

# A method for the microlensed flux variance of QSOs

Jeremy Goodman<sup>\*</sup> and Ai-Lei Sun

*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*

Accepted 2014 March 27. Received 2014 February 16; in original form 2013 September 9

## ABSTRACT

A fast and practical method is described for calculating the microlensed flux variance of an arbitrary source by uncorrelated stars. The required inputs are the mean convergence and shear due to the smoothed potential of the lensing galaxy, the stellar mass function, and the absolute square of the Fourier transform of the surface brightness in the source plane. The mathematical approach follows previous authors but has been generalized, streamlined, and implemented in publicly available code. Examples of its application are given for Dexter and Agol’s inhomogeneous-disc models as well as the usual Gaussian sources. Since the quantity calculated is a second moment of the magnification, it is only logarithmically sensitive to the sizes of very compact sources. However, for the inferred sizes of actual quasi-stellar objects (QSOs), it has some discriminatory power and may lend itself to simple statistical tests. At the very least, it should be useful for testing the convergence of microlensing simulations.

**Key words:** accretion, accretion discs – gravitational lensing: strong – quasars: general.

## 1 INTRODUCTION

Gravitational microlensing by stars along the line of sight to a QSO (quasi-stellar object) is sensitive to the size and structure of the optically luminous regions of the accretion disc, which are otherwise unresolvable at present: sources of angular size much smaller than the Einstein ring of a lensing star can be strongly amplified, whereas more extended sources cannot be (e.g. Young 1981). Recent quasar surveys have turned up a number of quasars suitably aligned with intervening galaxies, and the analysis of their light curves has yielded two principal results. First, in most cases the source size scales with wavelength approximately as expected for a steadily accreting, optically thick disc, namely  $\theta \propto \lambda^{4/3}$  (Anguita et al. 2008; Bate et al. 2008; Eigenbrod et al. 2008; Poindexter, Morgan & Kochanek 2008; Mosquera et al. 2011; Muñoz et al. 2011; but Floyd, Bate & Webster 2009; Blackburne et al. 2011 find otherwise). Secondly however, the absolute source size is too large by at least half an order of magnitude (Pooley et al. 2007; Blackburne et al. 2011; Jiménez-Vicente et al. 2012). Not only is the source larger than thin-disc theory predicts for likely ranges of black hole mass and accretion rate, but also, as Morgan et al. (2010) have emphasized, it is too large for any source that radiates locally as a blackbody unless the radial temperature profile is substantially shallower than  $\theta^{-3/4}$ .

If the latter conclusion is correct, then something is seriously wrong with steady-state thin-disc theory as applied to QSOs. The discrepancy is not small; stated in terms of the areas rather than the linear sizes of the sources, it is more than an order of mag-

nitude. Several physical possibilities would need to be explored, ranging from highly inhomogeneous discs (Dexter & Agol 2011), perhaps caused by thermal or viscous instabilities (Lightman & Eardley 1974; Syunyaev & Shakura 1975); to disc warps, perhaps driven by a radiation-pressure instability (Pringle 1996); to optically thick scattering-dominated winds; or even gap opening by embedded satellites (e.g. Armitage & Natarajan 2002; Goodman & Tan 2004). These modifications to the thin-disc model would affect the structure of the source in different ways, but microlensing fluctuations depend mainly on the half-light radius (Mortonson, Schechter & Wambsganss 2005). However, the error bars in the microlensing size estimates are still large; studies differ as to the magnitude of the size discrepancy, and sometimes even the sign (Rauch & Blandford 1991). The statistical methods used are sometimes opaque, especially when entire photometric time series are fitted to simulations. For these among other reasons, one worries that the overlarge inferred sizes may be due to inadequate statistics or systematic errors. Ultimately, this will be decided only by more data independently analysed.

In the meantime, having no data of our own to offer, we revive and extend a semi-analytic approach to the prediction of microlensing statistics pioneered by Katz, Balbus & Paczynski (1986) and Deguchi & Watson (1987), and improved by Seitz & Schneider (1994), Seitz, Wambsganss & Schneider (1994), and Neindorf (2003). The semi-analytic method adds no noise to the comparison between models and data and avoids possible biases due to choices of numerical parameters, such as the numbers of rays shot or the size of the regions simulated in the lens and source plane. It is easily extended to the autocorrelation of the magnification as a function of time lag provided that the velocity dispersion of the lensing stars is small compared to the motion of the lensing galaxy across the line

<sup>\*</sup>E-mail: [jeremy@astro.princeton.edu](mailto:jeremy@astro.princeton.edu)

of sight. The method is limited, however, to second moments of the light curve. Simulations are more flexible and can address higher moments of the magnification, and the structure of caustic crossings, effects that are more sensitive to the sizes of very compact sources. These are good scientific reasons to prefer simulations. Certain practical impediments to the use of the semi-analytic method, however, are removable. The method is perceived to be cumbersome and hard to use, involving as it does multiple integrals and expansions in special functions. In the efficient but restricted version implemented by Neindorf (2003), it is limited to Gaussian sources. In this paper, we develop an efficient and practical version of the semi-analytic formalism that can be applied to arbitrary source structures, not just Gaussian ones. In hopes that it will be more widely used, we have implemented the method in PYTHON code downloadable from a git repository <https://bitbucket.org/jjgoodman/mulensvar>.

The effort required by this method factors in a way well suited to exploring a large range of source structures and macrolensing parameters. The main part of the work is to calculate a kernel that depends upon dimensionless statistical properties of the lens such as optical depth, shear, and mean stellar mass or mass distribution. This kernel can be computed efficiently for almost any reasonable choice of these parameters, using a pre-computed special function (Section 3.1). The kernel so obtained is independent of the source structure. Second moments of magnification are obtained by convolving the kernel with the spatial autocorrelation function of the source. Hence, if one is reasonably confident of the average properties of the lenses, then a single computed kernel may quickly be applied to many possible sources. Even if simulations are necessary, one can at least use this method as partial check on convergence with respect to the purely numerical parameters of the simulations, such as the number of rays, stars, and realizations. Although agreement with the semi-analytic flux variance – using for example methods described in Lewis & Irwin (1995, 1996) and Refsdal & Stabell (1997) – does not guarantee that the simulations are fully converged for all purposes, one should not trust simulations that do not show such agreement.

The plan of this paper is as follows. Section 2 introduces the lensing kernel and associated notation. Section 3 describes our algorithms for evaluating the lensing kernel numerically and (in certain limits) analytically. Most of this quite technical section can safely be skipped if one wishes to use our codes as they are, but one should read it carefully before attempting modifications. The asymptotic analytic results (21a) and (21b) and their physical interpretation may be of more general interest, however. Interpolation between these formulae captures much of the behaviour of the microlensing kernel. Section 4 presents numerical examples and tests. These include applications to Huchra’s lens (Q2237+305, extensively studied by, e.g., Wambsganss, Paczynski & Schneider 1990; Jaroszynski, Wambsganss & Paczynski 1992; Mosquera et al. 2013) and to Dexter & Agol (2011)’s fragmented discs. The predictions of the method are compared with simulations made using a simple ray-shooting code.

## 2 DEFINITIONS

Macrolensing by the large-scale mean potential of the intervening galaxy splits the source into macroimages separated typically by arcseconds. Microlensing by stars divides each macroimage into a cluster of subimages with angular separations comparable to the Einstein ring of a single star,  $\theta_E = \sqrt{4GM_*D_{LS}/c^2D_S D_L}$ , which is  $\approx 2(M_*/M_\odot)^{1/2} \mu\text{as}$  ( $\approx 10^{-11}$  rad) for typical lens and source redshifts  $(z_L, z_S) = (0.5, 2.0)$ . At the angular-diameter distance of

the lens, this projects to  $D_L\theta_E \sim 10^{-2}$  pc. In theoretical studies of microlensing, it is often necessary to consider a region several tens of  $\theta_E$  across to allow for the interactions between caustics and for the motion of the line of sight across the lens plane (at, say,  $0.03\theta_E \text{ yr}^{-1}$ ). Even so, the region of interest is scarcely larger than a parsec. On such scales, the contribution of the large-scale smoothed galactic potential to the lens equation is well approximated by a locally constant linear transformation,

$$\mathbf{M} = \begin{pmatrix} 1 - \kappa - \gamma \cos 2\alpha & -\gamma \sin 2\alpha \\ -\gamma \sin 2\alpha & 1 - \kappa + \gamma \cos 2\alpha \end{pmatrix}. \quad (1)$$

The notation of Miralda-Escudé (1991) is used for the dimensionless convergence  $\kappa$ , total shear  $\gamma$ , and orientation  $\alpha$  of the principal axes of the shear. The macrolensing magnification is  $(\det \mathbf{M})^{-1} = [(1 - \kappa)^2 - \gamma^2]^{-1}$ . The local mean surface density in stars contributes a portion  $\kappa_*$  to the convergence. The remainder,  $\kappa' = \kappa - \kappa_*$ , is due to dark matter and gas and is assumed to be smoothly distributed. Typically,  $\kappa_*/\kappa \sim 0.05-0.1$  (Mediavilla et al. 2009), but stars may dominate the convergence when the macroimage lies in the central parts of the lensing galaxy, as in Q2237+305.

Within a single macroimage, the lens equation relating a line of sight from Earth in direction  $\theta$  to its unlensed counterpart  $\theta_S$  on the source plane becomes

$$\theta_S = \mathbf{M} \cdot \theta + \left( \theta_E^2 \sum_k \frac{\xi_k - \theta}{|\xi_k - \theta|^2} + \kappa_* \theta \right) \equiv \mathbf{M} \cdot \theta + \phi. \quad (2)$$

Here  $\xi_k$  is the angular position of the  $k$ th microlensing star, with the origin of coordinates for  $\theta$  and  $\xi_k$  taken at the centre of the macroimage. The term  $\kappa_* \theta$  has been grouped with the sum over stars to prevent double-counting the mean stellar convergence, since  $-\kappa = -\kappa_* - \kappa'$  in  $\mathbf{M}$ . For the moment, all stars have the same mass and therefore the same  $\theta_E^2$ , but a mass function will be introduced later (Section 3.2). Lewis & Irwin (1995) show that the probability distribution of amplification is insensitive to the mass function given a fixed mean mass. In the approximation that the sum above ranges over stars distributed throughout the lens plane with constant mean number per unit area, the net stellar deflection  $\phi$  is a random variable with zero mean and stationary statistics. In other words, the joint probability density for this variable to take on values  $\{\phi_1, \dots, \phi_n\}$  at positions  $\{\theta_1, \dots, \theta_n\}$  depends upon the differences  $\theta_i - \theta_j$  and not on the centroid  $\bar{\theta} = (\theta_1 + \dots + \theta_n)/n$ . Binary and higher multiplicity correlations among the stars would not be incompatible with stationary statistics but would make results along the lines of this paper almost impossible. Such correlations could be represented in simulations quite easily but rarely are.

An explicit expression is possible for Fourier transform  $P(\omega_1, \dots, \omega_n)$  of the  $n$ -point probability density  $P(\phi_1, \dots, \phi_n)$ . We call this the ‘characteristic function’:

$$\begin{aligned} & \langle \exp [i\omega_1 \cdot \phi(\theta_1) + \dots + i\omega_n \cdot \phi(\theta_n)] \rangle \\ &= \exp \left\{ \nu \int \left[ \exp \left( i \sum_{j=1}^n \frac{\theta_E^2 \omega_j \cdot (\xi - \theta_j)}{|\xi - \theta_j|^2} \right) - 1 - i \sum_{j=1}^n \frac{\theta_E^2 \omega_j \cdot (\xi - \theta_j)}{|\xi - \theta_j|^2} \right] d^2\xi \right\}. \quad (3) \end{aligned}$$

Here  $\langle \dots \rangle$  denotes expectation value, and  $\nu \equiv \kappa_*/\pi\theta_E^2$  is the mean number of stars per unit area. Chandrasekhar (1943) derived this for  $n = 1$ , and Deguchi & Watson (1987) stated the result for  $n = 2$

without derivation. Seitz & Schneider (1994) devote an entire paper to the discussion of the two-point function and its inverse Fourier transform. For completeness, we sketch a derivation here for general  $n$ : consider the counterpart to the left-hand side of equation (3) when  $\phi_j$  replaced by  $\phi'_j \equiv \phi_j - \kappa_* \theta_j$ . The primed variables do not have stationary statistics because their means are  $-\kappa_* \theta_j$ . However, the contribution of a small area  $A$  of the lens plane with centroid  $\xi_A$  to their characteristic function is, with  $f_A \equiv \sum_j \theta_j^2 \omega_j \cdot (\xi_A - \theta_j) / |\xi_A - \theta_j|^2$ ,

$$\sum_{N_A=0}^{\infty} \frac{(vA)^{N_A}}{N_A!} e^{-vA} e^{iN_A f_A} = \exp[(e^{i f_A} - 1) vA],$$

if the number of stars within area  $A$  ( $N_A$ ) is Poissonian with mean  $vA$ . Multiplying the independent contributions from all such areas in the lens plane gives equation (3) except for the final term within the square brackets, which can be explained as follows. If the stars were restricted to a circular region  $|\xi| < R$ , then the average deflection due to these stars at  $|\theta_j| < R$  would be

$$v\theta_e^2 \int \frac{\xi - \theta_j}{|\xi - \theta_j|^2} d^2\xi = -\kappa_* \theta_j.$$

Thus, the term in question removes the mean deflection from the characteristic function, as appropriate for the  $\phi_j$  rather than the  $\phi'_j$ .

## 2.1 Moments of the flux

For a transparent lens, the observed surface brightness  $I$  in direction  $\theta$  is equal to the unlensed surface brightness  $I_0$  at the position  $\theta_S = \mathbf{M} \cdot \theta + \phi(\theta)$ , where the lensed ray intercepts the source plane. Therefore, the flux of a macroimage summed over all of its microimages is

$$F = \int I(\theta) d^2\theta = \int I_0(\mathbf{M} \cdot \theta + \phi) d^2\theta.$$

The source and lensing variables can be separated by introducing the Fourier transform  $\hat{I}_0$  of the unlensed source,

$$I_0(\theta) = \int \frac{d^2\omega}{(2\pi)^2} \hat{I}_0(\omega) e^{i\omega \cdot \theta}, \quad (4)$$

so that

$$F = \int \frac{d^2\omega}{(2\pi)^2} \hat{I}_0(\omega) \int d^2\theta e^{i\omega \cdot (\mathbf{M} \cdot \theta + \phi)}. \quad (5)$$

The  $n$ th moment of the flux is therefore

$$\langle F^n \rangle = \left[ \prod_{j=1}^n \int \frac{d^2\omega_j}{(2\pi)^2} \hat{I}_0(\omega_j) \int d^2\theta_j e^{i\omega_j \cdot \mathbf{M} \cdot \theta_j} \right] \times \langle \exp[i\omega_1 \cdot \phi(\theta_1) + \dots + i\omega_n \cdot \phi(\theta_n)] \rangle.$$

The expectation value  $\langle \dots \rangle$  is independent of the centroid  $\bar{\theta} \equiv (\theta_1 + \dots + \theta_n)/n$ . This can be used to reduce the number of vectorial integrations by two. Setting  $\theta_j = \theta'_j + \bar{\theta}$  and integrating over  $\bar{\theta}$  produces (dropping the primes hereafter)

$$\langle F^n \rangle = |\det \mathbf{M}|^{-1} \left[ \prod_{j=1}^n \int \frac{d^2\omega_j}{(2\pi)^2} \hat{I}_0(\omega_j) \int d^2\theta_j e^{i\omega_j \cdot \mathbf{M} \cdot \theta_j} \right] \times (2\pi n)^2 \delta^2(\omega_1 + \dots + \omega_n) \delta^2(\theta_1 + \dots + \theta_n) \times \langle \exp[i\omega_1 \cdot \phi(\theta_1) + \dots + i\omega_n \cdot \phi(\theta_n)] \rangle. \quad (6)$$

For  $n = 1$ , the delta functions absorb both integrations, so that  $\langle F \rangle \rightarrow |\det \mathbf{M}|^{-1} \hat{I}_0(\mathbf{0})$ , which is the correct expression for the mean macrolensed flux.

As in previous works, our numerical methods are limited to the lowest non-trivial moment,  $n = 2$ . We set  $\omega_1 = -\omega_2 \equiv \omega$  and  $\theta_2 = -\theta_1 = \Delta\theta/2$ . Then

$$\langle F^2 \rangle = |\det \mathbf{M}|^{-2} \int \frac{d^2\omega}{(2\pi)^2} |\hat{I}_0(\omega)|^2 \hat{J}(\omega), \quad (7)$$

$$\hat{J}(\omega) = |\det \mathbf{M}| \int d^2\Delta\theta e^{i\omega \cdot \mathbf{M} \cdot \Delta\theta} \langle e^{i\omega \cdot [\phi(\theta) - \phi(\Delta\theta)]} \rangle. \quad (8)$$

Once the kernel  $\hat{J}(\omega)$  has been calculated, equation (7) can be used to find the second moment of the flux for a general source structure with spatial power spectrum  $|\hat{I}_0(\omega)|^2$ . Furthermore, if one adopts the ‘frozen-screen’ approximation in which the motions of the stars within the lensing galaxy are neglected compared to the transverse motion  $\mathbf{V}_\perp$  of the lensing galaxy across the line of sight, then one can use the same kernel to calculate the correlation between the microlensed flux at finite time lag:

$$\langle F(t)F(t + \tau) \rangle \approx |\det \mathbf{M}|^{-2} \times \int \frac{d^2\omega}{(2\pi)^2} |\hat{I}_0(\omega)|^2 \hat{J}(\omega) \exp\left(\frac{i\omega \cdot \mathbf{V}_\perp \tau}{D_L}\right). \quad (9)$$

As discussed by Kundic & Wambsganss (1993), however, the frozen-screen approximation is not very accurate, especially at high optical depth (Wyithe, Webster & Turner 2000).

As discussed in Section 3.3,  $\hat{J}(\omega)$  contains a term proportional to  $\delta^2(\omega)$  that accounts for the square of the mean flux in equations (7) and (9). If this term is subtracted from  $\hat{J}(\omega)$ , the right-hand sides of these equations yield the variance and covariance of the flux. In other words, subtracting the delta function from  $\hat{J}(\omega)$  has the same effect as removing the mean from  $F$  before its second moments are calculated. The numerical method that we use to evaluate  $\hat{J}(\omega)$  makes this subtraction automatic. At the risk of confusion, we hereafter interpret  $\hat{J}(\omega)$  as the modified kernel so that  $\langle F^2 \rangle \rightarrow \text{var}(F)$  in equation (7).

## 2.2 Transformation and scaling of the characteristic function

In the approach outlined above, the first step towards  $\langle F^n \rangle$  is to calculate the characteristic function (3). This involves integration over the vector  $\xi$ . The result will evidently be a function of the  $2n$  vector-valued parameters  $(\omega_1, \dots, \omega_n)$  and  $(\theta_1, \dots, \theta_n)$ . Because of the two delta functions in equation (6), only  $2n - 2$  of these parameters are independent, equivalent to  $4n - 4$  independent scalar parameters. Because the characteristic function is independent of the macrolensing shear, it is statistically isotropic and hence unaffected by an overall rotation of the lens plane. This reduces the number of essential scalar parameters by one. The following scaling property reduces the number by one more. Let  $\omega_j \rightarrow \sigma\omega_j$ ,  $\theta_j \rightarrow \sigma\theta_j$ , with the same factor  $\sigma$  for all  $j$ . Rescaling the dummy integration variable  $\xi \rightarrow \sigma\xi$  and the stellar number density  $\nu \rightarrow \sigma^{-2}\nu$  (equivalently  $\theta_e^2 \rightarrow \sigma^{-2}\theta_e^2$  or  $\kappa_* \rightarrow \sigma^{-2}\kappa_*$ ) restores the integration to its form for  $\sigma = 1$ .

As a result, the number of essential scalar parameters needed to specify the  $n$ -point characteristic function is  $4n - 6$ . (This assumes  $n > 1$ , else the characteristic function collapses to a constant.) For  $n = 2$ , this is manageable. The logarithm of the two-point characteristic function can be tabulated for a grid of values of two parameters, each entry in the table requiring a two-dimensional

quadrature. Fortunately, the calculation needs to be done only once, since it is independent of the macrolensing matrix ( $\mathbf{M}$ ), and even the dependence on  $v$  and  $\theta_E^2$  (equivalently,  $\kappa_* = \pi v \theta_E^2$ ) can be scaled out. This table can then be used to calculate  $\text{var}(F)$  for any combination of macrolensing, stellar density, and source structure.

Before going into those details, we generalize a change of variables introduced by Neindorf (2003). Associate with each real-valued vector  $\mathbf{v} = (v_x, v_y)$  a complex number  $\tilde{v} = v_x + i v_y$ . The argument of the inner exponential of equation (3) becomes  $i f_x$  in terms of the meromorphic function

$$\tilde{f}(\tilde{\xi}) \equiv \theta_E^2 \sum_{j=1}^n \frac{\tilde{\omega}_j}{\tilde{\xi} - \tilde{\theta}_j}. \quad (10)$$

The change of variables  $\tilde{\xi} \rightarrow \tilde{f}$  has Jacobian  $|d\tilde{f}/d\tilde{\xi}|^{-2}$ , so that the integral in equation (3) becomes

$$v \iint (e^{i f_x} - 1 - i f_x) |d\tilde{f}/d\tilde{\xi}|^2 d f_x d f_y. \quad (11)$$

The derivative  $d\tilde{\xi}/d\tilde{f}$  is needed as a function of  $\tilde{f}$ , which requires inverting the function (10). This leads to a polynomial of degree  $n$  in  $\tilde{\xi}$ . The Jacobian must be summed over all  $n$  roots. A slight simplification is that the coefficient of  $\tilde{\xi}^{n-1}$  in the polynomial vanishes because  $\sum_j \tilde{\omega}_j = \sum_j \tilde{\theta}_j = 0$ , so that the  $n$  roots for  $\tilde{\xi}$  sum to zero.

In particular, for  $n = 2$  the roots are

$$\tilde{\xi} = \pm \sqrt{\tilde{\theta}_1(\tilde{\theta}_1 + 2\theta_E^2 \tilde{\omega}_1 / \tilde{f})}.$$

Set  $\boldsymbol{\omega} = \boldsymbol{\omega}_1 = -\boldsymbol{\omega}_2$  and  $\Delta\boldsymbol{\theta} = 2\boldsymbol{\theta}_2 = -2\boldsymbol{\theta}_1$ . Rescale the integration variable to  $\tilde{r} \equiv u - i v \equiv \tilde{f} \Delta\tilde{\theta} / 4\theta_E^2 \tilde{\omega}$ . Finally, let  $\psi$  be the angle between  $\boldsymbol{\omega}$  and  $\Delta\boldsymbol{\theta}$ , and  $\Delta\theta \equiv |\Delta\boldsymbol{\theta}|$ . The two-point function is then

$$\langle e^{i\boldsymbol{\omega} \cdot [\boldsymbol{\phi}(0) - \boldsymbol{\phi}(\Delta\boldsymbol{\theta})]} \rangle = \exp \left[ \frac{\kappa_* (\Delta\theta)^2}{4\pi\theta_E^2} H \left( \frac{4\theta_E^2}{\Delta\theta} \boldsymbol{\omega} \right) \right], \quad (12a)$$

where  $H$  as a function of  $\boldsymbol{\beta} \Leftrightarrow (\beta \cos \psi, \beta \sin \psi)$  is

$$\begin{aligned} H(\boldsymbol{\beta}) &\Leftrightarrow H(\beta, \psi) \\ &\equiv \iint \frac{\exp[i\beta(u \cos \psi + v \sin \psi)] - 1 - i\beta(u \cos \psi + v \sin \psi)}{2(u^2 + v^2)^{3/2} \sqrt{(u-1)^2 + v^2}} \\ &\quad \times du dv. \end{aligned} \quad (12b)$$

### 3 NUMERICAL EVALUATION

Here we describe the methods used to compute the kernel  $\hat{J}(\boldsymbol{\omega})$  for the second moment of the flux (equation 8). The first step is to compute and tabulate the function  $H(\boldsymbol{\beta})$  (Section 3.1). It is quite easy to incorporate a distribution of stellar masses at this stage (Section 3.2). The final step is to carry out the integration over  $\Delta\boldsymbol{\theta}$  in equation (8), which is really a double integral since  $\Delta\boldsymbol{\theta}$  is a two-component vector (Section 3.3). As a check on the numerics, we derive asymptotic results for  $\omega_{\theta_E} \ll 1$  and  $\omega_{\theta_E} \gg 1$  and interpret these results physically (Section 3.4).

#### 3.1 The function $H$

Since  $u$  and  $v$  in equation (12b) are actually scaled versions of the ‘force’ (10), the singularity of the integrand at the origin represents the large area available to distant stars that exert weak deflections. Without the subtracted terms in the numerator, equation (12b) would be a two-dimensional Fourier transform. But without at least the  $-1$ , the singularity at the origin would not be integrable. We considered softening the singularities so that we could separate the three terms

in the numerator and evaluate the first by two-dimensional fast Fourier transforms. But instead, like Seitz & Schneider (1994), we chose to expand  $H(\beta, \psi)$  as a Fourier series in  $\psi$ :

$$\begin{aligned} H(\beta, \psi) &= \frac{1}{2} \pi \int_0^\infty \frac{dr}{r^2} [J_0(\beta r) - 1] b_{1/2}^{(0)}(r) \\ &\quad + i\pi (\cos \psi) \int_0^\infty \frac{dr}{r^2} \left[ J_1(\beta r) - \frac{1}{2} \beta r \right] b_{1/2}^{(1)}(r) \\ &\quad + \pi \sum_{n=2}^\infty i^n \cos(n\psi) \int_0^\infty \frac{dr}{r^2} J_n(\beta r) b_{1/2}^{(n)}(r) \\ &\equiv \sum_{n=0}^\infty i^n h_n(\beta) \cos(n\psi). \end{aligned} \quad (13)$$

This representation has two advantages: (i)  $H$  is dominated by the terms for  $n \leq 2$  at both large and small  $\beta$ ; (ii) convolution with a stellar mass spectrum, as described below, acts in the radial direction, i.e. along lines of constant polar angle  $\psi$ . The Bessel functions,  $J_n$ , result from expanding the exponential in (12b) as a Fourier series in  $\psi - \alpha$ , where  $\alpha = \tan^{-1}(v/u)$ . The Laplace coefficients (e.g. Brouwer & Clemence 1961),

$$\begin{aligned} b_{1/2}^{(n)}(r) &\equiv \frac{1}{\pi} \int_0^{2\pi} \frac{\cos n\phi d\phi}{\sqrt{1+r^2-2r\cos\phi}} \\ &= \frac{2}{\pi\sqrt{r}} Q_{n-1/2} \left( \frac{1+r^2}{2r} \right), \end{aligned} \quad (14)$$

result from expanding the second part of the denominator of equation (12b). Here  $Q_{n-1/2}$  are Legendre functions of the second kind and are calculated in our code by recursion on  $n$  starting from expressions for  $Q_{-1/2}$  and  $Q_{+1/2}$  in terms of complete elliptic integrals, or by hypergeometric series. The radial integrals are performed numerically in the complex  $r$  plane with due attention to the branch points at  $r = 1$ .

To facilitate convolution with a stellar mass function, it is convenient to tabulate the functions  $h_n(\beta)$  on a grid uniformly spaced in  $\ln \beta$ . We typically use  $\Delta \ln \beta = 0.01$  in the range  $10^{-3} \leq \beta \leq 10^2$ . The first three functions  $h_0, \dots, h_2$  dominate the series, though convergence with respect to  $n$  is slow at  $\beta \sim 1$ . We usually tabulate up to  $n_{\max} = 7$ , tapering the series to hasten convergence. These choices are input parameters to our code. The functions  $h_n(\beta)$  with  $n \geq 2$  oscillate in  $\beta$  with period  $\approx 2\pi$ ; this is due to logarithmic singularities of the Laplace coefficients at  $r = 1$ . In order to resolve these oscillations on a logarithmic grid, it is necessary that  $\beta_{\max} \Delta \ln \beta < \pi$ . The first few  $h_n(\beta)$  are shown in Fig. 1.

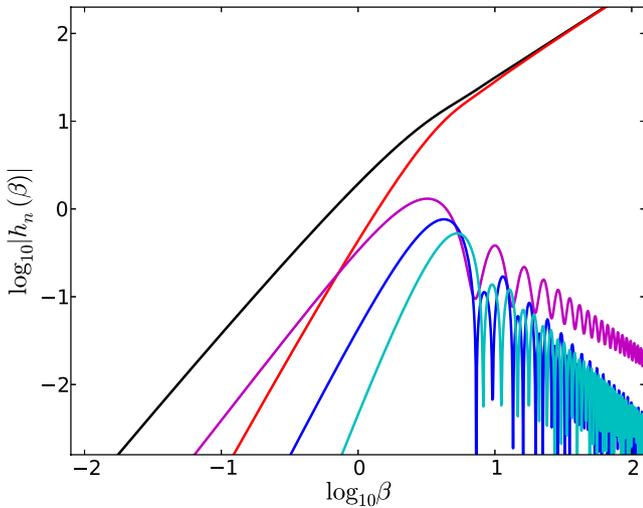
For extrapolation in  $\beta$  beyond the range of the table,

$$\begin{aligned} H(\beta, \psi) &\rightarrow \frac{\pi}{4} \beta^2 (\ln \beta + C_0) - \frac{\pi}{8} \beta^2 \cos 2\psi + O(\beta^3 \ln \beta), \\ &\quad \text{as } \beta \rightarrow 0; \\ H(\beta, \psi) &\rightarrow -\pi\beta - i\pi(\beta - 1) \cos \psi + O(\beta^{-1}), \\ &\quad \text{as } \beta \rightarrow \infty. \end{aligned} \quad (15)$$

The constant  $C_0 = \gamma - 1 - \ln 8 \approx -2.5022$ .

#### 3.2 Incorporating a stellar mass spectrum

The stellar mass ( $M_*$ ) enters the problem through the square of the Einstein-ring radius,  $\theta_E^2 \propto M_*$ , in equation (3) for the  $n$ -point characteristic function. This makes it easy to introduce a spectrum



**Figure 1.** The first few azimuthal harmonics  $h_n(\beta)$  of the function  $H(\beta)$  determining the two-point characteristic function (equations 12b and 13). From top to bottom at  $\beta = 1$ , these are  $n = 0, 1, 2, 3, 4$ .

of stellar masses. Let  $f(\log M_*) d \log M_*$  represent the fraction of the total number of stars that have masses in the logarithmic interval  $d \log M_*$ . (Here  $\log \equiv \log_{10}$ .) The mean mass is then

$$\bar{M}_* = \int_{-\infty}^{\infty} M_* f(\log M_*) d \log M_*.$$

Let  $\theta_E$  be the Einstein-ring radius based on the mean mass,  $\theta_E^2 \equiv 4G\bar{M}_*D_{LS}/c^2D_LD_S$ . To represent the mass spectrum  $f(\log M_*)$  in the two-point function  $H(\beta)$ , one needs to only make the replacement

$$h_n(\beta) \rightarrow \bar{h}_n(\beta) = \int_{-\infty}^{\infty} h_n(10^x \beta) f(x) dx \quad (16)$$

in the polar expansion (13). The asymptotics (15) need to be adjusted accordingly.

In fact, provided  $f(x)$  itself is reasonably smooth, integrating over a mass spectrum would accelerate the convergence of the radial integrals in (13), because  $\int J_n(10^x \beta r) f(x) dx$  decays exponentially rather than oscillates when  $\log(\beta r)$  is larger than the width of  $f(x)$ . This would allow us to integrate entirely on the real  $r$  axis. However, we choose to tabulate the  $h_n$  for a single mass. The smoothing (16) is performed very easily after the fact with any desired mass function. We adopt the log-normal form

$$f(\log M_*) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(\log M - \log M_c)^2/2\sigma^2]. \quad (17)$$

The replacement (16) should leave the mean mass unchanged; this requires that the characteristic mass  $M_c = \bar{M}_* \exp[-(\sigma \ln 10)^2/2]$ . In our code, by default,  $\sigma = 0.3$  in agreement with the initial mass function recommended by Chabrier (2003) for the spheroidal component of the Galaxy [ $\sigma = 0.33 \pm 0.03$ ,  $M_c = 0.22 \pm 0.05 M_\odot$ ], except that Chabrier replaces the tail of the log-normal function above  $0.7 M_\odot$  with a power law of roughly Salpeter slope ( $x = 1.3$ ). The assumption is that since the spheroid is an old population, most of the stars in the tail will have evolved off the main sequence, effectively truncating the present-day mass function at the turnoff ( $0.7 M_\odot$ ). Our code omits this refinement and uses the log-normal form without truncation.

### 3.3 The outer integral

Given the function  $H(\beta)$ , it remains to compute

$$\hat{J}(\omega) = |\det \mathbf{M}| \int d^2 \Delta \theta e^{-i\Delta \theta \cdot \mathbf{M} \cdot \omega} \exp \left[ \frac{\kappa_* \Delta \theta^2}{4\pi \theta_E^2} H \left( \frac{4\theta_E^2 \omega}{\Delta \theta} \right) \right]. \quad (18)$$

As written, however, this double integral is not convergent for small  $\omega$  and large  $\Delta \theta$ . Following (15),  $H(\beta) \propto \beta^2 \ln \beta$  as  $\beta \rightarrow 0$ , whence the second exponential above tends to unity as  $\omega \theta_E \rightarrow 0$ . The rest of the integrand oscillates with constant modulus, whence  $\hat{J}(\omega) \rightarrow (2\pi)^2 \delta^2(\omega)$  as  $\omega \rightarrow 0$ .

This is to be expected. A very extended source will have a Fourier transform  $\hat{I}_0(\omega)$  that decreases rapidly with increasing  $|\omega|$ . Such a source should suffer little microlensing, so that  $\langle F^2 \rangle = \langle F \rangle^2$ . But the mean macrolensed flux is  $\langle F \rangle = |\det \mathbf{M}|^{-1} \hat{I}_0(\mathbf{0})$ . Comparison with equation (8) shows that the microlensing kernel must contain exactly the delta function identified in the previous paragraph.

For numerical purposes, one must have a convergent expression. One option is to subtract unity from the second exponential in equation (18). This removes the problem at small  $\omega$  but creates a similar problem at large  $\omega$ . A better tactic is to multiply the integrand by a broad and smooth window function that gradually tapers to zero at large  $\Delta \theta$ ; this smears the delta function into a narrow but finite spike centred at  $\omega = 0$  without much changing the finite part of the integral at  $\omega > 0$ . Instead, we have chosen to evaluate equation (18) by a version of Euler summation, which works as follows (e.g. Hardy 1949). Let  $\sum_k (-1)^k a_k$  be an alternating series in which the terms  $\{a_k\}$  have constant sign but may increase, provided  $\lim_{k \rightarrow \infty} (a_k + 1/a_k) = 1$ . Let  $S_n^{(0)} = \sum_{k=0}^{n-1} (-1)^k a_k$  be the  $n$ th partial sum, and for  $m > 0$

$$S_n^{(m)} = \frac{1}{2} \left[ S_n^{(m-1)} + S_{n+1}^{(m-1)} \right] = 2^{-m} \sum_{j=0}^m \binom{m}{j} S_{n+j}^{(0)}.$$

If the sequence  $\{S_n^{(0)}\}$  converges, then  $\{S_n^{(m)}\}$  converges to the same limit. But  $\{S_n^{(m)}\}$  may converge when  $\{S_n^{(0)}\}$  does not. For example, if  $a_k = k$ , then  $\{S_n^{(m)}\} \rightarrow -1/4$  for  $m \geq 2$ , which is the ‘correct’ result if this series is regarded as the limit of  $\sum_k k(-x)^k = x(d/dx)(1-x)^{-1}$  as  $x \rightarrow 1^-$ .

To apply this, we set  $(-1)^k a_k$  equal to the integral (18) restricted to the annulus

$$k \frac{\pi}{\mu} \leq |\Delta \theta| \leq (k+1) \frac{\pi}{\mu}, \quad \text{where } \mu \equiv |\mathbf{M} \cdot \omega|.$$

For sufficiently large  $k$ ,

$$(-1)^k a_k \sim 2\pi |\det \mathbf{M}| \int_{k\pi/\mu}^{(k+1)\pi/\mu} d\Delta \theta J_0(\mu \Delta \theta) (\Delta \theta)^{1-\kappa_* \theta_E^2 \omega^2}.$$

The Bessel function  $J_0(z) \sim \sqrt{2/\pi z} \cos(z - \pi/4)$  for  $z \gg 1$ . Hence,  $a_k \sim k^\sigma$  with  $\sigma = 1/2 - \kappa_* \theta_E^2 \omega^2$ . After inspecting the smoothed partial sums  $S_n^{(m)}$  for residual oscillations, our code adds more annuli as needed to enable further smoothing. This works reasonably well and automatically discards the delta function at  $\omega = 0$ .

### 3.4 Limiting behaviours of the kernel and their interpretation

We can check the numerical results against analytical ones for  $\omega \theta_E \ll 1$  and for  $\omega \theta_E \gg 1$ . Let  $\omega$  be small enough so that it makes sense to replace the second exponential in (18) by the first two terms of its power series. Discard the leading term (unity), which gives the delta function. Also ignore the small contribution from the range

$0 \leq \Delta\theta \lesssim 4\theta_E^2\omega$ , so that  $H(\beta)$  may be replaced by the top line of equation (15). After integration over the azimuth of  $\Delta\theta$ ,

$$\hat{J}(\omega) \approx 2\pi\kappa_*(\omega\theta_E)^2 |\det \mathbf{M}| \int_0^\infty d\Delta\theta \times \Delta\theta \left[ \ln \left( \frac{4\theta_E^2\omega}{\Delta\theta} \right) J_0(\mu\Delta\theta) + \frac{1}{2} J_2(\mu\Delta\theta) \cos 2\alpha \right], \quad (19)$$

where  $\alpha$  is the angle between  $\omega$  and  $\mu \equiv \mathbf{M} \cdot \omega$ . The standard integral (Abramowitz & Stegun 1970, section 11.4.16)

$$\int_0^\infty t^\mu J_\nu(t) dt = 2^\mu \frac{\Gamma\left(\frac{\nu+\mu+1}{2}\right)}{\Gamma\left(\frac{\nu-\mu+1}{2}\right)} \quad (20)$$

converges only if  $\mu + \nu > -1$  and  $\mu < \frac{1}{2}$  but can be analytically continued via the right-hand side where that is finite and non-zero. Logarithms can be inserted by differentiation with respect to  $\mu$ . In particular,

$$\int_0^\infty dt J_0(t) t \ln t \rightarrow -1, \quad \int_0^\infty dt J_2(t) t \rightarrow +2.$$

Applying these rules to equation (19) yields

$$\hat{J}(\omega) \approx 4\pi\kappa_*\theta_E^2 |\det \mathbf{M}| \frac{(\omega \cdot \mathbf{M} \cdot \omega)^2}{|\mathbf{M} \cdot \omega|^4}, \quad 0 < \omega\theta_E \ll 1. \quad (21a)$$

On the other hand, when  $\omega\theta_E \gg 1$ , we may approximate  $H$  by the second line of equation (15), with the result

$$\hat{J}(\omega) \approx 2\pi |\det \mathbf{M}| \frac{\kappa_*\omega}{(\kappa_*^2\omega^2 + |\mathbf{M}' \cdot \omega|^2)^{3/2}} \quad \text{if } \omega\theta_E \gg 1. \quad (21b)$$

Here  $\mathbf{M}' \equiv \mathbf{M} + \kappa_*\mathbf{1}$ , i.e. the residual of the macrolensing matrix when the mean stellar convergence is removed.

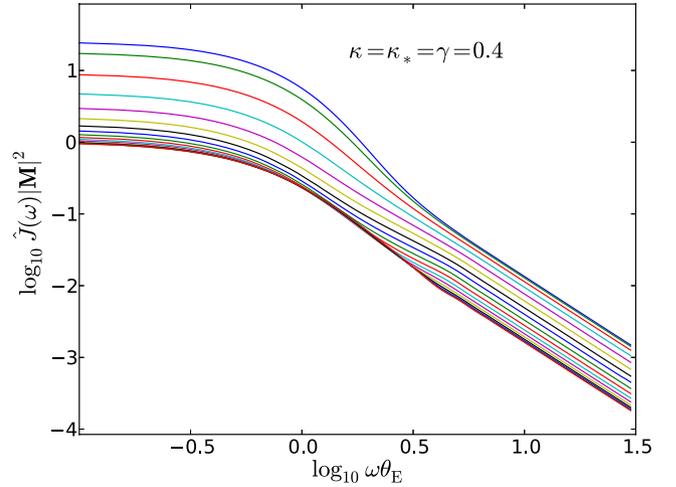
For a spectrum of stellar masses, equation (21b) is unchanged, but equation (21a) is multiplied by  $\bar{M}_*^2/M_*^2$  if  $\theta_E^2$  is based on the mean mass  $\bar{M}_*$ . For a log-normal mass function with dispersion  $\sigma$  in  $\log M_*$ , the correction factor is  $\exp[(\sigma \ln 10)^2]$ .

### 3.4.1 Physical interpretation

The Fourier transform  $\hat{I}_0(\omega)$  of a compact source has significant power at large values of its argument,  $\omega$ . Therefore, the variance of the microlensed flux from such a source should be dominated by the asymptotic behaviour (21b), which scales with  $|\omega|$  as  $\omega^{-2}$ . Suppose that the spatial power spectrum  $|\hat{I}_0(\omega)|^2 \sim \text{constant} \equiv C$  for  $\omega < \Theta_S^{-1} \equiv \omega_S$ , where  $\Theta_S$  is the angular size of the source, and that the power spectrum falls off rapidly for  $\omega > \omega_S$ . Then if  $\omega_S \gg \theta_E^{-1}$ , it follows from equation (21b) that  $\text{var}(F)/\langle F \rangle^2 \propto C \log \omega_S$ , with a constant of proportionality that depends upon  $\kappa_*$ ,  $\det \mathbf{M}$ , and the angular dependence of  $|\mathbf{M}' \cdot \omega|$ . Hence, the flux variance diverges logarithmically in the limit that the angular size of the source tends to zero, as expected.

In the opposite limit of an extended source,  $\Theta_S \gg \theta_E$ , similar reasoning based on the small- $\omega$  asymptotics (21a) leads to  $\text{var}(F)/F_0^2 \propto \omega_S^2 \propto \Theta_S^{-2}$ . Why should sources larger than the Einstein ring of an individual star show any microlensing fluctuations at all? The answer is that the stellar convergence, whose average value we denote by  $\kappa_*$ , is subject to Poisson fluctuations in the number of stars projected on to the source:

$$\frac{\Delta\kappa_*}{\kappa_*} \sim \frac{\Delta N_*}{N_*} \sim N_*^{-1/2} \sim (\pi\nu\Theta_S^2)^{-1/2}.$$



**Figure 2.** The numerically evaluated Fourier-transformed microlensing kernel  $\hat{J}(\omega)$  for the macrolensing parameters of Image A in Q2237+0305 as given by Poindexter & Kochanek (2010). Constant along each curve is the angle ( $\alpha$ ) between the spatial wavevector  $\omega$  and the major axis of the macrolensing matrix  $\mathbf{M}$  (equation 1), i.e. the eigenvector corresponding to the smaller eigenvalue,  $1 - \kappa - \gamma$ . These values are uniformly spaced from  $\alpha = 0$  (top curve) to  $\alpha = \pi/2$  (bottom).

Small fluctuations in the convergence translate linearly to fluctuations in magnification, whose variance (normalized by the mean magnification) therefore scales as  $\Theta_S^{-2}$ .

## 4 NUMERICAL EXAMPLES AND APPLICATIONS

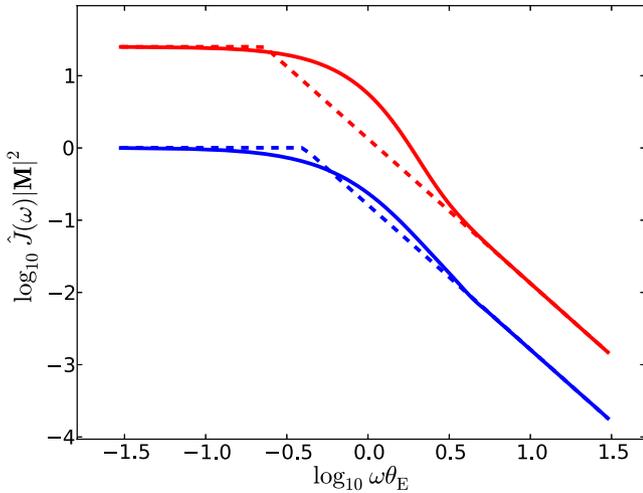
The poster child of microlensed quasars is Q2237+305, variously known as Huchra's lens after its discoverer, or the Einstein Cross. For the brightest of the four visible macroimages, Poindexter & Kochanek (2010) estimate that  $\kappa$ ,  $\kappa_*$ , and  $\gamma$  are all close to 0.4. Hence, the stars dominate the convergence, and the mean magnification of this image is  $[(1 - \kappa - \gamma)(1 - \kappa + \gamma)]^{-1} \approx 5$ .

Fig. 2 shows the microlensing kernel numerically computed for these parameters. The anisotropy of the kernel – its dependence upon the angle  $\alpha$  describing the direction of  $\omega$  – is due to the anisotropy of the macrolensing matrix,  $\mathbf{M}$ . The symmetries of the kernel are such that it suffices to calculate  $\hat{J}(\omega)$  for  $0 \leq \alpha \leq \pi/2$ , where  $\alpha$  is measured with respect to the eigenvector of  $\mathbf{M}$  corresponding to the eigenvalue of smaller absolute value.

The logarithmic axes influence the visibility of some details. There are small-amplitude wiggles in the curves in the decade  $1 < \omega\theta_E < 10$  at the level of tens of per cent, somewhat more pronounced for smaller magnifications. These are vestiges of much stronger oscillations in the azimuthal harmonics of the  $H$  function (Fig. 1). The curves in Fig. 2 for different  $\alpha$  are farther apart near  $\omega = 0$  than as  $\omega \rightarrow \infty$ . This shows that  $\hat{J}(\omega) = \hat{J}(\omega \cos \alpha, \omega \sin \alpha)$  is not separable in  $\omega$  and  $\alpha$ . Although we do not show it here, the inseparability is yet more striking for the more typical case that  $\kappa = \gamma = 0.45 = 10\kappa_*$ . The limiting behaviors of the kernel obtained by analytic means in Section 3.4 are compared with the numerical results for a representative case in Fig. 3.

### 4.1 Inhomogeneous discs

Dexter & Agol (2011, hereafter DA) have proposed toy models of inhomogeneous discs, in which the disc temperature fluctuates



**Figure 3.** Solid lines show the extreme values  $\alpha = 0$  (upper curve, in red) and  $\alpha = \pi/2$  (lower, blue) from Fig. 2. Broken dashed lines show the asymptotic predictions (21).

around its steady-state value. These models are an interesting test of our formalism. They have more structure than simple Gaussian source models, yet have spatial power spectra that are rather easily described. The fluctuations are spatially correlated within cells whose width is constant in azimuth and log radius. Each cell radiates as a blackbody at a temperature chosen from a log-normal distribution ( $\log \equiv \log_{10}$ ):

$$\exp \left\{ \left[ \log T_r - \log T_{r,0} + \sigma_T^2 \right]^2 / \sigma_T^2 \right\} \frac{d \log T_r}{\sqrt{\pi} \sigma_T}. \quad (22)$$

Here  $T_{0,r} \propto (r/r_s)^{-3/4}$  is the temperature profile of a homogeneous disc with a constant mass accretion rate, neglecting the inner and outer edges. The fiducial radius is defined so that  $T_{0,r} = h/k_b \lambda$  at  $r = r_s$  if  $\lambda$  is the wavelength of observation referred to the disc rest frame. Following equation (22), the variance of  $\ln T_r$  is  $(\sigma_T \ln 10)^2/2$ , and the mean of  $\ln T_r$  is less than  $\ln T_{r,0}$ . The bolometric flux is the same on average as for the homogeneous disc:  $\langle T_r^4 \rangle = T_{0,r}^4$ . The narrow-band luminosity, however, decreases with increasing temperature variance:  $L_\lambda(\sigma_T) = L_\lambda(0) \exp[-\frac{8}{9}(\sigma_T \ln 10)^2]$ . At the same time, the apparent size of the source increases: the  $n$ th radial moment of the light scales  $\propto L_\lambda r_s^n \exp[\frac{4}{9}n(n+1)(\sigma_T \ln 10)^2]$ .

The fiducial radius scales with the black hole mass and accretion rate as  $r_s \propto (M\dot{M})^{1/3}$ . However, without reference to  $M$  or  $\dot{M}$ , one can infer the half-light radius  $r_h(0)$  of a homogeneous disc from the observed narrow-band luminosity, corrected for lensing:  $r_h(0) \propto L_\lambda^{1/2} \lambda^{3/2}$ . The constant of proportionality depends only on fundamental constants if a point on the disc radiates as a blackbody with temperature scaling as  $r^{-3/4}$ .

Based on the temporal variability and optical-to-UV spectra of QSOs, as well as the microlensing observations, DA conclude that  $0.35 \lesssim \sigma_T \lesssim 0.5$ . Since, as they remark,  $r_h(\sigma_T)/r_h(0) \propto \exp[0.85(\sigma_T \ln 10)^2]$ , it follows that the half-light radii are larger than those of homogeneous discs by factors of 1.7 to 3 when referred to the same  $L_\lambda$ .

To apply our methods, we must estimate the spatial power spectra of these models. We take the disc to be viewed face-on and adopt angular coordinates  $\theta \equiv r/D_S$  in the source plane; in particular,  $\theta_s \equiv r_s/D_S$ . Polar coordinates  $(\theta, \phi)$  are defined so that  $\theta = (\theta \cos \phi, \theta \sin \phi)$ . We numerically average the Planck function at each radius in the disc over the temperature distribution (22), thus

obtaining the mean source  $\langle I(\theta) \rangle_S$ . This is axisymmetric, as is its Fourier/Hankel transform,

$$\langle \hat{I}(\omega) \rangle_S = 2\pi \int_0^\infty J_0(\omega\theta) \langle I(\theta) \rangle_S \theta d\theta. \quad (23)$$

The subscript  $S$  serves as a reminder that the average is taken over realizations of the source, not over the microlensing. The spatial correlation of the temperature fluctuations – the size of the cells – plays no role in equation (23), because the average  $\langle I(\theta) \rangle_S$  of the surface brightness is computed independently at each point on the disc.

Equation (7) for the microlensing flux variance involves  $|\hat{I}(\omega)|^2$ . Since this is a random variable in DA's inhomogeneous models, we must average it over realizations of the source:

$$\langle |\hat{I}(\omega)|^2 \rangle_S = \langle |\langle \hat{I}(\omega) \rangle_S|^2 + \text{var}_S[\hat{I}(\omega)] \rangle. \quad (24)$$

The last term is the Fourier transform of the two-point correlation of the surface brightness fluctuations,  $\langle \Delta I(\theta_1) \Delta I(\theta_2) \rangle$ , which vanishes unless the points  $\theta_1$  and  $\theta_2$  belong to the same cell. The correlation depends on  $\phi_1 - \phi_2$ ,  $\ln(\theta_1/\theta_2)$ , and  $\ln \bar{\theta} \equiv \frac{1}{2} \ln(\theta_1 \theta_2)$ , but varies more rapidly with the first two variables than the third if  $N_r$ , the number of cells per octave in radius, is large. Thus to an adequate approximation,

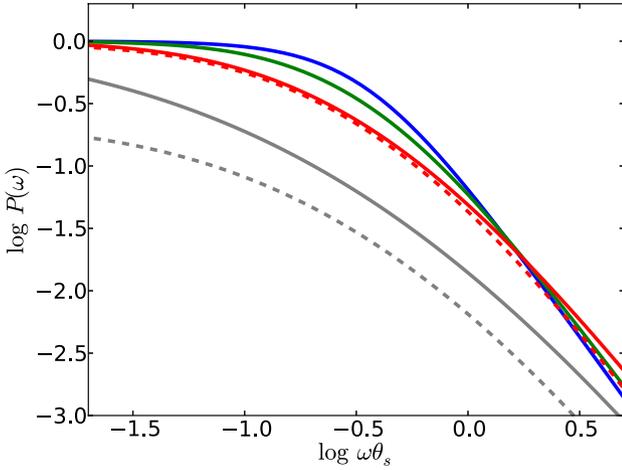
$$\langle \Delta I(\theta_1, \phi_1) \Delta I(\theta_2, \phi_2) \rangle_S \approx \text{var}[I(\theta_1)] W_\phi(\phi_1 - \phi_2) W_\theta(\ln \theta_1 - \ln \theta_2), \quad (25)$$

where  $W_\phi$  and  $W_\theta$  are triangular window functions of width  $2\pi/N_\phi$  and  $(\ln 2)/N_r$ , respectively,  $N_\phi$  being the number of cells in azimuth. Then<sup>1</sup>

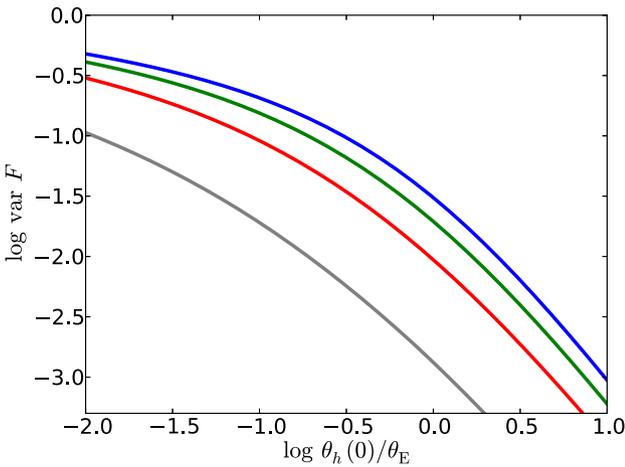
$$\begin{aligned} \text{var}_S[\hat{I}(\omega)] &\approx \frac{(2\pi)^2 \ln 2}{N_\theta N_\phi} \sum_{n=-N_\phi/2}^{N_\phi/2-1} \int_0^\infty \text{var}_S[I(\theta)] J_n^2(\omega\theta) \\ &\times \text{sinc}^2 \left( \frac{\omega\theta \ln 2}{2N_r} \right) \theta^3 d\theta. \end{aligned} \quad (26)$$

The strength of this term relative to the first term on the right-hand side of (24) increases with  $\sigma_T$  at fixed  $N_\theta N_\phi$ . DA state that they obtain the best match to the variability data if the number of cells per octave in radius is in the range  $100 \lesssim n \lesssim 1500$ . We interpret this  $n$  to correspond to  $N_\theta N_\phi$  in our notation. DA do not specify the aspect ratio of their cells. We presume that they are roughly square, although because of differential rotation, one might expect that the correlation length of disc inhomogeneities should be longer in azimuth than in radius. We take  $N_\theta = 10$ ,  $N_\phi = 55$ , so that  $\Delta\theta = 0.11$  rad and  $\Delta \ln r = 0.069$ . We normalize the power spectra to unity at  $\theta = 0$ :  $P(\omega) = \langle |\hat{I}(\theta)|^2 \rangle_S / \langle |\hat{I}(0)|^2 \rangle_S$ . This is not the same as normalizing by the square of the mean flux,  $\langle F \rangle_S = \langle \hat{I}(0) \rangle_S$ . When microlensing is measured via flux ratios of macroimages, after correction for time delays, intrinsic variations of the source flux cancel. Normalizing the power spectrum to unity at  $\theta = 0$  is therefore more appropriate than normalizing by the average flux. In practice, it does not much matter which normalization one uses for DA's preferred parameter range. The ratio  $\langle \hat{I}(0)^2 \rangle_S / \langle \hat{I}(0) \rangle_S^2$  increases with  $\sigma_T$  and decreases with  $N_\theta N_\phi$ . It is less than 1.04 for

<sup>1</sup> The use of  $\text{sinc } x \equiv (\sin x)/x$  is a further approximation but is accurate when  $\omega\theta \gg \max(|n|, 1)$ ; in the opposite limit, the integrand is negligible anyway. For numerical quadratures, we replace  $\text{sinc}^2(x/2)$  with  $\text{sech}^2(x/\pi)$ , which has the same equivalent width and the same effect of suppressing the integrand where  $\omega$  is larger than the reciprocal of the local cell size, but avoids the sidelobes.



**Figure 4.** Spatial power spectra of inhomogeneous discs with log-normal temperature variations, following DA. Power spectra are normalized to unity at wavenumber  $\omega = 0$ . Abscissa is scaled by the angular radius  $\theta_s$  at which  $hc/k_b T_{\text{eff}} = \lambda_{\text{rest}}$  in the homogeneous ( $\sigma_T = 0$ ) disc. From top to bottom at  $\log \omega \theta_s \approx -0.5$ , solid curves correspond to  $\sigma_T = \{0, 0.3, 0.5, 0.7\}$ . Dashed curves show  $\langle \hat{I}(\omega) \rangle_S^2 / \langle \hat{I}(0) \rangle_S^2$  rather than  $P(\omega) \equiv \langle \hat{I}^2(\omega) \rangle_S / \langle \hat{I}^2(0) \rangle_S$  for the most variable cases ( $\sigma_T = 0.5, 0.7$ ).

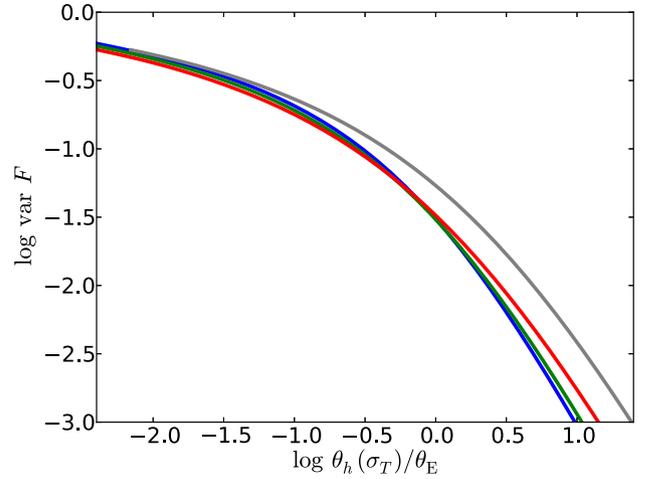


**Figure 5.** Microlensed flux variance of the inhomogeneous discs from Fig. 4 for the fiducial microlensing parameters  $\kappa = \gamma = 0.45$ ,  $\kappa_* = 0.045$  versus inferred half-light radii  $\theta_h(0)$  of homogeneous discs of the same  $L_\lambda$ . From top to bottom at  $\log \theta_h(0)/\theta_E \approx -1$ , the curves correspond to  $\sigma_T = 0, 0.3, 0.5, 0.7$ .

$\sigma_T \leq 0.5$  at  $N_\theta N_\phi = 550$ , but rises rapidly for larger temperature dispersions, reaching 4.53 at  $\sigma_T = 0.7$ .

Fig. 4 shows the source power spectra calculated as described for several values of  $\sigma_T$ . As expected, the general effect of increasing  $\sigma_T$  is to suppress the power spectrum at  $\omega \theta_s \lesssim 2$ , because the mean source size increases. However, at least for  $\sigma_T \lesssim 0.5$ , the power spectrum is actually enhanced at larger  $\omega$ , probably because the hottest individual cells, which are much smaller than  $\theta_s$ , increasingly dominate. Comparison of the solid and dashed curves shows that for  $\sigma_T \lesssim 0.5$ , the power spectrum is nearly equal to the square of the Fourier transform of the mean intensity, which is independent of the cell size. For  $\sigma_T \gtrsim 0.7$ , however, the cell-to-cell variance is more important, so that the second term in equation (24) cannot be neglected.

Fig. 5 shows the microlensing variance calculated by applying equation (7) to the power spectra shown in Fig. 4. On the abscissa,



**Figure 6.** Like Fig. 5, but plotted against the true half-light radius of the mean inhomogeneous source. From top to bottom at  $\log \theta_h(\sigma_T)/\theta_E \approx +1$ , the curves correspond to  $\sigma_T = 0.7, 0.5, 0.3, 0$ .

the size of the source is measured not by  $\theta_s$ , but rather by the half-light radius of a homogeneous disc of equivalent narrow-band luminosity:  $\theta_h(0) = 2.44 \exp[-(4/9)(\sigma_T \ln 10)^2] \theta_s$ . At fixed  $\theta_h(0)$ , larger  $\sigma_T$  makes for a larger true source size and therefore, as expected, smaller microlensing flux variations. To gauge the effect of the temperature fluctuations on the source size inferred from microlensing, it is more appropriate to consider the horizontal rather than vertical distance between curves, i.e. the dependence of  $\theta_h(0)$  on  $\sigma_T$  at fixed  $\text{var } F$ . For example, along the curve for  $\sigma_T = 0$ , variances of 0.05, 0.1, 0.2, and 0.4 are achieved for  $\theta_h(0)/\theta_E = 0.649, 0.303, 0.106$ , and 0.0191, whereas for  $\sigma_T = 0.5$ , these same variances require  $\theta_h(0)/\theta_E = 0.212, 0.0879, 0.0266$ , and 0.00421. The former sizes are larger than the latter by factors ranging from 3.1 to 4.5: these are the factors by which the size of a  $\sigma_T = 0.5$  source inferred from microlensing variability would exceed the size inferred by fitting a conventional disc to its narrow-band luminosity.

If one plots the microlensing variance against the *true* half-light radius of the corresponding mean source, then the curves lie almost on top of each other, at least for  $\theta_h \lesssim \theta_E$  (Fig. 6). This confirms the result from Mortonson et al. (2005) that microlensing variability is relatively insensitive to the details of the source structure at a given half-light radius. However, this rule of thumb breaks down for sufficiently wild sources, as the curve for  $\sigma_T = 0.7$  shows.

## 4.2 Comparison with simulations

We have tested the predicted flux variance against simulations carried out with an elementary inverse-ray-shooting code resembling that of Kochanek (2004) in its design. The source and lens plane are periodic in both  $x$  and  $y$  with periods  $(L_x, L_y)$  in the lens plane and  $(L'_x = |1 - \kappa - \gamma|, L'_y = |1 - \kappa + \gamma|)$  on the source plane. Stars with masses chosen from a log-normal distribution with  $\sigma_{\text{mass}} = 0.3 \text{ dex}$  are placed randomly on the lens plane with mean number density  $\kappa_*/\pi\theta_E^2$  and assigned to grid points via a cloud-in-cell technique. The deflections are computed by Fourier transforms using a particle-mesh (PM) method.<sup>2</sup> The numbers of cells along each dimension of lens domain are in the ratio  $N_x : N_y \approx \sqrt{L_x/L_y}$

<sup>2</sup> Kochanek (2004) used P<sup>3</sup>M to better resolve short-range forces. However, we typically have  $\lesssim 10^{-4}$  stars per grid point. Hence, the few rays that pass within a cell width of a star carry little light (all rays being weighted equally).

**Table 1.** Microlensed flux variances of Gaussian sources via simulations and the semi-analytic method. See the text in Section 4.2 for meaning of columns.

$\gamma$	$\kappa$	$\kappa_*$	$L_x$	$\Delta\theta'_x$	$\Delta\theta'_y$	$\sigma_S = 0.01$	0.04	0.16	0.64
<b>0.4</b>	<b>0.4</b>	<b>0.4</b>	–	–	–	<b>0.567</b>	<b>0.455</b>	<b>0.341</b>	<b>0.218</b>
			128	0.0125	0.0284	0.498(6)	0.395(6)	0.282(5)	0.155(4)
			256	0.025	0.0569	0.498(3)	0.410(3)	0.300(3)	0.172(3)
			256	0.0125	0.0284	0.515(4)	0.412(4)	0.300(4)	0.171(3)
			256	0.003 12	0.0070	0.528(8)	0.417(8)	0.303(8)	0.175(7)
			512	0.0125	0.0284	0.525(2)	0.422(2)	0.309(2)	0.179(2)
			1024	0.0125	0.0284	0.529(2)	0.426(2)	0.313(2)	0.182(2)
<b>0.45</b>	<b>0.45</b>	<b>0.045</b>	–	–	–	<b>0.529</b>	<b>0.360</b>	<b>0.197</b>	<b>0.069</b>
			128	0.006 25	0.0198	0.400(6)	0.252(4)	0.113(2)	0.025(1)
			256	0.006 25	0.0198	0.452(4)	0.294(3)	0.142(3)	0.036(1)
			512	0.0250	0.0791	0.455(3)	0.314(3)	0.160(2)	0.046(1)
			512	0.0125	0.0395	0.468(3)	0.315(2)	0.158(2)	0.044(1)
			512	0.006 25	0.0198	0.474(3)	0.312(2)	0.155(2)	0.043(1)
			512	0.003 12	0.0099	0.482(3)	0.316(3)	0.158(2)	0.045(2)
			1024	0.006 25	0.0198	0.479(3)	0.316(2)	0.159(2)	0.047(1)
<b>0.0</b>	<b>0.8</b>	<b>0.8</b>	–	–	–	<b>0.340</b>	<b>0.319</b>	<b>0.297</b>	<b>0.260</b>
			128	0.0187	0.0187	0.309(4)	0.286(4)	0.262(4)	0.214(4)
			256	0.0187	0.0187	0.328(4)	0.305(4)	0.281(4)	0.234(4)
			256	0.0093	0.0093	0.321(4)	0.300(4)	0.276(4)	0.228(4)
			256	0.0047	0.0047	0.316(9)	0.295(9)	0.272(9)	0.224(8)
			512	0.0187	0.0187	0.341(3)	0.318(3)	0.294(3)	0.245(3)
			1024	0.0187	0.0187	0.343(2)	0.320(2)	0.297(2)	0.248(2)
<b>0.55</b>	<b>0.55</b>	<b>0.055</b>	–	–	–	<b>2.518</b>	<b>1.608</b>	<b>0.753</b>	<b>0.169</b>
			256	0.006 25	0.0198	2.24(3)	1.37(2)	0.560(8)	0.093(3)
			256	0.001 56	0.0049	2.31(4)	1.39(3)	0.57(1)	0.096(4)
			1024	0.006 25	0.0198	2.33(3)	1.42(2)	0.586(7)	0.102(2)

on the theory that because of the macrolensing shear; the four microimages<sup>3</sup> split by an isolated on-axis star have separations along  $x$  and  $y$  that lie in this ratio. The rays strike a uniform mesh in the lens plane with 16 rays per grid cell, for a total of almost  $2 \times 10^9$  rays in our largest simulations (which have  $10\,796^2$  or  $16\,384 \times 7200$  pixels). It has been argued that a much larger number of rays per cell are needed for accurate results with the inverse-ray-shooting method (e.g. Mediavilla et al. 2006). However, that conclusion is reached on the basis of a pixel-by-pixel comparison with some analytic solution such as that for an isolated star. We smooth the source-plane magnification pattern with Gaussians before computing the flux variance, so it is the number of rays per smoothing length or the number per star that is relevant.

Table 1 compares the variances obtained from our simulations to the semi-analytic predictions for Gaussian sources<sup>4</sup> with dispersions ranging from  $\sigma_S = 0.01\theta_E$  to  $0.64\theta_E$  listed in the final four columns of the table. The simulations are grouped according to three combinations of lensing parameters ( $\gamma$ ,  $\kappa$ ,  $\kappa_*$ ), shown in the leading row of each group and highlighted in boldface, together with the variances predicted by the semi-analytic method for each source size. The following rows of each group describe simulations for the same ( $\gamma$ ,  $\kappa$ ,  $\kappa_*$ ) but different numerical parameters as given in columns 4–6.  $L_x$  (column 4) is the periodicity length in the  $x$  direction on the source plane, in units of  $\theta_E$ . The corresponding periodicity along  $y$

is  $L_y = L_x |1 - \kappa - \gamma| / |1 - \kappa + \gamma|$  so that  $L'_y = L'_x$ . Columns 5 and 6 list the pixel dimensions  $\Delta\theta'_x$  and  $\Delta\theta'_y$  on the source plane.<sup>5</sup> The final columns list the flux variances estimated from the simulations for four choices of source size ( $\theta_S$ ). For each row, 16 to 256 realizations are created with the same parameters but different random numbers. The flux variance is computed pixel by pixel and averaged over realizations, with statistical uncertainties in the final digit of the mean shown in parentheses.

The first set of lensing parameters ( $\gamma$ ,  $\kappa$ ,  $\kappa_*$ ) = (0.4, 0.4, 0.4) is believed appropriate for the brightest macroimage in Q2237+305 (Poindexter & Kochanek 2010). The second set, (0.45, 0.45, 0.045), is adopted as typical of the brighter image in a generic image pair by Jiménez-Vicente et al. (2012). The third set was suggested by our anonymous referee. Whereas the first three parameter sets describe macroimages that lie at minima of the time-delay surface, the fourth is a saddle, with the parameters typical of the second macroimage according to Jiménez-Vicente et al. (2012).

The Gaussian source sizes ( $\sigma_S$ ) chosen for the variance estimates represent a compromise between what is thought relevant to lens systems and what can easily be simulated on a present-day desktop computer, i.e. no more than  $\sim 10^8$  pixels and  $\sim 10^9$  rays per simulation. (Altogether, Table 1 required  $\sim 3 \times 10^{11}$  rays.) Mosquera & Kochanek (2011) find that most of the ‘flux sizes’ ( $R_S$ ) of the sources in a sample of 87 quasar systems lie between  $0.01R_E$  and  $0.1R_E$ , whereas the sizes inferred from microlensing are typically two to three times larger. The linear half-light radius (in a given narrow spectral band) of a thin disc is  $R_{1/2} \approx 2.44R_S$ , and the angular

These rays contribute even less to caustics because large deflections entail strong demagnification.

<sup>3</sup> The softened potentials of our simulated stars create an odd number of images, but the central image is demagnified.

<sup>4</sup> That is, the unlensed surface brightness is  $I_0(\theta) \propto \exp(-\theta^2/2\sigma_S^2)$ .

<sup>5</sup> From these relations, it is possible to deduce the numbers of cells in each dimension,  $N_x = L'_x/\Delta\theta'_x$  and  $N_y = L'_y/\Delta\theta'_y$ .

half-light radius of a Gaussian source is  $\theta_{1/2} = \sigma_S \sqrt{2 \ln 2} \approx 1.18 \sigma_S$ , so that  $\sigma_S / \theta_E \Leftrightarrow 2.07 R_S / R_E$  insofar as a Gaussian can stand in for a thin-disc source.

The agreement between the flux variances estimated from the simulations and the semi-analytic theory tends to improve both with resolution and with periodicity length, although the former is systematically lower than the latter by 10–20 per cent even in the largest simulations, except in the third group (where  $\gamma = 0$ ). The improvement with periodicity length suggests that long-range forces are important. In a periodic domain, the forces exerted by each star must also be periodic and hence cannot follow the correct scaling  $\phi \propto |\Delta\theta|^{-1}$  at separations  $\Delta\theta$  larger than about half the periodicity length. We have found that the variances are sensitive to the approximation chosen for the force kernel. The results shown here were obtained by constructing the force kernel in coordinate space and windowing the exact kernels for the  $x$  and  $y$  components of the deflection with  $\cos(\pi\Delta\theta_x/L_x)$  and  $\cos(\pi\Delta\theta_y/L_y)$ , respectively. More sophisticated simulation algorithms, such as the inverse-polygon method of Mediavilla et al. (2006), would likely improve the rate of convergence with resolution. It would be interesting to see whether they also help the convergence with domain size.

Evidently, the flux variance and its sensitivity to source size depend strongly on the shear and convergence. In particular, the second and fourth cases have the same mean magnification except for sign ( $\det \mathbf{M} = \pm 0.1$ ) and the same stellar fraction ( $\kappa_*/\kappa = 0.1$ ), but very different variances.

## 5 SUMMARY

Motivated in part by discrepancies between the angular sizes of QSOs inferred from microlensing and those expected from disc theory, we have developed a practical method for computing the variance of the microlensed fluxes of angularly extended sources of arbitrary structure. The method requires as inputs the spatial power spectrum of the unlensed source, the shear and convergence provided by the smooth mass distribution of the lensing galaxy, and the mean number density and mass function of the lensing stars. We have written and made available for download a small suite of codes that accept these inputs and calculate the microlensing variance. The mathematical formulation of the method is described above in sufficient detail to allow an interested user to understand the workings of our codes. Further improvements in efficiency and accuracy are doubtless possible. The semi-analytic method has been tested against a simple inverse-ray-shooting simulation code and compared with the published results of DA for their toy models of highly inhomogeneous discs.

Perhaps the most important lesson learned from these tests is that significant numerical effort – in terms of the numbers of pixels and simulated stars, as well as the number of independent trials – is necessary to obtain good agreement between the flux variance estimated from the simulations and that calculated semi-analytically by the method developed here. Doubtless a cleverer simulation method could get by with fewer rays, but the need to simulate a large region of the lens planes,  $\gtrsim 10^2 \theta_E$  on a side, seems inescapable, at least when the macrolensing magnification is large ( $\sim 10$ ), as it typically is in present lensing surveys.

## ACKNOWLEDGEMENTS

We thank Jason Dexter for help with understanding Dexter & Agol (2011).

## REFERENCES

- Abramowitz M., Stegun I. A., 1970, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. U.S. Government Printing Office, Washington, DC
- Anguita T., Schmidt R. W., Turner E. L., Wambsganss J., Webster R. L., Loomis K. A., Long D., McMillan R., 2008, *A&A*, 480, 327
- Armitage P. J., Natarajan P., 2002, *ApJ*, 567, L9
- Bate N. F., Floyd D. J. E., Webster R. L., Wyithe J. S. B., 2008, *MNRAS*, 391, 1955
- Blackburne J. A., Pooley D., Rappaport S., Schechter P. L., 2011, *ApJ*, 729, 34
- Brouwer D., Clemence G. M., 1961, *Methods of celestial mechanics*. Academic Press, New York
- Chabrier G., 2003, *PASP*, 115, 763
- Chandrasekhar S., 1943, *Rev. Mod. Phys.*, 15, 1
- Deguchi S., Watson W. D., 1987, *Phys. Rev. Lett.*, 59, 2814
- Dexter J., Agol E., 2011, *ApJ*, 727, L24 (DA)
- Eigenbrod A., Courbin F., Meylan G., Agol E., Anguita T., Schmidt R. W., Wambsganss J., 2008, *A&A*, 490, 933
- Floyd D. J. E., Bate N. F., Webster R. L., 2009, *MNRAS*, 398, 233
- Goodman J., Tan J. C., 2004, *ApJ*, 608, 108
- Hardy G. H., 1949, *Divergent Series*. Clarendon Press, Oxford
- Jaroszynski M., Wambsganss J., Paczynski B., 1992, *ApJ*, 396, L65
- Jiménez-Vicente J., Mediavilla E., Muñoz J. A., Kochanek C. S., 2012, *ApJ*, 751, 106
- Katz N., Balbus S., Paczynski B., 1986, *ApJ*, 306, 2
- Kochanek C. S., 2004, *ApJ*, 605, 58
- Kundic T., Wambsganss J., 1993, *ApJ*, 404, 455
- Lewis G. F., Irwin M. J., 1995, *MNRAS*, 276, 103
- Lewis G. F., Irwin M. J., 1996, *MNRAS*, 283, 225
- Lightman A. P., Eardley D. M., 1974, *ApJ*, 187, L1
- Mediavilla E., Muñoz J. A., Lopez P., Mediavilla T., Abajas C., Gonzalez-Morcillo C., Gil-Merino R., 2006, *ApJ*, 653, 942
- Mediavilla E. et al., 2009, *ApJ*, 706, 1451
- Miralda-Escudé J., 1991, *ApJ*, 370, 1
- Morgan C. W., Kochanek C. S., Morgan N. D., Falco E. E., 2010, *ApJ*, 712, 1129
- Mortonson M. J., Schechter P. L., Wambsganss J., 2005, *ApJ*, 628, 594
- Mosquera A. M., Kochanek C. S., 2011, *ApJ*, 738, 96
- Mosquera A. M., Muñoz J. A., Mediavilla E., Kochanek C. S., 2011, *ApJ*, 728, 145
- Mosquera A. M., Kochanek C. S., Chen B., Dai X., Blackburne J. A., Chartas G., 2013, *ApJ*, 769, 53
- Muñoz J. A., Mediavilla E., Kochanek C. S., Falco E. E., Mosquera A. M., 2011, *ApJ*, 742, 67
- Neindorf B., 2003, *A&A*, 404, 83
- Poindexter S., Kochanek C. S., 2010, *ApJ*, 712, 658
- Poindexter S., Morgan N., Kochanek C. S., 2008, *ApJ*, 673, 34
- Pooley D., Blackburne J. A., Rappaport S., Schechter P. L., 2007, *ApJ*, 661, 19
- Pringle J. E., 1996, *MNRAS*, 281, 357
- Rauch K. P., Blandford R. D., 1991, *ApJ*, 381, L39
- Refsdal S., Stabell R., 1997, *A&A*, 325, 877
- Seitz C., Schneider P., 1994, *A&A*, 288, 1
- Seitz C., Wambsganss J., Schneider P., 1994, *A&A*, 288, 19
- Syunyaev R. A., Shakura N. I., 1975, *Sov. Astron. Lett.*, 1, 158
- Wambsganss J., Paczynski B., Schneider P., 1990, *ApJ*, 358, L33
- Wyithe J. S. B., Webster R. L., Turner E. L., 2000, *MNRAS*, 312, 843
- Young P., 1981, *ApJ*, 244, 756

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.