

Modelling microlensing events

Martin Dominik^{*†}

*SUPA, University of St Andrews – School of Physics & Astronomy, North Haugh, St Andrews,
KY16 9SS, United Kingdom*

E-mail: md35@st-andrews.ac.uk

Modelling microlensing events is a complex job extending far beyond the optimization of a function of adopted model parameters by means of standard techniques. Non-linearity of the underlying models leads to parameter ambiguities and degeneracies, and already the most simple case of a single point-like observed source star and a single point-mass lens as well as a static observer is substantially less trivial than one might think, while a proper treatment of the various microlensing anomalies requires tailored approaches to ensure that all regions in a high-dimensional intricate model parameter space are identified that could provide a suitable description of the nature of the observed event. Catastrophic behaviour leading to caustics results in the need to optimize functions that are not smooth, which causes a significant problem to standard algorithms. As if this were not enough, the photometric measurements frequently show complex statistics, varying amongst different sites, and involving outliers, an increased abundance of larger deviations as compared to a Gaussian distribution, as well as effects depending on the observing conditions. It requires advanced statistical methods and numerical strategies to account for the encountered difficulties. In order to be able to identify deviations from an ordinary light curve of ongoing microlensing events, a real-time modelling is required, and for realizing an optimal monitoring strategy, accurate predictions of model parameters are a valuable ingredient. The bias of maximum-likelihood estimates however frequently leads to substantial mispredictions, whereas maximum a-posteriori estimates with an appropriate prior are doing a better job. Multivariate parameter probability densities of the model parameters for observed events contain all the relevant information from which any interval estimates can be derived. Moreover, by means of Bayes' theorem, these can be converted to yield the probabilistic distributions of any physical properties with the assumption of respective priors. The identification of characteristic features in the observed light curve provides the clue to a successful modelling, and tailored artificial neural networks might replace human intuition, eventually leading to fully-automated real-time modelling systems.

The Manchester Microlensing Conference: The 12th International Conference and ANGLES Microlensing Workshop

January 21-25 2008

Manchester, UK

^{*}Speaker.

[†]Royal Society University Research Fellow

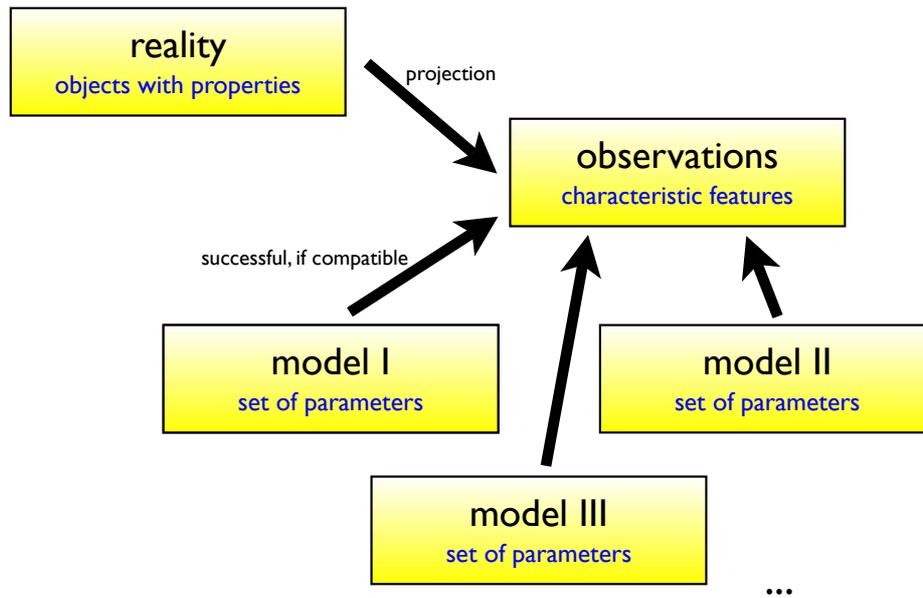


Figure 1: The relations between reality, observations, and models.

1. Introduction

Covering every aspect of modelling microlensing events would have required a 2-day workshop on its own (at least), and the confinement of this topic to a single lecture therefore calls for severe compromises. Presenting a little bit of everything would have meant to discuss nothing to an extent that allows to gain a satisfactorily level of understanding. Therefore, the numerical techniques applied to evaluate the model function, to search parameter space, and finally to converge at proper estimates of model parameters, are deliberately skipped. Moreover, optimal parametrizations of all the diverse microlensing anomalies with their manifold degeneracies and ambiguities have also escaped consideration, and only references are being provided. Instead, the lecture concentrates on the most fundamental concepts, illustrated by some examples, and in particular stresses various points that are typically not found in the literature, or are even frequently overlooked, despite the fact that many of these are essential for drawing the proper conclusions from the acquired data.

With the vast amount of data being collected right now, and the data acquisition rate increasing over the next few years, efforts need to be invested into the implementation of a comprehensive automated modelling strategy, whereas with thousands of events per year, an analysis of the data based on substantial human intervention is not feasible.

2. Data and modelling

2.1 Reality, observations, and models

The limitations of human mind do not allow us to determine the reality of the world. To the best of our abilities, we can only try to fit it into our imagination. It should therefore be distinguished

what the true nature of the objects under study is, what is being revealed about them, and how we describe them. Any observations are only a projection of reality expressed in measurable features.

Modelling now means to try to match the measurements with a set of model parameters. A model is deemed successful, if it is compatible with the observations. While models are never unique, some of them may be equivalent, but in general there can be non-equivalent competing models. The chosen model parameters are a representation of what we imagine to be properties of real-world objects, but our picture will always remain incomplete, given that the observations themselves are only a projection. Nevertheless, the general aim of science is to move towards a larger degree of completion, while reducing the number of parameters, thereby increasing the predictability of the theory.

2.2 Measurements and their uncertainties

Microlensing observations provide us with a set of data points $(t_i^{(k)}, F_i^{(k)}, \sigma_{F_i}^{(k)})$, consisting of the times of observation $t_i^{(k)}$, the measured fluxes $F_i^{(k)}$, and their uncertainties $\sigma_{F_i}^{(k)}$ for each of the observing sites and filters, denoted by the multi-index k , with n_k measurements, respectively.

Drawing the right conclusions from small deviations in the light curves requires a proper understanding of the distribution of the measurement uncertainties [1–4].¹ It is a quite common assumption that the reported error bars can be interpreted as the standard deviation of a Gaussian distribution

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (2.1)$$

of the measured value around the true value, i.e. all

$$x_i^{(k)} = \frac{F^{(k)}(t_i^{(k)}) - F_i^{(k)}}{\sigma_{F_i}^{(k)}} \quad (2.2)$$

are distributed as $f(x_i^{(k)})$.

In order to test this hypothesis, let us consider a recent event from the OGLE (Optical Gravitational Lensing Experiment) sample, which comprises calibrated data on thousands of events [5, 6], namely OGLE-2007-BLG-185, of which the data together with a model light curve are shown in Fig. 2. Rather than the measured fluxes, the OGLE team report the corresponding magnitudes m_i along with their uncertainties σ_{m_i} .

One of the popular standard procedures is to perform a Kolmogorov-Smirnow test of the residuals against a Gaussian distribution. Let us get around any dependencies on the magnitude, and just consider the data at event baseline, where the baseline magnitude m_{base} simply follows as the weighted average

$$m_{\text{base}} = \sum_{i=1}^n \frac{m_i}{\sigma_{m_i}^2} / \sum_{i=1}^n \frac{1}{\sigma_{m_i}^2} \quad (2.3)$$

over the n acquired data points, so that the residuals read

$$x_i = \frac{m_{\text{base}} - m_i}{\sigma_{m_i}}. \quad (2.4)$$

¹Deviations that by far exceed the measurement uncertainties can however be discussed without detailed knowledge of the measurement statistics, given that the principal behaviour is obvious.

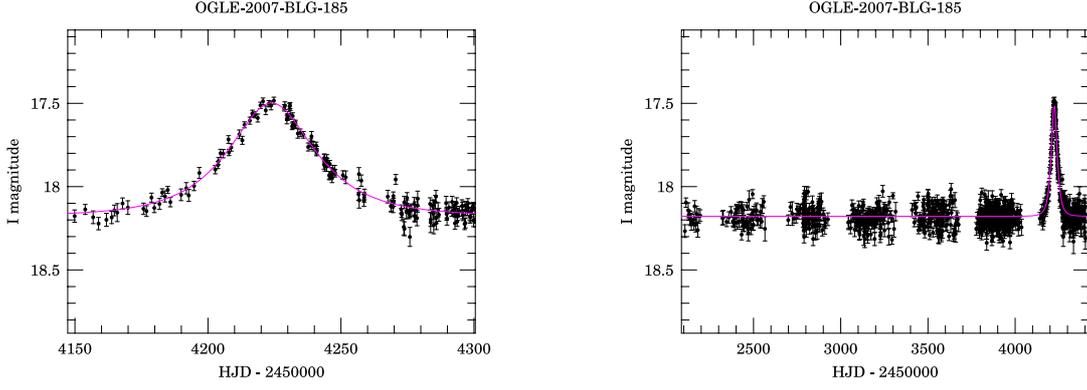


Figure 2: Data collected by the OGLE (Optical Gravitational Lensing Experiment) collaboration on the event OGLE-2007-BLG-185, illustrating the scatter around the model light curve [6]. – with kind permission from Andrzej Udalski for the OGLE team.

α	K_α
20 %	1.07
15 %	1.14
10 %	1.22
5 %	1.36
2 %	1.52

Table 1: Critical values K_α of the Kolmogorow-Smirnow test statistic that correspond to selected significance levels α .

With the distribution functions

$$\Phi_n(x) = \frac{1}{n} \sum_{i=1}^n \Theta(x - x_i) \quad (2.5)$$

for the residuals x_i , and

$$\Phi(x) = \int_{-\infty}^x f(x') dx' \quad (2.6)$$

for the Gaussian distribution $f(x)$, one finds the test statistic

$$K = \sqrt{n} \sup_x |\Phi_n(x) - \Phi(x)|, \quad (2.7)$$

which depends on the maximal absolute deviation and the number of data points. The stochastic distribution of the Kolmogorow-Smirnow test statistic K then defines a critical value K_α for any chosen significance level α , some of which are displayed in Table 1 for selected α .

If one adopts the reported error bars for OGLE-2007-BLG-185 for the residuals, a pretty good mismatch of their distribution with a Gaussian is found, as illustrated in Fig. 3 on the upper left. Most prominently, the true scatter turns out to be substantially larger than the reported uncertainties indicate. Therefore, one needs to account for more sophisticated error bar models.

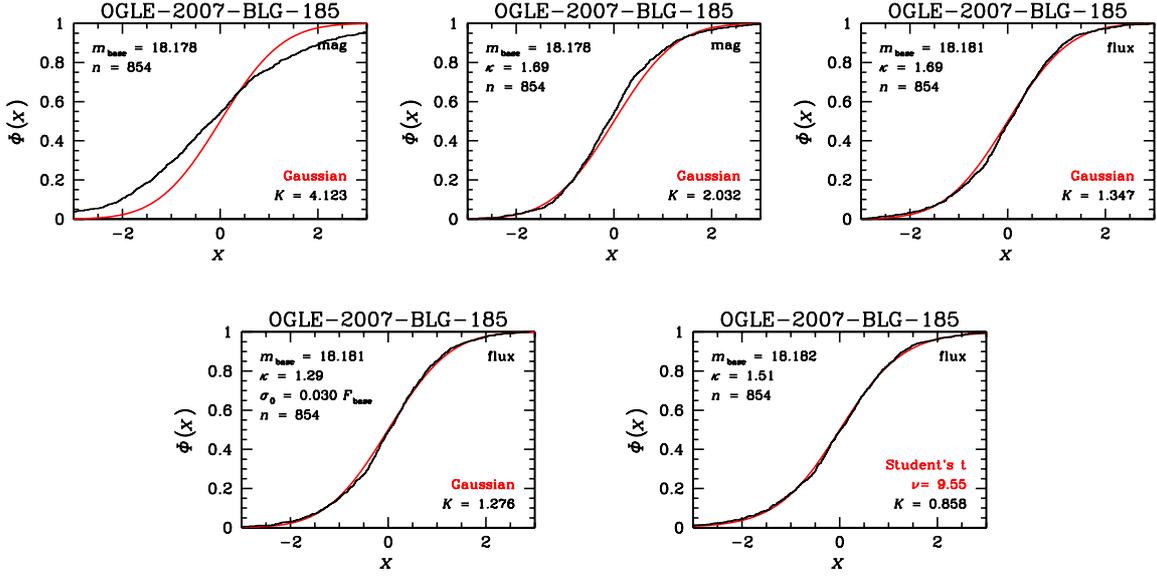


Figure 3: Comparison between the distribution functions $\Phi_n(x)$ of the data residuals x_i of the OGLE baseline measurements on event OGLE-2007-BLG-185 and $\Phi(x)$ of a Gaussian distribution $f(x)$ or $\Phi_\nu(x)$ of Student's t-distribution $f_\nu(x)$ to the parameter ν . While n denotes the number of considered data points, m_{base} is the estimated baseline magnitude, κ stands for a uniform scaling factor, and σ_0 refers to a systematic error added in quadrature, with the error bar model defined by Eq. (2.11). The quoted K is the Kolmogorow-Smirnow test statistic (Eq. (2.7)), providing a measure of the disagreement between the proposed and actual distributions.

In a first step to correct for the discrepancy, let us scale the error bars with a factor κ , so that

$$\sigma'_{m_i} = \kappa \sigma_{m_i}. \quad (2.8)$$

A maximum-likelihood estimate (see Sect. 3.2) yields $\kappa = 1.69$, and the corresponding distribution function is shown in the upper middle panel of Fig. 3. Still, we are being left with an asymmetry of the deviations. As an alternative, let us look into the statistics of the observed fluxes F_i rather than the magnitudes m_i , where

$$F_i = F_{\text{ref}} 10^{-\frac{m_i - m_{\text{ref}}}{2.5}}, \quad (2.9)$$

with the pair $(F_{\text{ref}}, m_{\text{ref}})$ determining the calibration, and the measurement uncertainties being

$$\sigma_{F_i} = \frac{\ln 10}{2.5} F_i \sigma_{m_i}. \quad (2.10)$$

As Fig. 3 shows in the upper right panel, the discrepancies now look smaller, reflected in the more favourable K , and the asymmetry appears to have gone. However, the Kolmogorow-Smirnow test leads to acceptance only at significance levels of 5% and below. If one adds a systematic error in quadrature to the uncertainties, i.e.

$$\sigma'_{F_i} = \sqrt{(\kappa \sigma_{F_i})^2 + \sigma_0^2}, \quad (2.11)$$

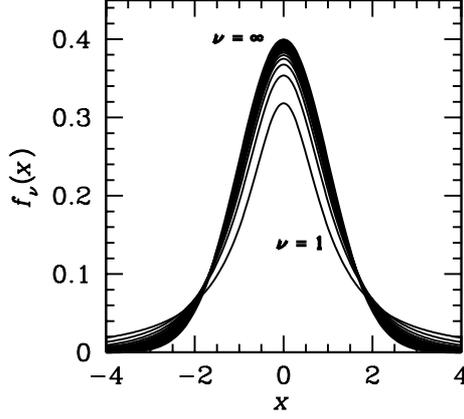


Figure 4: Student's t-distribution $f_\nu(x)$ for all natural ν . While $f_1(x)$ with its wide tail equals the Cauchy-distribution, $f_\nu(x)$ approaches a Gaussian distribution for $\nu \rightarrow \infty$.

a maximum-likelihood estimate gives $\sigma_0 \sim 0.03 F_{\text{base}}$, while $\kappa = 1.29$. A further improvement on K would however still lead to rejection at 10% significance level, and the respective plot, shown in the lower left panel of Fig. 3 suggests that there are systematically wider tails.

An alternative generic distribution accounting for wider tails takes the shape of Student's t-distribution

$$f_\nu(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad (2.12)$$

shown in Fig. 4, which depends on an order parameter ν . While $f_1(x)$ is identical to a Cauchy distribution, $f_\nu(x)$ approaches a Gaussian $f(x)$ for $\nu \rightarrow \infty$. The natural origin of Student's t-distribution is the estimation of the variance from a finite number of measurements rather than an infinite one. As shown in the lower right panel of Fig. 3, the adoption of Student's t-distribution $f_\nu(x)$ with a free parameter ν estimated from the sample of data leads to a far better match even without assuming a systematic error as compared to a Gaussian distribution $f(x)$ with this further scaling freedom, while both comprise the same number of parameters.

In conclusion, one should be careful about the assumption of Gaussianity. While some rescaling might seem to provide a rough match, one does not overcome the issue that larger deviations appear to occur more frequently in the data than a Gaussian distribution would suggest.

3. Estimating model parameters

3.1 Properties of estimators

Rather than by the large number of data itself, one would like to describe the observations by means of a much smaller number of model parameters. This requires to adopt a suitable model and to find parameters that match the data, while one should abstain from trying to match data to an adopted favourite model. Amongst an infinite number of values for any continuous model parameter, which one should be quoted? Despite the fact that a "most-probable value" is sometimes

referred to, such does not exist. For a continuous variable, the probability for each specific realization is exactly the same, namely zero, while probabilities larger than zero can only be attached to parameter ranges of non-vanishing extent.

Ideally, an estimator $\langle \mathbf{p} \rangle$ of the parameter vector \mathbf{p} should have the following properties: (1) as the number of data points approaches infinity, it should converge in probability (or better: almost surely) to the true value; (2) its expectation value should equate the true value, $\mathcal{E}(\langle \mathbf{p} \rangle) = \mathbf{p}$, at least asymptotically, i.e. the estimator should be unbiased (or asymptotically unbiased); (3) the variance on $\langle \mathbf{p} \rangle$, i.e. $\mathcal{E}(\langle \mathbf{p} \rangle^2) - [\mathcal{E}(\langle \mathbf{p} \rangle)]^2$ should be as small as possible. In fact, rather than being able to optimize on all of the desired properties, in general a compromise between these is required.

3.2 Maximum-likelihood estimate

A popular and very convenient choice is the maximum-likelihood (ML) estimator $\langle \mathbf{p} \rangle_{\text{ML}}$, for which the so-called likelihood function $\mathcal{L}(\text{data}; \mathbf{p})$ assumes a maximum. The likelihood function is given by the product of the probability densities for the distribution of the measured data around the corresponding model values at the time of measurement with the measurement uncertainties, so that for the microlensing observations $(t_i^{(k)}, F_i^{(k)}, \sigma_{F_i}^{(k)})$, one finds

$$\mathcal{L}(t_i^{(k)}, F_i^{(k)}, \sigma_{F_i}^{(k)}; \mathbf{p}) = \prod_{k=1}^s \prod_{i=1}^{n_k} f\left(\frac{F^{(k)}(t_i^{(k)}; \mathbf{p}) - F_i^{(k)}}{\sigma_{F_i}^{(k)}}\right). \quad (3.1)$$

Some care is required with regard to the interpretation of the likelihood function $\mathcal{L}(\text{data}; \mathbf{p})$. In particular, it is *not* a probability density in model parameter space, i.e. amongst the realizations of the parameter vector \mathbf{p} .

Asymptotically, and *only* asymptotically, the maximum-likelihood estimator $\langle \mathbf{p} \rangle_{\text{ML}}$ converges almost surely to the true value, becomes unbiased, and even approaches a Gaussian distribution. A disadvantage is that $\langle \mathbf{p} \rangle_{\text{ML}}$ is biased, which can be seen as follows. If one reparametrizes the parameter vector \mathbf{p} with a bijective function $\varphi(\mathbf{p})$, every realization of the parameters carries its likelihood along, i.e. the likelihood at the function value is equal to that of the original value

$$\mathcal{L}_{\varphi(\mathbf{p})}[\text{data}; \varphi(\mathbf{p}_0)] = \mathcal{L}_{\mathbf{p}}(\text{data}; \mathbf{p}_0), \quad (3.2)$$

which trivially results from the definition of the likelihood function, given that the model value of the observable remains unchanged. As a consequence, the ML-estimate of a function of the parameter vector is equal to the function value of its ML-estimate,

$$\langle \varphi(\mathbf{p}) \rangle_{\text{ML}} = \varphi(\langle \mathbf{p} \rangle_{\text{ML}}) \quad (3.3)$$

Since such a relation holds for the expectation value only as long as $\varphi(\mathbf{p})$ is a linear function of \mathbf{p} , one finds that $\mathcal{E}(\langle \mathbf{p} \rangle_{\text{ML}}) \neq \mathbf{p}$ in general. The most advantageous property of the maximum-likelihood estimator $\langle \mathbf{p} \rangle_{\text{ML}}$ is that there is no unbiased estimator with a lower variance [7].

In the case of Gaussian error bars, the maximization of the likelihood function $\mathcal{L}(\text{data}; \mathbf{p})$ becomes equivalent to minimizing the sum of squared residuals

$$\chi^2(t_i^{(k)}, F_i^{(k)}, \sigma_{F_i}^{(k)}; \mathbf{p}) = \sum_{k=1}^s \sum_{i=1}^{n_k} \left(\frac{F^{(k)}(t_i^{(k)}; \mathbf{p}) - F_i^{(k)}}{\sigma_{F_i}^{(k)}} \right)^2, \quad (3.4)$$

given that

$$-2 \ln \mathcal{L}(\text{data}; \mathbf{p}) = \chi^2(\text{data}; \mathbf{p}) + \text{const.} \quad (3.5)$$

If the model function is linear in the parameters \mathbf{p} , χ^2 is parabolic and has a unique minimum. However, this does not hold in general, and in fact, the parameters to describe microlensing events make χ^2 highly non-linear in these. As a consequence, there can be lots of minima. Unfortunately, there is no algorithm known that is able to find all local minima or the global minimum of a continuous function. This means that one needs to adopt a tailored approach for identifying all regions of parameter space that could host a minimum.

3.3 Goodness-of-fit

Once a minimum of χ^2 with respect to the model parameters is found, one would like to assess whether the corresponding \mathbf{p}_{\min} represents a model that is in "agreement" with the data. This "agreement" can be quantified by means of deriving information based on the distribution of test statistics. One such statistic is χ^2 itself, which happens to be " χ^2 -distributed" with the number of "degrees of freedom" $N = n - \lambda$ as a parameter, where n denotes the number of data points and λ the number of ML-estimated parameters, i.e. λ is usually equal to the dimension of the vector \mathbf{p}_{\min} . The evaluation of critical values is simplified by the fact that

$$\lim_{N \rightarrow \infty} \sqrt{2\chi^2} = \mathcal{N}(\sqrt{2N-1}, 1), \quad (3.6)$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 , so that

$$P(\chi^2 \geq \chi_{\min}^2) \simeq 1 - \Phi\left(\sqrt{2\chi_{\min}^2} - \sqrt{2N-1}\right) \quad (N \geq 30). \quad (3.7)$$

This means that the probability for χ^2 to exceed the obtained minimal value χ_{\min}^2 can be extracted from the distribution function $\Phi(x)$ of a unit Gaussian, where the deviation in units of the Gaussian standard deviation σ reads

$$\delta_{\text{Gauss}}/\sigma = \sqrt{2\chi_{\min}^2} - \sqrt{2N-1}. \quad (3.8)$$

Following the standard procedure of testing a hypothesis, a model is then rejected at a significance level α , if $P(\chi^2 \geq \chi_{\min}^2) \leq \alpha$. In contrast, χ^2/N is *not* an appropriate measure for the quality of a model-parameter fit.

In principle, one may choose *any* test statistic for assessing the goodness-of-fit, where for each the probabilistic distribution and critical values need to be worked out. This means in particular that there is no unique statistical test. One needs to be aware of the fact that actually none of these tests can provide evidence for a model being "correct", they can only fail to provide evidence that leads to its rejection, and their power for rejecting models can be substantially different. In particular, a χ^2 -test only considers the absolute amount of the deviations, but not their signs, whereas a run test is just working the other way round, by only considering the sign, but neglecting the size.

3.4 Interval estimation

Having obtained a maximum-likelihood estimate, we are just given a single point \mathbf{p}_{\min} in parameter space, but as pointed out before, the probability that it reflects the 'true' value is zero.

Therefore, rather than just a point estimate, one needs to determine parameter confidence regions that refer to a defined probability. All the required information is contained in the multivariate parameter probability density, and in order to reduce the amount of information, one is free to choose arbitrary confidence regions of any position, size, and shape. There is just one essential requirement, namely that any quoted region has to contain the quoted probability. Actually, the literature is full of examples where this necessary condition has been violated.²

For linear models *and* Gaussian error bars, we are in a lucky situation that makes life quite easy. First, the parameter probability density becomes a multivariate Gaussian centered at the best-fitting value, i.e. the maximum-likelihood estimate. Second, depending on the dimension, contours of χ^2 define symmetric intervals, ellipses, ellipsoids, hyperellipsoids, around \mathbf{p}_{\min} . Third, and finally,

$$\Delta\chi^2 \equiv \chi^2(\text{data}, \mathbf{p}) - \chi^2(\text{data}, \mathbf{p}_{\min}) \quad (3.9)$$

is χ^2 -distributed with λ degrees of freedom, and any projection of $\Delta\chi^2$ onto a parameter sub-space is also χ^2 -distributed with a correspondingly lower number of degrees of freedom. Therefore, the enclosed probability is easily calculated by means of the χ^2 -distribution. As long as an applied non-linear model can be locally transformed into a linear model, one can still profit from these properties. However, such a procedure cannot succeed once it becomes necessary to consider more than a single local minimum. In particular, $\Delta\chi^2$ is no longer χ^2 -distributed over a region covering more than one minimum, nor is there a linear transformation that could achieve this.

In the more general case of non-linear models or non-Gaussian error bars, one is still free to adopt confidence regions that are defined by χ^2 contours, given that one can choose *any* region, but one needs to properly evaluate the enclosed probability, which can be done in general by means of Monte-Carlo simulations [8]. Using relations between χ^2 and the enclosed probabilities that only hold for Gaussian error bars and (locally) linear models will not yield the correct result, and moreover may yield results that are not even close to the correct one.

3.5 Physical properties

Just from the model parameters, it is not always possible to infer directly the properties of the underlying physical system that one is primarily interested in. However, a probability density for the latter given the obtained model parameters can be derived by means of Bayes' theorem, which means to convolve the likelihood $\mathcal{L}(\mathbf{p}|\psi)$ for an observed event described by model parameters \mathbf{p} to arise from a configuration having property ψ with the prior probability density $f_\psi(\psi)$ [9], i.e.

$$f_\psi(\psi|\mathbf{p}) = \frac{\mathcal{L}(\mathbf{p}|\psi) f_\psi(\psi)}{\int \mathcal{L}(\mathbf{p}|\psi') f_\psi(\psi') d\psi'} \quad (3.10)$$

is the probability density for the property ψ given the model parameters \mathbf{p} and the prior $f_\psi(\psi)$. A deeply non-trivial point in Bayes' theorem for continuous variables is on where probability densities and where likelihood functions are referred to.

A worked example for determining the probabilistic distribution of physical properties formed part of the analysis of the OGLE-2005-BLG-390 event [10], with its cool rocky/icy planet, where probability densities for the lens mass, lens distance, orbital radius, and orbital period have been derived. These are shown in Fig. 5.

²Please allow me to abstain from listing respective references.

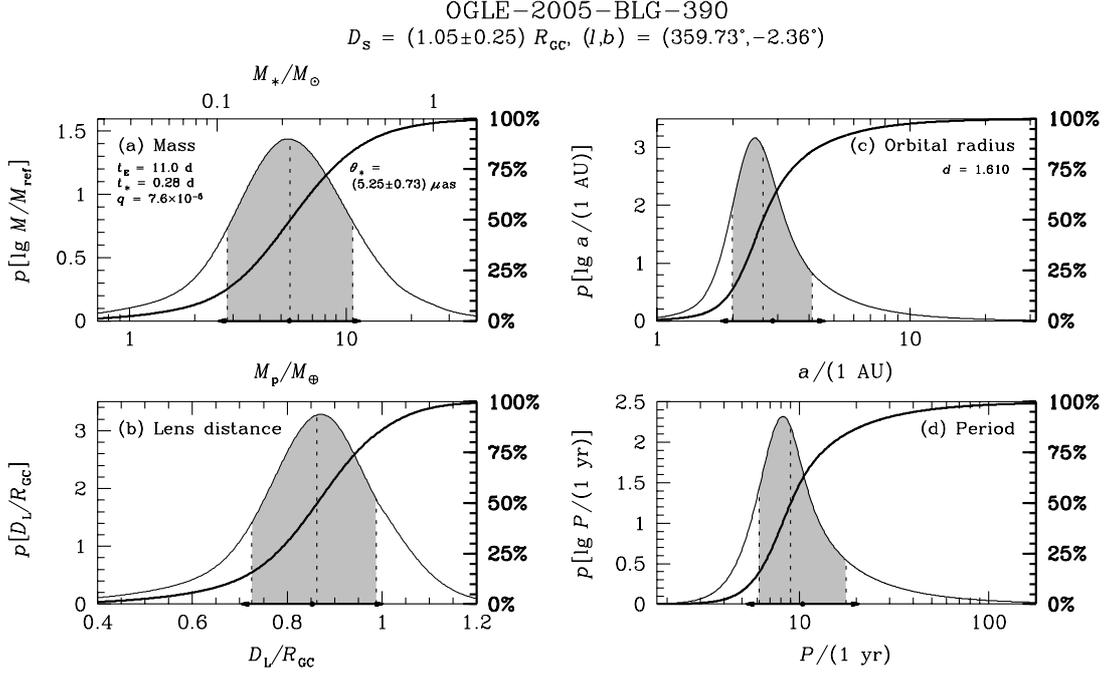


Figure 5: Probability densities for the masses of the planet detected in event OGLE-2005-BLG-390 [10] and its host star, their distance from the observer, the orbital radius (assuming a circular orbit), and the orbital period. These have been obtained by applying Bayes' theorem and assuming mass and velocity distributions of Galactic bulge and disk stars and well as a stellar and sub-stellar mass function [9].

4. Parametrizing gravitational microlensing events

4.1 Observed flux and source star magnification

The phenomenon of gravitational microlensing leads to a characteristic magnification $A(t; \hat{\mathbf{p}})$ as a function of time t and the model parameters $\hat{\mathbf{p}}$ that describe the lens-observer-source geometry, the lens properties, and the source brightness profile. The observed flux $F^{(k)}(t; \mathbf{p})$ however, furthermore is a linear function of the intrinsic flux of the observed source star $F_S^{(k)}$, which is magnified, and a background flux $F_B^{(k)}$, where

$$F^{(k)}(t; \mathbf{p}) = F_S^{(k)} A(t; \hat{\mathbf{p}}) + F_B^{(k)}. \quad (4.1)$$

This allows us to isolate these two parameters from the remaining parameter space, so that

$$\mathbf{p} = (F_S^{(k)}, F_B^{(k)}, \hat{\mathbf{p}}), \quad (4.2)$$

and for every value of $\hat{\mathbf{p}}$, minimizing χ^2 means that the most-likely source and background fluxes can be expressed in closed analytical form

$$F_S = \frac{\sum \frac{A(t_i) F_i}{\sigma_{F_i}^2} \sum \frac{1}{\sigma_{F_i}^2} - \sum \frac{A(t_i)}{\sigma_{F_i}^2} \sum \frac{F_i}{\sigma_{F_i}^2}}{\sum \frac{[A(t_i)]^2}{\sigma_{F_i}^2} \sum \frac{1}{\sigma_{F_i}^2} - \left(\sum \frac{A(t_i)}{\sigma_{F_i}^2} \right)^2}, \quad F_B = \frac{\sum \frac{[A(t_i)]^2}{\sigma_{F_i}^2} \sum \frac{F_i}{\sigma_{F_i}^2} - \sum \frac{A(t_i)}{\sigma_{F_i}^2} \sum \frac{A(t_i) F_i}{\sigma_{F_i}^2}}{\sum \frac{[A(t_i)]^2}{\sigma_{F_i}^2} \sum \frac{1}{\sigma_{F_i}^2} - \left(\sum \frac{A(t_i)}{\sigma_{F_i}^2} \right)^2}, \quad (4.3)$$

while the non-linear minimization process can be restricted to

$$\chi^2(t_i^{(k)}, F_i^{(k)}, \sigma_{F_i}^{(k)}; \hat{\mathbf{p}}) = \sum_{k=1}^s \sum_{i=1}^{n_k} \left(\frac{A(t_i^{(k)}; \hat{\mathbf{p}}) - A_i^{(k)}}{\sigma_{A_i}^{(k)}} \right)^2, \quad (4.4)$$

expressed by means of the magnification rather than the flux, where

$$A_i^{(k)} = \frac{F_i^{(k)} - F_B^{(k)}}{F_S^{(k)}} \quad (4.5)$$

and

$$\sigma_{A_i}^{(k)} = \frac{\sigma_{F_i}^{(k)}}{|F_S^{(k)}|}. \quad (4.6)$$

4.2 Ordinary microlensing events

The most simple case, namely ordinary microlensing light curves, which means that a point-like source star is magnified due to the gravitational bending of its light by an intervening single point-lens star, is far less boring and trivial than commonly thought.

The unique characteristic scale of gravitational microlensing is given by the angular Einstein radius [11]

$$\theta_E = \sqrt{\frac{4GM}{c^2} \frac{\pi_{LS}}{1 \text{ AU}}}, \quad (4.7)$$

where M denotes the mass of the foreground lens star, G is the universal gravitational constant, c is the vacuum speed of light, and

$$\pi_{LS} = 1 \text{ AU} (D_L^{-1} - D_S^{-1}) \quad (4.8)$$

stands for the relative source-lens parallax, with D_L and D_S being the lens distance or source distance, respectively.

With the angular separation between lens and source star being $u \theta_E$, the magnification is an analytic function of u , namely

$$A[u(t; u_0, t_0, t_E)] = \frac{u^2 + 2}{u \sqrt{u^2 + 4}}. \quad (4.9)$$

For uniform proper motion, the dimensionless separation u can be expressed by means of three quantities that form the parameter vector $\hat{\mathbf{p}} = (u_0, t_0, t_E)$, so that

$$u(t; u_0, t_0, t_E) = \sqrt{u_0^2 + \left(\frac{t - t_0}{t_E} \right)^2}, \quad (4.10)$$

where $u_0 \theta_E$ is the smallest angular separation, encountered at epoch t_0 , while the source moves by an angular Einstein radius relative to the lens within the time-scale $t_E \equiv \theta_E / \mu$ [12]. Whereas u_0 and t_0 do not carry any information about the relevant physical properties that determine the magnification, namely the relative parallax π_{LS} , the relative proper motion μ , and the lens mass M , all these are convolved into the event time-scale t_E . As illustrated in Fig. 6, ordinary light curves: (1) are symmetric with respect to a peak at epoch t_0 , (2) reach a peak flux there, (3)

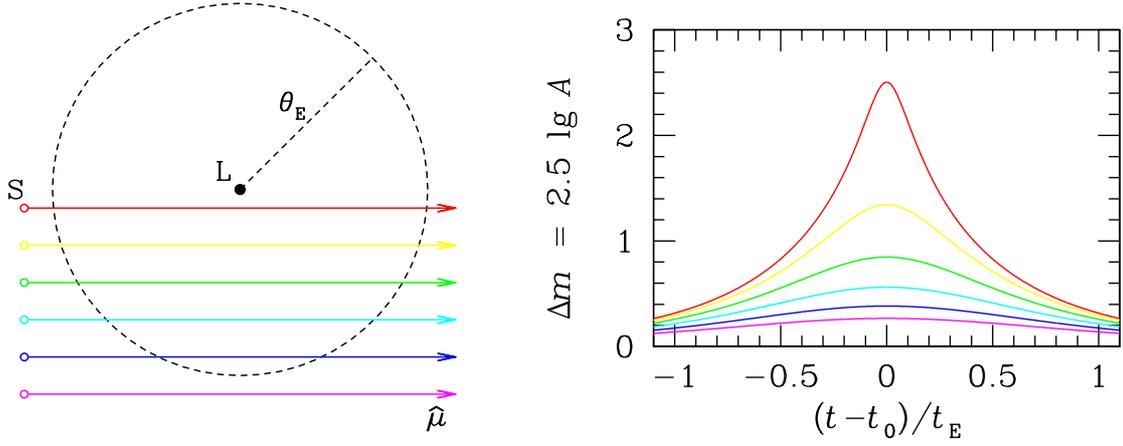


Figure 6: Source trajectories of different impact parameters $u_0 \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1.1\}$ (left) and the corresponding magnitude shift $\Delta m = 2.5 \lg A$ as a function of $(t - t_0)/t_E$ (right), where the magnification $A[u(t; t_0, t_E, u_0)]$ is given by Eq. (4.9), t_0 is the peak epoch, and $t_E = \theta_E/\mu$ is the event time-scale. The source trajectory relative to the lens, assuming uniform relative proper motion μ , is described by $u(t; t_0, t_E, u_0)$, as given by Eq. (4.10), with $u\theta_E$ being the angular separation between lens L and source S. The angular Einstein radius θ_E , defined by Eq. (4.7), marks the unique characteristic scale of gravitational microlensing, indicated as a circle around the lens star in the left panel.

approach a baseline flux for times far away from the peak, and (4) show characteristic inflection points. However, the parameter t_0 is the only one that directly relates to the characteristic features of the light curve, while all the others are not a proper reflection. This is not at all favourable for modelling, and instead being able to essentially read off the model parameters from the collected data would ease life a lot.

4.3 Features and asymptotics

Fortunately, the observed flux can be rewritten by means of a different parametrization. With the baseline flux $F_{\text{base}}^{(k)}$ and the peak flux $F_0^{(k)}$ given by

$$\begin{aligned} F_{\text{base}}^{(k)} &= F_S^{(k)} + F_B^{(k)}, \\ F_0^{(k)} &= F_S^{(k)} A[u(t_0; \hat{\mathbf{p}})] + F_B^{(k)}, \end{aligned} \quad (4.11)$$

the maximal flux difference $\Delta F^{(k)}$ reads

$$\Delta F^{(k)} = F_0^{(k)} - F_{\text{base}}^{(k)} = F_S^{(k)} [A(u_0) - 1]. \quad (4.12)$$

Not only in $F_S^{(k)}$ and $F_B^{(k)}$, but also in $F_{\text{base}}^{(k)}$ and $\Delta F^{(k)}$, the observed flux $F^{(k)}(t; \mathbf{p})$ is a linear function, namely

$$F^{(k)}(t; \mathbf{p}) = \Delta F^{(k)} \frac{A[u(t; \hat{\mathbf{p}})] - 1}{A(u_0) - 1} + F_{\text{base}}^{(k)}, \quad (4.13)$$

so that these parameters can be separated by a linear fit as well. One can then define a half-maximum time $t_{1/2}$, so that at time $t_0 \pm t_{1/2}$, half of the flux offset is encountered, i.e.

$$F^{(k)}(t_0 \pm t_{1/2}; u_0, t_0, t_E) - F_{\text{base}}^{(k)} = \frac{1}{2} \Delta F^{(k)}. \quad (4.14)$$

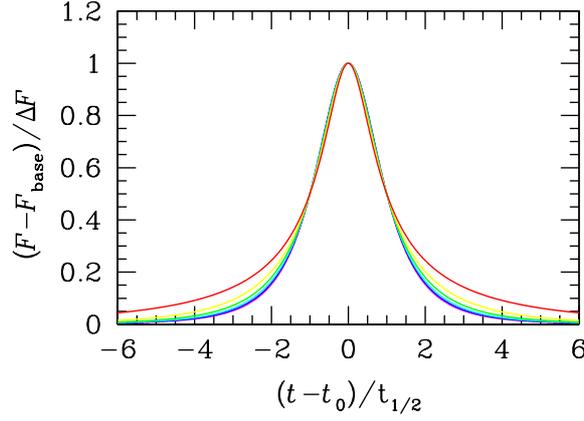


Figure 7: Light curves for the same impact parameters as adopted for Fig. 6, but now the relative flux difference $(F - F_{\text{base}})/\Delta F$ with respect to its peak is plotted as a function of $(t - t_0)/t_{1/2}$, where the half-maximum time-scale $t_{1/2}$ is defined by Eq. (4.14).

This means that the epochs $t_0 \pm t_{1/2}$ correspond to a magnification

$$A_{1/2}(u_0) \equiv A[u(t_0 \pm t_{1/2}; u_0, t_0, t_E)] = \frac{A(u_0) + 1}{2}. \quad (4.15)$$

With $u_{1/2} \equiv u(t_0 \pm t_{1/2}; u_0, t_0, t_E)$ and by means of Eq. (4.10), one finds

$$t_{1/2} = t_E \sqrt{u_{1/2}^2 - u_0^2}, \quad (4.16)$$

which leads to

$$u(t; u_0, t_0, t_{1/2}) = u_0 \sqrt{1 + \left(\frac{u_{1/2}^2}{u_0^2} - 1 \right) \left(\frac{t - t_0}{t_{1/2}} \right)^2}. \quad (4.17)$$

If one now plots the offset flux $F^{(k)}(t; \mathbf{p}) - F_{\text{base}}^{(k)}$ in units of the difference $\Delta F^{(k)}$ between peak and baseline, and scales $t - t_0$ with $t_{1/2}$ rather than t_E , as illustrated in Fig. 7, one finds that light curves with different u_0 nearly coincide around their peaks.

Moreover, if one looks more closely at the behaviour in the limits of small or large separations [13], one finds that the wing of the light curve also shows a convergence towards a single characteristic curve, albeit different ones for small or large u_0 , as can be seen in Fig. 7, indicating a degeneracy in this parameter for these limits. This results from the fact that for $u \ll 1$ and for $u \gg 1$, the magnification can reasonably well be approximated by a more simple expression, which leads in both cases to the half-peak time-scale $t_{1/2}$ becoming proportional to t_E , however with different proportionality factors. With the relative excess magnification $[A(u) - 1]/[A(u_0) - 1]$ approaching u_0/u for $u \ll 1$, and $(u_0/u)^4$ for $u \gg 1$, the observed flux becomes independent of u_0 in these two limits, but the functional dependence on $(t - t_0)/t_{1/2}$ differs, which explains convergences to two different curves. The detailed asymptotic behaviour of the relevant quantities has been summarized in Table 2.

		$u \ll 1$	$u \gg 1$
$A(u)$	\simeq	u^{-1}	$1 + 2u^{-4}$
A_0	\simeq	$\frac{1}{u_0}$	$1 + \frac{2}{u_0^4}$
$A_{1/2}$	\simeq	$\frac{1}{2u_0}$	$1 + \frac{1}{u_0^4}$
$u_{1/2}$	\simeq	$2u_0$	$\sqrt[4]{2}u_0$
$t_{1/2}$	\simeq	$\sqrt{3}u_0 t_E$	$(\sqrt{2}-1)u_0 t_E$
$\frac{A[u(t; \hat{\mathbf{p}})] - 1}{A(u_0) - 1}$	\simeq	$\frac{u_0}{u(t; \hat{\mathbf{p}})}$	$\left(\frac{u_0}{u(t; \hat{\mathbf{p}})}\right)^4$
$F^{(k)}(t; \Delta F^{(k)}, F_{\text{base}}^{(k)}, t_0, t_{1/2}) - F_{\text{base}}^{(k)}$	\simeq	$\frac{\Delta F^{(k)}}{\sqrt{1 + 3 \left(\frac{t-t_0}{t_{1/2}}\right)^2}}$	$\frac{\Delta F^{(k)}}{\left[1 + (\sqrt{2}-1) \left(\frac{t-t_0}{t_{1/2}}\right)^2\right]^2}$

Table 2: Asymptotic behaviour of the observed flux $F^{(k)}$ in the limits $u \ll 1$ or $u \gg 1$, and its matching parametrization.

5. Predicting microlensing light curves

One might think that 3 parameters (like u_0 , t_0 , and t_E) can be obtained straightforwardly by means of regression from as few as 4 data points, but this does not work if the observable does not significantly change with a variation of the considered parameter. Reviewing the asymptotics for $u \ll 1$ and $u \gg 1$ from a different perspective leads to the conclusion that there is no interpretability between the peak and wing regions of the observed light curve, and in particular, early observations give poor estimates of the peak magnification, as well as on the time-scale t_E . Let us look at a practical example, and see how the estimation of the peak magnification changes between the first issue of an alert by OGLE and the final value obtained after the event has been observed completely.

For 386 events alerted by OGLE between 1 Apr 2007 and 15 Oct 2007 that were over on 14 Nov 2007, with suspected anomalies removed from sample, Fig. 8 shows the difference in the prediction of the peak magnitude shift $(\Delta m)_0 = 2.5 \lg A_0$ between the initial and the final value. If one just considers differences by two magnitudes or less (i.e. about a factor 6 in A_0), it looks like the shifts for the majority of events are less than 0.3, i.e. A_0 is mispredicted by less than 30%, while there are larger effects on some of the events. However, the full picture shows that there is another substantial group of events for which the peak magnification is initially overpredicted by 6-10 magnitudes.

These systematic overpredictions are known to be the result of the maximum-likelihood estimator being biased, or in other words: "What you get is not what you expect". In fact, a "better" prediction arises from a posterior-mode estimate [14], also known as maximum a-posterior or "MAP" estimate $\langle \mathbf{p} \rangle_{\text{MAP}}$, which is a rather simple modification of the maximum-likelihood estimate $\langle \mathbf{p} \rangle_{\text{ML}}$, where the likelihood $\mathcal{L}(\text{data}; \mathbf{p})$ is multiplied by a prior probability density $f_{\mathbf{p}}(\mathbf{p})$ of

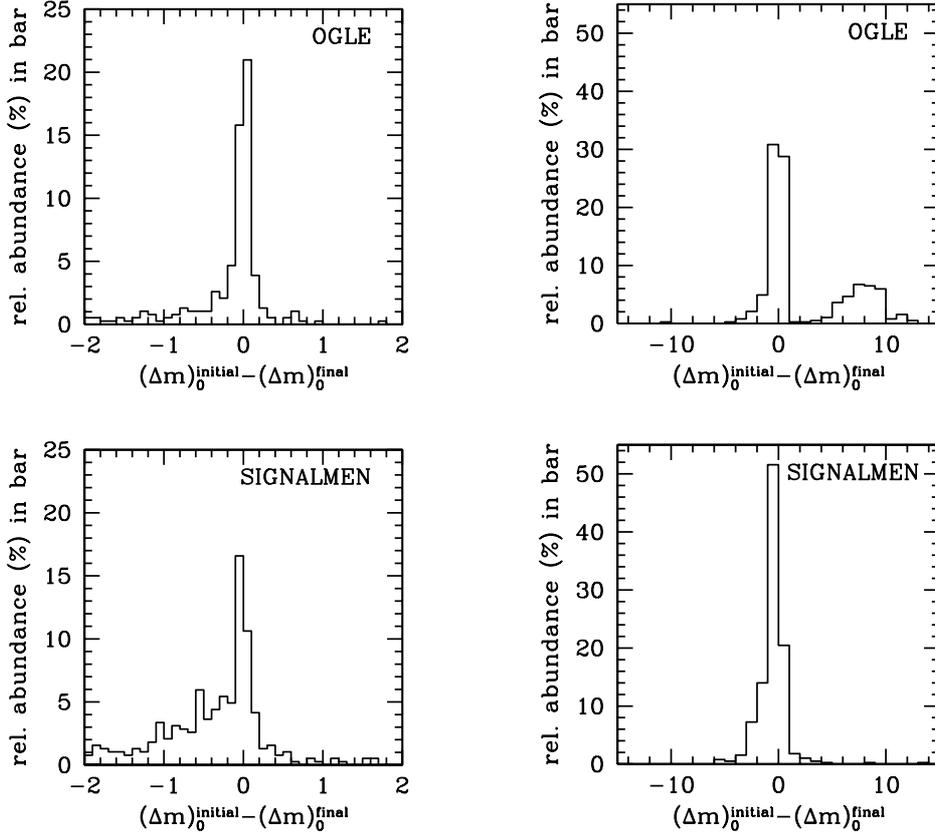


Figure 8: Difference in the prediction of the peak magnitude shift $(\Delta m)_0 = 2.5 \lg A_0$ between the first announcement of the ongoing event by the OGLE team ('initial') and the event being over ('final') for 386 non-anomalous events alerted by OGLE between 1 Apr 2007 and 15 Oct 2007 that were completed on 14 Nov 2007 [5, 6]. While the upper panels refer to the maximum-likelihood estimates carried out by OGLE, the results shown in the lower panels correspond to the posterior-mode (or MAP) estimates as obtained by SIGNALMEN [15].

the model parameters \mathbf{p} , so that

$$\langle \mathbf{p} \rangle_{\text{MAP}} = \arg \max_{\mathbf{p}} \frac{\mathcal{L}(\text{data}; \mathbf{p}) f_{\mathbf{p}}(\mathbf{p})}{\int \mathcal{L}(\text{data}; \mathbf{p}') f_{\mathbf{p}}(\mathbf{p}') d\mathbf{p}'}. \quad (5.1)$$

For the various model parameters, such priors can be obtained from the sample of events, where Fig. 9 shows the result of such an analysis based on the OGLE 2002 season data, where fits to simple functions have been carried out [14], leading to

$$\begin{aligned} f(\lg A_0) &= 0.660 \exp(-1.289 \lg A_0), \\ f\{\lg[t_E/(1 \text{ d})]\} &= 0.476 \exp\left(-\frac{\{\lg[t_E/(1 \text{ d})] - 1.333\}^2}{0.330}\right). \end{aligned} \quad (5.2)$$

Applying such priors, successive MAP estimates along with the corresponding ML estimate for the event OGLE-2003-BLG-208 as it evolves are shown in Fig. 10. Initially, the MAP estimate underpredicts the peak magnification, whereas the ML estimate overpredicts it. At this stage, we

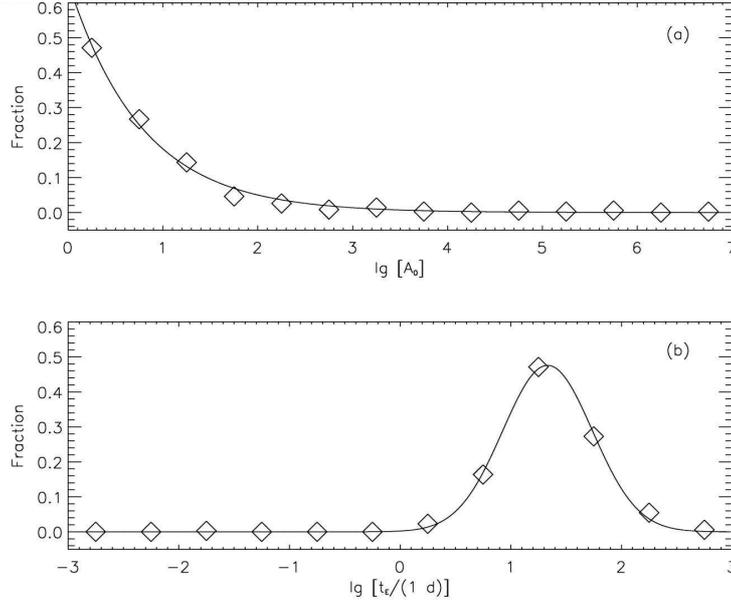


Figure 9: Fitted parametric prior functions for the peak magnification A_0 and the event time-scale t_E , listed in Eq. (5.2), that resulted from an analysis of the 2002 season OGLE data [14] – reproduced by permission of the AAS.

essentially do not have any idea about where the event will go. Apparently, the MAP estimate approaches the correct epoch and correct order of peak magnification substantially faster than the ML estimate. It is also interesting to see that the future development of the event becomes fairly predictable, once the inflection point is reached. Since the light curve is symmetric, it is basically determined once the peak has been passed.

As can be seen from Fig. 8, the MAP estimation carried out by the SIGNALMEN anomaly detector [15] on the 2007 OGLE season data with the respective adopted prior functions leads to a substantially larger number of early underpredictions, but a more appropriate value is approached quite soon with just a few further data points being reported. On the other hand, and more importantly, the vast overpredictions disappear almost entirely.

6. Inhomogeneous data and outliers

Further difficulties arise from outliers in the data, which will definitely show up from real-time photometry and can have disastrous effects if not properly accounted for. Robust-fitting techniques ensure that the model follows the bulk of the data, and thereby practically eliminate the influence of data outliers. This is achieved by using weights $w_i^{(k)}(\mathbf{p})$ that depend on the deviation of data from the considered model, and by minimizing a modified χ^2 , namely

$$\chi_{\text{mod}}^2(t_i^{(k)}, F_i^{(k)}, \sigma_{F_i}^{(k)}; \mathbf{p}) = \sum_{k=1}^s \sum_{i=1}^{n_k} w_i^{(k)}(\mathbf{p}) \left(\frac{F^{(k)}(t_i^{(k)}; \mathbf{p}) - F_i^{(k)}}{\sigma_{F_i}^{(k)}} \right)^2. \quad (6.1)$$

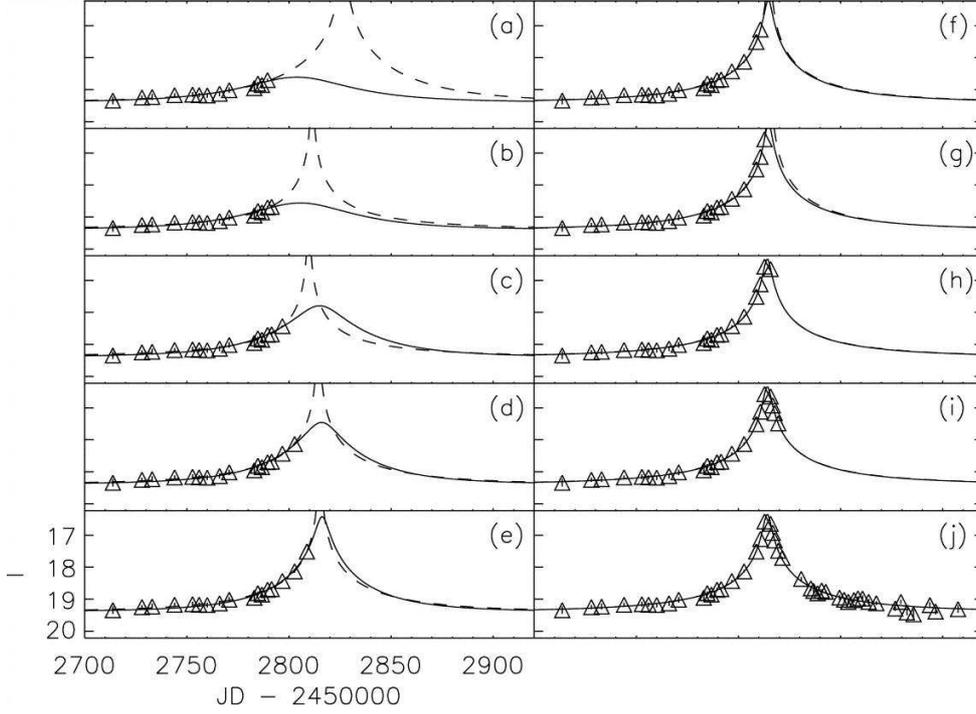


Figure 10: Prediction of the light curve of event OGLE-2003-BLG-208 at various successive stages resulting from a maximum-likelihood (ML) estimate of the model parameters (dashed line) or a posterior-mode (or MAP) estimate (solid line) [14]. – reproduced by permission of the AAS.

While the literature is full of proposed choices for suitable weights [8], the SIGNALMEN anomaly detector [15] in particular determines the absolute residuals

$$r_i^{(k)}(\mathbf{p}) = \left| \frac{F^{(k)}(t_i^{(k)}; \mathbf{p}) - F_i^{(k)}}{\sigma_{F_i}^{(k)}} \right| \quad (6.2)$$

and their median value $\tilde{r}^{(k)}(\mathbf{p})$, and defines the weights $w_i^{(k)}(\mathbf{p})$ to be a bisquare function of the ratio between the absolute residual and its median, i.e.

$$w_i^{(k)}(\mathbf{p}) = \begin{cases} \left[1 - \left(\frac{r_i^{(k)}(\mathbf{p})}{K \tilde{r}^{(k)}(\mathbf{p})} \right)^2 \right]^2 & \text{for } |r_i^{(k)}(\mathbf{p})| < K \tilde{r}^{(k)}(\mathbf{p}) \\ 0 & \text{for } |r_i^{(k)}(\mathbf{p})| \geq K \tilde{r}^{(k)}(\mathbf{p}) \end{cases} \quad (6.3)$$

with the further choice $K = 6$. It is important to realize that it is not just the deviation in units of the reported measurement uncertainty that is being used, but that a scaling with the median residual is made.

The procedure of minimizing and reweighting can be iterated until a desired degree of convergence is reached. Within the same iterative process, further relative weights between different sites and filters can also be incorporated. In particular, forcing all

$$(\chi_{\text{mod}}^2)^{(k)} = \sum_{i=1}^{n_k} w_i^{(k)}(\mathbf{p}) \left(\frac{F^{(k)}(t_i^{(k)}; \mathbf{p}) - F_i^{(k)}}{\sigma_{F_i}^{(k)}} \right)^2 \quad (6.4)$$

Microlensing anomalies

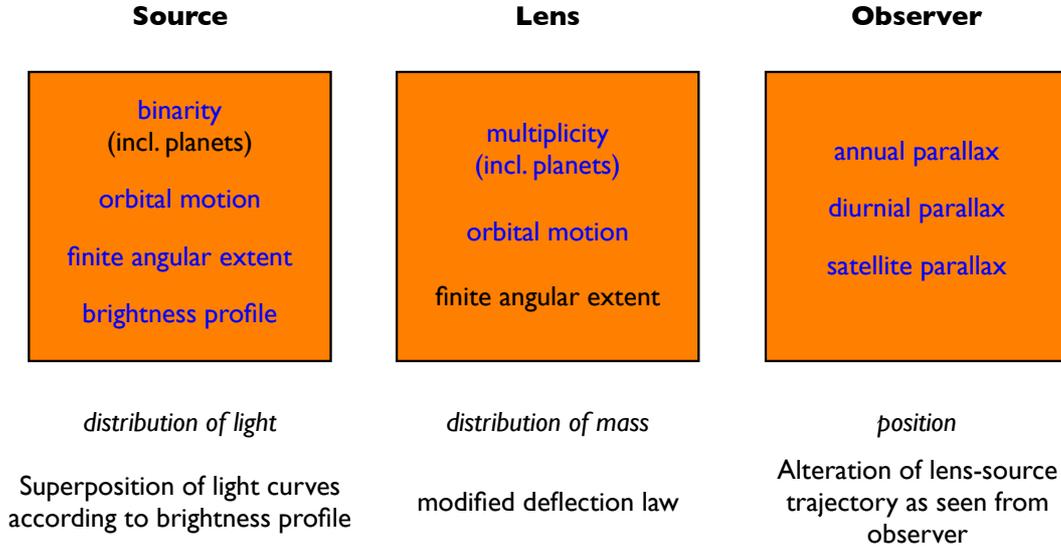


Figure 11: Microlensing anomalies related to the source, lens, or observer. Those shown in blue have been identified in observed events.

to be proportional to the number of respective data points n_k normalizes the data to the actual scatter, and avoids the estimation of model parameters being dominated or (mis-)driven by data sets that show typical deviations that by far exceed the reported uncertainties. However, given that we are dealing with non-linear models, it can happen that one ends up switching between different local minima, where the convergence happens only within each of the subsequences. In fact, even ordinary events come with model ambiguities.

7. Microlensing anomalies

7.1 Overview of anomalies

Beyond point-like sources and lenses, and a static observer, there are a variety of effects that create so-called anomalies in microlensing light curves, where Fig. 11 presents an overview. Those can be grouped into alterations with respect to the source, the lens, or the observer. The source affects the microlensing magnification by means of the distribution of its light, where one observes simply the superposition of light rays. In contrast, it is the distribution of mass that matters for the lens. Here, things become far more sophisticated, given that the deflection law is altered. The observer has an effect on the light curve only by means of its position, which effectively alters the lens-source trajectory, so that one finds parallax effects.

Three different effects shifting the position of the observer have been observed by now. An annual parallax [16–19] results from the revolution of the Earth around the Sun, a diurnal parallax [20] results from the rotation of the Earth around its axis, and a displacement of the position such as achieved by observations from an artificial satellite [21–23] marks the remaining type of parallax.

Anomalies caused by the structure of the gravitational microlens arise from it being composed of several components, including double stars and planetary systems [24–27], and such systems in general show an orbital motion [17, 28, 29], which however frequently goes undetected. Given that the angular Einstein radius θ_E usually by far exceeds the angular radius of the lens star, the finite angular extent is unlikely to play a role, and so far it can safely be neglected for all known events.

Similar to the lens, the source might be composed of more than one object. Stellar binaries have first been convincingly identified with systems that allow their characterization by means of wiggles resulting from their orbital motion [30–32], whereas the characterization of practically static systems [33] suffers both from the lack of suitable signatures or the existence of alternative interpretations. Revealing planets around the source stars prove difficult due to their small contribution to the total light, but they might be detectable during caustic passages [34] or indirectly from the periodic motion of their host star resulting from the gravitational pull of the orbiting planet [35]. The finite angular size of the source regularly plays a role in a substantial fraction of events, in particular if the source passes over a caustic (including the point-like caustic at the position of a single point-mass lens) or comes close to it [36, 37]. This makes gravitational microlensing one of the few successful techniques for measuring brightness profiles of stars [38, 39], including starspots [40, 41], and the only one known that works for quite distant ones (several kpc away), offering a wide choice of stellar types for study.

7.2 Binary point-mass lenses

Binary point-mass lenses constitute a frequent scenario that is full of intricacies. Moreover, it comprises star-planet systems, and these need to be distinguished from stellar binaries.

With $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ denoting the two-dimensional position angles of the source star and its image resulting from the gravitational bending of its light, respectively, the corresponding dimensionless $\mathbf{y} = \boldsymbol{\beta}/\theta_E$ and $\mathbf{x} = \boldsymbol{\theta}/\theta_E$ are related for a single point-mass lens by the gravitational lens equation

$$\mathbf{y} = \mathbf{x} - \frac{\mathbf{x}}{|\mathbf{x}|^2}. \quad (7.1)$$

For weak gravitational fields, the deflections by the individual components of a binary lens can be superposed, and provided that the orbital period of this system by far exceeds the event time-scale t_E , lens binarity enters the lens equation by means of only two further dimensionless parameters, which can be taken as d , where d/θ_E is the instantaneous angular separation, and the secondary-to-primary mass ratio q , so that

$$\begin{aligned} y_1 &= x_1 - \frac{1}{1+q} \frac{x_1 - \frac{qd}{1+q}}{\left(x_1 - \frac{qd}{1+q}\right)^2 + x_2^2} - \frac{q}{1+q} \frac{x_1 + \frac{d}{1+q}}{\left(x_1 + \frac{d}{1+q}\right)^2 + x_2^2} \\ y_2 &= x_2 - \frac{1}{1+q} \frac{x_2}{\left(x_1 - \frac{qd}{1+q}\right)^2 + x_2^2} - \frac{q}{1+q} \frac{x_2}{\left(x_1 + \frac{d}{1+q}\right)^2 + x_2^2}, \end{aligned} \quad (7.2)$$

where a centre-of-mass system has been adopted, and the binary system has been aligned with the y_1 -axis. Given that the binary-lens axis defines a reference system for the source trajectory, the trajectory angle α is a further model parameter, where the separation parameter in units of θ_E is

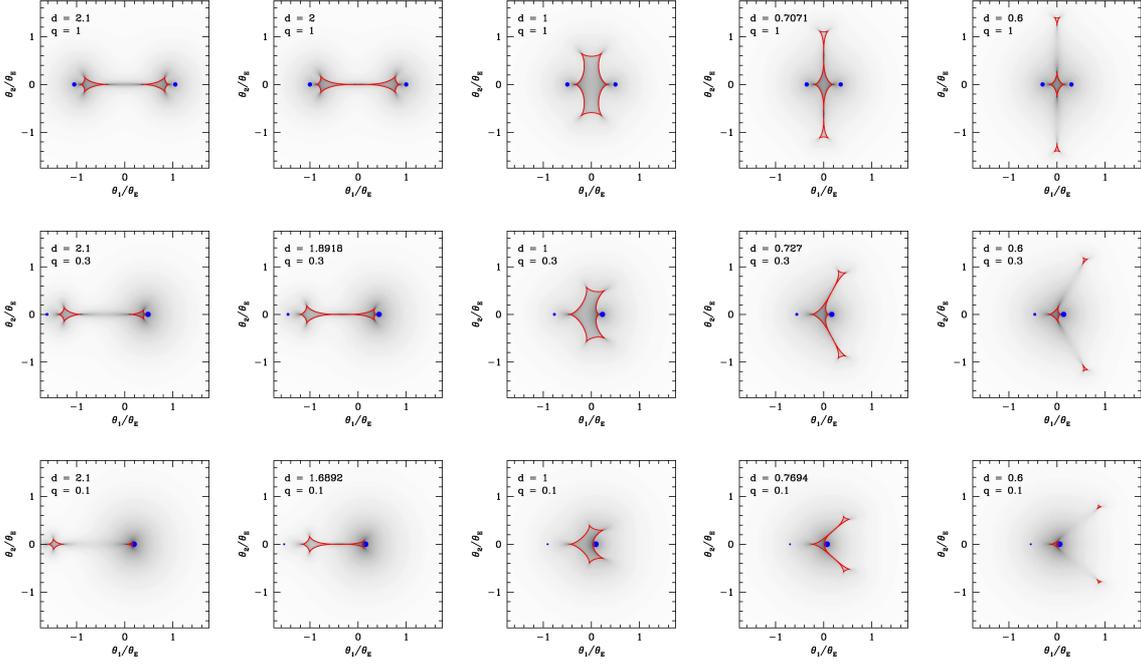


Figure 12: Caustics (red) and magnification (greyscale) for binary lenses of three mass ratios, namely $q \in \{1, 0.3, 0.1\}$ (top to bottom), and angular separations $d \theta_E$ that correspond to the wide, intermediate, and close topologies (from left to right) as well as the transitions. The positions of the lens objects are indicated by filled blue circles with their size reflecting the respective mass. The centre-of-mass of the lens system marks the origin of the coordinate system.

given by

$$\mathbf{u}(t; u_0, t_0, t_E, \alpha) = u_0 \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix} + \frac{t - t_0}{t_E} \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}. \quad (7.3)$$

Finally, the arising finite caustics frequently reveal the angular source radius θ_* , providing another parameter $\rho_* \equiv \theta_*/\theta_E$, so that one finds the 7-dimensional parameter vector $\hat{\mathbf{p}} = (u_0, t_0, t_E, \alpha, d, q, \rho_*)$.

The caustics correspond to those loci, for which the magnification of a point source tends to infinity, which means that the Jacobian determinant of the lens mapping, Eq. (7.2), vanishes. The caustics for a binary point-mass lens match one of just three different topologies [42, 43] depending on the separation parameter d , where there are two transitions from wide to intermediate, and from intermediate to close binaries. As Fig. 12 illustrates, there are two diamond-shaped caustics on the binary axis for a wide binary. On the transition to the intermediate binary, the two cusps facing each other on the binary axis touch, and a single six-cusp caustic results. At the transition to a close binary, the caustic curve touches itself above and below the binary-lens axis, on detachment leaving two cusps facing each other, so that for close binaries, there is a single central diamond-shaped caustic, and two off-axis triangular-shaped caustics. This classification transits smoothly from the equal-mass case to extreme mass ratios such as those representative of a star with a tiny planet, with however the range for intermediate binaries becoming smaller and vanishing as $q \rightarrow 0$ [44]. The shape and the position of the caustics are very important characteristics of the magnification pattern. Given that the magnification along an observed light curve is a one-dimensional cut through this

pattern, the morphology of light curves can be classified based on the topology of the caustics and the source trajectory, and seed solutions can be inferred from a *finite* number of matching configurations [45].

8. Devising a modelling strategy

Partly as part of a review of what was mentioned before, let us put together some essential ingredients for building a successful modelling strategy. As a starting point, it is valuable to identify characteristic features of the collected data. These features then need to be matched to the model parameters. The easiest way for achieving this is to adopt model parameters that are directly related to the observed features. If that can be realized, the full parameter space decouples into subspaces with reduced dimensionality, where parameter searches can be carried out inside each of them independently. Furthermore, a feature-oriented parametrization avoids strong correlations between model parameters, while degeneracies are absorbed into a single parameter, as has been demonstrated in the case of ordinary light curves in Sect. 4.3.

An explicit particular recipe has been derived for finding all possible descriptions of the nature of events for which the source passes over a line-shaped (fold) caustic [46], illustrated in Fig. 13, which marks a worked example of the more general concepts. In this specific case, the fold-caustic passage marks a characteristic feature, and just this part of the light curve can be modelled easily with a semi-analytic function of 5 parameters. With these parameters already determined, a search for the full matching set of parameters only needs to be carried out in a lower-dimensional subspace.

Strategies of such a kind had in fact been adopted in several earlier modelling efforts [47, 48]. For an automated modelling system, artificial neural networks might replace part of the human intuition in identifying the characteristic features [49]. Relating these to a complete set of initial seeds for finding adequate model parameters can then be done by means of event libraries [50] or analytic results [45]. It is hoped that one day it will be possible to constrain the nature of ongoing microlensing events almost in real time as the data are being acquired.

References

- [1] A. Udalski, M. Szymański, K.Z. Stanek, J. Kałużny, M. Kubiak, M. Mateo, W. Krzemiński, B. Paczyński and R. Venkat, *The Optical Gravitational Lensing Experiment. The Optical Depth to Gravitational Microlensing in the Direction of the Galactic Bulge*, *Acta Astronomica* **44** (1994) 165 [astro-ph/9407014].
- [2] R. Ansari, *Une méthode reconstruction photométrique pour l'expérience EROS*, report LAL 94-10, Laboratoire de l'Accélérateur Linéaire d'Orsay 1994.
- [3] M. Dominik, *Galactic Microlensing beyond the Standard Model*, PhD thesis, Universität Dortmund 1996.
- [4] M. Albrow, J.-P. Beaulieu, P. Birch, J.A.R. Caldwell, S. Kane, R. Martin, J. Menzies, R.M. Naber, J.-W. Pei, K. Pollard, P.D. Sackett, K.C. Sahu, P. Vreesswijk, A. Williams and M.A. Zwaan, *The 1995 Pilot Campaign of PLANET: Searching for Microlensing Anomalies through Precise, Rapid, Round-the-Clock Monitoring*, *ApJ* **509** (1998) 687 [astro-ph/9807299].
- [5] A. Udalski, *The Optical Gravitational Lensing Experiment. Real Time Data Analysis Systems in the OGLE-III Survey*, *Acta Astronomica* **53** (2003) 291 [astro-ph/0401123].

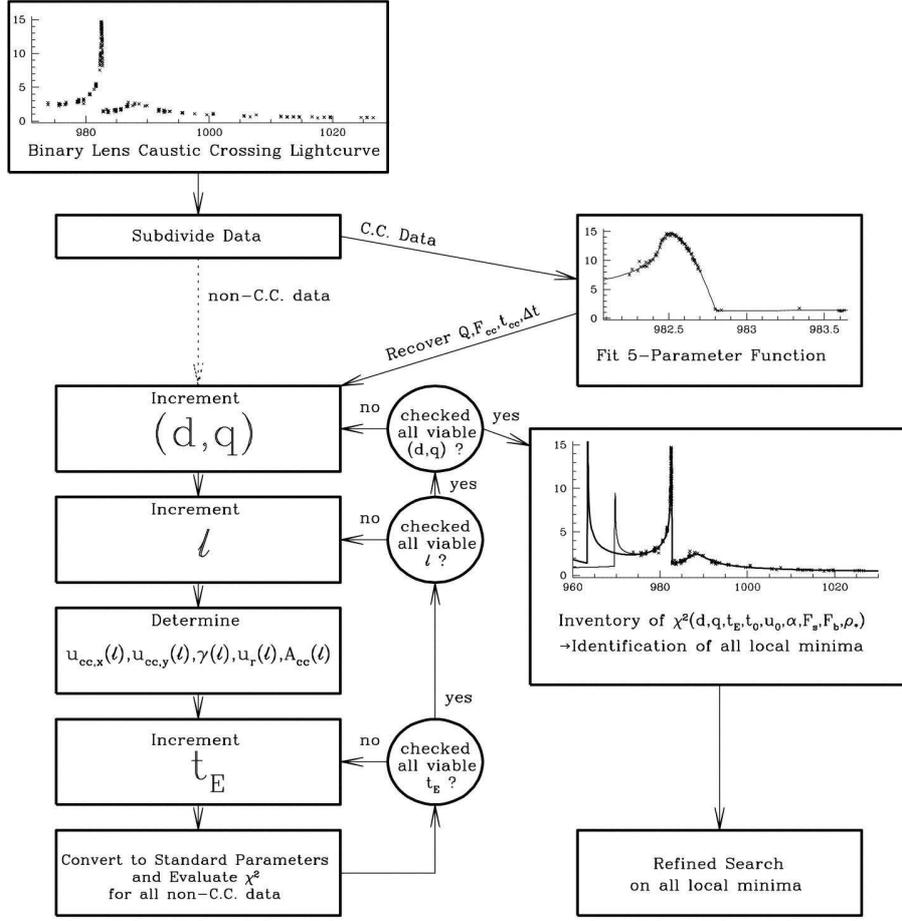


Figure 13: Recipe for finding all regions of suitable model parameters for events that involve an observed passage of the source star over a line-shaped caustic [46], illustrating the more general approach of identifying characteristic features, linking them to model parameters, and thereby decomposing the parameter space. – reproduced by permission of the AAS.

[6] *OGLE-III Early Warning System*, <http://ogle.astrouw.edu.pl/ogle3/ews/ews.html>.

[7] W.T. Eadie, *Statistical Methods in Experimental Physics*, 2nd edition, Elsevier 1983.

[8] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press 2007.

[9] M. Dominik, *Stochastic distributions of lens and source properties for observed galactic microlensing events*, *MNRAS* **367** (2006) 669 [astro-ph/0507540].

[10] J.-P. Beaulieu, D.P. Bennett, P. Fouqué, A. Williams, M. Dominik, U.G. Jørgensen, D. Kubas, A. Cassan, C. Coutures, J. Greenhill, K. Hill, J. Menzies, P.D. Sackett, M. Albrow, S. Brillant, J.A.R. Caldwell, J.J. Calitz, K.H. Cook, E. Corrales, M. Desort, S. Dieters, D. Dominis, J. Donatowicz, M. Hoffman, S. Kane, J.-B. Marquette, R. Martin, P. Meintjes, K. Pollard, K. Sahu, C. Vinter, J. Wambsganss, K. Woller, K. Horne, I. Steele, D.M. Bramich, M. Burgdorf, C. Snodgrass, M. Bode, A. Udalski, M.K. Szymański, M. Kubiak, T. Wieckowski, G. Pietrzyński, I. Soszyński, O. Szweczyk, Ł. Wyrzykowski, B. Paczyński, F. Abe, I.A. Bond, T.R. Britton, A.C. Gilmore,

- J.B. Hearnshaw, Y. Itow, K. Kamiya, P.M. Kilmartin, A.V. Korpela, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, C. Okada, K. Ohnishi, N.J. Rattenbury, T. Sako, S. Sato, M. Sasaki, T. Sekiguchi, D.J. Sullivan, P.J. Tristram, P.C.M. Yock and T. Yoshioka, *Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing*, *Nature* **439** (2006) 437 [astro-ph/0601563].
- [11] A. Einstein, *Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field*, *Science* **84** (1936) 506.
- [12] B. Paczyński, *Gravitational microlensing by the galactic halo*, *ApJ* **304** (1986) 1.
- [13] M. Dominik, *Parameter degeneracies and unpredictability of gravitational microlensing events*, *MNRAS* submitted.
- [14] M.D. Albrow, *Early Estimation of Microlensing Event Magnifications*, *ApJ* **607** (2004) 821 [astro-ph/0402323].
- [15] M. Dominik, N.J. Rattenbury, A. Allan, S. Mao, D.M. Bramich, M.J. Burgdorf, E. Kerins, Y. Tsapras and Ł. Wyrzykowski, *An anomaly detector with immediate feedback to hunt for planets of Earth mass and below by microlensing*, *MNRAS* **380** (2007) 792 [arXiv:0706.2566].
- [16] C. Alcock, R.A. Allsman, D. Alves, T.S. Axelrod, D.P. Bennett, K.H. Cook, K.C. Freeman, K. Griest, J. Guern, M.J. Lehner, S.L. Marshall, B.A. Peterson, M.R. Pratt, P.J. Quinn, A.W. Rodgers, C.W. Stubbs, and W. Sutherland, *First Observation of Parallax in a Gravitational Microlensing Event*, *ApJ* **454** (1995) L125 [astro-ph/9506114].
- [17] M. Dominik, *Galactic microlensing with rotating binaries*, *A&A* **329** (1998) 361.
- [18] M.C. Smith, S. Mao and B. Paczyński, *Acceleration and parallax effects in gravitational microlensing*, *MNRAS* **339** (2003) 925 [astro-ph/0210370].
- [19] A. Gould, *Resolution of the MACHO-LMC-5 Puzzle: The Jerk-Parallax Microlens Degeneracy*, *ApJ* **606** (2004) 319 [astro-ph/0311548].
- [20] J.H. An, M.D. Albrow, J.-P. Beaulieu, J.A.R. Caldwell, D.L. DePoy, M. Dominik, B.S. Gaudi, A. Gould, J. Greenhill, K. Hill, S. Kane, R. Martin, J. Menzies, R.W. Pogge, K.R. Pollard, P.D. Sackett, K.C. Sahu, P. Vermaak, R. Watson and A. Williams, *First Microlens Mass Measurement: PLANET Photometry of EROS BLG-2000-5*, *ApJ* **572** (2002) 521 [astro-ph/0110095].
- [21] S. Refsdal, *On the possibility of determining the distances and masses of stars from the gravitational lens effect*, *MNRAS* **134** (1966) 134.
- [22] A. Gould, *MACHO velocities from satellite-based parallaxes*, *ApJ* **421** (1994) L75.
- [23] S. Dong, A. Udalski, A. Gould, W.T. Reach, G.W. Christie, A.F. Boden, D.P. Bennett, G. Fazio, K. Griest, M.K. Szymański, M. Kubiak, I. Soszyński, G. Pietrzyński, O. Szewczyk, Ł. Wyrzykowski, K. Ulaczyk, T. Wieckowski, B. Paczyński, D.L. DePoy, R.W. Pogge, G.W. Preston, I.B. Thompson, and B.M. Patten, *First Space-Based Microlens Parallax Measurement: Spitzer Observations of OGLE-2005-SMC-001*, *ApJ* **664** (2007) 862 [astro-ph/0702240].
- [24] S. Mao and B. Paczyński, *Gravitational microlensing by double stars and planetary systems*, *ApJ* **374** (1991) L37.
- [25] M. Dominik and A.C. Hirshfeld, *The binary nature of an observed dark galactic object*, *A&A* **289** (1994) L31.

- [26] A. Udalski, M. Szymański, S. Mao, R. Di Stefano, J. Kałużny, M. Kubiak, M. Mateo and W. Krzemiński, *The Optical Gravitational Lensing Experiment: OGLE no. 7: Binary microlens or a new unusual variable?*, *ApJ* **436** (1994) L103 [astro-ph/9407084].
- [27] I.A. Bond, A. Udalski, M. Jaroszyński, N.J. Rattenbury, B. Paczyński, I. Soszyński, Ł. Wyrzykowski, M. Szymański, M. Kubiak, O. Szewczyk, K. Żebruń, G. Pietrzyński, F. Abe, D.P. Bennett, S. Eguchi, Y. Furuta, J. Hearnshaw, K. Kamiya, P.M. Kilmartin, Y. Kurata, K. Masuda, Y. Matsubara, Y. Muraki, S. Noda, K. Okajima, T. Sako, T. Sekiguchi, D.J. Sullivan, T. Sumi, P.J. Tristram, T. Yanagisawa and P.C.M. Yock, *OGLE 2003-BLG-235/MOA 2003-BLG-53: A Planetary Microlensing Event*, *ApJ* **606** (2004) L155 [astro-ph/0404309].
- [28] M.D. Albrow, J.-P. Beaulieu, J.A.R. Caldwell, M. Dominik, B.S. Gaudi, A. Gould, J. Greenhill, K. Hill, S. Kane, R. Martin, J. Menzies, R.M. Naber, K.R. Pollard, P.D. Sackett, K.C. Sahu, P. Vermaak, R. Watson, A. Williams, H.E. Bond, and I.M. van Bemmell, *Detection of Rotation in a Binary Microlens: PLANET Photometry of MACHO 97-BLG-41*, *ApJ* **534** (2000) 894 [astro-ph/9910307].
- [29] B.S. Gaudi, D.P. Bennett, A. Udalski, A. Gould, G.W. Christie, D. Maoz, S. Dong, J. McCormick, M.K. Szymański, P.J. Tristram, S. Nikolaev, B. Paczyński, M. Kubiak, G. Pietrzyński, I. Soszyński, O. Szewczyk, K. Ulaczyk, Ł. Wyrzykowski, D.L. DePoy, C. Han, S. Kaspi, C.-U. Lee, F. Mallia, T. Natusch, R.W. Pogge, B.-G. Park, F. Abe, I.A. Bond, C.S. Botzler, A. Fukui, J.B. Hearnshaw, Y. Itow, K. Kamiya, A.V. Korpela, P.M. Kilmartin, W. Lin, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, T. Okumura, K. Ohnishi, N.J. Rattenbury, T. Sako, To. Saito, S. Sato, L. Skuljan, D.J. Sullivan, T. Sumi, W.L. Sweatman, P.C.M. Yock, M.D. Albrow, A. Allan, J.-P. Beaulieu, M.J. Burgdorf, K.H. Cook, C. Coutures, M. Dominik, S. Dieters, P. Fouqué, J. Greenhill, K. Horne, I. Steele, Y. Tsapras, B. Chaboyer, A. Crocker, S. Frank, and B. Macintosh, *Discovery of a Jupiter/Saturn Analog with Gravitational Microlensing*, *Science* **319** (2008) 927 [arXiv:0802.1920].
- [30] N. Palanque-Delabrouille, C. Afonso, J.N. Albert, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, F. Bauer, J.P. Beaulieu, A. Bouquet, S. Char, X. Charlot, F. Couchot, C. Coutures, F. Derue, R. Ferlet, J.F. Gliçenstein, B. Goldman, A. Gould, D. Graff, M. Gros, J. Haissinski, J.C. Hamilton, D. Hardin, J. de Kat, É. Lesquoy, C. Loup, C. Magneville, B. Mansoux, J.B. Marquette, É. Maurice, A. Milsztajn, M. Moniez, O. Perdereau, L. Prevot, C. Renault, J. Rich, M. Spiro, A. Vidal-Madjar, L. Vigroux, and S. Zylberajch, *Microlensing towards the Small Magellanic Cloud EROS 2 first year survey*, *A&A* **332** (1998) 1 [astro-ph/9710194].
- [31] B. Paczyński, *Binary Source Parallax Effect in Gravitational Micro-lensing* [astro-ph/9711007].
- [32] F. Derue, C. Afonso, C. Alard, J.N. Albert, A. Amadon, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, F. Bauer, J.-P. Beaulieu, A. Bouquet, S. Char, X. Charlot, F. Couchot, C. Coutures, R. Ferlet, J.-F. Gliçenstein, B. Goldman, A. Gould, D. Graff, M. Gros, J. Haissinski, J.-C. Hamilton, D. Hardin, J. de Kat, A. Kim, T. Lasserre, É. Lesquoy, C. Loup, C. Magneville, B. Mansoux, J.-B. Marquette, É. Maurice, A. Milsztajn, M. Moniez, N. Palanque-Delabrouille, O. Perdereau, L. Prevot, N. Regnault, J. Rich, M. Spiro, A. Vidal-Madjar, L. Vigroux, and S. Zylberajch, *Observation of microlensing towards the galactic spiral arms. EROS II. 2 year survey*, *A&A* **351** (1999) 87 [astro-ph/9903209].
- [33] K. Griest and W. Hu, *Effect of binary sources on the search for massive astrophysical compact halo objects via microlensing*, *ApJ* **397** (1992) 362.

- [34] D.S. Graff and B.S. Gaudi, *Direct Detection of Large Close-in Planets around the Source Stars of Caustic-crossing Microlensing Events*, *ApJ* **538** (2000) L133 [astro-ph/0004089].
- [35] S. Rahvar and M. Dominik, *Detecting exoplanets with the xallarap microlensing effect*, in proceedings of *The Manchester Microlensing Conference: The 12th International Conference and ANGLES Microlensing Workshop*, PoS (GMC8) 037.
- [36] H.J. Witt and S. Mao, *Can lensed stars be regarded as pointlike for microlensing by MACHOs?*, *ApJ* **430** (1994) 505.
- [37] C. Alcock, W.H. Allen, R.A. Allsman, D. Alves, T.S. Axelrod, T.S. Banks, S.F. Beaulieu, A.C. Becker, R.H. Becker, D.P. Bennett, I.A. Bond, B.S. Carter, K.H. Cook, R.J. Dodd, K.C. Freeman, M.D. Gregg, K. Griest, J.B. Hearnshaw, A. Heller, M. Honda, J. Jugaku, S. Kabe, S. Kaspi, P.M. Kilmartin, A. Kitamura, O. Kovo, M.J. Lehner, T.E. Love, D. Maoz, S.L. Marshall, Y. Matsubara, D. Minniti, M. Miyamoto, J.A. Morse, Y. Muraki, T. Nakamura, B.A. Peterson, M.M. Phillips, M.R. Pratt, P.J. Quinn, I.N. Reid, M. Reid, D. Reiss, A. Retter, A.W. Rodgers, W.L.W. Sargent, H. Sato, M. Sekiguchi, P.B. Stetson, C.W. Stubbs, D.J. Sullivan, W. Sutherland, A. Tomaney, T. Vandehei, Y. Watase, D.L. Welch, T. Yanagisawa, M. Yoshizawa, and P.C.M. Yock, *MACHO Alert 95-30: First Real-Time Observation of Extended Source Effects in Gravitational Microlensing*, *ApJ* **491** (1997) 436 [astro-ph/9702199].
- [38] M.D. Albrow, J.-P. Beaulieu, J.A.R. Caldwell, M. Dominik, J. Greenhill, K. Hill, S. Kane, R. Martin, J. Menzies, R.M. Naber, J.-W. Pel, K. Pollard, P.D. Sackett, K.C. Sahu, P. Vermaak, R. Watson, A. Williams and M.S. Sahu, *Limb Darkening of a K Giant in the Galactic Bulge: PLANET Photometry of MACHO 97-BLG-28*, *ApJ* **522** (1999) 1011 [astro-ph/9811479].
- [39] D. Heyrovský, D. Sasselov and A. Loeb, *Probing Red Giant Atmospheres with Gravitational Microlensing*, *ApJ* **543** (2000) 406 [astro-ph/9902273].
- [40] D. Heyrovský and D. Sasselov, *Detecting Stellar Spots by Gravitational Microlensing*, *ApJ* **529** (2000) 69 [astro-ph/9906024].
- [41] M.A. Hendry, H.M. Bryce and D. Valls-Gabaud, *The microlensing signatures of photospheric starspots*, *MNRAS* **335** (2002) 539 [astro-ph/0210208].
- [42] P. Schneider and A. Weiß, *The two-point-mass lens – Detailed investigation of a special asymmetric gravitational lens*, *A&A* **164** (1986) 237
- [43] H. Erdl and P. Schneider, *Classification of the multiple deflection two point-mass gravitational lens models and application of catastrophe theory in lensing*, *A&A* **268** (1993) 453.
- [44] M. Dominik, *The binary gravitational lens and its extreme cases*, *A&A* **349** (1998) 108 [astro-ph/9903014].
- [45] V. Bozza, *Trajectories of the images in binary microlensing*, *A&A* **374** (2001) 13 [astro-ph/0105268].
- [46] M.D. Albrow, J.-P. Beaulieu, J.A.R. Caldwell, D.L. Depoy, M. Dominik, B.S. Gaudi, A. Gould, J. Greenhill, K. Hill, S. Kane, R. Martin, J. Menzies, R.M. Naber, R.W. Pogge, K.R. Pollard, P.D. Sackett, K.C. Sahu, P. Vermaak, R. Watson, A. Williams, *A Complete Set of Solutions for Caustic Crossing Binary Microlensing Events*, *ApJ* **522** (1999) 1022 [astro-ph/9903008].
- [47] M. Dominik and A.C. Hirshfeld, *Evidence for a binary lens in the MACHO LMC No. 1 microlensing event*, *A&A* **313** (1996) 841.
- [48] M. Dominik, *Ambiguities in fits of observed binary lens galactic microlensing events*, *A&A* **341** (1999) 943 [astro-ph/9703003].

- [49] P. Vermaak, *Fitting Binary Lens Gravitational Microlensing Events with Example-based Algorithms*, PhD thesis, University of Cape Town 2007.
- [50] S. Mao and R. Di Stefano, *Interpretation of gravitational microlensing by binary systems*, *ApJ* **440** (1995) 22.