may be utilized to accelerate the analysis of observed events. A Python implementation of the two approaches is provided.\footnote{https://github.com/kmzzhang/analytic-lensing}

2. THE PURE-SHEAR LENS

Let us begin by re-examining the various forms of shear-based approximations used GL92, GG97, and D99, using a uniform notation of the complex two-body lens equation (Equation 2). Given that all three of the previous works consider the lensing behavior near the planetary lens companion (see Section 1), we should first transform the complex lens equation from the primary frame \((\zeta, z)\) to the planetary frame \((\zeta^2, z^2)\), which has coordinate origins at the location of the planet \((z = s)\) and the planetary caustic \((\zeta = s - 1/s)\), and units of the planetary Einstein radius \(\theta_{E,p} = \sqrt{q}\theta_{E,*}\). Applying the
coordinate transformation
\[ \zeta = \sqrt{q}z^{[2]} + s - 1/s \]
\[ z = \sqrt{q}z^{[2]} + s, \]
and rearranging, the lens-equation becomes
\[ \zeta^{[2]} = z^{[2]} - \frac{1}{z^{[2]}} + \frac{z^{[2]} - \bar{z}^{[2]} + \gamma \bar{z}^{[2]}}{s \sqrt{q}z^{[2]} + s}. \] (4)

In the limit of \( q \to 0 \), the above equation is reduced to the Chang-Refsdal lens equation (e.g., An & Evans 2006),
\[ \zeta^{[2]} = z^{[2]} - \frac{1}{z^{[2]}} + \frac{\bar{z}^{[2]}}{s}, \] (5)
with constant shear \( \gamma = 1/s^2 \) (Dominik 1999). For finite \( q \ll 1 \), the Chang-Refsdal lens represents the first order Taylor expansion of Equation 4 around \( \bar{z}^{[2]} = 0 \) (Dominik 1999; Bozza 2000), which also results in a power-series in \( \sqrt{q} \),
\[ \zeta^{[2]} = z^{[2]} - \frac{1}{z^{[2]}} + \sum_{i=1}^{\infty} (-1)^{i+1} \cdot q^{(i-1)/2} \cdot \frac{(\bar{z}^{[2]})^i}{s^i}. \] (6)

On the other hand, rather than the location of the planet \( \bar{z}^{[2]} = 0 \), GL92 considered Taylor expansions at the unperturbed location of the image being perturbed, \( z_0^{[2]} \equiv (\bar{z}_0 - s)/\sqrt{q} \), where \( z_0 \) is analytically known as a function of the true source location. This is based on the assumption that “the perturbed images lie near the unperturbed image.” However, this assumption is valid for at most one of the two or four resulting images — the one that is closest to the unperturbed location. In this case, the unperturbed location is also required to be far from the planetary critical curve (\( |z_0 - s| \gg \theta_{\text{E.P.}} \)), which is equivalent to the source being away from the planetary caustic (GG97). Additionally, the pair of images created for sources inside of caustics is also generally far from the unperturbed image location in units of \( \theta_{\text{E.P.}} \).

With these limitations in mind, the Taylor expansion at \( z_0^{[2]} \) results in
\[ \zeta^{[2]} = z^{[2]} - \frac{1}{z^{[2]}} + \frac{\bar{z}^{[2]} - \bar{z}_0^2 + (s - \bar{z}_0)^2}{\sqrt{q}s\bar{z}_0^2} + \mathcal{O}\left( \left( \bar{z}^{[2]} - \bar{z}_0^{[2]} \right)^2 \right). \] (7)

The resulting expression describes a modified pure-shear lens for which the shear varies as \( \gamma = 1/\bar{z}_0^2 \), but with an additional term that is independent of \( \bar{z}^{[2]} \) and \( \zeta^{[2]} \). This extra term can be interpreted as a positive offset to the origin of \( \zeta^{[2]} \), and only vanishes for \( \bar{z}_0 = s \), i.e. the Taylor expansion at \( \bar{z}^{[2]} = 0 \) (D99). The necessity of this extra term is clear for a hypothetical source at infinity. Here, the shear goes to zero, and \( \zeta^{[2]} \to z^{[2]} \) without the extra term. However, the origins for \( \zeta^{[2]} \) and \( z^{[2]} \) differ by \( 1/s \), which is exactly accounted for by the offset term.

While the residual term \( \mathcal{O}(\bar{z}^{[2]} - \bar{z}_0^{[2]}))^2 \) is almost always greater than order-unity for reasons discussed, the overall magnification contour and planetary caustic shapes remain roughly correct (c.f. footnote 3 in GL92).

We suggest that this should be attributed to the equivalence between the power series in \( \bar{z}^{[2]} \) and the power series in \( \sqrt{q} \) in the limit of \( |z_0 - s| \to 0 \) (Equation 6), which is essentially the planetary caustic limit.

GG97 further simplified the GL92 approximation by adapting a Chang-Refsdal lens with fixed shear evaluated at the mid-point of the perturbation, i.e., on the lens-axis. Nevertheless, the mathematical correspondence to the full planetary lens was not studied in GG97, and the derivations in D99 did not directly support the GG97 formalism either, as the shear is evaluated at the location of the planet \( (\gamma = 1/s^2) \) instead.

### 2.1. The Proposed Formalism

The proposed pure-shear formalism is partly driven by the observation that vertical trajectories under the offset degeneracy result in nearly identical light-curves. Recall that the GG97 degeneracy results from a simple symmetry of the Chang-Refsdal lens along the real-axis, i.e., under the exchange \( \zeta^{[2]} \to -\zeta^{[2]} \) and \( z_1^{[2]} \to -z_1^{[2]} \) in Equation 5. For any variable-shear formalism to maintain this symmetry and thus allow for the offset degeneracy as a direct consequence, it must at the minimum restrict the shear to be real, which is not satisfied by the variable-shear formalism of GL92. The simplest alternative, is then a pure-shear lens whose line-of-constant-shear (LCS) follows perpendicular to the lens-axis.

Let us define a pure-shear lens (Equation 5) with real positive shear \( \gamma = 1/\bar{z}_0^2 \), where
\[ z_+ = \frac{\sqrt{\bar{z}^2 + 4} + \bar{\zeta}}{2}. \] (8)

For sources on the real axis \( (\zeta = \bar{\zeta}) \), \( z_+ \) corresponds to the unperturbed location of the image that is assumed to be perturbed by the planet, which is the major image for \( \xi > 0 \) and the minor image for \( \xi < 0 \). This image is always in the positive lens plane with \( z_1 > 0 \) and thus the “+” subscript. For sources off the real axis \( (\zeta = \bar{\zeta} + i\bar{\eta}) \), the shear is evaluated at the projection of the source location on the real-axis. Since \( z_+ \) does not depend on \( \eta \), the LCS is indeed perpendicular to the real-axis.

The fact that the image being perturbed is always in the same half of the lens plane as the planet becomes intuitive by considering the effect of the planet on the
critical curve, as illustrated in Figure 1. In the non-resonant regime, there is an isolated planetary critical curve(s) centered on the planet that scales as $\theta_{E,p}$. Parts of the primary critical curve ($\theta_{E,p}$) near its intersection with the positive real-axis is elongated, which accounts for the existence of central caustics. In the resonant regime, the primary and planetary critical curves merge together. In all of the cases above, parts of the critical curve in the negative lens plane ($z_1 < 0$) is only weakly affected by the existence of the planet, indicating that the image there is also largely unperturbed. As we will see, this recognition alone also allows for exact semi-analytic solutions to be obtained (Section 5). Note that the example shown in Figure 1 results in the major-image perturbed by a $s < 1$ planet, whereas the previous perturbative picture is restricted to major-image perturbations by $s > 1$ planets and vice versa (see Appendix in Han et al. 2018).

Nevertheless, the absence of the additional offset term in Equation 7 indicates that the pure-shear lens does not correctly recover the image locations, even though the source plane behavior is accurately approximated (Sections 3 & 4). Therefore, the pure-shear lens is degenerate with the exact planetary lens in that they share the same source-plane but not lens-plane behavior. This prevents the application of the pure-shear lens to astrometric microlensing, for which one should turn instead to the exact semi-analytic approach (Section 5). The possibility of obtaining a pure-shear formalism that also recovers the correct image locations, as well as detailed mathematical considerations, should be explored in future work.

To acquire closed-form magnification solutions, one first take the complex conjugate of Equation 5, and substitute the expression for $\bar{z}$ back into Equation 5 itself. After clearing fractions, we arrive at a quartic polynomial,

$$p(z^{[2]}_i) = \sum_{i=0}^{4} a_i(z^{[2]}_i, \gamma, \mu) \cdot (z^{[2]}_i)^i = 0, \quad (9)$$

where, with $\gamma = 1/2s^2_+$ (Equation 8),

$$a_0 = \gamma$$
$$a_1 = -\zeta^{[2]} + 2\gamma\zeta^{[2]}$$
$$a_2 = -2\gamma^2 - \zeta^{[2]}_2 + \gamma\zeta^{[2]}_2$$
$$a_3 = \gamma\zeta^{[2]}_2 + \zeta^{[2]}_2 - 2\gamma^2\zeta^{[2]}_2$$
$$a_4 = -\gamma + \gamma^4.$$

Note that not all roots of the quartic polynomial are solutions to the original lens equation, and each solution should be verified with Equation 5. The total magnification is the sum of the magnification of each individual image, which is given by the absolute value of the inverse Jacobian determinant,

$$\mu_\gamma = \sum_j \left| 1 - \frac{\partial \zeta^{[2]}_j}{\partial \zeta^{[2]}_j} \right|^{-1}, \quad (10)$$

where the derivatives are evaluated using Equation 5 at the valid image solutions $z^{[2]}_j$.

Note that the pure-shear lens only provides the magnification perturbation through,

$$\Delta \mu = \mu_\gamma - \mu_\infty \quad (11)$$
$$\mu_\infty = \frac{1}{|\gamma^2 - 1|}, \quad (12)$$

where $\mu_\infty$ is the terminal magnification ($|\zeta^{[2]}| \to \infty$). The full magnification for the total of three or five images can be found by adding back the single-lens magnifications,

$$\mu = \frac{u^2 + 2}{uvu^2 + 4} + \Delta \mu. \quad (13)$$

Although lengthy when expressed as a function of $(\zeta, s, q)$, Equation 13 is indeed closed-form and thus
orders-of-magnitude faster to evaluate compared to numerically solving quintic polynomials. Let us now turn our attention to the accuracy of the closed-form magnification.

3. MAGNIFICATION

Let us compare pure-shear and exact magnification maps for three empirical classes of planetary microlensing events. Firstly, resonant-topology lenses have a relatively large sphere of influence in the source plane, as the size of the resonant caustic scales as $q^{1/3}$. In the case of semi-resonant lenses where central and planetary caustics are in close proximity to each other but do not merge into a single resonant caustic, there are extended regions of excess magnification or suppressed magnification between the central and planetary caustics (e.g., Abe et al. 2013; Yee et al. 2021). These scenarios with $s \to 1$ have become increasingly common pathways to planet detections with pure-survey strategies (e.g., Zang et al. 2022). Second, high-magnification events have previously been the linchpin for planet discovery with large-scale follow-up programs (e.g., Gould et al. 2010) and have continued to see great utility with pure-survey experiments (e.g., Yang et al. 2022). The third class of perturbations associated with isolated planetary caustics is already well studied in previous works, and is therefore omitted in the current discussion.

3.1. Resonant and Semi-Resonant Topology Events

Figure 2 shows pure-shear and exact calculations of magnification maps for lenses in or near the resonant regime. They appear qualitatively the same. The magnification difference maps reveal two major regimes where the two calculations differ by $>1\%$. First, there is a dumbbell-shaped structure along the imaginary axis of size $\Delta \eta \sim 0.1$. Since planetary perturbations are typically weak along the imaginary axis away from caustics, this “type I” discrepancy is more relevant for high-magnification events where the source crosses the imaginary-axis closer to caustics (see Section 3.2).

Deviations greater than 1% also occur along the excess magnification ridge towards the backside of the central/resonant caustic, which we refer to as the “type II” discrepancy. For close topology lenses (top panel in Figure 2), the excess magnification ridge can be seen as the extension of the two under-fold of the planetary caustics, towards the two off-axis central-caustic cusps. For resonant or wide topology lenses, this excess magnification ridge also originates from the leftmost off-axis cusps, but is more limited in spacial scale. In both cases, de-
Figure 3. Vertical magnification slices associated with Figure 2. Dotted-lines show pure-shear calculations whereas solid lines shows exact calculations. A legend for the color-scheme is provided in the bottom subplot. The intercept on the real-axis is indicated in the upper left corner of each subplot. In each of the top and bottom subplots, two light-curves are shown to be degenerate because the values of $\xi = \pm 0.1$ coincides with those predicted from the offset degeneracy. Here, the pure-shear calculations of the degenerate light-curves are exactly identical, due to a symmetry in the pure-shear lens equation.

Figure 4 shows magnification maps produced with the pure-shear and exact calculations, zoomed into the central region to the size of the central caustic that scales as $q$. One immediately notices the discontinuity across the imaginary axis for the pure-shear calculation, but not for the exact calculation. For sources near the imaginary axis and the primary star, the unperturbed major/minor images are also close to the imaginary axis and about equidistant to the planet, which causes the perturbative assumption to break down. As seen in Figure 4, the magnification near the imaginary axis is overestimated for minor-image perturbations and underestimated for major-image perturbations. Nevertheless, this discontinuity is known in advance, and does not usually coincide with true planetary features. For modeling of observed events, the type I discrepancy may be alleviated through post-hoc treatments, for example, by reducing the weights of photometric data-points near the imaginary axis.

The type II discrepancy, on the other hand, does correlate with planetary features themselves, and becomes increasingly significant for high-magnification minor-image perturbation events. Figure 5 shows light-curves for vertical trajectories with various impact parameters in units of the central caustic size. The left two columns show that the pure-shear calculation overestimates the strength of the excess magnification ridge, or the “exit-wings” of the suppressed magnification zone (“dip”), which itself remains well approximated. Note that the type II discrepancy is much less severe for major-image perturbation events. The extent to which biases will be induced in the inference of $(s,q)$ parameters remains to be seen in future work in the context of finite source calculations (Section 5).

The middle column of Figure 4 shows that the type II discrepancy diverges in the immediate vicinity of the central caustic back-end. There are regions with an unphysical negative magnification under the primary folds of the central/resonant caustic (Figure 4). The size of the negative magnification zone remains size of order $q$, and becomes slightly larger for $s = 1$. This artifact needs to be taken into account for events with $u_0 \approx q$, or $u_0 + \rho \approx q$ when considering finite source effects. The interpretation of these effects requires considerations of the detailed caustic structure, which we will turn our attention to in the next section. Lastly, despite the type I & II discrepancies, Figure 6 shows that the pure-shear calculation is nearly identical to the exact calculation on the lens-axis. This suggests a deeper mathematical connection between the pure-shear and exact planetary lens exists on the real-axis.

3.2. High-Magnification Events
Figure 4. Magnification maps similar to Figure 2, but zoomed into a central region near the primary star. The x-axis for pure-shear is re-parameterized with shear ($\gamma$) using the definition in Equation 8. Prominent artifacts can be seen in the pure-shear magnifications, including the discontinuity between $\gamma > 1$ and $\gamma < 1$. The $\xi = 0$ line is shown in white for the exact calculation for comparison. The color coding in the difference maps indicates differences of less than 0.1%, 1%, 10%, for purple, dark-green, and light-green. The uncolored regions primarily show differences of greater than 10%.

Figure 5. Light-curves resulting from five different trajectories passing close to two different central caustics. Red shows exact calculations and dashed-lines show pure-shear calculations. The impact parameter is shown at the top in units the central caustic size ($\Delta\xi_{caus}$), where a negative value means passing towards the back-end. The x-axis scale is also shown in units of $\Delta\xi_{caus}$. 
the primary cusp and the location of the primary star under the exact calculation. Moreover, the two folds originating from the primary cusp are restricted to the negative source-plane for the pure-shear case as they are locally part of $\gamma > 1$ Chang-Refsdal caustics. However, the primary folds are allowed to traverse into the positive source plane under the exact calculation. The two types of magnification discrepancies discussed in Section 3 can be seen as extensions to the effects above, all of which are tied to the breaking-down of the perturbative assumption.

Elsewhere, the pure-shear approximation accurately recovers known caustic properties. Cusp locations for pure-shear caustics in any topology can be derived by recognizing that Chang-Refsdal cusp locations can be expressed as simple analytic expressions of the shear (e.g., An & Evans 2006), which is related back to the cusp location via Equation 8. Here, I will examine two cases: the central caustic and the $s = 1$ resonant caustic.

Given that the primary cusp is always located at the primary, the central caustic length is given by the location of its other on-axis cusp at $\xi_c \ll 1$. To first-order in $\xi_c$, the shear at $\xi_c$ is,

$$\gamma_c = \frac{1}{z^2(\xi_c)} \simeq 1 - \xi_c,$$

(14)

which is illustrated in the axes labels in Figure 4. The real-axis cusps for $\gamma < 1$ Chang-Refsdal caustics are located at $\pm 2\gamma/\sqrt{1 - \gamma}$, which results in

$$\frac{2\gamma_c}{\sqrt{1 - \gamma_c}} = \pm \sqrt{q} \left( s - \frac{1}{s} - \xi_c \right),$$

(15)

where the plus sign corresponds to the wide topology and minus sign for the close topology. This equation can be rearranged into a cubic polynomial in $\sqrt{\xi_c}$. We may then Taylor expand the valid cubic root in $q$ and acquire

$$\xi_c = \frac{4q}{(s - 1/s)^2} - \frac{32q^2s^3(s^2 - s - 1)}{(s^2 - 1)^3} + \mathcal{O}(q^3).$$

(16)

The first order $q$ term is invariant under $s \leftrightarrow 1/s$ and is in agreement with the exact planetary lens (e.g., Chung et al. 2005). However, the second order term disagrees (e.g., An 2021; Eq. 13), indicating higher-order differences. Note that since the lowest order term in Equation 15 is $\xi_c^{-1/2}$, by dropping the $\xi_c$ term on the right hand side, Equation 15 itself becomes invariant under $s \leftrightarrow 1/s$, allowing one to acquire directly the first order $q$ term without the Taylor expansion.

For the $s = 1$ resonant caustic, the planetary coordinate origin coincides with the primary star $(s - 1/s = 0)$, indicating that the “planetary cusps” are exactly on the

4. CAUSTICS

As seen in the previous section, caustics under the pure-shear approximation are largely in agreement with the exact calculation, but substantial differences exist off the real-axis near the primary star. Recognizing that the pure-shear lens is essentially a Chang-Refsdal lens with variable shear, we may acquire an intuitive interpretation of the pure-shear caustics with a “Chang-Refsdal” (uniform-shear) decomposition along the vertical LCS (line of constant shear). Let us first re-examine the behavior of caustics in Figure 4.

Under the pure-shear formalism, caustics in the positive source plane ($\xi > 0$) are strictly part of $\gamma < 1$ Chang-Refsdal caustics, whereas caustics in the negative source plane are part of $\gamma > 1$ Chang-Refsdal caustics. Crossing from the positive real-axis to the negative real-axis at the primary star, the $\gamma \to 1$ Chang-Refsdal caustic splits into two and the real-axis transitions from the inside to the outside of caustics, thus resulting in a cusp exactly at the primary star. For comparison, there is always a small offset between the location of

![Figure 6: Magnifications on the real-axis for the three cases in Figure 2 (top panel) and the two cases in Figure 4 (bottom panel). Solid lines are exact calculations and dotted lines are pure-shear calculations. For the top panel, the two arrows mark the locations that would result in the offset degeneracy for $(s_A, s_B) = (1/1.1, 1)$ and $(1, 1.1)$, as illustrated in Figure 3.](image-url)
imaginary axis with shear $\gamma = 1$ and thus cusp locations $\eta[2] = \pm 2\gamma/\sqrt{1+\gamma} = \pm \sqrt{2}$. Therefore, the locations of the planetary cusps in the primary frame are simply $\eta_r = \pm \sqrt{2q}$, and the vertical size of the $s = 1$ caustic is thus $\Delta \eta_r = 2\sqrt{2q}$.

As for the cusp on the real-axis towards the planet, we may equate the cusp location

$$\frac{2\gamma}{\sqrt{1-\gamma}} = \sqrt{q} \cdot \xi_r.$$  

Substituting $\xi_r$ for $\gamma$ using the shear definition (Equation 8), and expanding up to first order in $\xi_r$, we have

$$\frac{2}{\sqrt{\xi_r}} - \frac{3\sqrt{\xi_r}}{2} - \xi_r = 0.$$  

For $\xi_r \ll 1$ and $q \ll 1$, the $\sqrt{\xi_r}$ term may be dropped, which result in $\xi_r = \sqrt{2q}$, and this is the length of the resonant caustic. The above results also show that the vertical-to-horizontal width ratio of the resonant caustic scales as $\eta_r/\xi_r \propto q^{1/6}$.

5. DISCUSSION

I have introduced a pure-shear approximation to planetary microlensing, which not only allows for closed-form magnification solutions, but also generalizes the perturbative picture (GL92, GG97, D99) to all three caustic topologies by relaxing the assumed proximity between the image being perturbed and the planet. By observing the effect of the planet on the critical curve, I argued that regardless of the caustic topology, the single-lens image in the negative source plane ($\xi < 0$) is almost always largely unperturbed by the planet, whereas the other image in the positive source plane is split into two or four images. Based on this recognition alone, I now discuss an alternative method to acquire semi-analytic solutions that are exact.

Given that the image in the negative source plane is only weakly affected by the planet, its unperturbed location

$$z_{\text{PSPL}} = \frac{\xi}{2} \left( 1 \pm \sqrt{1 + 4|\xi^2} \right)$$

can be used as an initial guess to Newton’s or Laguerre’s method to quickly solve for its exact location, where PSPL refers to point-source point-lens. In the above equation, the plus sign represents the major image location that is chosen for minor image perturbations, and vice versa. Once one quintic root is found and divided out, the resulting quartic polynomial can be solved in closed-form. The quartic roots can then be verified with the full quintic equation and depending on the requested precision, may be optionally refined by Newton’s method to reduce the numerical noise from the initial root-division, which is nevertheless expected to be small for well-isolated roots (e.g. Skowron & Gould 2012). In fact, the weakly perturbed image is also generally the most isolated image (Figure 1).

There are several practical aspects that need to be considered. First, the closed-form quartic solution discovered by Lodovico Ferrari is known to suffer from certain round-off errors for cases with large root-spread (e.g. Stroback 2010), defined as ratio between the largest and smallest root magnitudes. One should then be careful with defining the origin at the planetary location because one image is usually very close to the planet, resulting in larger root spread. Primary-centered coordinates may be preferable, as the minor image only becomes close to the origin for very faraway sources. Alternatively, an improved quartic solver proposed by Orellana & Michele (2020) may be used instead, which is robust against these errors but only costs twice the computational time as Ferrari’s solution.

Second, the initial root-refinement step would benefit from a combination of Newton’s and Laguerre’s methods depending on the polynomial residual (Skowron & Gould 2012), which is larger for sources near the imaginary axis in our case. On the other hand, iterative methods may also benefit from using an initial guess of the two unperturbed PSPL image locations (Equation 19) plus the planet location, where the remaining two roots could then be solved in closed-form, similar to Skowron & Gould (2012). As an illustration, for the $s = 1$ magnification map in Figure 2, it only took 2/4 iterations with Laguerre’s/Newton’s method to refine the weakly affected image location to a polynomial residual less than $10^{-14}$ for 99.9% of the locations. Using the other PSPL location, the planet location, and the primary location takes 7/16, 5/16, and 9/23 iterations with Newton’s/Laguerre’s method to locate one root subject to the same precision requirements.

The last aspect concerns how the semi-analytic method proposed here should integrate with finite-source algorithms such as contour integration (e.g., Gould & Gaucheler 1997; Dominik 1998; Bozza 2010) or multi-pole expansion (e.g., Gould 2008; Pejcha & Heyrovský 2008), where the lens-equation is solved repeatedly for nearby points on the same extended source. Here, the existence of approximate quintic roots for subsequent evaluations on the same source surface will much accelerate iterative methods but less so for the semi-analytic method. An optimized implementation of the proposed algorithm to integrate with existing finite-source codes (e.g. VBBL; Bozza et al. 2018) will reveal both the optimal finite-source strategy — which may be
a hybrid of the two depending on the finite-source size — and the exact factor of speed-up.

Moving forward, it is beneficial for the two analytic approaches together with finite-source algorithms to be implemented in automatic-differentiation frameworks such as jax (Bradbury et al. 2018) or julia (Bezanson et al. 2017), which allows the gradient of the likelihood function to be acquired without deriving explicit expressions. This allows for the use of gradient-based inference algorithms, particularly Hamiltonian Monte Carlo (HMC) methods including the No-U-Turn Sampler (NUTS; Hoffman et al. 2014), which utilize gradient information to avoid the random walking behavior of common Markov chain Monte Carlo (MCMC) samplers such as Metropolis-Hastings. For the exact semi-analytic approach, it is also not necessary for gradient to be “back-propagated” through the root-refinement step for planetary mass-ratios, where the location of the weakly affect image is insensitive to the planetary parameters.

Practical utilities aside, the pure-shear approximation also demonstrates that planetary lenses can always be approximately decomposed into Chang-Refsdal lenses in the source plane, which offers an interesting conceptual explanation as to why there exists a single unified regime of magnification degeneracy (Zhang et al. 2022) across all three caustic topologies. Although intuition is presented as to why the pure-shear lens should be formulated as such, further mathematical considerations as well as potential corrections to enable the pure-shear approximation for the lens plane, remain open questions for future work.

I would like to thank Scott Gaudi, Jin An, and Frac Bartolić for helpful discussions, and Scott Gaudi and Joshua Bloom for comments on the manuscript. K.Z. thanks the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining Grant #1829740, the Brinson Foundation, and the Moore Foundation: his participation in the program has benefited this work. This work is partially supported by funding from a Two Sigma faculty fellowship and the Moore Foundation.

REFERENCES

Abe, F., Airey, C., Barnard, E., et al. 2013, Monthly Notices of the Royal Astronomical Society, 431, 2975. http://academic.oup.com/mnras/article/431/4/2975/1134263/

An, J. 2021, arXiv:2102.07950 [astro-ph], arXiv:2102.07950. http://arxiv.org/abs/2102.07950

An, J. H., & Evans, N. W. 2006, Monthly Notices of the Royal Astronomical Society, 369, 317. https://doi.org/10.1111/j.1365-2966.2006.10303.x

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. 2017, SIAM Review, 59, 65. https://epubs.siam.org/doi/10.1137/141000671

Bozza, V. 1999, Astronomy and Astrophysics, 348, 311. http://adsabs.harvard.edu/abs/1999A%26A...348..311B

—. 2000, Astronomy and Astrophysics, 355, 423. http://adsabs.harvard.edu/abs/2000A%26A...355..423B

—. 2010, Monthly Notices of the Royal Astronomical Society, 408, 2188, eprint: https://academic.oup.com/mnras/article-pdf/408/4/2188/4222270/mnras0408-2188.pdf. https://doi.org/10.1111/j.1365-2966.2010.17265.x

Zhang et al. 2022

I note that one such differentiable microlensing code named caustics (Bartolić & Dominik, in prep), is currently under development.
Gould, A., & Loeb, A. 1992, The Astrophysical Journal, 396, 104.
http://adsabs.harvard.edu/abs/1992ApJ...396..104G

Gould, A., Dong, S., Gaudi, B. S., et al. 2010, The Astrophysical Journal, 720, 1073, publisher: American Astronomical Society.
https://doi.org/10.1088/0004-637x/720/2/1073

Griest, K., & Safizadeh, N. 1998, The Astrophysical Journal, 500, 37, publisher: IOP Publishing.
https://iopscience.iop.org/article/10.1086/305729/meta

Han, C., Bond, I. A., Gould, A., et al. 2018, The Astronomical Journal, 156, 226, publisher: American Astronomical Society.
https://doi.org/10.3847/1538-3881/aac38e

Hoffman, M. D., Gelman, A., et al. 2014, J. Mach. Learn. Res., 15, 1593

Mao, S., & Paczynski, B. 1991, The Astrophysical Journal, 374, L37, aDS Bibcode: 1991ApJ...374L..37M.
https://ui.adsabs.harvard.edu/abs/1991ApJ...374L..37M

Orellana, A. G., & Michele, C. D. 2020, ACM Transactions on Mathematical Software, 46, 1.
https://dl.acm.org/doi/10.1145/3386241

Pejcha, O., & Heyrovský, D. 2008, The Astrophysical Journal, 690, 1772.
https://doi.org/10.1088/0004-637x/690/2/1772

Skowron, J., & Gould, A. 2012, General Complex Polynomial Root Solver and Its Further Optimization for Binary Microlenses, arXiv, number: arXiv:1203.1034
arXiv:1203.1034 [astro-ph].
http://arxiv.org/abs/1203.1034

Zhang, K., & Gaudi, B. S. 2022, A Mathematical Treatment of the Offset Microlensing Degeneracy, arXiv, number: arXiv:2205.05085 arXiv:2205.05085 [astro-ph].
http://arxiv.org/abs/2205.05085

Zhang, K., Gaudi, B. S., & Bloom, J. S. 2022, Nature Astronomy, 6, 782.
https://www.nature.com/articles/s41550-022-01671-6