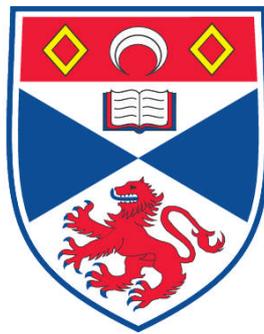


**THE SOLAR SYSTEM IN PERSPECTIVE : FROM DEBRIS DISCS
TO EXTRASOLAR PLANETS**

Noé Kains

**A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews**



2010

**Full metadata for this item is available in the St Andrews
Digital Research Repository
at:**

<https://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/1030>

This item is protected by original copyright

**This item is licensed under a
Creative Commons License**

The Solar System in Perspective

From Debris Discs to Extrasolar Planets

by

Noé Kains

Submitted for the degree of Doctor of Philosophy in Astrophysics

July 2010



University
of
St Andrews

Declaration

I, Noé Kains, hereby certify that this thesis, which is approximately 25,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date

Signature of candidate

I was admitted as a research student in September 2006 and as a candidate for the degree of PhD in September 2006; the higher study for which this is a record was carried out in the University of St Andrews between 2006 and 2010.

Date

Signature of candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date

Signature of supervisor

Copyright Agreement

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis: Access to Printed copy and electronic publication of thesis through the University of St Andrews.

Date

Signature of candidate

Date

Signature of supervisor

Abstract

The last twenty-five years have seen our understanding of the formation and abundance of planets revolutionised, thanks to the first detections of debris discs, and, a decade later, of the first extrasolar planets. Hardly a week now goes by without a planet discovery, and the range of methods used to search for planets has expanded to include techniques that are efficient at detecting different types of planets. By combining the discoveries of the various methods, we therefore have the opportunity to build a picture of planet populations across the Galaxy. In this thesis, I am presenting work done as a basis towards such an effort: first I present work carried out to improve modelling methods for gravitational microlensing events.

Since the first microlensing observing campaigns, the amount of data of anomalous events has been increasing ever faster, meaning that the time required to model all observed anomalous events is putting a strain on available computational and human resources. I present work to develop a method to fit anomalous microlensing events automatically and show that it is possible to conduct a thorough and unbiased search of the parameter space, illustrating this by analysing an event from the 2007 observing season. I then discuss the possible models found with this method for this event and their implication (Kains et al., 2009), and find that this algorithm locates good-fit models in regions of parameters that would have been very unlikely to be found using standard modelling methods. Results indicate that it is necessary to use a full Bayesian approach, in order to include prior information on the parameters. I discuss the analytical priors calculated by Cassan et al. (2009), and suggest a possible form of an automatic fitting algorithm by incorporating these priors in the algorithm used by Kains et al. (2009).

Another topic with which this thesis is concerned is the evolution of debris discs around solar-type stars. Late-type stars are expected to be the most numerous host stars of planets detected with the microlensing technique. Understanding how their debris discs evolve equates to understanding the earliest stages of planet formation around these stars, allowing us to truly put our Solar System in perspective.

Using the analytical model of Wyatt et al. (2007a), I modelled the evolution of infrared excess flux at 24 and 70 μm using published data of debris discs around solar-type (spectral

types F, G and K) stars from the *Spitzer Space Telescope*. By comparing the results of this study to an analogous study carried out for A stars by Wyatt et al. (2007b), I find that although best-fit parameters are significantly different for solar-type stars, this may be due to the varying number of inefficient emitters around stars of different spectral types. I suggest that although effective properties are different by an order of magnitude or more, intrinsic properties, while still different, are so by a much smaller factor. These differences may be due to the longer timescales over which solar-type stars evolve, which allow for the formation of larger and stronger planetesimals.

Acknowledgements

There are many people whom I have to thank for helping me finish this Ph.D. Firstly, I thank my supervisor Keith Horne, for taking me on as his student, for giving me so much research and geographical freedom in the last four years, and for being so understanding of me disappearing to prepare recitals or go on orchestra and choir tours! I am also very grateful to Jane Greaves for accepting to supervise me on a second project working on debris discs when my brain needed to take a step away from microlensing. Next, I would like to thank Arnaud Cassan, friend and collaborator, with whom I had countless discussions on work, music, and many other subjects which were crucial to keeping me motivated and focused. I am also indebted to many other research collaborators: Martin Dominik, Kailash Sahu, John Greenhill, Pascal Fouqué, Virginie Batista, Paul Browne, Yiannis Tsapras, and Mark Wyatt, and to Joachim Wambsganss and Andrew Cameron for agreeing to be my thesis examiners.

Next I would like to thank my friends who coped with the inevitable emotional rollercoaster that makes a Ph.D. such a unique experience! Special prize to Alexis Smith, who not only shared an office with me for three years, but also shared a flat with me for a year, and lived to tell the tale. Thanks to Nick Dunstone for the many games of tennis and photography tips, Lee, gwa2 and ew60 for general banter, Grant for more games of tennis, badminton, and for a memorable road trip around Scotland, and David for being a great office mate. Thanks also to all my musical friends, who probably saw me more than my office mate did, particularly everyone in the Madrigal Group, especially my academic mother Charly. Thanks also to Stéphane and Tim for being exceptional chamber music partners. Thanks to Ryan, Jake and other close friends who provided me with an escape from St Andrews whenever I needed contact with the real world.

Last but not least, I thank my family for their love and support in the last four years: my mother for teaching me to seize all opportunities, my grandmother for her astrophysical expertise, my grandfather for his interest in my field of research, my father for teaching me to aim high, my sister Elise for keeping me updated with the latest in developments of French soaps and Belgian politics, and my brother Benjamin who inspired me to become a scientist in the first place.

*We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.*

T.S. Eliot, *Four Quartets*

To my mother.

Contents

Declaration	i
Copyright Agreement	iii
Abstract	v
Acknowledgements	vii
1 Searching for other worlds	1
1.1 First detections and methods	2
1.2 A brief history of microlensing	4
1.2.1 The first microlensing campaigns	6
2 Microlensing formalism	9
2.1 The single-lens case	9
2.2 Multiple lens case	13
2.2.1 Deriving the multiple lens equation	13
2.2.2 Critical curves and caustics	15
2.2.3 Final binary-lens parametrisation	18
2.3 Second-order effects	21
2.4 Extracting physical parameters from microlensing modelling	25
2.5 Binary sources	26
3 Markov Chain-Monte Carlo methods	29
3.1 Motivation and Formalism	29
3.1.1 Sample autocorrelation	33
3.2 Application to microlensing	34
3.3 Fitting lightcurves with MCMC	35
3.3.1 PSPL events	35

3.3.2	Parameter correlations	37
3.3.3	Orthogonalising the parameter space	37
3.3.4	Binary lens events	41
3.3.5	A test binary-lens event: OGLE-2005-BLG-390	47
4	The basis for an automatic binary-lens fitting algorithm	51
4.1	Introduction	52
4.2	Binary-lens events fitting scheme	53
4.2.1	Parameterisation of binary lens lightcurves	53
4.2.2	Exploration of the parameter space	54
4.3	Application to OGLE-2007-BLG-472	56
4.3.1	Alert and photometric follow-up	56
4.3.2	Data reduction	57
4.3.3	Modelling OGLE-2007-BLG-472	58
4.3.4	Excluding minima	61
4.3.5	Refining local minima	61
4.3.6	Parameter correlations	65
4.3.7	Physical properties of the models	66
4.4	Summary and prospects	71
5	Towards a full Bayesian approach for automatic fitting of binary-lens events	73
5.1	Introduction	73
5.2	Analytical priors	74
5.3	Additional prior information	76
5.4	An exhaustive modelling strategy	78
6	Modelling the evolution of debris discs around solar-type stars	87
6.1	Introduction	87
6.2	Debris discs around solar-type stars	89
6.3	Debris disc model	90
6.4	Modelling procedure	92
6.5	Data	94
6.5.1	Trilling et al. (2008) sample	95

6.5.2	Beichman et al. (2006) sample	95
6.5.3	Hillenbrand et al. (2008) sample	98
6.6	Best-fit models	98
6.6.1	Fit to the $70\mu\text{m}$ statistics	98
6.6.2	Fit to the $24\mu\text{m}$ statistics	99
6.7	Discussion of results	103
6.7.1	Significance of best-fit parameters	103
6.7.2	Effective and real parameters	106
6.7.3	Fractional luminosity vs. radius	108
6.7.4	Fractional luminosity vs. age	110
6.7.5	Radius vs. age	112
6.7.6	Histogram of f/f_{max}	114
6.8	Conclusions	115
7	Conclusions and outlook	117
7.1	Automatic fitting: the future for microlensing	118
7.2	Towards a census of debris discs around all types of stars	118
7.3	Conclusion	119
	References	121

List of Figures

1.1	Extrasolar planet discoveries as of 11 March 2010, plotted in the (semi-major axis, mass) plane. Planets discovered with the radial velocity (RV), transit, microlensing, imaging and pulsar timing methods are plotted with different symbols and colours, detailed in the bottom right box. The Solar System's planets are plotted in blue and labeled with their initial. Method sensitivities are given for the RV, transit and microlensing methods, with a distinction between space-based and ground-based sensitivities for the latter two. Data from the Exoplanet Encyclopedia (http://exoplanet.eu), sensitivities from Keith Horne (private communication).	5
2.1	Geometry of a point-source point-lens microlensing event. Labeled quantities are the deflection angle α , the closest approach r_{\min} , the angular closest approach on the lens plane θ , the angular source position β , as well as the distances to the lens and to the source, D_L and D_S	11
2.2	The PSPL model parameter: the source takes a time t_E to cross an Einstein ring radius θ_E , while it reaches closest approach at time t_0 , at which point it is an angular distance u_0 from the lens.	13
2.3	Separation lines in the logarithmic (d, q) plane between the caustic regimes, after Erdl & Schneider (1993). Each region is labeled with the corresponding regime name and number of caustics.	16
2.4	Caustic topology for the wide (upper panel), intermediate (middle panel) and close (lower panel) regimes. The d and q values corresponding to each panel are inset in the top left corner. Crosses indicate the position of each lens component, with the more massive component closer to the origin (which corresponds to the centre of mass of the lens system).	17
2.5	The trajectory angle parameter α , defined as the angle between the source trajectory and the line joining the two lens components on the lens plane. . . .	18
2.6	An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.04, \alpha = 1.0, t_E = 15, d = 0.7$ and $q = 0.1$. The source crosses both the central and the upper secondary caustic, leading to clear caustic-crossing magnification patterns.	19
2.7	An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.3, \alpha = 0.1, t_E = 15, d = 0.7$ and $q = 0.1$. In this case, the source approaches both the central and secondary caustics without crossing them.	19

2.8	An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.46, \alpha = 1.3, t_E = 15, d = 1.3$ and $q = 10^{-3}$. In this case, the source approaches a secondary caustic but doesn't cross it.	19
2.9	An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.02, \alpha = 1.0, t_E = 15, d = 0.8$ and $q = 0.1$. The source crosses an elongated intermediate caustic, exits it and crosses it again.	20
2.10	An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.2, \alpha = 1.6, t_E = 15, d = 1.1$ and $q = 0.5$. The source crosses an intermediate caustic.	20
2.11	An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.2, \alpha = 1.7, t_E = 15, d = 0.5$ and $q = 0.1$. The source approaches a central caustic, which results in a subtle departure from the PSPL magnification pattern (shown as a dotted line here for comparison).	20
2.12	The effects of having an extended source in a single-lens model. The extended source-point lens (ESPL) model is shown with a solid line, and the PSPL model is shown with a dotted line. The ESPL model has a source size $\rho_* = 0.2$, and both models have an impact parameter $u_0 = 0.02$	22
2.13	The source size effects for binary-lens models. <i>Left:</i> The source trajectory and caustics. The positions of the lens components are indicated with crosses. <i>Right:</i> The resulting lightcurve. The extended source-binary lens (ESBL) model is shown with a solid line, and the point source-binary lens (PSBL) model is shown with a dotted line. The ESBL model has a source size $\rho_* = 0.04$, and both models have parameters $u_0 = 0.02, \alpha = 1.0, d = 1.5$ and $q = 0.5$	22
2.14	<i>Top:</i> the limb darkening law, showing the decrease in intensity as one moves away from the centre of its disc, for different values of $\gamma_1 (\equiv \Gamma)$. The x -axis is the distance from the centre of the source in units of the source radius. <i>Bottom:</i> the effect of an extended source having the different limb-darkening coefficient γ_1 shown in the top panel, in a single-lens model (bottom left) and on the shape of the photometric signal of a caustic crossing (bottom right). The x -axis is the source-lens separation in units of source radii. Taken from Cassan (2005).	24
2.15	<i>Top:</i> the best-fit binary-source model for OGLE-2004-BLG-347, corresponding to a source with a primary component ~ 2 times brighter than the secondary. <i>Bottom:</i> the best-fit binary-lens model ($d \sim 3, q \sim 0.75$) for that event. Neither model is significantly better than the other, so the nature of this event is ambiguous.	27
3.1	The value of χ^2 against the number of iterations of the MCMC chain for a fit to a PSPL microlensing event. The burn-in and stationary periods are labeled and a dotted line indicates where the chain starts sampling for the posterior distribution.	31
3.2	Example pairwise correlations showing the progress of the Markov chain for four different starting points (shown with filled circles). The chains take a few steps during the burn-in period before converging to the stationary distribution (dense area in the middle of the plot).	31

3.3	MCMC fit and residuals for OGLE-2006-BLG-180	36
3.4	MCMC fit and residuals for OGLE-2006-BLG-397 (note the worse data quality than OGLE-2006-BLG-180).	36
3.5	Correlations between standard PSPL model parameters for OGLE-2006-BLG-180. 38	
3.6	Correlations between t_0, u_0 and chosen model parameters ω, ϵ and f_0 for OGLE-2006-BLG-180.	38
3.7	<i>Left:</i> Schematic representation of the “shearing” of the parameter space for a pairwise parameter correlation pattern, which becomes circular rather than elliptical, i.e. less correlated. <i>Right:</i> Gram-Schmidt orthogonalisation for the first two vectors of an N-dimensional parameter space.	39
3.8	Correlations between orthogonalised parameters β_i for OGLE-2006-BLG-180 . .	42
3.9	Model derivatives for the PSPL case, with respect to standard parameters t_0, t_E, u_0, f_s and f_b (top panel) and the orthogonal parameter β_i (bottom panel), for OGLE-2006-BLG-180 (Fig. 3.3).	43
3.10	Incomplete PSBL fit to the anomalous event OGLE-2008-BLG-513. This fit was made while the event was ongoing, but subsequent observations showed that the magnification pattern became significantly affected by the finite source size. Because this event is also one in which the source is highly magnified, this meant that modelling of this event could not be pursued with limited computational resources.	44
3.11	The effect of diagonalising the covariance matrix on the acceptance rate of the MCMC algorithm. The burn-in and stationary phases are separated by a vertical dashed line, while a dotted line shows where the covariance matrix is first diagonalised.	47
3.12	MCMC fit and residuals for the planetary binary-lens event OGLE-2005-BLG-390. The best-fit parameters are given in Table 3.3.	48
3.13	Correlations between the 7 model parameters and 2 flux parameters (in this case for the OGLE telescope) for the binary-lens event OGLE-2005-BLG-390. . .	49
4.1	The alternative binary-lens parameterisation used in this discussion. <i>Top:</i> The s coordinate, which runs along the caustic folds from $s = 0$ to $s = 2$. <i>Bottom:</i> The source enters the caustic at time $t = t_{\text{in}}$; its location on the caustic fold at that time is $s(t_{\text{in}}) = s_{\text{in}}$. Similarly, the source exits the caustic at $t = t_{\text{out}}$, at which point it is located at $s = s_{\text{out}}$. The example caustic used here is for parameters $d = 1.5, q = 0.5$	54
4.2	The relation between the s coordinate and the components of the source’s position ζ, x (top left) and y (top right), as well as the derivatives of s with respect to each of these components (bottom panels). These plots correspond to an intermediate caustic of a binary-lens with $d = 1.1, q = 0.1$, which is plotted on Fig. 4.1.	55
4.3	OGLE, UTas and Danish data sets for OGLE-2007-BLG-472. Data points are plotted with $1-\sigma$ error bars. The x -axis is time in HJD-2450000.	57

4.4	$\chi^2(d, q)$ map for the intermediate and central caustic configurations. Contour lines and minima regions (in blue shades) are plotted at $\Delta\chi^2 = 5, 20, 50, 100, 250$. The two dashed curves are the separation between the close, intermediate and wide regimes. The best-fit models are labelled and marked with white filled circles.	59
4.5	Same as Fig. 4.4 for the the intermediate and secondary caustic configuration. .	59
4.6	Map of the value of t_E in the (d, q) plane for converged models at each grid point, superimposed on the χ^2 map, zoomed in on the close regime part of parameter space. Contours lines (orange) are labeled with their corresponding value of t_E while χ^2 contour lines are plotted at $\Delta\chi^2 = 5, 20, 50, 100, 250$ and filled with gradual shades of blue. The dashed curve is the separation between the close and intermediate regimes. The models of Table 4.2 are labelled and marked with white filled circles.	60
4.7	Best-fitting binary lens model C_s with residuals and a zoom on the anomaly (left inset). Data points are plotted with $1-\sigma$ error bars. The trajectory of the source in the lens plane with the caustics is plotted as an inset in the top right corner of the figure, with the primary lens component located at the coordinate system's origin. Axes for the insets are x and y positions in the lens plane, in units of the Einstein radius. Caustics are very small for the values of d and q of model C_s and are difficult to see on the plot.	63
4.8	Same as Fig. 4.7 for model I	64
4.9	Same as Fig. 4.7 for model C_c	64
4.10	Same as Fig. 4.7 for model W_c	65
4.11	Colour - Magnitude diagram of the field. The target OGLE-2007-BLG-472 is shown as a black triangle at $(V - I, I) = (2.43, 15.61)$. The position of the deblended source for each model is labeled and indicated by a solid line and a coloured diamond, with the blend for each model also plotted as a diamond in the same colour.	67
4.12	Probability densities for the mass of the primary lens star and the fractional distance D_L/D_S , for a lens in the disc (left side) and a lens in the bulge (right side). The values quoted in Tables 4.2 & 4.2 are the median value and the limits of the 68.3% confidence interval. On each plot, the probability densities are plotted for model C_s (red), model C_c (green), model I (dark blue), and model W_c (light blue).	69
4.13	Mass-distance diagram showing the constraint on the lens mass from the source size, given by Eq. (4.2), for each model. The curves are labeled with the name of the model to which they correspond.	70
5.1	Prior maps $P(s_{in}, s_{out})$ and caustics for a close configuration (here $d = 0.5, q = 0.1$), with an uninformative prior on all other parameters. <i>Top</i> : Prior map for the central caustic. <i>Bottom</i> : Prior map for the (upper) secondary caustic (upper or lower secondary caustics have identical prior maps). Whiter shades represent a higher prior probability density.	77

5.2	Same as Fig. 5.1 for an intermediate caustics configuration ($d = 1.1, q = 0.1$).	78
5.3	Same as Fig. 5.1 for a wide caustics configuration ($d = 2.0, q = 0.1$). <i>Top</i> : Priors for the central caustic. <i>Bottom</i> : Prior for the upper secondary caustic. The prior map for the lower secondary caustic is identical.	79
5.4	A histogram of microlensing event timescales for the OGLE 2006 and 2007 observing seasons, with the model distribution of Wood & Mao (1995) shown as a solid black line. Taken from Cassan et al. (2009).	80
5.5	Maps of the prior $P(s_{in}, s_{out})$, for close caustics (here $d = 0.5, q = 0.1$) with the prior on t_E advocated by Wood & Mao (2005). <i>Top</i> : Priors for the central caustic. <i>Bottom</i> : Prior for the secondary caustic (upper or lower secondary caustics have identical prior maps). Whiter shades represent a higher prior probability density.	81
5.6	Same as Fig. 5.6 for an intermediate caustic configuration ($d = 1.1, q = 0.1$).	82
5.7	Same as Fig. 5.5 for a wide caustics configuration ($d = 2.0, q = 0.1$). <i>Top</i> : Priors for the central caustic. <i>Bottom</i> : Prior for the secondary caustic.	83
5.8	Flow chart of our proposed automatic fitting algorithm for caustic-crossing binary-lens microlensing events. The example here is the flow for a close caustics case. Each caustic is explored separately, and each sub-box of the $P(s_{in}, s_{out})$ map is labeled with a number (if the sub-box has regions of non-zero prior), or N (if the prior is null everywhere in the sub-box). In this case, the central caustic prior map has 10 non-zero prior sub-boxes, while the secondary caustic prior maps have 6.	84
6.1	Histogram for the distribution of disc radii of the combined observational samples (diamonds) for discs detected at both 24 and 70 μm , shown with \sqrt{N} error bars, and the distributions for the entire model population (red, dashed line) and the model population that could be detected at 24 and 70 μm (blue, dotted line).	99
6.2	<i>Top</i> : total 70 μm flux divided by the photospheric flux against age. The model population is shown with small dots, and the observations of Trilling, Hillenbrand and Beichman are shown as filled dots, triangles and squares respectively. Disks detected at 70 μm with $< 3\sigma$ confidence are shown with open diamonds. Red symbols indicate disc emission detected at $> 3\sigma$ at both 24 μm and 70 μm . The horizontal dashed line separates the populations with small and large excess flux. <i>Bottom</i> : fractional populations of stars with different flux ratios for different age bins (0.03 – 0.3 Gyr, 0.3 – 2 Gyr, 2 – 4 Gyr, 4 – 10 Gyr). Small ($R_{70} = F_{70,tot}/F_{70,phot} < 15$) and large excesses ($R_{70} > 15$) are shown as crosses and squares respectively. Model predictions are connected with lines and observed values (Trilling et al., 2008; Hillenbrand et al., 2008; Beichman et al., 2006a) are plotted with \sqrt{N} error bars. On both panels, the dotted vertical lines indicate the limits of the age bins.	100

- 6.3 *Top*: total $24\mu\text{m}$ flux divided by the photospheric flux against age. Symbols are the same as for Fig. 6.2. Horizontal dashed lines separate the populations with small and large excess flux. *Bottom*: fractional populations of stars with different flux ratios for different age bins (0.03 – 1 Gyr, 1 – 3 Gyr, 3 – 13 Gyr). Small ($R_{24} = F_{24,tot}/F_{24,phot} < 1.25$), and large excesses ($R_{24} > 1.25$) are shown as crosses, triangles and squares respectively. Model predictions are connected with lines and observed values (Trilling et al., 2008; Hillenbrand et al., 2008; Beichman et al., 2006a) are plotted with \sqrt{N} error bars. On both panels, the dotted vertical lines indicate the limits of the age bins. 101
- 6.4 Same as Fig. 6.2, but with the parameters found for A stars (Wyatt et al. 2007). This illustrates the significant difference between the evolutionary timescales corresponding to the models found in this analysis compared to those found for A stars; this significant time evolution is not seen in observations of discs around FGK stars. 104
- 6.5 Same as Fig. 6.3, but with the parameters found for A stars (Wyatt et al. 2007). This illustrates the significant difference between the evolutionary timescales corresponding to the models found in this analysis compared to those found for A stars; the lack of a significant time evolution, not seen in observations of discs around FGK stars, is most obvious at $24\mu\text{m}$, where no significant excess is predicted beyond ~ 500 Myr. 105
- 6.6 Dust luminosity $f = L_{IR}/L_*$ plotted as a function of disc radius R_{dust} . The top panel shows the whole model population, while the bottom panel only has the model population that could be detected at 24 and $70\mu\text{m}$. The model population is shown with small dots, and data symbols are the same as for Fig. 6.3. Detection thresholds at 24 and $70\mu\text{m}$ for an F0 dwarf are indicated by solid lines. Lines showing the maximum possible fractional luminosity for an F0 dwarf are shown as dashed lines for 2 and 10 Gyr. The dotted line indicates the limit, given by Eq. (6.11), where Poynting-Robertson drag becomes important for an F0 dwarf. 109
- 6.7 Dust luminosity $f = L_{IR}/L_*$ plotted as a function of age. The model population is shown with small dots, and data symbols are the same as for Fig. 6.3. For $< 3\sigma$ detections, the fluxes plotted are maximum values. *Top*: Entire model population, and all the $70\mu\text{m}$ excess detections, with objects detected at both 24 and $70\mu\text{m}$. Also shown are theoretical evolutionary tracks for discs of radii 10 (solid line) and 100 (dashed line) AU, and for initial disc mass of 3, 30 and $300 M_{\oplus}$ (with the higher masses corresponding to the lines with higher early-age fractional luminosities), around an F0 star. *Bottom*: Same with only the populations that could be detected at both 24 and $70\mu\text{m}$ plotted. Also plotted on the bottom panel are power-law fits to the model (dashed line) and the data (solid line). 111
- 6.8 Disk radius plotted as a function of age. The model population is shown with small dots, and data symbols are the same as for Fig. 6.3. Only the populations which could be detected at both 24 and $70\mu\text{m}$ are plotted. Also shown are power-law fits to the model (dashed line) and observed (solid line) populations. 112

6.9 Histogram of the dust fractional luminosity for the observed population with an excess at both 24 and 70 μm (solid line), the model population (dashed, red) and the model population with a Gaussian smoothing of 1 dex (dotted, blue). Only the model population which could be detected at both 24 and 70 μm is plotted. 114

List of Tables

3.1	OGLE-2006-BLG-180: MCMC best fit parameters compared to the OGLE published values	35
3.2	OGLE-2006-BLG-397: MCMC best fit parameters compared to the OGLE published values	37
3.3	MCMC best fit parameters compared to the published values of Beaulieu et al. (2006); all are in good agreement.	48
4.1	OGLE-2007-BLG-472 datasets and error bar rescaling factors.	58
4.2	Best-fitting binary lens model parameters for OGLE-2007-BLG-472. The blending factor $g(I) = F_B(I)/F_S(I)$ is given for the OGLE data (I -band). The error bars were rescaled for each telescope by the factor given in Table 4.1, which lead to the rescaled χ^2 indicated here. Physical parameters are also given for each model, for the case of a lens in the disc, and a lens in the bulge. These were calculated using the procedure detailed in Sec. 4.3.7	62
6.1	List of sources with $70\mu\text{m}$ excess only. Flux ratios and limits are given as total flux divided by photospheric flux. $R24$ is defined as $R24 = F24/F24_*$, and similarly, $R70 = F70/F70_*$. $C24$ and $C70$ are the $3 - \sigma$ calibration limits at 24 and $70\mu\text{m}$, defined as $C_\lambda = 1 + 3\sigma_{*,\lambda}/F_{*,\lambda}$	96
6.2	List of sources with 24 and $70\mu\text{m}$ excesses. Flux ratios and limits are given as total flux divided by photospheric flux. $R24$, $R70$, $C24$ and $C70$ as defined as in Table 6.1. Stellar properties are either published values found by fitting Kurucz models to the data when these were available, or calculated using Schmidt-Kaler relations for main sequence stars.	97

1

Searching for other worlds

In the history of humanity, the past century and a half will no doubt be remembered as a turning point in the development of scientific knowledge. In that short time, we discovered that everything is made up of little things called atoms, that the Universe is expanding, that genetic information is passed on in tiny double helices located in almost each cell of living organisms, and countless other things which revolutionised our understanding of the Universe and its origins. These are questions which have been pondered by humans for millennia, yet in 150 years or so, more answers than ever before were found. But with these answers came even more questions. In the mid-1990s, the first discovery of a planet outside our Solar System provided the first glimpse of an answer to one of the most basic questions in science: is there life beyond planet Earth? By allowing us to see how common or how rare planets like ours are, we can perhaps understand just how special, or common, life on Earth is.

Since this first discovery of an exoplanet orbiting a main-sequence star in 1995 (Mayor & Queloz, 1995), over 450 more extrasolar planets have been discovered, and new ones are being found at an increasing rate, allowing us to slowly build an answer to that question, and

to determine whether Earth-like rocky and cool planets are common, or if they are rare and perhaps most planets are Jupiter-like gaseous giants. In essence, the fundamental question extrasolar planet searches are trying to answer is *How special are we?*

The first person to suggest that the planets of the Solar System might not be the only ones in the Galaxy was Giordano Bruno, who realised that each of the stars in the night sky was in fact not very different from the Sun. For this revolutionary belief, as well as his firm defense of the Copernican heliocentric model, Bruno was declared a heretic by the Church and burnt at the stake in 1600, a fate only narrowly avoided by Galileo and Copernicus. Just nine years later, Galileo, using his refracting telescope, observed the Milky Way and found that it wasn't nebulous as previously thought, but that it was in fact made up of a myriad of stars, suggesting that not only could there be planets orbiting other stars, but with stars being so numerous, such planets could exist in high numbers in our Galaxy. Over the next few centuries, many scientists and philosophers were to give serious consideration to the possible existence of other worlds, some taking care to distance themselves from Bruno's theories, either for fear of landing in trouble with religious authorities, or from a desire to not be associated with Bruno's other ideas. In parallel to the debate on the plurality of worlds, the possible existence of extra-terrestrial life was also discussed by many thinkers, with Thomas Dick even going so far as to estimate the population of the different planets in the Solar System, basing his calculation on the population density of 18th-century England (Crowe, 2008). Four centuries would see a complete paradigm shift from the ancient principles of Aristotelian cosmology to a heliocentric view of the Solar System.

1.1 First detections and methods

Although some claims of planet detections were made from the middle of the 19th century, mostly based on orbital anomalies of stars, the first hint that extrasolar planets could really be detected came with the detection of the first circumstellar discs in the 1980s, around Vega (Harvey et al., 1984) and β Pictoris (Smith & Terrile, 1984). The first serious planet discovery is generally held to be that of planets around pulsar PSR 1257+12 Wolszczan & Frail (1992). Three years later, in 1995, the first planet around a main-sequence star was detected around 51 Peg (Mayor & Queloz, 1995), opening the door to a complete revolution of our view of planetary systems and indeed of our own place in the Universe. This discovery was made using the radial velocity method, which measures Doppler shifts of stellar spectral lines to infer the

presence of a secondary companion orbiting the star. To this day, this method accounts for the largest number of planet discoveries, with over two thirds of all known exoplanets found with this technique. Alongside this method, there are several other ways of finding extrasolar planets. The **transit** method relies on a planetary companion causing a drop in the observed brightness of its stellar host as the planet passes in front of it; **astrometry** planet searches look for shifts in the position of stars as they orbit the centre of mass of a star-planet(s) system. **Direct imaging** can lead to the detection of planets by looking at their infrared radiation, while **timing** examines pulses of radio emission from rotating pulsars and other suitable clocks, and looks for light travel time delays caused by the presence of planetary companions to the pulsar.

Gravitational microlensing, the method with which this thesis is concerned, relies on an effect predicted by general relativity: when a background *source* star passes behind a foreground *lens* star, the light of the source is deflected by the gravitational field of the lens, causing the latter to act a bit like a focusing lens in classical optics. The important difference is that for a gravitational lens, light is bended more the closer it passes to the lens; this means that a gravitational lens doesn't have a single focal point, but rather a focal line. From an observer's point of view, this means that several images of the source are formed. Although the separation is generally too small to be resolved with current technology (it is usually \sim a few mas), one can observe the effect of the light deflection by monitoring the source's brightness. As the source, lens and observer move in and out of alignment, the source appears to brighten and fade, peaking when the alignment is optimal. When the lens is made of more than one object (e.g. in the case of a binary star or of a planetary system), the light deflection is perturbed by the gravitational field of these additional objects. Monitoring the brightness of the source star as it is being lensed by these systems, and careful modelling of the resulting lightcurve, allows us to infer the presence of planets and to determine some of their properties.

Many of the different methods are complementary in their capabilities, with radial velocity and transit searches being most sensitive to large planets with small orbital radii (so-called *Hot Jupiters*), although recent missions, in particular the space telescope *Kepler* launched in March 2009, will extend this sensitivity range down to smaller planets at larger orbital radii. Astrometry is sensitive to large planets at large orbital radii, as is direct imaging, since the infrared emission of a planet at a small orbital radius would be hard to detect due to the glare of its host star. Gravitational microlensing, on the other hand, is currently the only method

capable of detecting low-mass planets at large orbital radii from the ground. The limits and complementarity of the various methods are obvious when looking at a plot of discovered extrasolar planets (Fig. 1.1).

1.2 A brief history of microlensing

After initial work going back as early as Newton's *Opticks* (1704), the deflection angle of light due to the presence of a massive body on its path was first calculated by the German physicist Johann von Soldner in the early 19th century (Soldner, 1804). Using Newtonian mechanics and assuming that light was made of corpuscles, he obtained an expression for the deflection angle in terms of the speed of light, the mass of the deflecting body and the impact parameter. In 1911, Einstein derived the same expression using his theory of special relativity, this time with no assumptions regarding the nature of light itself (Einstein, 1911). However in 1916, in light of his recently developed theory of General Relativity, Einstein realised that the deflection angle should be twice the value that he had found using special relativity, due to the curvature of spacetime in the vicinity of a massive body (Einstein, 1916). This result was famously confirmed by Eddington when he measured the positions of stars during a solar eclipse. Subsequent work showed the variety of objects that could act as gravitational lenses, from galaxies (Zwicky, 1937) to quasars and stars. In 1969, Byalko first suggested looking for lensing of quasars by stars in the Milky Way (Byalko, 1969). However it was not until 1979 that the first observations of lensing effects occurred, with lensing of extra-galactic quasars observed by astronomers at the Kitt Peak National Observatory (Walsh et al., 1979). It would take another 10 years for the first Galactic lensing observations to be made.

Lensing of a star by another star was discussed as early as in 1920 by Eddington (1920), and then by Chwolson (1924) and Einstein (1936) but Einstein concluded that the phenomenon could never be observed due to the very small deflection angle that would be caused by the lens star, meaning that the different images of the source star could not be resolved. Although Einstein noted that lensing of the source star should lead to an observed magnification of its luminosity, potentially allowing us to observe microlensing after all, he concluded that the phenomenon still would never be observed because of the high degree of alignment necessary for microlensing to occur (Einstein, 1936). What even Einstein did not foresee was the extraordinary technological developments that would take place in the next few decades, which would eventually allow us to observe microlensing despite the probability of the re-

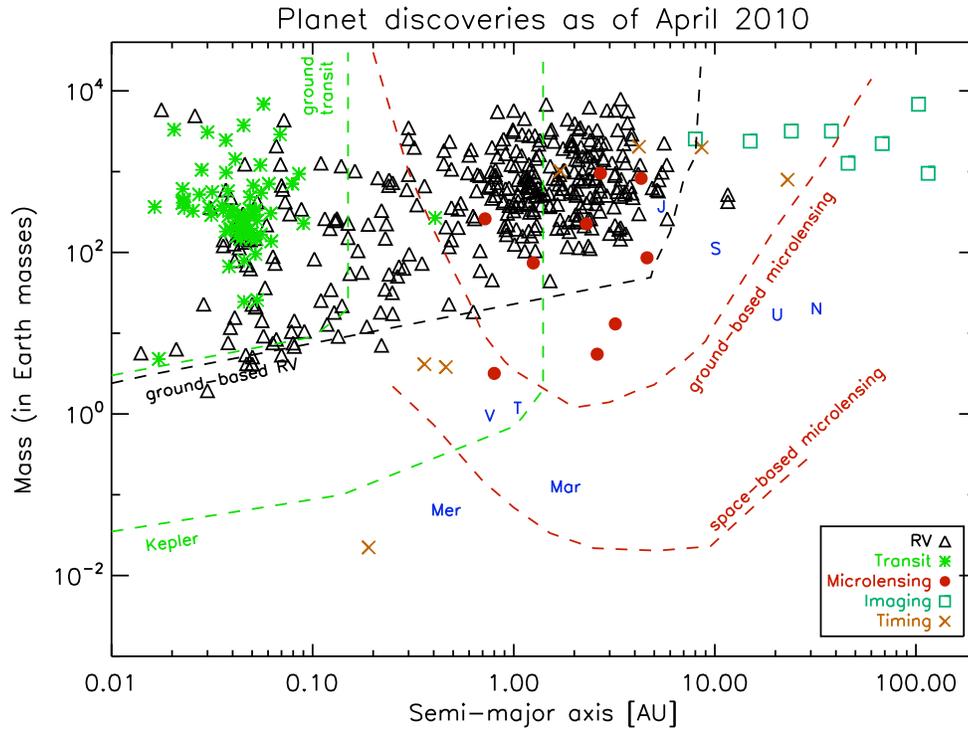


Figure 1.1: Extrasolar planet discoveries as of 11 March 2010, plotted in the (semi-major axis, mass) plane. Planets discovered with the radial velocity (RV), transit, microlensing, imaging and pulsar timing methods are plotted with different symbols and colours, detailed in the bottom right box. The Solar System's planets are plotted in blue and labeled with their initial. Method sensitivities are given for the RV, transit and microlensing methods, with a distinction between space-based and ground-based sensitivities for the latter two. Data from the Exoplanet Encyclopedia (<http://exoplanet.eu>), sensitivities from Keith Horne (private communication).

quired alignment occurring only being around one in a million: as technology developed, our capability to monitor the brightness of tens of millions of stars meant that observing such rarely-occurring events became possible. Chang and Refsdal observed the effect in 1979, the first observation of a gravitational microlensing event (Chang & Refsdal, 1979).

1.2.1 The first microlensing campaigns

In 1986, Bohdan Paczyński proposed that microlensing be used to look for baryonic dark matter in the galactic halo (Paczynski, 1986). From this idea, two microlensing collaborations were formed by astronomers and particle physicists: the French collaboration EROS (*Expérience de Recherche d'Objets Sombres*) in 1990 and the British-Australian collaboration MACHO (*Massive Compact Halo Objects*) in 1991. Both collaborations observed the Magellanic Clouds, the location of which, near the Galactic Halo, provided a good supply of source stars whose magnification by MACHO lens objects could be detected. This led to the detection of three microlensing events, reported separately by the two groups in *Nature* in 1993 (Alcock et al. 1993; Aubourg et al. 1993). In 1991, Mao & Paczyński showed that microlensing observations could help detect low-mass objects such as brown dwarfs and extrasolar planets by monitoring the Galactic Bulge (Mao & Paczyński, 1991). In 1992, Udalski founded the OGLE (*Optical Gravitational Lensing Experiment*, Udalski et al. 1992) collaboration, which started routine monitoring of the Galactic Bulge. MACHO also started observing towards the Galactic Bulge, as did the DUO experiment (Alard et al., 1995) and EROS-II in 1995 and 1996 respectively. Microlensing then appeared to be a good method to use to try to detect the first extrasolar planet. Gould & Loeb (1992) described the potential of microlensing to detect extrasolar planets and made estimates of the detection probabilities for different star-planet systems. Their conclusions were encouraging enough to justify considerable efforts to find extrasolar planets by microlensing. However, it was also clear that the ongoing microlensing monitoring campaigns could not serve this purpose as they were, because of the high photometric accuracy and intensive observations needed to detect a photometric perturbation due to a planet orbiting a lens star.

In order to solve this problem, it was decided in 1995 to use a number of small telescopes, which would be located at different longitudes so as to enable continuous photometric observations, and in the Southern Hemisphere in order to observe fields in the Galactic Bulge. These small photometric *follow-up* telescopes were to complement the wide-field *alert* telescopes,

which were to monitor many stars and refer events of interest to the follow-up telescopes. In this context, the PLANET (Probing Lensing Anomalies NETWORK) and GMAN (Global Microlensing Alert Network) collaborations were founded in 1995, followed in 1997 by MPS (Microlensing Planet Search), and the Japan-New Zealand collaboration MOA (Microlensing Observations in Astrophysics) in 1998. Furthermore, μ FUN (Microlensing Follow-Up Network) was formed in 2001 after a schism within PLANET. At present, the OGLE and MOA collaborations still lead the wide-field survey effort, while current follow-up teams include PLANET and μ FUN, the MiNDSTEP consortium, as well as the robotic telescopes of RoboNet and LCOGT (Las Cumbres Observatory Global Telescopes Network).

The first extrasolar planet discovered using gravitational microlensing, a planet of 2.6 ± 0.8 Jupiter masses, was found in 2003 (Bond et al., 2004). Since then, several planets have been found using the method, including the first cool, rocky extrasolar planet ever found, with a median mass estimate of 5.5 times the mass of the Earth, and a median orbital radius estimate of 2.1 AU (Beaulieu et al., 2006). Other discoveries have included several sub-Neptune-mass planets, and a 2-planet Jupiter/ Saturn analog (Gaudi et al., 2008).

These are exciting times for planet-hunting: during my Ph.D., the number of known extrasolar planets has doubled, plans have been made for the deployment of several networks of robotic telescopes, including LCOGT and KMTNet (Korea Microlensing Telescope Network), the launch of planet-hunting space missions such as CoRoT and Kepler, and of The Herschel Space Observatory, which is expected to revolutionise our view of the Solar System. Alongside these missions, extrasolar planet research is coming of age, and the focus is slowly moving from discovering single planets to using the information collected over the last fifteen years to consider the bigger picture: understanding planet populations, and our place in the Universe.

2

Microlensing formalism

2.1 The single-lens case

I start by deriving the main equations for a case where a single (source) star is lensed by another (lens) star, leading in general to the formation of two images of the source. By using the Schwarzschild metric to solve the Einstein Field Equations in a vacuum

$$ds^2 = c^2 \left(1 + \frac{2\phi}{c^2}\right) dt^2 - \left(1 + \frac{2\phi}{c^2}\right)^{-1} dr^2 - r^2 d\Omega, \quad (2.1)$$

where ϕ is the Newtonian gravitational potential,

$$\phi = -\frac{GM}{r}, \quad (2.2)$$

one can derive the deflection angle due to light passing in the vicinity of a massive body. In the weak-field, small angle limit, which is a good approximation in the framework of

microlensing, the deflection angle α is found to be

$$\alpha = \frac{4GM}{r_{\min}c^2} = \frac{2R_S}{r_{\min}}, \quad (2.3)$$

where $R_S = \frac{2GM}{c^2}$ is the Schwarzschild radius, and r_{\min} is the closest approach to the massive body. A diagram of such a microlensing event is shown on Fig.2.1, with α, r_{\min} and other relevant quantities labeled. Using this diagram and simple geometrical considerations, along with Eq.(2.3), we can write down the lens equation for a point-source-point-lens (PSPL) microlensing event, i.e. for the case where the light from one point-like object is deflected by the gravitational field of one other point-like object. With the small angle approximation, valid when $\alpha \ll \pi$, we have

$$\theta D_S = \alpha(D_S - D_L) + \beta D_S, \quad (2.4)$$

where D_S and D_L are the distance to the source and the lens respectively, and the angles α, θ and β are the angles labeled on Fig. 2.1.

Rearranging this equation, we find

$$\beta = \theta - \left(\frac{D_S - D_L}{D_S} \right) \alpha, \quad (2.5)$$

which we can in turn rewrite as

$$\beta = \theta - \frac{\theta_E^2}{\theta}, \quad (2.6)$$

where we have used Eq.(2.3) and the relation $r_{\min} = \theta D_L$ (Fig. 2.1). The *angular Einstein ring radius* θ_E is then defined as

$$\theta_E = \sqrt{\frac{4GM}{D_S c^2} \frac{D_S - D_L}{D_L}}. \quad (2.7)$$

If we then define the normalised image and source positions x and u as

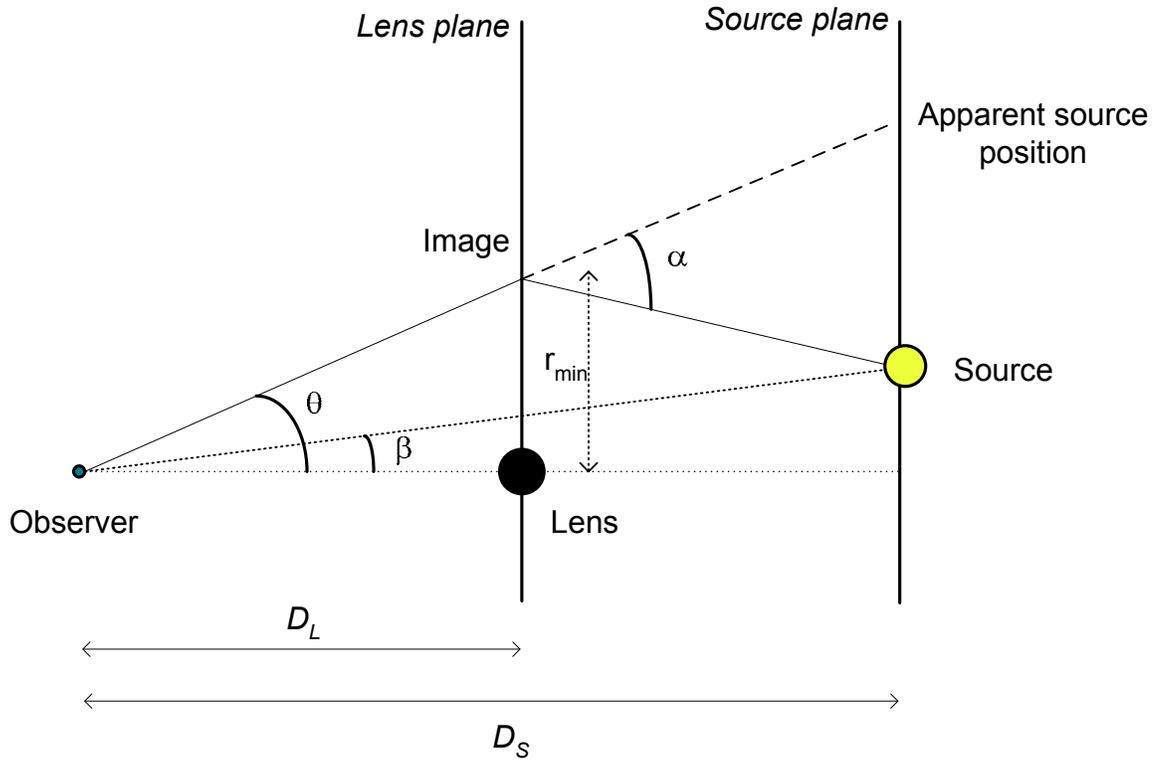


Figure 2.1: Geometry of a point-source point-lens microlensing event. Labeled quantities are the deflection angle α , the closest approach r_{\min} , the angular closest approach on the lens plane θ , the angular source position β , as well as the distances to the lens and to the source, D_L and D_S .

$$x = \theta/\theta_E, \quad (2.8)$$

$$u = \beta/\theta_E, \quad (2.9)$$

this is equivalent to considering the positions of the source and the images on the lens-plane (defined on Fig.2.1). Eq.(2.6) can then be rewritten as

$$u = x - \frac{1}{x}. \quad (2.10)$$

We can solve this equation to find the positions of the images produced by the lens. For the case where $u=0$, i.e. when the source and the lens are both perfectly aligned with the observer, we see that Eq.(2.10) becomes $x - \frac{1}{x} = 0$, which has the solution $x^2 = 1$. This is the equation of circle of radius 1, meaning that the image will be a circle of radius θ_E (since x is normalised by θ_E), called an *Einstein Ring*. When the observer, lens and source are not perfectly aligned, $u \neq 0$ and Eq.(2.10) can be rewritten as $x^2 - ux - 1 = 0$ which has two solutions, corresponding to the two image positions:

$$x_{\pm} = \frac{1}{2} \left(u \pm \sqrt{u^2 + 4} \right). \quad (2.11)$$

The magnification of the source is the sum of the magnifications due to each separate image, so $A = A_+ + A_-$. The magnification of each image is given by

$$A_{\pm} = \left| \frac{x_{\pm}}{u} \frac{\partial x_{\pm}}{\partial u} \right|, \quad (2.12)$$

and hence the total magnification of the source is

$$A(u) = \frac{u^2 + 2}{u\sqrt{u^2 + 4}}. \quad (2.13)$$

Since $u \equiv u(t)$ as u describes the trajectory of the source in the lens plane with time, $A(u)$ describes the magnification of the source star changing over time with u . Indeed, u can be written as

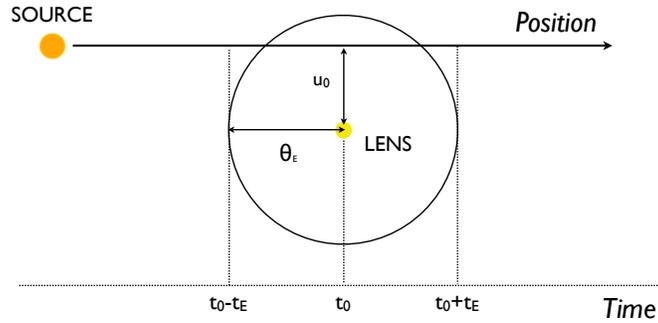


Figure 2.2: The PSPL model parameter: the source takes a time t_E to cross an Einstein ring radius θ_E , while it reaches closest approach at time t_0 , at which point it is an angular distance u_0 from the lens.

$$u(t) = \sqrt{u_0^2 + \left(\frac{t - t_0}{t_E}\right)^2}, \quad (2.14)$$

where u_0 is the angular separation of the source and the lens at closest approach in units of θ_E , t_0 is the time of closest approach and t_E is the Einstein Ring radius crossing time (the time it takes the source to move by an angular distance θ_E), which provides a characteristic timescale for the event. These three parameters are illustrated on Fig. 2.2; they completely describe the magnification of the source when its light rays are being deflected by the lens, as they move in and out of alignment.

In order to translate this magnification, which depends only on the event's geometry, into an observed flux, one must take into account light contamination from neighbouring sources (*blending*) by adding additional parameters. These will differ for each telescope tracking a given microlensing event. The total measured flux at time t at a telescope i is then given by $F(t, i) = f_{s,i}A(t) + f_{b,i}$, where $f_{s,i}$ and $f_{b,i}$ are the source and blend flux respectively.

2.2 Multiple lens case

2.2.1 Deriving the multiple lens equation

In the case of multiple lenses, it is convenient to use the complex formalism developed by Witt (1990), which describes the whole microlensing event in the lens plane. The positions of the images and source in the lens plane can then be expressed by complex numbers \mathbf{x} and \mathbf{y}

respectively. Schneider & Weiss (1992) derived the multiple lens equation by minimising the Fermat potential,

$$\phi(\mathbf{x}; \mathbf{y}) = \frac{1}{2}(\mathbf{x} - \mathbf{y})^2 - \varphi(\mathbf{x}), \quad (2.15)$$

where $\varphi(\mathbf{x})$ is the *deflection potential*, defined by its relation to the deflection angle at \mathbf{x} , $\boldsymbol{\alpha}(\mathbf{x}) = \nabla\varphi(\mathbf{x})$. For a collection of N masses at position \mathbf{d}_i of mass $\mu_i M$, where $\mu_i = m_i/M$ is the fractional mass of the i^{th} component of the lens, the deflection potential is

$$\varphi(\mathbf{x}) = \sum_{i=1}^N m_i \ln|\mathbf{x} - \mathbf{d}_i|. \quad (2.16)$$

Using Eqs.(2.15) and (2.16), minimising the Fermat potential, in other words solving $\nabla_{\mathbf{x}}\phi(\mathbf{x}; \mathbf{y}) = 0$, yields the multiple lens equation:

$$\mathbf{x} - \mathbf{y} - \sum_{i=1}^N m_i \frac{\mathbf{x} - \mathbf{d}_i}{|\mathbf{x} - \mathbf{d}_i|^2} = 0. \quad (2.17)$$

Unlike the PSPL lens equation, in general this equation cannot be solved analytically. There are various ways to solve it, including turning it into a complex polynomial equation. This can be done by eliminating \mathbf{x}^* , the complex conjugate of \mathbf{x} , from Eq.(2.17), and keeping the solutions of this equation which are also solutions of the lens equation. For a binary-lens ($N=2$) case, this yields a fifth-order complex polynomial equation, as shown by Witt & Mao (1994). In the PSPL case, the magnification could simply be calculated through Eq. (2.12), and for multiple lenses, we can simply extend this to express the mapping between image and source positions \mathbf{x} and \mathbf{y} . This is given by the Jacobi matrix, the elements of which are

$$J_{ij} = \frac{\partial y_j}{\partial x_i}, \quad (2.18)$$

and the magnification A_n at image n located at x_n is given by

$$A_n = \frac{1}{|J_{ij}(x_n)|}, \quad (2.19)$$

where $|J_{ij}(x_n)|$ is the determinant of the Jacobi evaluated at x_n . The total magnification is just the sum of the magnifications of each image,

$$A_{\text{tot}} = \sum_{n=1}^{N_i} A_n, \quad (2.20)$$

where N_i is the number of images (see below). The main difficulty lies in calculating the image positions, as this requires solving the multiple lens equation (Eq.2.17).

2.2.2 Critical curves and caustics

Looking at Eq.(2.19), it is clear that when the determinant of the Jacobi J_{ij} vanishes, the magnification diverges. The loci of infinite magnification are called *critical curves* (in the lens plane) or *caustics* (in the source plane). In the single lens case, we can see from Eq.(2.13) that this happens when $u = 0$, so the caustic is a single point, and it corresponds to a perfect alignment of the observer, the lens and the source. The critical curve is found by setting $u = 0$ in Eq.(2.11), which yields $x^2 = 1$, meaning that the critical curve is the Einstein Ring in the PSPL case.

In multiple lens cases, caustics and critical curves are much more complicated and are essential to determining the features of the magnification curve of the source. They determine the number of images of the source, which depends on the source's position relative to the region enclosed by the caustic curve. Witt (1990) showed that the number of images is always odd and at most equal to $N^2 + 1$ in the case of a lens with N components. There is a large magnification gradient near the location of the caustic lines, meaning that a source approaching or crossing a caustic will be highly magnified as it moves in the vicinity of the caustics. In theory, since the magnification diverges near caustics, a perfect point source would be infinitely magnified when it crosses a caustic line; in practice a source is never perfectly point-like, and the magnification, while it can be high, is lessened due to finite source size effects (see below).

Much work has been done on the topology and properties of caustics for the case of a binary lens (i.e. the $N=2$ case), notably by Erdl & Schneider (1993), who studied the different possible cases for caustics topology and determined the locus of separation between the different regimes.

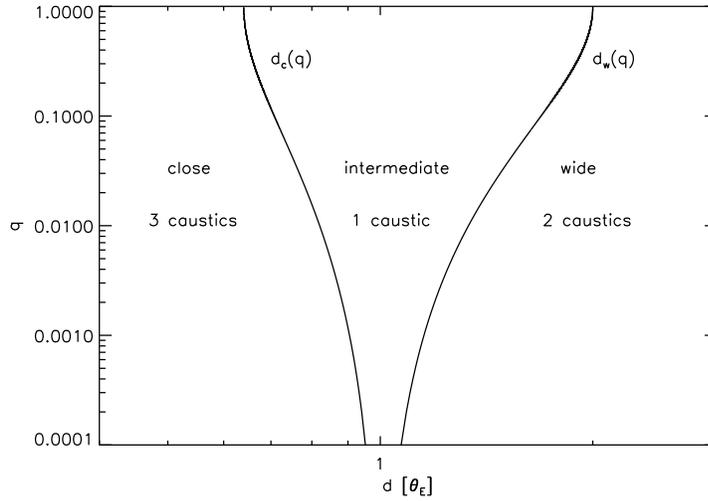


Figure 2.3: Separation lines in the logarithmic (d, q) plane between the caustic regimes, after Erdl & Schneider (1993). Each region is labeled with the corresponding regime name and number of caustics.

Caustics and critical curves produced by binary lenses can be described by the mass ratio of the two lens components q and their separation d . Three regimes can be distinguished, characterised by ranges of values for q and d : *close*, *intermediate* and *wide* binaries. Erdl & Schneider (1993) derived expressions for the limiting values of d separating the different regimes as a function of the mass ratio q :

$$d_c^8 = \frac{(1+q)^2}{27q} (1-d_c^4)^3, \quad (2.21)$$

$$d_w = \sqrt{\frac{(1+q^{\frac{1}{3}})^3}{1+q}}, \quad (2.22)$$

where d_c is the separation value between the close and intermediate regimes, while d_w is the separation value between the intermediate and wide regimes.

A plot of the separation lines between the different regimes (Eqs.2.21-2.22) is shown on Fig. 2.3. The number of caustics, as well as their position, shape and size depends solely on the values of d and q . An example set of caustics for each regime is shown on Fig. 2.4. In the wide configuration, there are two caustics, both on the line joining the lens components; in the intermediate configuration there is one caustic, while in the close regime, there are three caustics: a *central* caustic and two *secondary* caustics off the lens system's axis of symmetry.

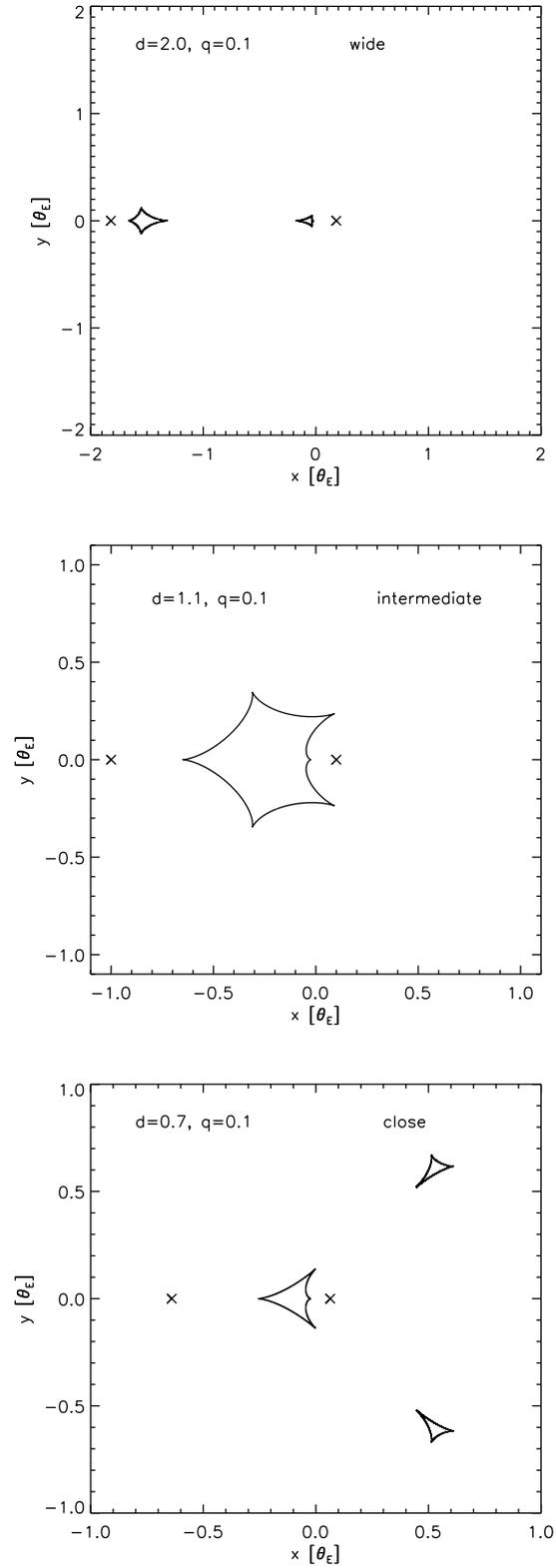


Figure 2.4: Caustic topology for the wide (upper panel), intermediate (middle panel) and close (lower panel) regimes. The d and q values corresponding to each panel are inset in the top left corner. Crosses indicate the position of each lens component, with the more massive component closer to the origin (which corresponds to the centre of mass of the lens system).

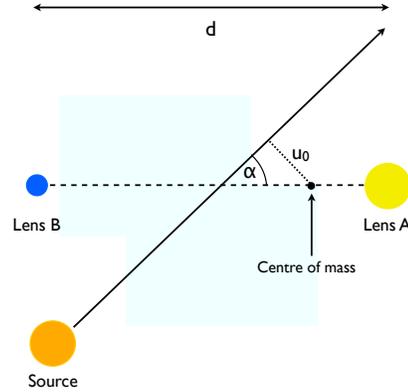


Figure 2.5: The trajectory angle parameter α , defined as the angle between the source trajectory and the line joining the two lens components on the lens plane.

2.2.3 Final binary-lens parametrisation

So far we have encountered three PSPL parameters (t_0 , t_E and u_0) and two parameters used to describe basic properties of binary lenses (d and q). In the case of multiple lenses, the parameter u_0 is the closest approach to the *barycentre* of the lens system (in the case of a single lens, this is identical to the location of the lens). One more parameter is needed to characterise a point-source binary-lens microlensing event: the source trajectory angle α (not to be confused with the deflection angle α encountered previously), measured between the source trajectory and the line joining the two lens components on the lens plane (see Fig. 2.5).

Therefore, if second-order effects such as finite source size effects, parallax and orbital motion are ignored (see below), we can characterise a static binary-lens event by a set of 6 parameters: t_0 , t_E , α , u_0 , d and q . This is the standard binary-lens parametrisation, and is not without its problems, as will be discussed later. Example lightcurves described by this 6-parameter model are shown on Figs. 2.6-2.11. It is clear from the way these parameters are defined that modelling lightcurves with these only will not yield direct properties of the lensing system. Apart from q , the only other parameter that is directly related to physical properties is t_E , which is a degenerate function of D_S , D_L and M (see Eq. 2.7). However, this degeneracy can be broken under certain conditions, namely when certain second-order effects can be measured.

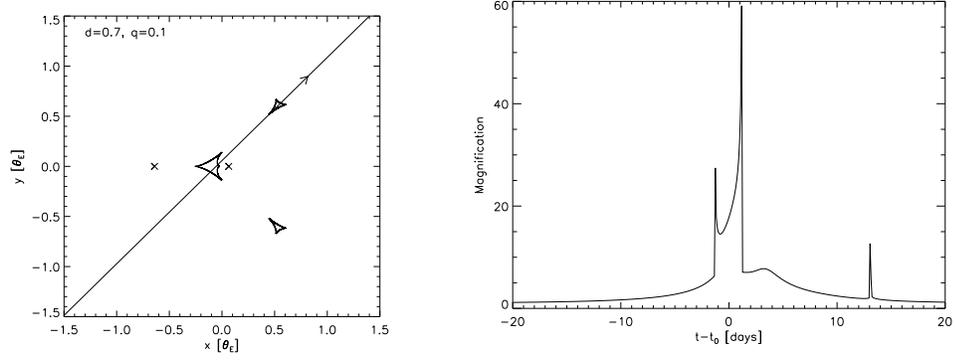


Figure 2.6: An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.04$, $\alpha = 1.0$, $t_E = 15$, $d = 0.7$ and $q = 0.1$. The source crosses both the central and the upper secondary caustic, leading to clear caustic-crossing magnification patterns.

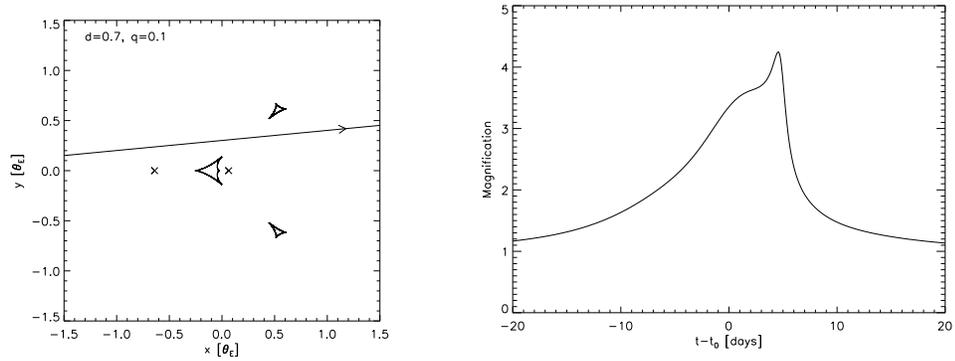


Figure 2.7: An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.3$, $\alpha = 0.1$, $t_E = 15$, $d = 0.7$ and $q = 0.1$. In this case, the source approaches both the central and secondary caustics without crossing them.

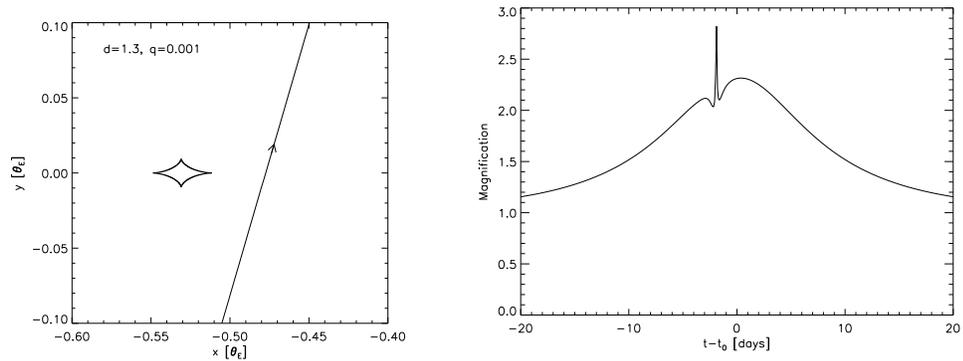


Figure 2.8: An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.46$, $\alpha = 1.3$, $t_E = 15$, $d = 1.3$ and $q = 10^{-3}$. In this case, the source approaches a secondary caustic but doesn't cross it.

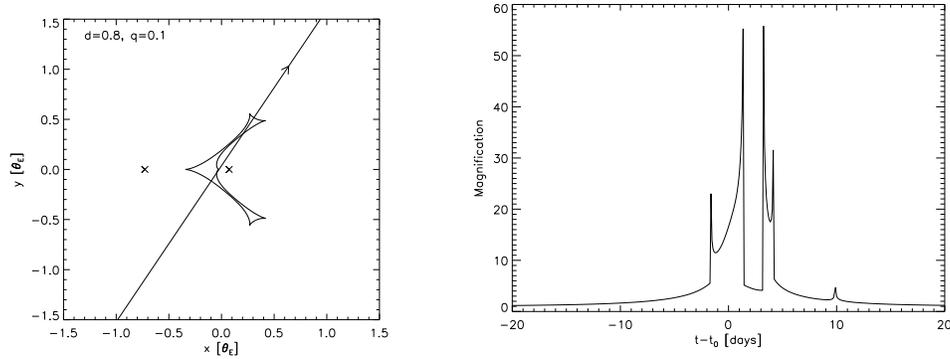


Figure 2.9: An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.02, \alpha = 1.0, t_E = 15, d = 0.8$ and $q = 0.1$. The source crosses an elongated intermediate caustic, exits it and crosses it again.

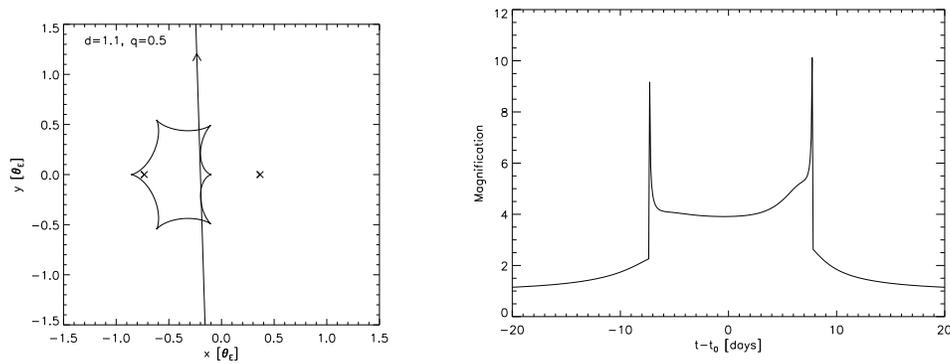


Figure 2.10: An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.2, \alpha = 1.6, t_E = 15, d = 1.1$ and $q = 0.5$. The source crosses an intermediate caustic.

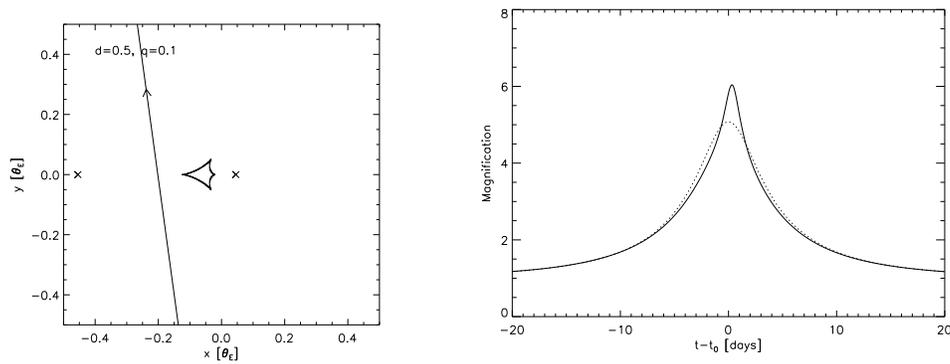


Figure 2.11: An example point source-binary lens (PSBL) lightcurve, with parameters $u_0 = 0.2, \alpha = 1.7, t_E = 15, d = 0.5$ and $q = 0.1$. The source approaches a central caustic, which results in a subtle departure from the PSPL magnification pattern (shown as a dotted line here for comparison).

2.3 Second-order effects

Beside binary-lens effects when the source approaches caustics, the magnification pattern can deviate from a PSPL-like pattern when it is affected by other second-order effects: finite source-size effects, limb-darkening, parallax due to the orbital motion of the Earth, orbital motion of the source (also called *xallarap*) and orbital motion within the lens itself, e.g. the motion of a planet orbiting the lens star in a binary lens.

Source size effects

We can take the effect of having an extended source into account by adding a parameter $\rho_* = \theta_*/\theta_E$, i.e. the angular size of the source, θ_* , in units of θ_E . The result of this is that the magnification will not tend to infinity when the source crosses a caustic, because in the case of a finite source size, the magnification is found by integrating a differential version of Eq.(2.19) over the source, and this integral will not diverge. In practice, this effect needs to be taken into account when the source is located in the vicinity of a large magnification gradient. This happens frequently when the source approaches caustics in binary-lens cases. When this occurs, microlensing essentially allows us to spatially resolve stars in the Galactic Bulge, a remarkable feat which means outdoing the best available telescopes by several orders of magnitude. Taking finite source size effects into account is currently one of the main computational issues in microlensing modelling, as integrating the differential equivalent of Eq.(2.19) is extremely expensive computationally, despite the existence of several approximations designed to speed up calculations (e.g. Gould 2008). Moreover, for most anomalous microlensing events, finite source size effects are so important that they must be taken into account even in the early stages of modelling an event.

Source size effects are illustrated for a PSPL case on Fig. 2.12 and for a binary-lens case on Fig. 2.13.

Limb-darkening effects

Limb-darkening, the radial decrease in intensity of a star as one moves away from the centre of its disc on the sky, can significantly affect the photometry, and hence the modelling, when the magnification gradient is large, i.e. when the source approaches caustics. Indeed, if the source is not uniformly bright, then the differential magnification that takes place as the

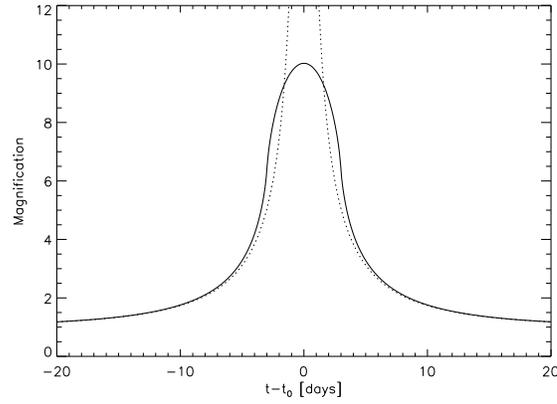


Figure 2.12: The effects of having an extended source in a single-lens model. The extended source-point lens (ESPL) model is shown with a solid line, and the PSPL model is shown with a dotted line. The ESPL model has a source size $\rho_* = 0.2$, and both models have an impact parameter $u_0 = 0.02$.

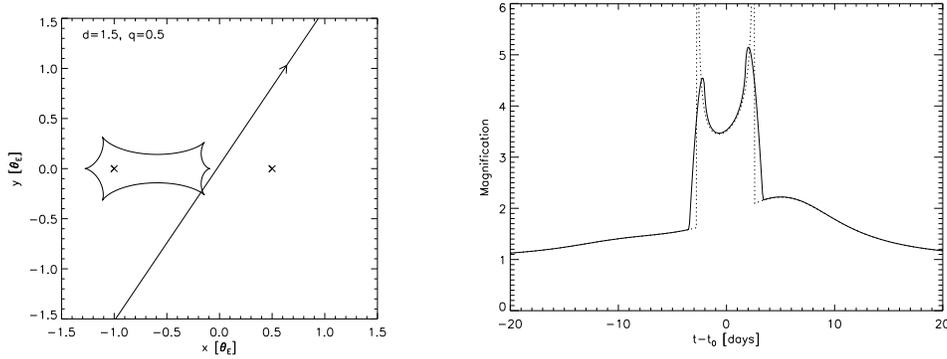


Figure 2.13: The source size effects for binary-lens models. *Left:* The source trajectory and caustics. The positions of the lens components are indicated with crosses. *Right:* The resulting lightcurve. The extended source-binary lens (ESBL) model is shown with a solid line, and the point source-binary lens (PSBL) model is shown with a dotted line. The ESBL model has a source size $\rho_* = 0.04$, and both models have parameters $u_0 = 0.02$, $\alpha = 1.0$, $d = 1.5$ and $q = 0.5$.

source approaches the caustic is strongly influenced by the limb-darkening coefficients of the source: clearly, for a star with significant limb darkening, when the edge of the source is near the caustic, less light is being magnified than when the centre of the source approaches it. For microlensing events, it is generally sufficient (Dominik, 2004) and most informative (Kubas et al., 2005) to assume a linear limb-darkening law, which, for normalised flux can be expressed as (Albrow et al., 1999)

$$I(r) = \frac{1}{\pi} \left[1 - \Gamma \left(1 - \frac{3}{2} \sqrt{1 - r^2} \right) \right], \quad (2.23)$$

where r is the fractional radius of the source ($r \in [0, 1]$), and Γ is the linear limb-darkening coefficient; this expression for I yields a flux of unity when integrated over a disc of unit radius. This has been used to measure linear limb-darkening coefficients for several stars, mostly for giant stars located in the Galactic Bulge (e.g. Albrow et al. 2000a), but also several other types of stars (e.g. Albrow et al. 2001). The effect of different limb-darkening coefficients on a point-lens and binary-lens lightcurve is shown on Fig. 2.14.

Parallax and orbital motion

The parallax effect was first measured in a lightcurve by Alcock et al. (1995). The amplitude of this effect depends on the length of the observed microlensing event, becoming more pronounced when the timescale of the event is a significant fraction of the Earth's orbital period; it also depends on the position of the observed event in the sky. The net effect of parallax is that the apparent relative trajectory of the source and the lens deviates from a straight line, leading to additional variation in the observed lightcurve. The amplitude of the effect is given by

$$\pi_E = \frac{(D_L^{-1} - D_S^{-1})}{\theta_E} = \frac{\pi_{LS}}{\theta_E}, \quad (2.24)$$

in other words, is equal to the lens-source relative parallax $\pi_{LS} = (D_L^{-1} - D_S^{-1})$ in units of θ_E . In addition to π_E , a second parameter is needed to completely characterise parallax; this second parameter, generally called ψ , is necessary to transpose the trajectory of the Earth's orbit into the source plane.

When possible, measuring both the source size and parallax allows us to calculate directly

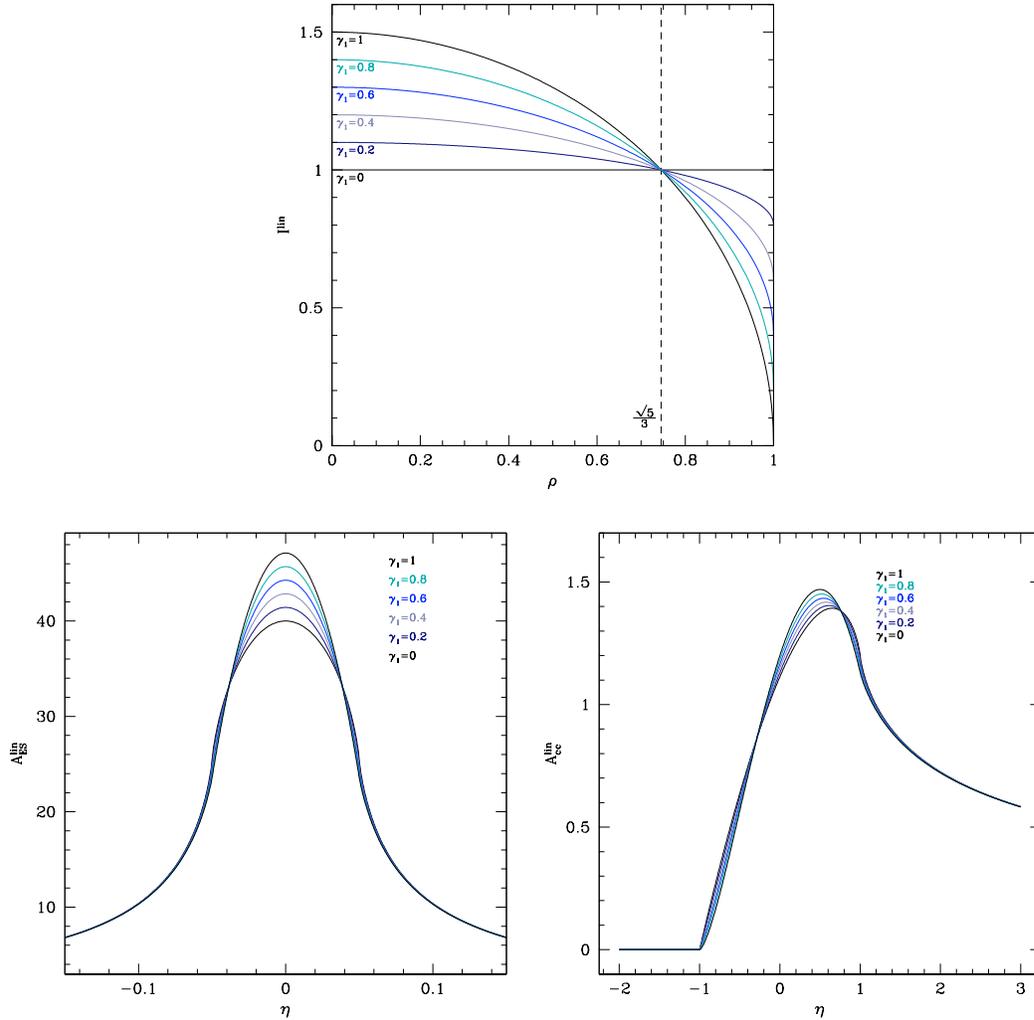


Figure 2.14: Top: the limb darkening law, showing the decrease in intensity as one moves away from the centre of its disc, for different values of $\gamma_1 (\equiv \Gamma)$. The x-axis is the distance from the centre of the source in units of the source radius. Bottom: the effect of an extended source having the different limb-darkening coefficient γ_1 shown in the top panel, in a single-lens model (bottom left) and on the shape of the photometric signal of a caustic crossing (bottom right). The x-axis is the source-lens separation in units of source radii. Taken from Cassan (2005).

the lens mass. Indeed, measuring the source size allows us to determine the value of θ_E (Gould et al., 1994; Nemiroff & Wickramasinghe, 1994; Witt & Mao, 1994). If θ_E is known, we can use Eq. (2.7), re-expressed as

$$\theta_E = \sqrt{\frac{4GM}{c^2} \pi_{LS}}, \quad (2.25)$$

to see that if the parallax parameter π_E is measured as well, Eq. (2.25) can be rearranged to obtain an expression for the mass of the lens M . This means that microlensing can be used as a technique to measure the mass of single stars as well (e.g. Ghosh et al. 2004); this was first achieved by Alcock et al. (2001).

Sometimes orbital motion of the source, also called *xallarap*, can mimic the effect of parallax (e.g. Bennett 1998). When that is the case, one must check whether the detected xallarap signal is consistent with the signal that could be produced by parallax for the event timescale, and which effect results in a better fit to the available data (e.g. Poindexter et al. 2005).

Finally, one can also characterise a lensing system through the detection of orbital motion of the lens, as suggested by Dominik (1998b); indeed this effect was first incorporated into the modelling by Albrow et al. (2000b).

In general, with the exception of finite source size effects, I will ignore second-order effects in this thesis, unless otherwise stated.

2.4 Extracting physical parameters from microlensing modelling

As discussed in the previous section, measuring both the parallax effect and finite source size effects allows us to determine exact properties of the lensing system. This, however, can only be done for very few events. In the vast majority of events for which this is not possible, we can only derive probabilistic distributions of lens properties, based on a chosen Galactic Model. One then has to make assumption about the location of the source (generally, this is taken to be the Galactic Bulge), and to consider separately cases where the lens resides in the Galactic Bulge or in the Disc. Using the distributions of stellar masses, velocities and distances given by the Galactic models (e.g. Han & Gould 2003, Dominik 2006), we can then derive corresponding probability density functions of the lens properties. Such an analysis was performed for the modelling of OGLE-2007-BLG-472, for which only the source size was

measured; this analysis is presented in Chapter 4.

2.5 Binary sources

Another potential cause for deviations from single-lens magnification patterns is multiple components in the *source* system. It is estimated that roughly half the stars in our Galaxy are part of a binary or multiple star system, yet few cases of lensing of binary stars, i.e. binary *source* events have been reported in the literature (e.g. Gaudi & Han 2004; Collinge 2004; Jaroszynski et al. 2004). The reasons for this dearth of binary-source events have been the subjects of several studies. Griest & Hu (1992) found that up to 95% of binary-source events have a lightcurve which is affected by only one of the source's components, a result confirmed by Dominik (1998c), who found that most of the binary source events have lightcurves that can be fitted with a PSPL model. Dominik (1998c) explained this by the fact that for a typical binary source, the separation between the source components is large, resulting in a large difference between the impact parameters of the two components. However, Han & Jeong (1998) estimated that there should still be up to 8% of events in which the components have a separation small enough that they would both affect the event's lightcurve. They found that more reasons contribute to the lack of detected binary-source events, including binary systems in which one component is much brighter than the other, i.e. with a large flux ratio.

These events are particularly interesting in the context of planet searches, because binary source lightcurves can sometimes mimic binary lens lightcurves, leading to an ambiguity in the nature of an observed event, and most of the reported binary-source events can also be fitted with binary-lens models (see Fig. 2.15). Gaudi (1998) presented a detailed discussion of the ambiguity between high flux ratio binary-source and low mass-ratio binary-lens models, and suggested taking a spectrum of the source while the event is ongoing, or very dense photometric coverage, as a way to break this ambiguity. For the planetary event OGLE-2005-BLG-390 (Beaulieu et al., 2006), dense data coverage indeed meant that the planetary model was strongly favoured over the binary-source model.

During my Ph.D., I have developed a fitting algorithm for binary source - single lens events, and worked on a method to resolve the ambiguity between some binary-lens and binary-source lightcurves that would not require as much observational effort as is needed to obtain dense photometric coverage or to carry out spectroscopy. The method is based on previous

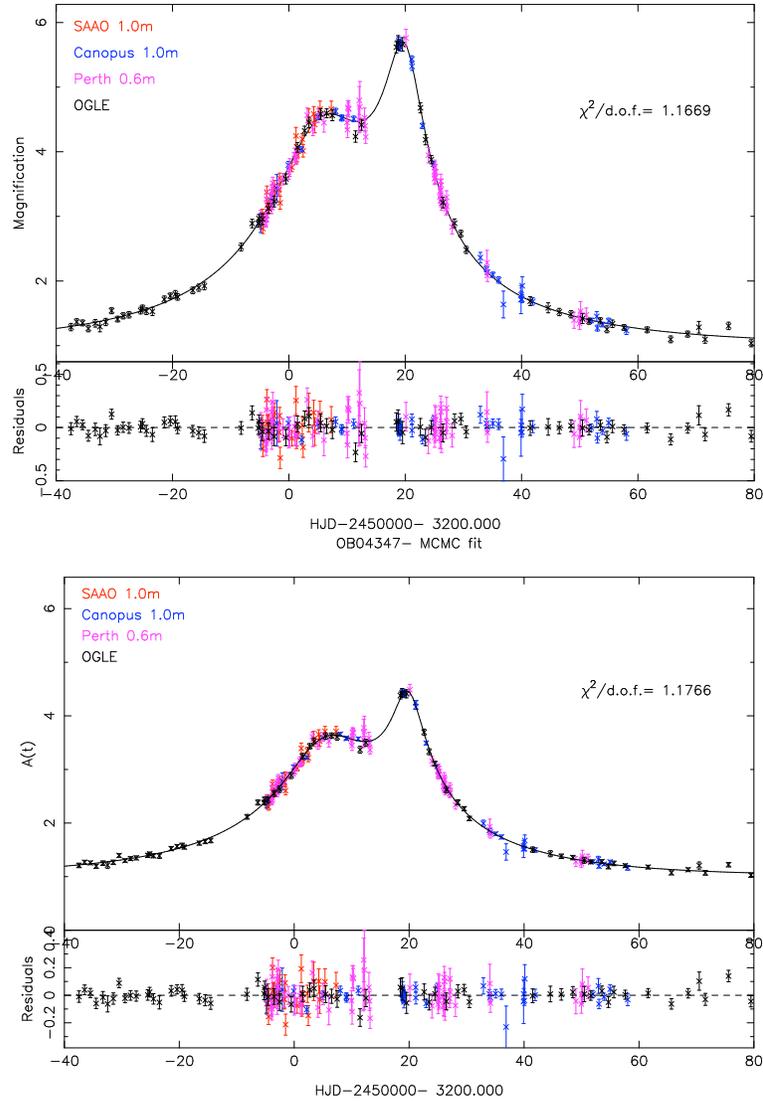


Figure 2.15: *Top:* the best-fit binary-source model for OGLE-2004-BLG-347, corresponding to a source with a primary component ~ 2 times brighter than the secondary. *Bottom:* the best-fit binary-lens model ($d \sim 3, q \sim 0.75$) for that event. Neither model is significantly better than the other, so the nature of this event is ambiguous.

work by Dominis (2006), and relies on the fact that in most cases, binary source events are chromatic, i.e. the total colour of the binary-source will change as each component is lensed, whereas binary-lens events are not. This work, however is still in progress and is therefore not part of this thesis.

3

Markov Chain-Monte Carlo methods

This Chapter outlines the development of my Markov Chain-Monte Carlo (MCMC) algorithm to fit microlensing events. I start by recalling the basic mathematical principles of MCMC methods, and detail various issues associated with that class of algorithms. Finally, I describe the algorithm I have produced, illustrating this with a few examples of event fits.

3.1 Motivation and Formalism

MCMC methods are a class of random walk algorithms that allow us to sample from an unknown target distribution. This is made possible by the fact that the MCMC random walk chain can generally be made to converge to a stationary distribution that is the desired target distribution. In Bayesian terms, if the MCMC steps in the parameter space θ and we are trying to model data y , the MCMC algorithm converges to the posterior distribution $p(\theta|y)$ for any "reasonable" prior distribution. Several methods exist to construct MCMC chains, all of them based on the work by Metropolis et al. (1953) and by Hastings (1970). From Bayes' theorem, given a prior distribution $p(\theta)$ and a likelihood $p(y|\theta)$, the posterior distribution is given by

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}. \quad (3.1)$$

This posterior distribution can be used for Bayesian inference, e.g. to calculate the posterior expectation value of a function of θ ,

$$E_{p(\theta|y)}(f(\theta)|y) = \int f(\theta)p(\theta|y)d\theta. \quad (3.2)$$

The difficulty lies in evaluating the integrals in Eqs. (3.1) and (3.2). One way to do this is by Monte Carlo integration, of which MCMC is a class. Monte Carlo integration allows one to draw a large number of samples θ_i from the posterior distribution $p(\theta|y)$ without the need to know its exact form. This therefore allows us to approximate the integral on the right hand side of Eq. (3.2) as

$$E_{p(\theta|y)}(f(\theta)|y) \simeq \frac{1}{n} \sum_{i=1}^n f(\theta_i), \quad (3.3)$$

i.e. by approximating the expected value as a sample mean. In the case of MCMC integration, the key feature is that the generated samples form a Markov Chain, that is, each sample x_{i+1} that is generated depends on the current state of the chain x_i . The corresponding equation to Eq. (3.3) will have a denominator $n - m$, where m is the number of iterations needed for the chain to reach the stationary distribution, starting from an initial point in parameter space; the period during these m iterations is called the *burn-in* of the chain. This is illustrated on Fig. 3.1 and Fig. 3.2. Since the burn-in points are not drawn from the posterior distribution, they are not used to compute statistics such as the sample means, standard deviations or the covariance matrix. Once the chain is past the burn-in phase, the ergodic theorem ensures that the expectation value estimator converges to the real expectation value (Roberts, 1992).

Several different types of MCMC algorithms exist, characterised by the way the Markov Chain is generated. The most general form of MCMC method is the Metropolis-Hastings algorithm, in which we seek to generate, from a starting point x_0 , a chain of samples x_i from a target distribution p , using a proposal distribution q . This is achieved by drawing a candidate point ξ from q , which then gets accepted or rejected with probability $\alpha(x_i, \xi)$ where

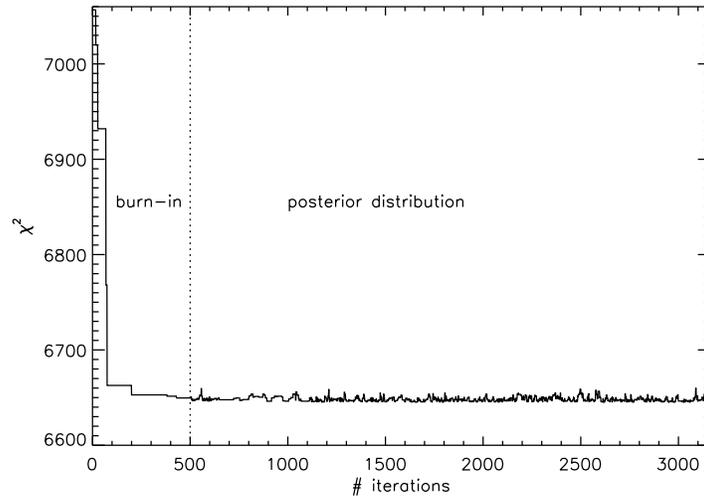


Figure 3.1: The value of χ^2 against the number of iterations of the MCMC chain for a fit to a PSPL microlensing event. The burn-in and stationary periods are labeled and a dotted line indicates where the chain starts sampling for the posterior distribution.

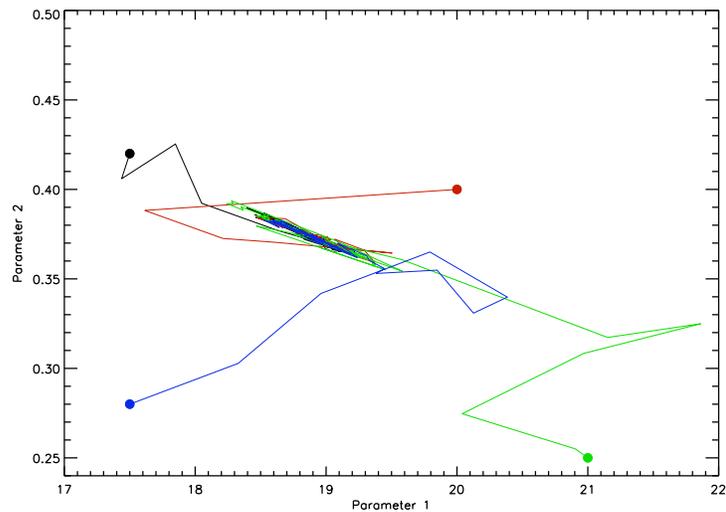


Figure 3.2: Example pairwise correlations showing the progress of the Markov chain for four different starting points (shown with filled circles). The chains take a few steps during the burn-in period before converging to the stationary distribution (dense area in the middle of the plot).

$$\alpha(x_i, \xi) = \min\left(1, \frac{p(\xi)q(x_i|\xi)}{p(x_i)q(\xi|x_i)}\right). \quad (3.4)$$

If ξ is accepted, the chain moves on to that place in parameter space, i.e. $x_{i+1} = \xi$, otherwise it remains where it was, i.e. $x_{i+1} = x_i$. Remarkably, it has been shown that the chain will converge to the target distribution for almost any form of the proposal distribution q with only a few restrictions (Roberts, 1992). The *transition kernel* of the chain is defined as $P(\xi|x_i) = q(\xi|x_i)\alpha(\xi, x_i)$ and is a defining feature of a given MCMC algorithm.

When the proposal distribution q is symmetric, the method is known simply as the Metropolis algorithm. Since in that special case $q(x|y) = q(y|x)$, the acceptance probability α reduces to

$$\alpha(x_i, \xi) = \min\left(1, \frac{p(\xi)}{p(x_i)}\right). \quad (3.5)$$

An important issue in MCMC algorithms is to decide when the chain has converged to the target distribution, i.e. to determine the length of the burn-in period mentioned above. As mentioned above, the samples drawn from the burn-in stage of the chain will clearly need to be thrown out when using the output of an MCMC run, since they are not samples from the posterior distribution. Factors that will affect the length of the burn-in period include the choice of starting values of the chain and the proposal distribution q . If these are poorly chosen, the burn-in time can be much increased, wasting computational time on irrelevant regions of the parameter space. If starting values fall in a local minimum of the target distribution, the chain will be stuck in the minimum for a long time, and it will be said to be *poorly mixing* as it will not be exploring the parameter space efficiently. Another important factor is the size of the steps taken by the algorithm: if the steps are too large, too many steps will be rejected and the chain will be inefficient, whereas if the steps are too small, the chain will take too much time moving in the parameter space, or will only explore a small region around the starting value. When the target distribution has many local minima, one sensible option is to use scattered initial values for the parameters and to run many chains in parallel, as described by Gelman & Rubin (1992). These show the importance of a good choice of initial conditions for the MCMC algorithm. Most of the difficulty when using MCMC algorithms resides in determining ideal starting values and proposal distributions in order to

avoid wasting valuable computational time on the burn-in stages of MCMC runs.

Once the burn-in stage is passed, a number of different criteria for convergence can be applied. Geweke (1992) devised a simple test: two subsamples are taken from the (post burn-in) sample and the mean of parameters within these two subsamples is calculated. By comparing these two numbers for each parameter, the convergence can be quantified: if they are approximately equal (within a desired tolerance interval), then the chain has converged, otherwise the sequence is still converging. Other convergence tests include those developed by Hanson (2002) or Raftery & Lewis (1992).

3.1.1 Sample autocorrelation

Typically, the samples representative of the posterior distribution will be used to calculate the standard error in a parameter. However, due to the Markov Chain nature of the process used, successive samples will be correlated with one another. This can be quantified by evaluating the (first-order) autocorrelation function ρ of the sample α ,

$$\rho_{i,i-1} = \frac{\sum_{i=1}^{N_s-1} (\alpha_i - \bar{\alpha})(\alpha_{i-1} - \bar{\alpha})}{\sum_{i=1}^{N_s-1} (\alpha_i - \bar{\alpha})^2}, \quad (3.6)$$

where N_s is the length of the sample. The standard error in the mean $\bar{\alpha}$ of the sample is then affected by the autocorrelation and is given by:

$$s_{\bar{\alpha}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{1+\rho}{1-\rho}}, \quad (3.7)$$

where σ is the standard deviation of α calculated from the sample sample.

For high autocorrelation, ρ is close to 1 and the standard error is greatly increased compared to the case where autocorrelation is not taken into account ($\rho = 0$). An efficient way to reduce autocorrelation is *thinning* the sample, so that we only store one in every n points of the chain (Walsh, 2002). This effectively reduces the autocorrelation by a factor ρ^{n-1} . In order to minimise the effect of autocorrelation, I will also draw a random sample from the final chain to derive statistics from the posterior distribution.

3.2 Application to microlensing

When using an MCMC algorithm to fit a microlensing lightcurve to photometric data, as with most other methods, the aim is to minimise the badness-of-fit of the model lightcurve, generally χ^2 . In the case of a microlensing model parametrised by n parameters, this means exploring an n -dimensional parameter space to find the global minimum. An additional complication comes from the fact that data are taken at various observing sites, which have different source and blending fluxes f_S and f_B . The flux measured at a telescope i is then given by $F^i(t) = f_S^i A(t) + f_B^i$. There are two ways to treat these two additional parameters: treat them as free parameters like the other n parameters of the magnification model, or determine their optimal value by linear regression for each new set of parameters. The former solution means increasing the number of parameters by $2m$, where m is the number of observing sites, which rapidly becomes problematic in terms of computing time. Therefore we usually choose the latter method.

For a data set consisting of N data points, the χ^2 statistic for a given set of parameters is defined as

$$\chi^2 = \sum_{k=1}^N \left(\frac{F(t_k) - \mu(t_k)}{\sigma_k} \right)^2, \quad (3.8)$$

where t_k is the time at which the k^{th} data point was taken, $F(t_k)$ is the observed flux, $\mu(t)$ is the theoretical flux at time t and σ_k is the size of the error bar for the k^{th} data point. In order to minimise the χ^2 of the model fit, we choose a jumping distribution so that the acceptance probability will be given by

$$\alpha(x_i, x_{i+1}) = \min(1, e^{-(\chi_{i+1}^2 - \chi_i^2)}), \quad (3.9)$$

that is, if the proposed step x_{i+1} corresponds to an increase in χ^2 , it will be accepted with a probability $e^{-\Delta\chi^2}$ where $\Delta\chi^2 = \chi_{i+1}^2 - \chi_i^2$, while if a step goes toward a point in parameter space corresponding to a lower χ^2 , it will always be accepted. The ability to move ‘‘uphill’’ in the χ^2 landscape ensures that the chain does not get trapped in local minima of the parameter space, provided that the chain is long enough.

The chain is initialised at given starting values which only need to be "reasonable" to achieve convergence to the posterior parameter distribution, as mentioned above, and I use the convergence criterion of Geweke (1992).

Something that is obvious from Eq. (3.8) is the importance of the size of the error bars. In particular, when there are data available from multiple sites, as is generally the case, it is important not to give excessive or insufficient weight to data from a particular site, which can happen if the size of error bars are under- or over-estimated. For this reason many people rescale error bars by a site-dependent factor in order to obtain a final χ^2 per degree of freedom of 1 for the best-fit model. The models presented in this thesis are found without error bar rescaling, unless explicitly mentioned.

One of the most interesting features of using an MCMC algorithm to fit data is that in addition to a best-fit model, it computes the parameter covariance matrix, and values for the error bars on parameters, at no extra computational cost. These features are examined in the next section.

3.3 Fitting lightcurves with MCMC

3.3.1 PSPL events

Fits for two PSPL events are presented of Figs. 3.3 and 3.4, and their best-fit parameters as well as the values published by the OGLE collaboration¹ (for comparison) and MCMC output data are given in Tables 3.1 and 3.2. The agreement in best-fit parameters is excellent, but for OGLE-2006-BLG-397, the error bars found by the MCMC algorithm are significantly larger than those reported by OGLE. Since it is not known (to the author) how OGLE error bars are calculated, it is possible that this discrepancy is due to a difference in fitting methods.

Parameter	MCMC best fit value	OGLE published value	Units
t_0	78.178 ± 0.022	78.168 ± 0.025	HJD-2453800
t_E	18.75 ± 0.48	18.456 ± 0.408	days
u_0	0.378 ± 0.014	0.385 ± 0.014	—

Table 3.1: OGLE-2006-BLG-180: MCMC best fit parameters compared to the OGLE published values

¹<http://ogle.astrouw.edu.pl/ogle3/ews/ews.html>

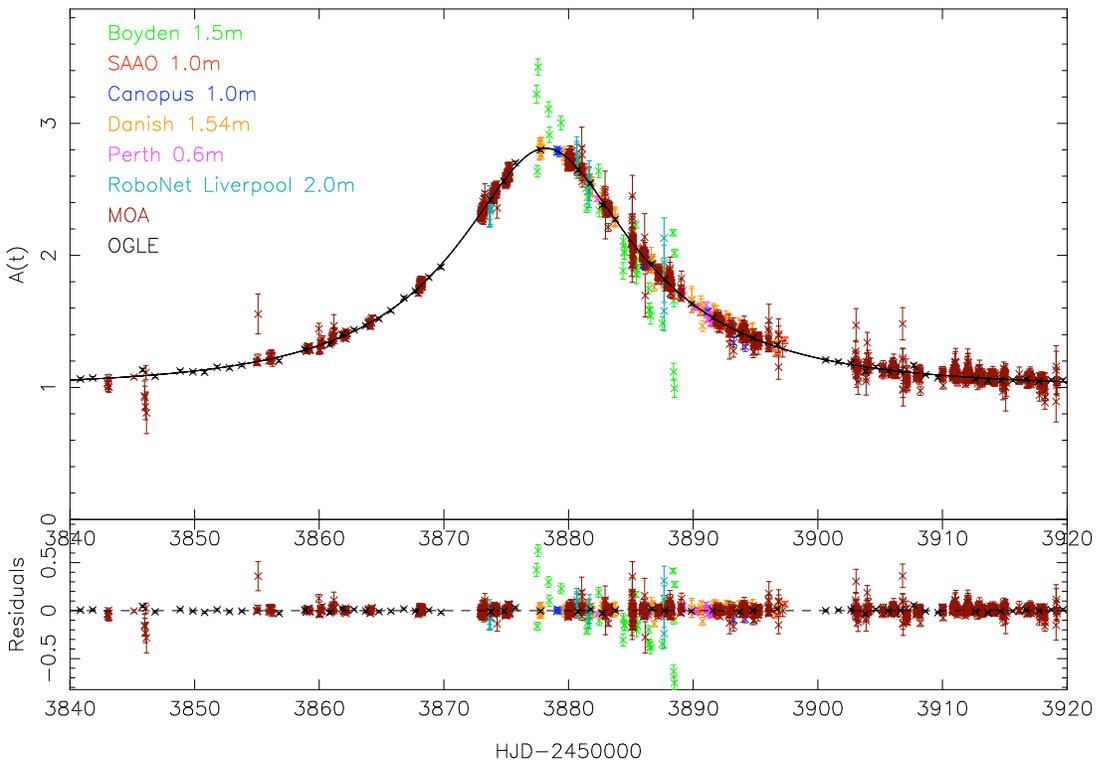


Figure 3.3: MCMC fit and residuals for OGLE-2006-BLG-180

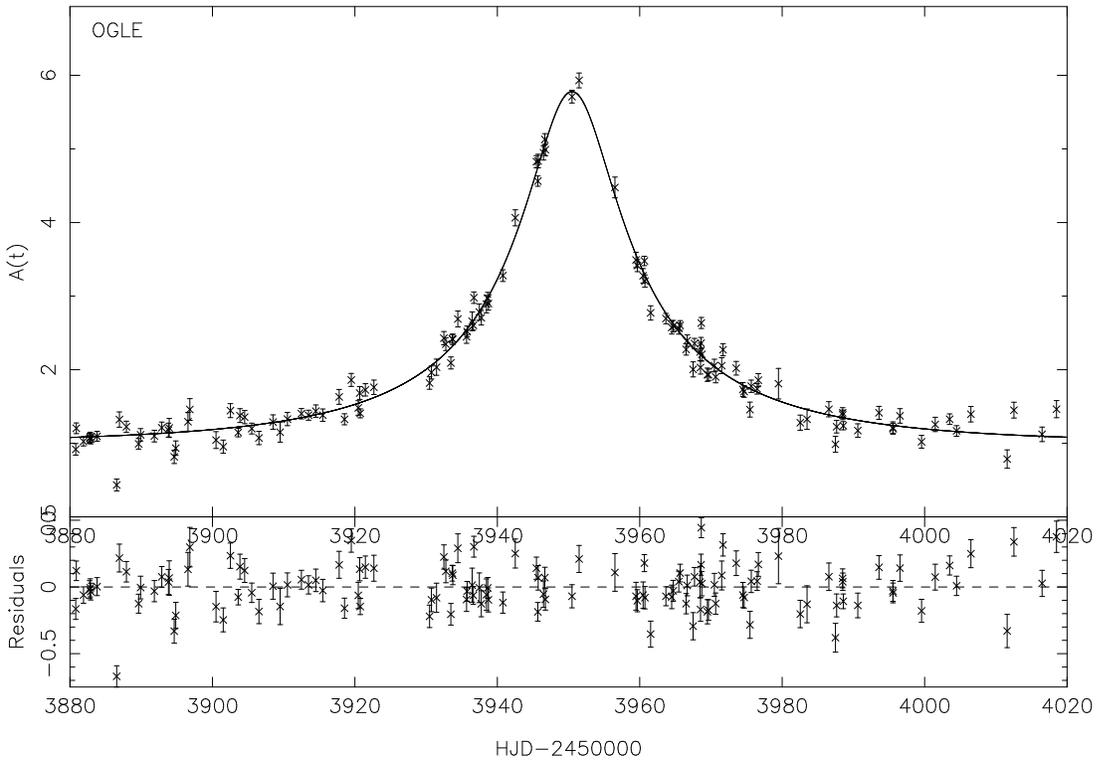


Figure 3.4: MCMC fit and residuals for OGLE-2006-BLG-397 (note the worse data quality than OGLE-2006-BLG-180).

Parameter	MCMC best fit value	OGLE published value	Units
t_0	50.494 ± 0.073	50.513 ± 0.073	HJD-2453900
t_E	39.00 ± 1.98	33.265 ± 0.276	days
u_0	0.174 ± 0.013	0.211 ± 0.002	—

Table 3.2: OGLE-2006-BLG-397: MCMC best fit parameters compared to the OGLE published values

3.3.2 Parameter correlations

In Sec. 3.1, we discussed factors that could make a MCMC algorithm slower or less efficient. One of the main factors affecting the efficiency of a MCMC run exploring a parameter space is strong correlations between the parameters. The correlations for one of the PSPL events shown above, OGLE-2006-BLG-180, are shown on Fig. 3.5. While t_0 is not strongly correlated with any other parameters, we observe that all the other pairwise correlations are very strong. This means that completely random steps are very unlikely to be accepted since one of the parameters in these pairwise correlation plots will be likely to wander far off their "1- σ zone", where it will cause the whole proposed parameter vector to be rejected.

One potential solution to this issue is to reparametrise the model used with uncorrelated (or weakly correlated) parameters. One example of an attempt at such a reparametrisation is shown on Fig. 3.6. This uses $\omega = \frac{1}{u_0 t_E}$, $\epsilon = \frac{f_s}{u_0}$ and the baseline flux $f_0 = f_s + f_b$. Although some of the more severe correlations have become weaker, some strong correlations remain. In the case of a complex parameter space it will be almost impossible to find such "replacement" parameters with which the model can be re-expressed.

3.3.3 Orthogonalising the parameter space

A more complete solution to the problem of parameter correlations is to completely orthogonalise the parameter space. One approach, using a Gram-Schmidt orthogonalisation process is presented here. The idea is to "shear" the parameter space so that all the parameters become mutually orthogonal. In a non-orthogonalised parameter space of a model μ , the steps for each parameter α_i are taken in the direction of the vector $P_i = \frac{\partial \mu}{\partial \alpha_i}$, which are not orthogonal to one another.

With the dot product defined as

$$P_a \cdot P_b = \sum_{i=1}^N \frac{P_a(t_i)P_b(t_i)}{\sigma_i^2} = \|P_a\| \|P_b\| \cos \theta, \quad (3.10)$$

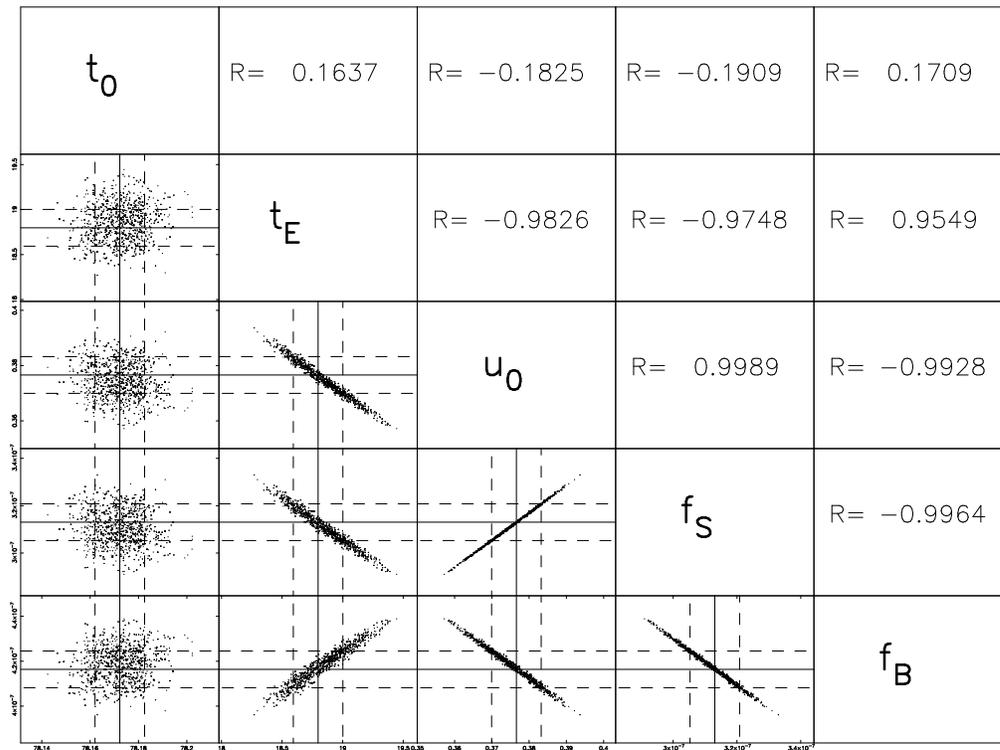


Figure 3.5: Correlations between standard PSPL model parameters for OGLE-2006-BLG-180.

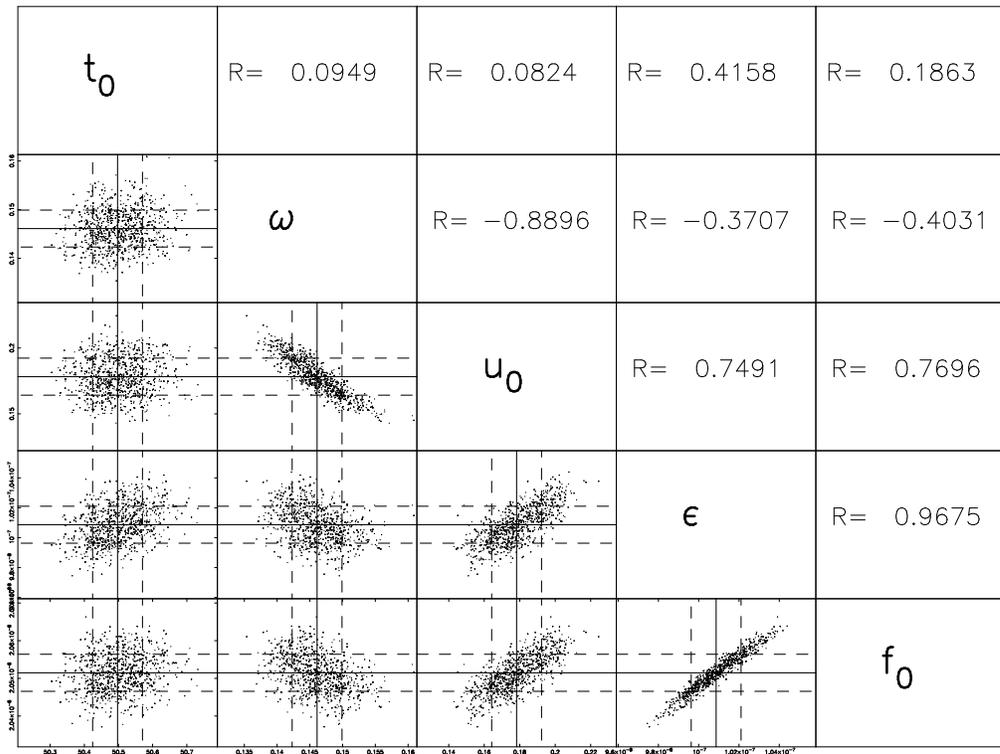


Figure 3.6: Correlations between t_0, u_0 and chosen model parameters ω, ϵ and f_0 for OGLE-2006-BLG-180.

where θ is the angle between the two vectors, we can use a Gram-Schmidt orthogonalisation to construct orthogonal vectors B_i as

$$B_i = P_i - \sum_{j=1}^{i-1} \left(\frac{P_i \cdot B_j}{B_j \cdot B_j} \right) B_j, \quad (3.11)$$



Figure 3.7: *Left:* Schematic representation of the “shearing” of the parameter space for a pairwise parameter correlation pattern, which becomes circular rather than elliptical, i.e. less correlated. *Right:* Gram-Schmidt orthogonalisation for the first two vectors of an N-dimensional parameter space.

with $B_1 = P_1$. This operation is shown on Fig. 3.7 The vectors satisfy $B_i = \frac{\partial \mu}{\partial \beta_i}$ for orthogonal parameters β_i . By defining the transformation matrices S and T as satisfying

$$B_i = \sum_{j=1}^i S_{ij} P_j \quad (3.12)$$

$$P_i = \sum_{j=1}^i T_{ij} B_j, \quad (3.13)$$

and comparing Eq. 3.13 to Eq. 3.11, we see that

$$T_{ij} = \begin{cases} \frac{P_i \cdot B_j}{B_j \cdot B_j} & i \geq j \\ 0 & i < j, \end{cases} \quad (3.14)$$

and noting that the diagonal elements of S and T are equal to 1 and that the two matrices are lower triangular matrices, we can calculate all the other non-zero non-diagonal elements of the transformation matrices by induction.

We start by rewriting Eq. (3.11) as

$$B_i = P_i - \sum_{j=1}^{i-1} T_{ij} B_j, \quad (3.15)$$

and by rewriting B_j in this equation in terms of components along P_k to obtain

$$B_i = P_i - \sum_{j=1}^{i-1} T_{ij} \sum_{k=1}^j S_{jk} P_k. \quad (3.16)$$

Since $S_{jk} = 0$ for $k > j$,

$$B_i = P_i - \sum_{j=1}^{i-1} T_{ij} \sum_{k=1}^n S_{jk} P_k, \quad (3.17)$$

and reversing the order of sums yields

$$B_i = P_i - \sum_{k=1}^n \left(\sum_{j=1}^{i-1} T_{ij} S_{jk} \right) P_k. \quad (3.18)$$

Since $S_{jk} = 0$ for $j < k$,

$$B_i = P_i - \sum_{k=1}^n \left(\sum_{j=k}^{i-1} T_{ij} S_{jk} \right) P_k, \quad (3.19)$$

and since $j < i$ and $S_{jk} = 0$ for $k > j$, it follows that $k < i$, and therefore

$$B_i = P_i - \sum_{k=1}^{i-1} \left(\sum_{j=k}^{i-1} T_{ij} S_{jk} \right) P_k. \quad (3.20)$$

Therefore, by comparing this expression to Eq. (3.12), we find that the elements of S are

$$S_{ik} = \begin{cases} -\sum_{j=1}^{i-1} T_{ij} S_{jk} & i > k \\ 1 & i = k \\ 0 & i < k. \end{cases} \quad (3.21)$$

These matrices are then used to transform steps so that they match the correlation of

parameters. For a random gaussian step $\Delta\beta_j$ in the orthogonal parameter space, the corresponding step for the i^{th} α parameter is given by

$$\Delta\alpha_i = \sum_{j=1}^i S_{ij} \Delta\beta_j. \quad (3.22)$$

The n -parameter model can then be written as

$$\mu = \mu_0 + \sum_{i=1}^n \Delta\alpha_i P_i = \mu_0 + \sum_{i=1}^n \Delta\beta_i B_i. \quad (3.23)$$

Given a dataset \mathbf{x} , the best fit $\hat{\beta}$ parameters and $1\text{-}\sigma$ error bars are then given by

$$\hat{\beta}_i = \frac{\mathbf{x} \cdot B_i}{B_i \cdot B_i} \quad (3.24)$$

$$\sigma_{\hat{\beta}_i}^2 = \frac{1}{B_i \cdot B_i}, \quad (3.25)$$

with the random step $\Delta\beta_i$ for each orthogonal parameter drawn from a gaussian distribution with $\sigma = \sigma_{\hat{\beta}_i}$. Derivatives for Gram-Schmidt orthogonalisation can then be re-calculated as the chain moves in parameter space.

The much weaker correlations between these β parameters are shown on Fig. 3.8, and model derivatives for the PSPL case are shown for standard and orthogonal parameters on Fig. 3.9.

3.3.4 Binary lens events

Fitting point source-binary lens (PSBL) events

As discussed in Chapter 2, the most demanding part in calculating binary-lens lightcurves consists of evaluating extended source effects. In most cases, these effects are crucial to modelling anomalous events, as they modify the magnification pattern significantly compared to point-source models with other parameters being equal. In some cases, however, lightcurves are not significantly affected by finite source size effects. Most of the time, this is true when an anomalous event is ongoing before the source is close to caustics, where finite source size

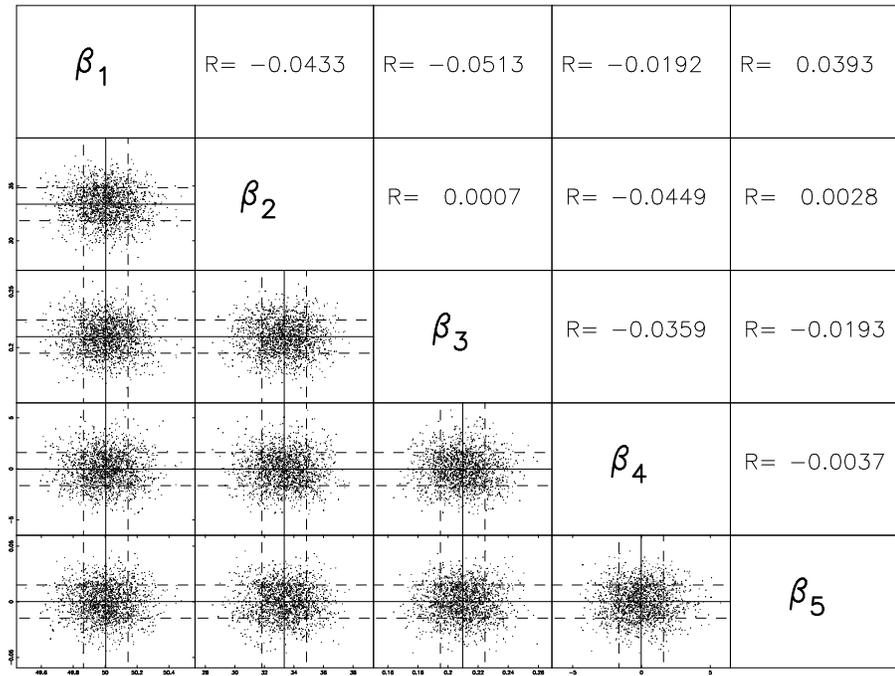


Figure 3.8: Correlations between orthogonalised parameters β_i for OGLE-2006-BLG-180

effects are most important. This means that until this happens, we can approximate the true lightcurve with a PSBL model. The advantage of this is obvious: while extended source-binary lens (ESBL) models take hours to compute, or even days for high magnification events, PSBL lightcurves generally take seconds to calculate, which allows us to extract information on the nature of an event from its best-fit parameters even while the event is still ongoing. An example of such a case is the partial PSBL fit of OGLE-2008-BLG-513 shown on Fig. 3.10. Data taken after this fit was computed began to exhibit pronounced finite source size effects, which, combined with a very high magnification event, made this event impractical to fit with my algorithm and limited computational resources.

Using adaptive contouring to calculate finite-source effects

The procedure used here to calculate binary-lens model lightcurves for each set of parameters in the MCMC relies on adaptive contouring to efficiently calculate lightcurves for extended source cases (Dominik, 2007). Adaptive contouring works by building a grid of squares representing the lens plane and using a ray-shooting approach to determine the corresponding true source position (using the lens equation). According to whether the ray hits the source or misses it, the grid is adapted to improve the determination of the source contour lines, following a hierarchical set of rules Dominik (2007).

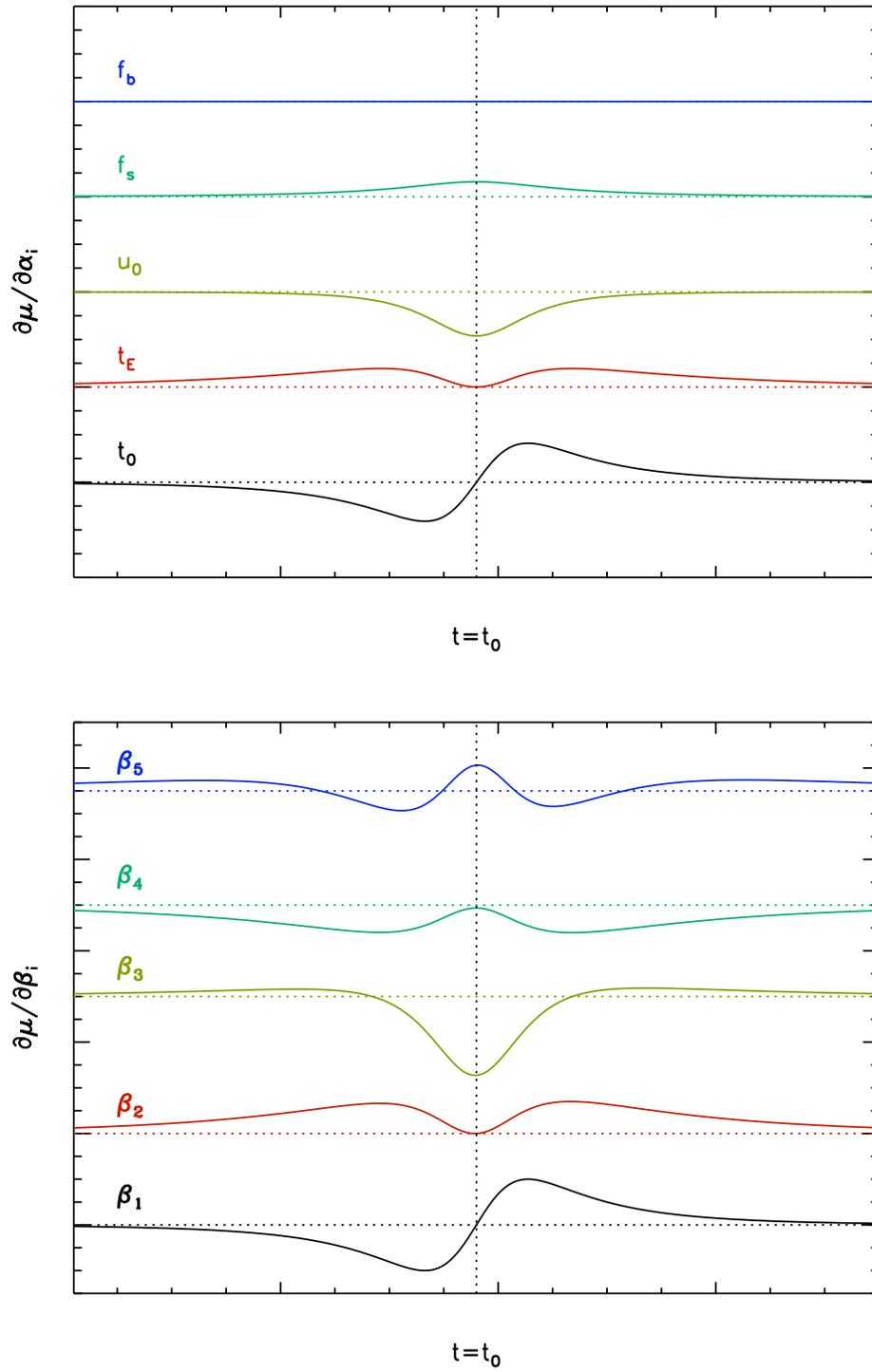


Figure 3.9: Model derivatives for the PSPL case, with respect to standard parameters t_0 , t_E , u_0 , f_s and f_b (top panel) and the orthogonal parameter β_i (bottom panel), for OGLE-2006-BLG-180 (Fig. 3.3).

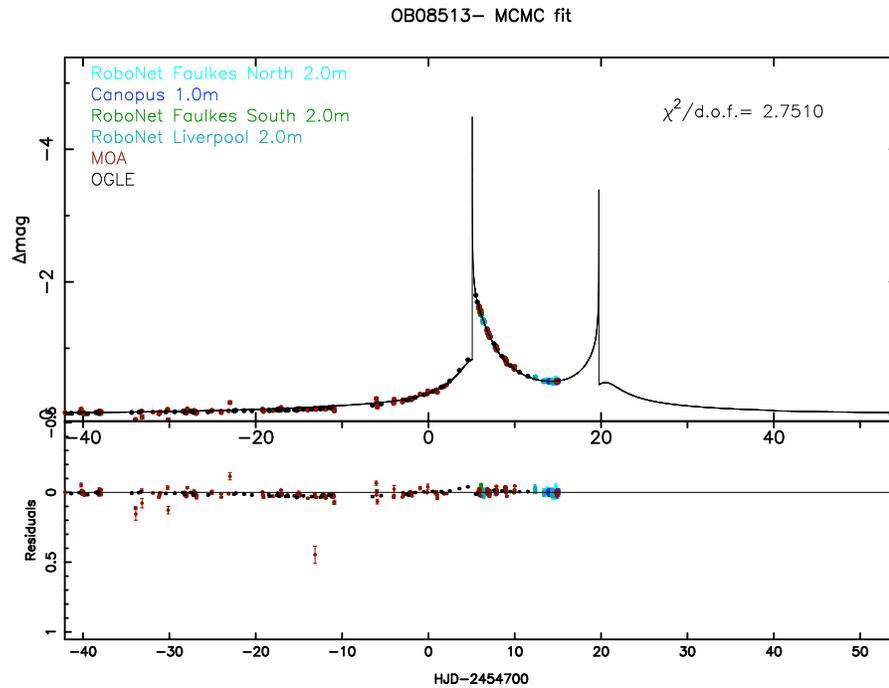


Figure 3.10: Incomplete PSBL fit to the anomalous event OGLE-2008-BLG-513. This fit was made while the event was ongoing, but subsequent observations showed that the magnification pattern became significantly affected by the finite source size. Because this event is also one in which the source is highly magnified, this meant that modelling of this event could not be pursued with limited computational resources.

For binary lens events, calculating lightcurves is much more computationally demanding than for single lenses. This is mainly because the lens equation cannot be solved analytically (as discussed in section 2.2.1) and therefore must be solved by numerical methods. For the same reasons, orthogonalising the parameter space requires calculating numerical derivatives of the binary-lens model, which is also computationally demanding. Using the Gram-Schmidt orthogonalisation procedure is further complicated by the fact that some of the correlations of the 7 binary-lens model parameters are highly non-linear, meaning that if the parameter starting values of the MCMC run are too far off the best-fit values, so will be the derivatives, and the stepping procedure will be too influenced by the geometry of the point in parameter space where the derivatives are calculated. A way to minimise this issue is to run the chain without orthogonalising the parameters until a set number $nkeep$ of steps have been kept. The derivatives can be recalculated as the chain moves in parameter space. In principle the chain should then have moved closer to the global minimum, so the derivatives evaluated at this point in parameter space can then be assumed to be roughly equal to those evaluated at the best fit values. Of course this approach is only interesting if the chain is rapidly mixing and $nkeep$ can be set to be reasonably low. If $nkeep$ needs to be large, the benefits of calculating derivatives (higher acceptance rate and faster convergence to the target distribution) are lost since the chain spends most of its time stepping in a non-orthogonalised parameter space.

Because calculating derivatives becomes impractical for complex cases such as binary-lens events, we need to resort to other methods to minimise the effects of parameter correlations on the modelling process. One such method is to diagonalise the covariance matrix, which allows to rotate, rather than de-shear, the parameter space.

Diagonalising the covariance matrix

In this section, we outline a way of stepping from the current parameter vector α to the proposed vector α_p , taking the correlations into account to minimise their effect on the acceptance rate of proposed steps. This is similar to approaches detailed by Tegmark et al. (2004) and Burke et al. (2007).

Any non-singular square matrix \mathbf{M} can be decomposed in the form

$$\mathbf{M} = \mathbf{PDP}^{-1}, \quad (3.26)$$

where \mathbf{P} is the square matrix formed by columns from the right eigenvectors of \mathbf{M} , \mathbf{P}^{-1} is the matrix inverse of \mathbf{P} , and \mathbf{D} is a diagonal matrix with non-zero elements being the corresponding eigenvalues of \mathbf{M} .

Therefore any matrix \mathbf{M} can be diagonalised through

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{M}\mathbf{P}. \quad (3.27)$$

We can use this to diagonalise the covariance matrix. The matrix \mathbf{P} will then contain a set of rotated axes for a new basis \mathbf{B} in which parameters described by the covariance matrix are orthogonal.

In order to take correlations into account when stepping in an MCMC algorithm, we can simply transform the set of original parameters α we are stepping from into a new set of decorrelated parameters $\beta = \mathbf{P}^{-1}\alpha$. Since the covariance matrix is diagonal in \mathbf{B} -space, and the variances of the orthogonal parameters β_i are simply given by the diagonal elements of the diagonal covariance matrix (themselves the eigenvalues λ_i of the original covariance matrix), we just take $1\text{-}\sigma$ Gaussian steps in \mathbf{B} space to propose a parameter vector β_p , the elements of which are

$$\begin{pmatrix} \beta_1 + \Delta\beta_1 \\ \dots \\ \beta_n + \Delta\beta_n \end{pmatrix}.$$

We can then transform these parameters back to the original parameter space, which has basis vectors \mathbf{A} , to calculate our model for the corresponding proposed parameter vector α_p

$$\begin{pmatrix} \alpha_1 + \Delta\alpha_1 \\ \dots \\ \alpha_n + \Delta\alpha_n \end{pmatrix}.$$

As when using Gram-Schmidt orthogonalisation, the n -parameter model μ can be expressed both in \mathbf{A} and \mathbf{B} space, since

$$\mu = \mu_0 + \sum_{i=1}^n \alpha_i \mathbf{A}_i = \mu_0 + \sum_{i=1}^n \beta_i \mathbf{B}_i. \quad (3.28)$$

Using the diagonalised covariance matrix to calculate steps in parameter space ensures that the random steps are taken along parameter correlations. The effect of the diagonalisation process on the acceptance rate of the MCMC algorithm is shown on Fig. 3.11: after diagonalisation, the acceptance rate increases significantly.

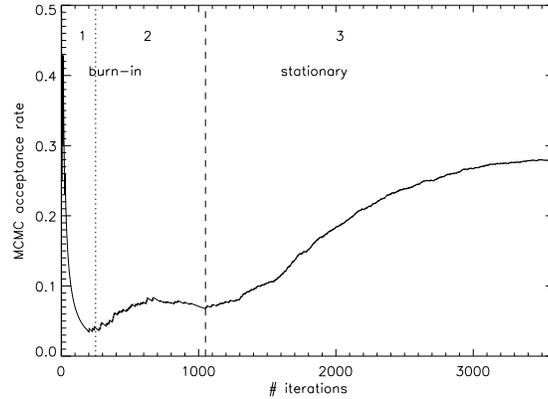


Figure 3.11: The effect of diagonalising the covariance matrix on the acceptance rate of the MCMC algorithm. The burn-in and stationary phases are separated by a vertical dashed line, while a dotted line shows where the covariance matrix is first diagonalised.

3.3.5 A test binary-lens event: OGLE-2005-BLG-390

OGLE-2005-BLG-390, an anomalous event observed in 2005, allowed the discovery of a planetary companion to the lens, OGLE-2005-BLG-390Lb, which was at the time the first cool Super-Earth (with a median mass of $5.5 M_{\oplus}$) detected (see Beaulieu et al. 2006). I use this as a test event for my binary-lens fitting algorithm. The best-fit model is shown on Fig. 3.12, parameter correlations are plotted on Fig. 3.13, and Table 3.3 lists the best-fit parameters produced by this algorithm and the published parameter values.

As is clear from Fig. 3.12 and Table 3.3, the model found by my algorithm is in excellent agreement with the published model (Beaulieu et al., 2006). Over the past observing seasons, I have used this algorithm to participate in real-time modelling of microlensing events, contributing to inform observers as to the possible nature of some of these. In the next section, we use this algorithm as a basis for conducting a systematic search of the parameter space for a particular class of binary-lens microlensing events.

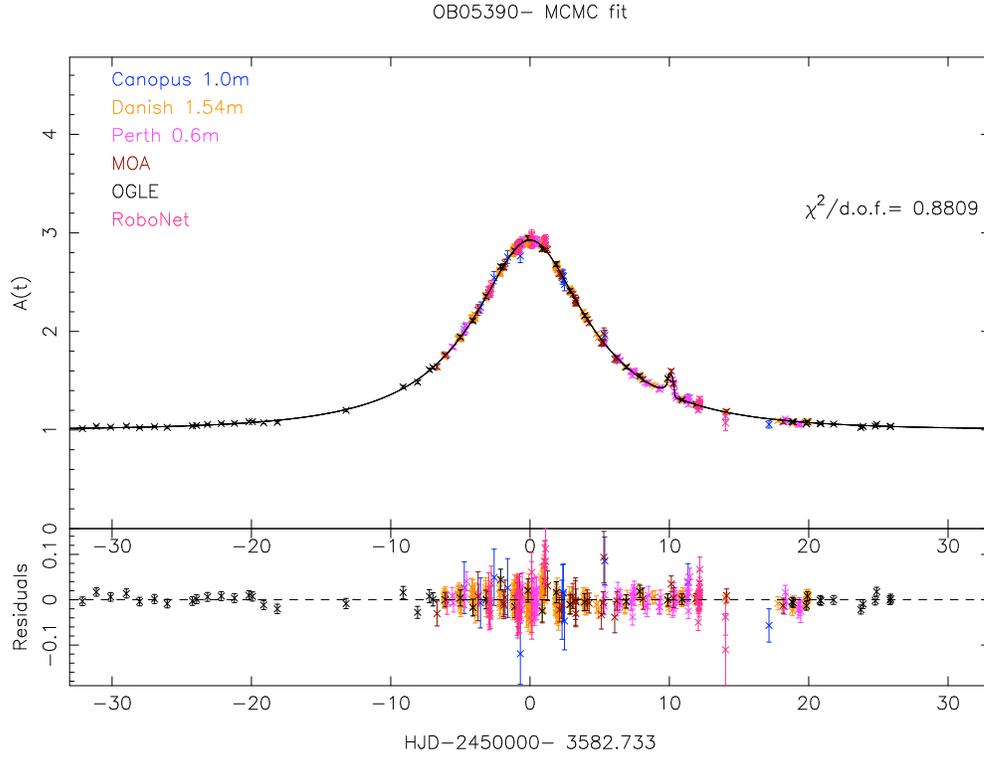


Figure 3.12: MCMC fit and residuals for the planetary binary-lens event OGLE-2005-BLG-390. The best-fit parameters are given in Table 3.3.

Parameter	MCMC best fit value	Published value	Units
t_0	3582.728 ± 0.007	3582.755 ± 0.006	HJD-2450000
t_E	11.17 ± 0.10	11.03 ± 0.11	days
α	2.759 ± 0.003	2.756 ± 0.003	rad
u_0	0.353 ± 0.004	0.359 ± 0.005	—
ρ_*	0.0233 ± 0.0013	0.0255 ± 0.0009	—
d	1.599 ± 0.006	1.610 ± 0.008	R_E
q	$(7.2 \pm 0.3) \times 10^{-5}$	$(7.6 \pm 0.7) \times 10^{-5}$	—

Table 3.3: MCMC best fit parameters compared to the published values of Beaulieu et al. (2006); all are in good agreement.

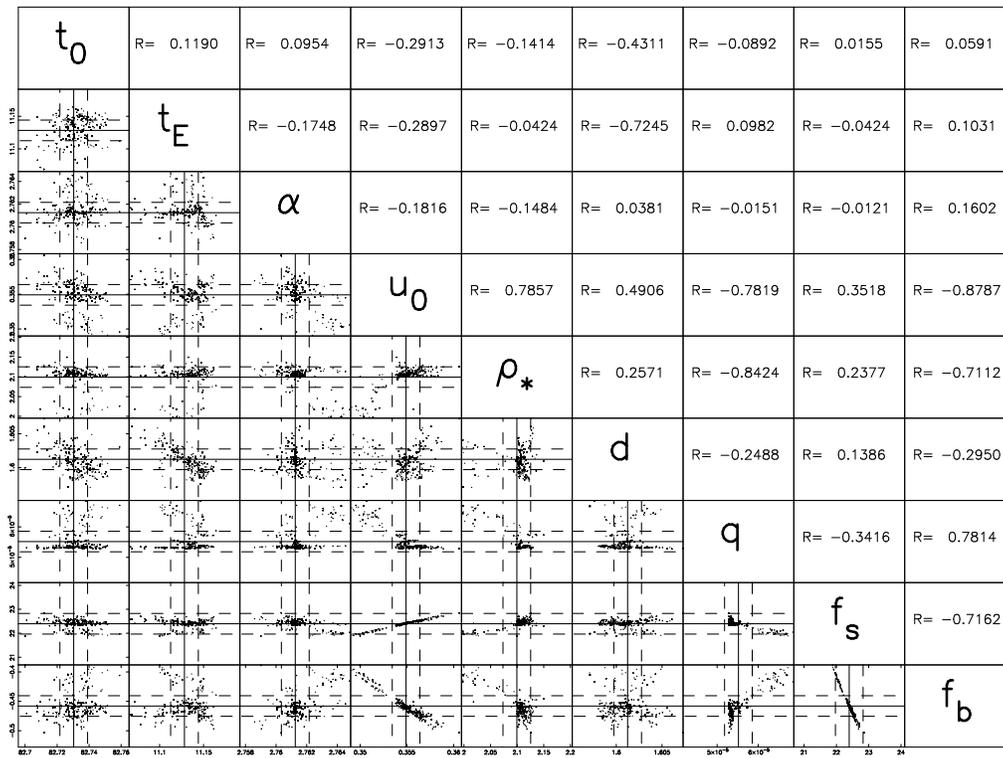


Figure 3.13: Correlations between the 7 model parameters and 2 flux parameters (in this case for the OGLE telescope) for the binary-lens event OGLE-2005-BLG-390.

4

The basis for an automatic binary-lens fitting algorithm

This chapter is based on Kains et al. (2009), MNRAS, 395, 2, 787-796

In this chapter, I outline a method for fitting binary-lens caustic-crossing microlensing events based on the alternative model parameterisation proposed and detailed by Cassan (2008). As an illustration of this methodology, I present an analysis of OGLE-2007-BLG-472, a double-peaked Galactic microlensing event with a source crossing the whole caustic structure in less than three days. In order to identify all possible models we conducted an extensive search of the parameter space, followed by a refinement of the parameters with a Markov Chain-Monte Carlo algorithm. We find a number of low- χ^2 regions in the parameter space, which lead to several distinct competitive best models. We examine the parameters for each of them, and estimate their physical properties. We find that our fitting strategy locates several minima that are difficult to find with other modelling strategies and is therefore a

more appropriate method to fit this type of event.

The work I contributed to the paper on which this chapter is based consisted observing (with Keith Horne), data reduction (with Michael Albrow, Pascal Fouqué), modelling and discussion of results (with Arnaud Cassan). In particular, I received help from Arnaud Cassan to implement the alternative binary-lens parameterisation in my MCMC algorithm.

4.1 Introduction

Anomalous microlensing events usually require very detailed analysis for a full characterisation of their nature to be possible. This applies in particular to a class of microlensing events which display caustic crossing features in their lightcurves. These events are of primary interest, because they account for around ten percent of the overall number of detected microlenses, and they represent an important source of information on physical properties of binary stars (Jaroszynski et al., 2006). However there exist several degeneracies that affect the modelling of this type of event. Without a robust modelling scheme and a full exploration of the parameter space, it is impossible to pin down the true nature of a given event. In addition to this, calculations of anomalous microlensing models for extended sources are very demanding computationally.

Given these issues, brute force is not an option when modelling caustic-crossing events, and one has to devise ways of speeding up calculations, for example by excluding regions of parameter space which cannot reproduce features that appear in datasets. A way to achieve this is to use a non-standard parameterisation of the binary-lens models that ties them directly to data features, as proposed by Cassan (2008), which we recall below.

In this chapter, I present our method for exploring the parameter space, and describe our approach to find all possible models for a given event (Sec. 4.2). We then use OGLE-2007-BLG-472, a microlensing event observed in 2007 by the OGLE and PLANET collaborations, as an illustration of our methodology applied to a binary lens event which intrinsically harbors many ambiguities (Sec. 4.3). We finally discuss the implications of the individual competitive models that we find in order to discriminate between realistic microlensing scenarios.

4.2 Binary-lens events fitting scheme

4.2.1 Parameterisation of binary lens lightcurves

A static binary lens is usually described by the mass ratio $q < 1$ of the two lens components and by their separation d , expressed in units of the angular Einstein radius (Einstein, 1936),

$$\theta_E = \sqrt{\frac{4GM}{c^2} \left(\frac{D_S - D_L}{D_S D_L} \right)}, \quad (4.1)$$

where M is the mass of the lens, and D_L and D_S are the distances to the lens and the source respectively. Such a lens produces caustics where the magnification of the source diverges to infinity for a perfect point source. The positions, sizes and shapes of the caustics depend on d and q . For the binary lens case, caustics can exist in three different topologies, as was discussed in Chapter 2 usually referred as *close*, *intermediate* and *wide*, each of which features caustics in different numbers and shapes. The limits between these configurations are indicated as the dashed lines in e.g. Fig. 4.4 (see also Fig. 2.4).

The description of the lightcurve itself requires four more geometrical parameters in addition to d and q , as discussed previously. Finally for a uniformly bright finite size source star, we add a further parameter, the source radius ρ_* in units of θ_E . However, and as discussed in Cassan (2008), this parameterisation is not well adapted to conducting a full search of the parameter space, because the value of the parameters are not closely related to features present in the lightcurve, namely caustic crossings for the type of event we are discussing in this chapter. Consequently, most of the probed models in a given fitting process do not exhibit the most obvious features in the lightcurve, leading to very inefficient modelling.

To avoid this drawback, Cassan (2008) introduced a new parameterisation in place of α, t_0, u_0 and t_E which is closely related to the appearance of caustic crossing features in the lightcurve. The caustic entry is then defined by a date t_{in} when the source center crosses the caustic¹ and its corresponding (two-dimensional) coordinate ζ_{entry} on the source plane. However, since by definition this point is located on a caustic line, Cassan (2008) introduced a (one-dimensional) caustic arc length s which locates the crossing point directly on the caustic, so that $\zeta_{\text{entry}} \equiv \zeta(s_{\text{in}})$. A given caustic structure is fully parameterised by $0 \leq s \leq 2$. The caustic entry is then characterised by a pair of parameters $(t_{\text{in}}, s_{\text{in}})$, and in the same way the

¹Alternatively, any other point at a fixed position from the source center can be defined as a reference.

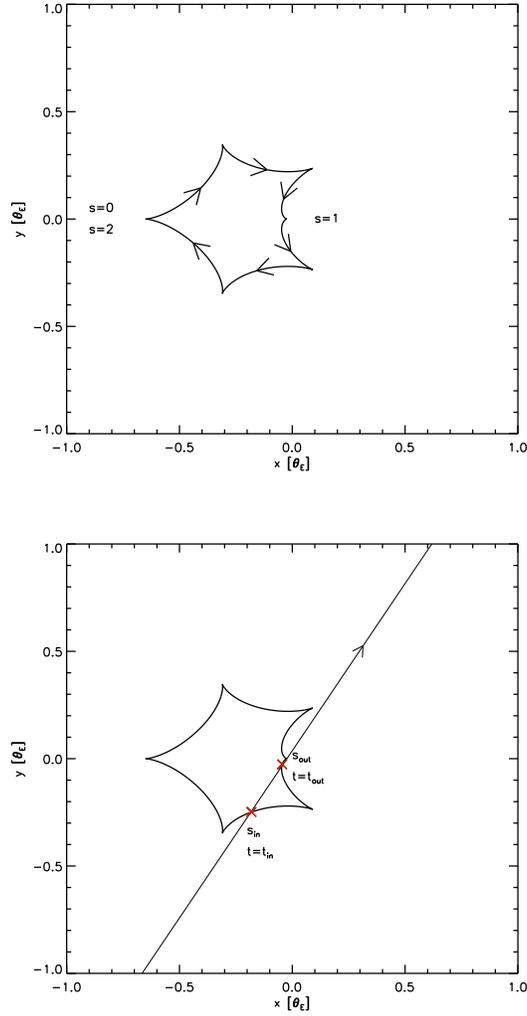


Figure 4.1: The alternative binary-lens parameterisation used in this discussion. *Top:* The s coordinate, which runs along the caustic folds from $s = 0$ to $s = 2$. *Bottom:* The source enters the caustic at time $t = t_{in}$; its location on the caustic fold at that time is $s(t_{in}) = s_{in}$. Similarly, the source exits the caustic at $t = t_{out}$, at which point it is located at $s = s_{out}$. The example caustic used here is for parameters $d = 1.5, q = 0.5$.

caustic exit by (t_{out}, s_{out}) . This parameterisation is illustrated on Fig. 4.1 and Fig. 4.2. These four parameters (in addition to d, q and ρ_*) which describe the caustic crossings therefore also define an alternative parameterisation of the binary lens, far better suited to describing the problem at hand.

4.2.2 Exploration of the parameter space

We start by exploring a wide region of the parameter space with a (d, q) grid regularly sampled on a logarithmic scale. This choice comes from the fact that the size of the caustic struc-

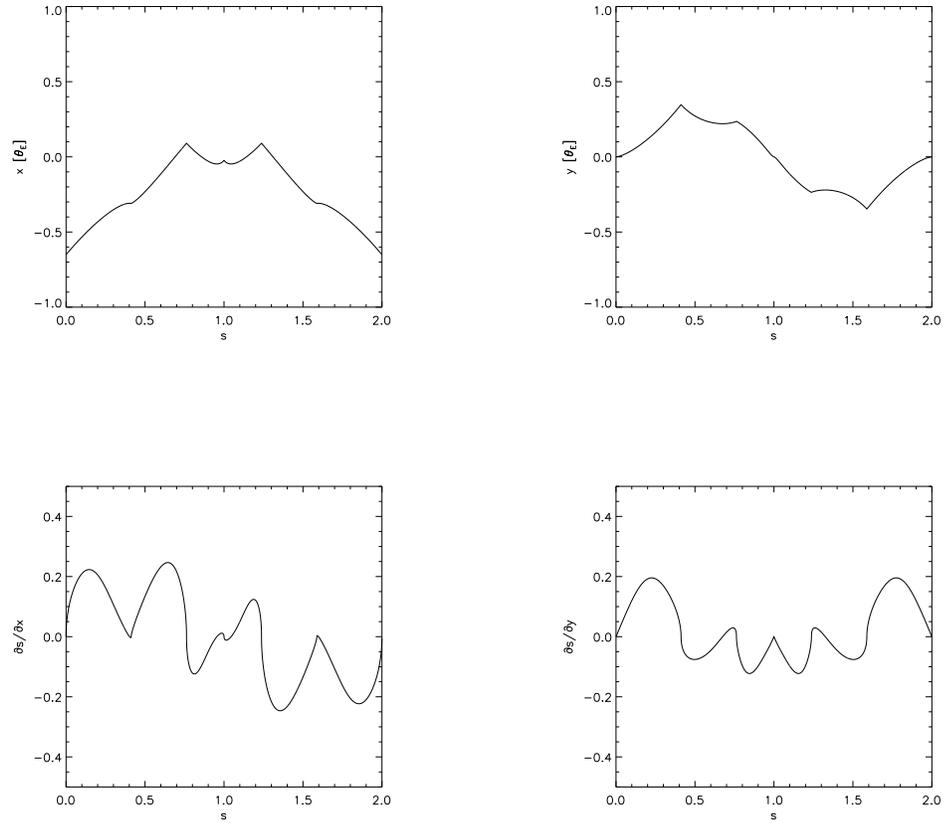


Figure 4.2: The relation between the s coordinate and the components of the source's position ζ , x (top left) and y (top right), as well as the derivatives of s with respect to each of these components (bottom panels). These plots correspond to an intermediate caustic of a binary-lens with $d = 1.1, q = 0.1$, which is plotted on Fig. 4.1.

tures behave like power-laws of the lens separation and mass ratio, as do the corresponding lightcurve anomalies. We fit for the remaining model parameters t_{in} , t_{out} , s_{in} , s_{out} and ρ_* , with (d, q) being held fixed. From this, we then build a $\chi^2(d, q)$ map that we use to locate the best-fit (d, q) -regions. In the wide and close binary cases and following Cassan (2008), we study separately models where the source crosses the central or the secondary caustic by building two $\chi^2(d, q)$ maps, corresponding to each configuration.

In order to sample efficiently and extensively s_{in} and s_{out} (which determine the source trajectory), we use a genetic algorithm (e.g. Charbonneau, 1995) that always retains the best model from one generation to the next (*elitism*). In fact, since we consider only models displaying caustics at the right positions, there are a couple of local minima associated with different $(s_{\text{in}}, s_{\text{out}})$ pairs. These would usually be missed by other minimisation methods, but a genetic algorithm naturally solves this problem in an efficient way. However, since it is difficult to make such an algorithm converge to the best model, we finally refine the model by performing a Markov-Chain Monte-Carlo (hereafter MCMC) fit, using the algorithm detailed in Chapter 3.

From the χ^2 maps, we then identify all the local minima regions and use the corresponding best models found on the (d, q) grid as starting points to refine the parameters, including (d, q) that we now allow to vary. Since the fit is performed within a minimum χ^2 region, the fitting process is very stable and fast.

4.3 Application to OGLE-2007-BLG-472

4.3.1 Alert and photometric follow-up

On 19 August 2007, the OGLE Early Warning System (Udalski, 2003) flagged microlensing candidate event OGLE-2007-BLG-472 at right ascension $\alpha_{2000.0} = 17^{\text{h}}57^{\text{m}}04^{\text{s}}.34$, and declination $\delta_{2000.0} = -28^{\circ}22'02''.1$ or $l = 1.77^{\circ}$, $b = -1.87^{\circ}$.

The OGLE lightcurve has an instrumental baseline magnitude $I=16.00$, which may differ from the calibrated magnitude by as much as 0.5 magnitudes. Lensing by the star in the point source-point lens (hereafter PSPL) approximation accounts for the broad rise and fall in the lightcurve, peaking around $\text{MHJD}^2=4334.0$ with an apparent half-width at half-peak of about 10 days (Fig. 4.3). Although the observed OGLE flux rises only by 0.06 mag in the

²MHJD=HJD-2450000

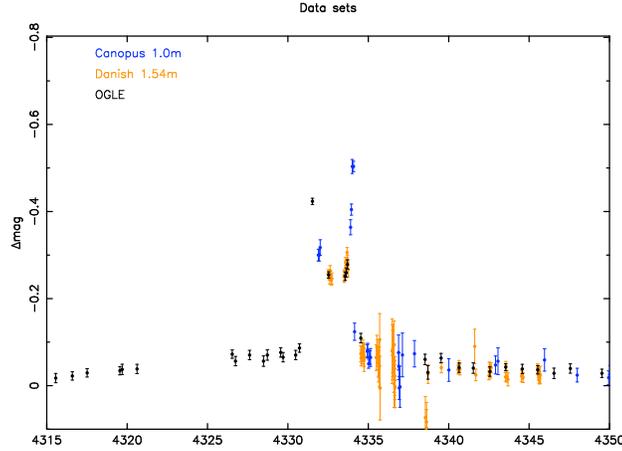


Figure 4.3: OGLE, UTas and Danish data sets for OGLE-2007-BLG-472. Data points are plotted with $1\text{-}\sigma$ error bars. The x -axis is time in HJD-2450000.

non-anomalous part of the lightcurve, the shape of the curve hints that blending is important for this target, with only $\sim 12\%$ of the baseline flux due to the un-magnified source.

On 19 August (MHJD=4331.5) an OGLE data point showed a sudden brightening of the source, with subsequent PLANET (UTas Mt. Canopus 1.0m telescope in Tasmania and Danish 1.54m telescope at La Silla, Chile, where I was observing at the time) and OGLE data indicating what appears to be a fold caustic crossing by the source, ending with a PLANET UTas data point on August 21 (MHJD=4334.1). The caustic entry is covered by a single OGLE point, while the caustic exit is well covered by our UTas data set (Fig. 4.3). Treating the lightcurve as the addition of an anomaly to a PSPL lightcurve, the underlying PSPL curve then apparently reaches peak magnification on August 22 (MHJD=4335.45). Particularly crucial in our data set is the UTas observation taken within a few hours of the caustic exit, which tightly constrains the position of the caustic exit on the lightcurve, and the size of the source. Although V-band observations were taken, the V lightcurve of this event does not sample the time when the source was magnified significantly, and therefore does not provide us with constraints on the properties of the source.

4.3.2 Data reduction

We reduced the PLANET data for this event using the data reduction pipeline pysis 3.0 (Albrow et al., 2009). This pipeline uses a kernel as a discrete pixel array, as proposed by Bramich (2008), rather than a linear combination of basis functions. This has the advantage that it removes the need for the user to select basis functions manually, which can lead to problems if inappropriate functions are chosen. In addition to this, the pixel array kernel copes better with

Table 4.1: OGLE-2007-BLG-472 datasets and error bar rescaling factors.

Telescope	Data	Error bar rescaling factor
UTas 1.0m	34	1.79
Danish 1.54m	84	1.55
OGLE	857	1.21

images that are not optimally aligned. The result of using this pipeline is a better reduction than was obtained with other methods. We kept all points with seeing < 3.5 arcseconds. Although some dubious points remain with this simple cut, the size of their associated error bars reflects their lack of certainty and ensures their weight in any modelling procedures is appropriately reduced. Our final data set consists of 34 UTas data points, 84 points from the Danish 1.54m telescope, and 857 points from OGLE (Table 4.1).

4.3.3 Modelling OGLE-2007-BLG-472

After a first exploration of the parameter space, we find a best model (close to model C_c , see below) which we use as a basis to rescale our error bars. In fact, these can vary rather widely from one telescope to another and are often underestimated by photometry software. Ignoring this effect would misrepresent the relative importance of the datasets. From this step, we choose the rescaling factors shown in Table 4.1, obtained by setting $\chi^2/\text{d.o.f.} \simeq 1$ for each data set. We then use the rescaled data to perform a new parameter space exploration.

We then apply the fitting scheme detailed in section 4.2 to our datasets. In particular, we choose a spacing between the (d, q) grid points of 0.070 in $\log d$ and 0.275 in $\log q$. For the genetic algorithm fit, we use a model population of 200 individuals evolving over 40 generations, which has proven to be enough to safely locate the regions of minimum χ^2 . Finite source effects are computed using the adaptive contouring method of Dominik (2007).

The final $\chi^2(d, q)$ maps that we obtain are plotted in Fig. 4.4 for the intermediate and central caustic configurations, and Fig. 4.5 for the intermediate and secondary caustic. The red crosses show the underlying (d, q) grid, and the blue shaded contours indicate values of $\Delta\chi^2 = 5, 20, 50, 100, 250$, where the reference model is C_s , the global best-fitted model (as obtained in Section 4.3.5).

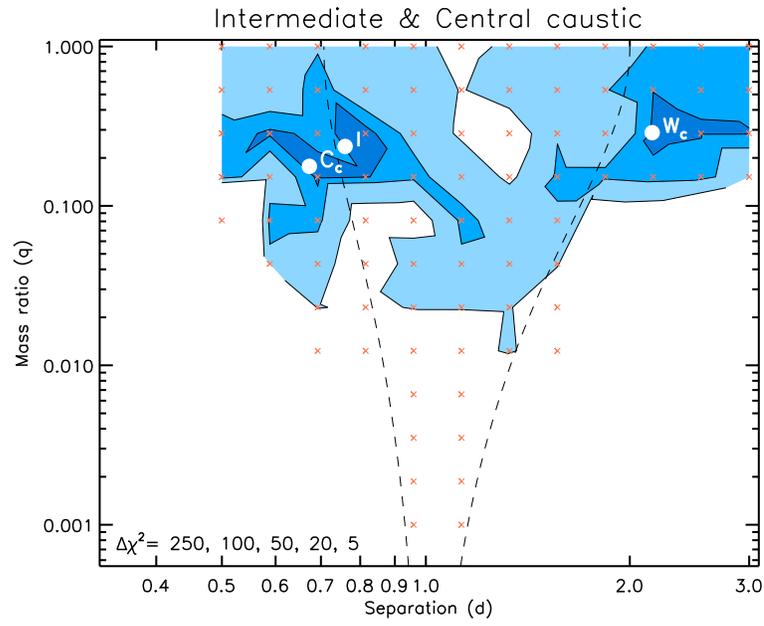


Figure 4.4: $\chi^2(d, q)$ map for the intermediate and central caustic configurations. Contour lines and minima regions (in blue shades) are plotted at $\Delta\chi^2 = 5, 20, 50, 100, 250$. The two dashed curves are the separation between the close, intermediate and wide regimes. The best-fit models are labelled and marked with white filled circles.

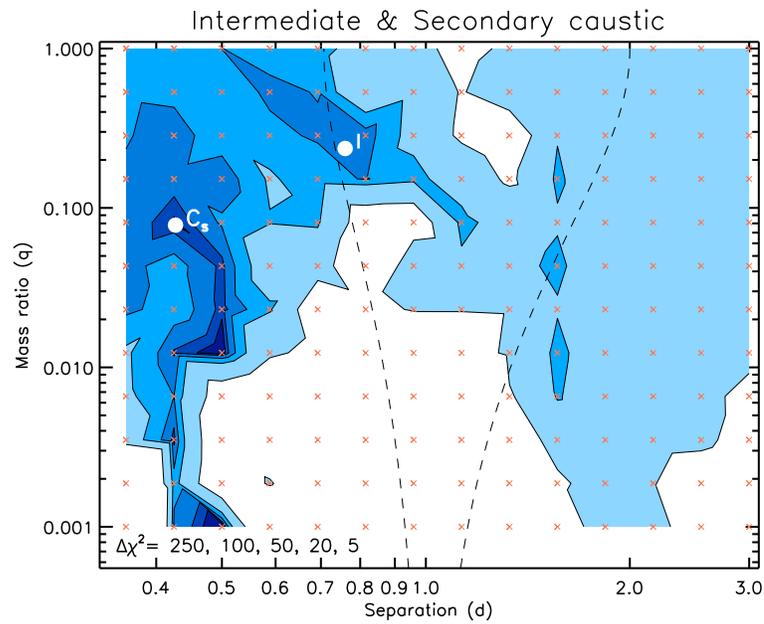


Figure 4.5: Same as Fig. 4.4 for the the intermediate and secondary caustic configuration.

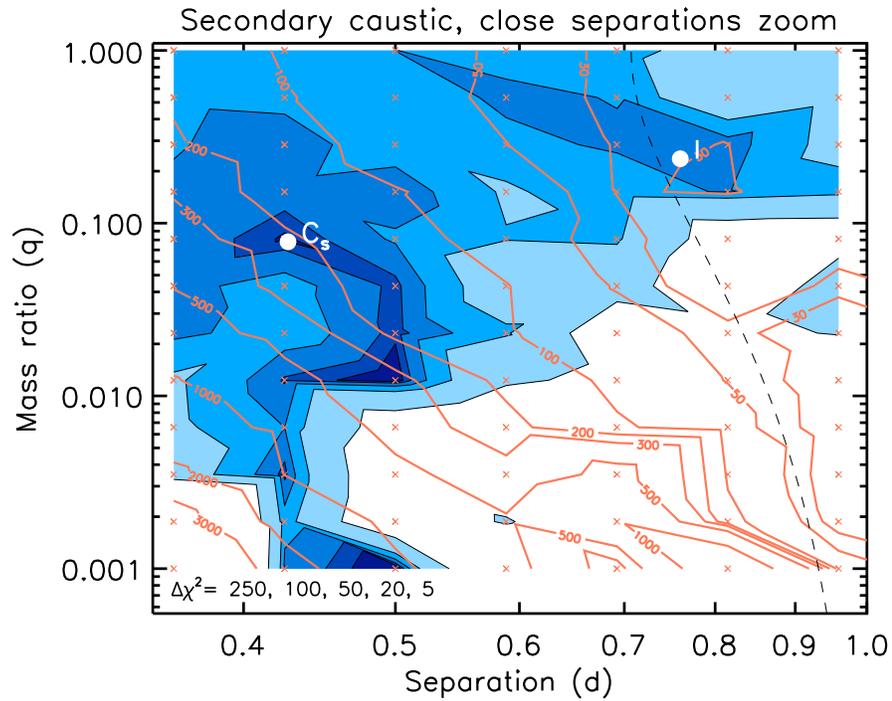


Figure 4.6: Map of the value of t_E in the (d, q) plane for converged models at each grid point, superimposed on the χ^2 map, zoomed in on the close regime part of parameter space. Contours lines (orange) are labeled with their corresponding value of t_E while χ^2 contour lines are plotted at $\Delta\chi^2 = 5, 20, 50, 100, 250$ and filled with gradual shades of blue. The dashed curve is the separation between the close and intermediate regimes. The models of Table 4.2 are labelled and marked with white filled circles.

4.3.4 Excluding minima

Fig. 4.6 shows a zoom on the $d < 1$ region of the χ^2 map for a source crossing a secondary caustic, with an overplot of t_E isocontours (orange lines) roughly equally spaced on a logarithmic scale. With this fitting approach, we put no initial constraints on the Einstein time t_E , though it will always remain physical ($t_E > 0$). Very good fits to the data are obtained with values of $t_E > 300$ days, which correspond to the minimum region in the left lower parts of Fig. 4.5 and Fig. 4.6. Such long Einstein times are unlikely, and some of the values found for t_0 correspond to a lightcurve that reaches its peak well in the future; these are very unlikely to be acceptable solutions. If we adopt the *posterior* t_E distribution of Dominik (2006), then $t_E > 400$ days is well in the tail of the distribution. Thus in the following, we will not consider solutions with t_E greater than 400 days. This means that we will exclude the low- q ($q \sim 0.001$) minima in the following discussion.

Although a very well-covered lightcurve generally enables a good characterisation of the deviation caused by the caustic approach or crossing, degeneracies make finding a unique best-fitting model difficult. In particular, Griest & Safizadeh (1998) and Dominik (1999) identified a two-fold degeneracy in the projected lens components separation parameter d , under the change $d \leftrightarrow 1/d$, when $q \ll 1$. Moreover, Kubas et al. (2005) showed that very similar lightcurves could arise for a source crossing the secondary caustic of a wide binary system and for the central caustic of a close binary system. These degeneracies cause widely separated χ^2 minima in the parameter space, which must then be located by exploring the parameter space thoroughly. In addition to these degeneracies, imperfect sampling can increase the number of local χ^2 minima; short events in particular are prone to under-sampling, leading to difficulties in modelling. OGLE-2007-BLG-472 is an example of this, as shown in the next section.

4.3.5 Refining local minima

We see from Fig. 4.4 (intermediate and central caustic) that there are three broad local minima in the region around the white filled circles marked as C_c , I and W_c (“I”, “C” and “W” for intermediate, close and wide models respectively, and subscript “c” for central caustic). In Fig. 4.5 (intermediate and secondary caustic), a best-fit region can easily be located around the region marked C_s (subscript “s” for secondary caustic), besides region I .

Now allowing for the parameters d and q to vary as well, we use our MCMC algorithm

Table 4.2: Best-fitting binary lens model parameters for OGLE-2007-BLG-472. The blending factor $g(I) = F_B(I)/F_S(I)$ is given for the OGLE data (I -band). The error bars were rescaled for each telescope by the factor given in Table 4.1, which lead to the rescaled χ^2 indicated here. Physical parameters are also given for each model, for the case of a lens in the disc, and a lens in the bulge. These were calculated using the procedure detailed in Sec. 4.3.7

Parameter	Model C_s	Model C_c	Model I	Model W_c	Units
χ^2 (rescaled σ)	949.00	963.16	972.48	988.55	—
$\Delta\chi^2$	—	13.2	23.5	39.8	—
χ^2_{UTas}	23.79	24.83	26.41	28.86	—
χ^2_{Danish}	79.77	79.60	80.75	88.93	—
χ^2_{OGLE}	845.50	858.77	865.24	870.55	—
t_0	4587.18 \pm 0.80	4332.27 \pm 0.29	4332.10 \pm 0.27	4334.99 \pm 0.28	MJD
t_E	213.82 \pm 1.04	52.00 \pm 3.63	38.32 \pm 2.60	53.46 \pm 0.81	days
α	2.810 \pm 0.006	3.227 \pm 0.030	3.305 \pm 0.037	4.570 \pm 0.018	rad
u_0	-1.573 \pm 0.013	0.091 \pm 0.005	0.164 \pm 0.019	0.277 \pm 0.010	—
$d^*/10^{-3}$	0.34 \pm 0.01	0.98 \pm 0.19	1.55 \pm 0.16	1.33 \pm 0.05	—
d	0.427 \pm 0.002	0.673 \pm 0.011	0.760 \pm 0.015	2.158 \pm 0.0169	—
q	0.078 \pm 0.001	0.177 \pm 0.017	0.236 \pm 0.024	0.288 \pm 0.0096	—
$g(I) = F_B(I)/F_S(I)$	7.15 \pm 0.013	68.11 \pm 0.013	40.13 \pm 0.09	56.98 \pm 0.019	—
I_s	17.89 \pm 0.01	20.21 \pm 0.01	19.65 \pm 0.09	20.02 \pm 0.01	—
I_b	15.75 \pm 0.01	15.63 \pm 0.01	15.64 \pm 0.09	15.63 \pm 0.01	—
$(V-I)_s$	1.80 \pm 0.10	1.93 \pm 0.11	1.91 \pm 0.11	1.92 \pm 0.12	—
θ^*	1.18 \pm 0.24	0.46 \pm 0.09	0.59 \pm 0.12	0.50 \pm 0.10	μas
Lens in the Disc					
M_1	1.50 ^{+1.85} _{-0.58}	0.42 ^{+0.40} _{-0.22}	0.34 ^{+0.36} _{-0.18}	0.34 ^{+0.37} _{-0.18}	M_\odot
M_2	0.12 ^{+0.14} _{-0.05}	0.07 ^{+0.07} _{-0.04}	0.08 ^{+0.08} _{-0.04}	0.10 ^{+0.11} _{-0.05}	M_\odot
D_L	1.00 ^{+0.95} _{-0.36}	5.7 ^{+1.1} _{-1.5}	6.1 ^{+1.1} _{-1.5}	6.1 ^{+1.0} _{-1.5}	kpc
v	25 ⁺²⁴ ₋₉	80 ⁺¹⁵ ₋₂₁	93 ⁺¹⁶ ₋₂₂	67 ⁺¹¹ ₋₁₆	km s^{-1}
Lens in the Bulge					
M_1	41 ⁺¹⁴ ₋₁₄	1.25 ^{+1.47} _{-0.59}	0.79 ^{+0.93} _{-0.35}	0.79 ^{+0.94} _{-0.36}	M_\odot
M_2	3.2 ^{+1.1} _{-1.1}	0.22 ^{+0.26} _{-0.10}	0.19 ^{+0.22} _{-0.08}	0.23 ^{+0.27} _{-0.10}	M_\odot
D_L	6.7 ^{+0.4} _{-0.7}	7.3 ^{+0.6} _{-0.8}	7.3 ^{+0.6} _{-0.8}	7.3 ^{+0.6} _{-0.8}	kpc
v	167 ⁺¹⁰ ₋₁₇	102 ⁺⁸ ₋₁₂	111 ⁺¹⁰ ₋₁₂	79 ⁺⁶ ₋₈	km s^{-1}

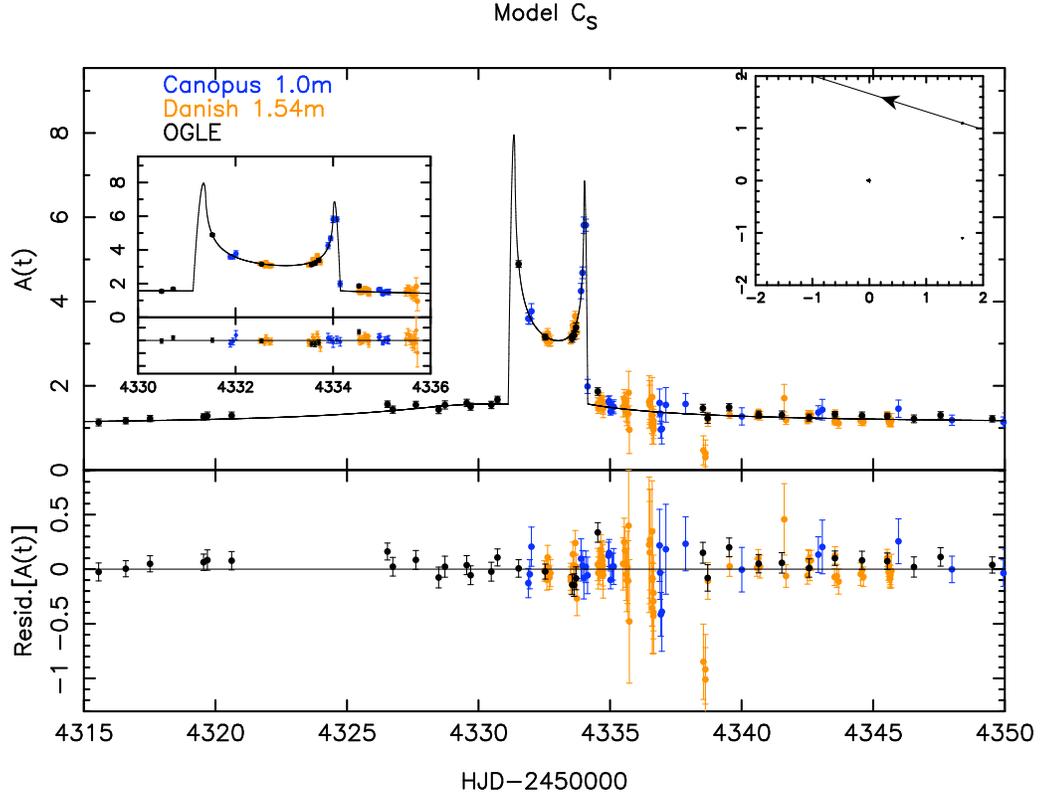


Figure 4.7: Best-fitting binary lens model C_s with residuals and a zoom on the anomaly (left inset). Data points are plotted with $1\text{-}\sigma$ error bars. The trajectory of the source in the lens plane with the caustics is plotted as an inset in the top right corner of the figure, with the primary lens component located at the coordinate system’s origin. Axes for the insets are x and y positions in the lens plane, in units of the Einstein radius. Caustics are very small for the values of d and q of model C_s and are difficult to see on the plot.

to find the best solutions in each of these local minimum regions. These are identified with white filled circles on Fig. 4.4 and 4.5 and correspond to the models listed in Table 4.2, and shown in Fig. 4.7, 4.8, 4.9 and 4.10. The best model lightcurve is dominated by strong caustics, which all viable models must reproduce, with the low-magnification base PSPL curve barely noticeable. All models have the first anomalous OGLE points on the descending side of the caustic entry except for the worst model, model W_c , which has this OGLE point on the ascending part of the caustic entry. Statistically, the former case is more likely to be observed since the ascending part of the caustic entry happens much more rapidly than the descending side.

Our best model, C_s , has $\chi^2 = 949$ for 975 data points, with the other competitive models at $\Delta\chi^2 = 13.2$ (model C_c), $\Delta\chi^2 = 23.5$ (model I) and $\Delta\chi^2 = 39.6$ (model W_c).

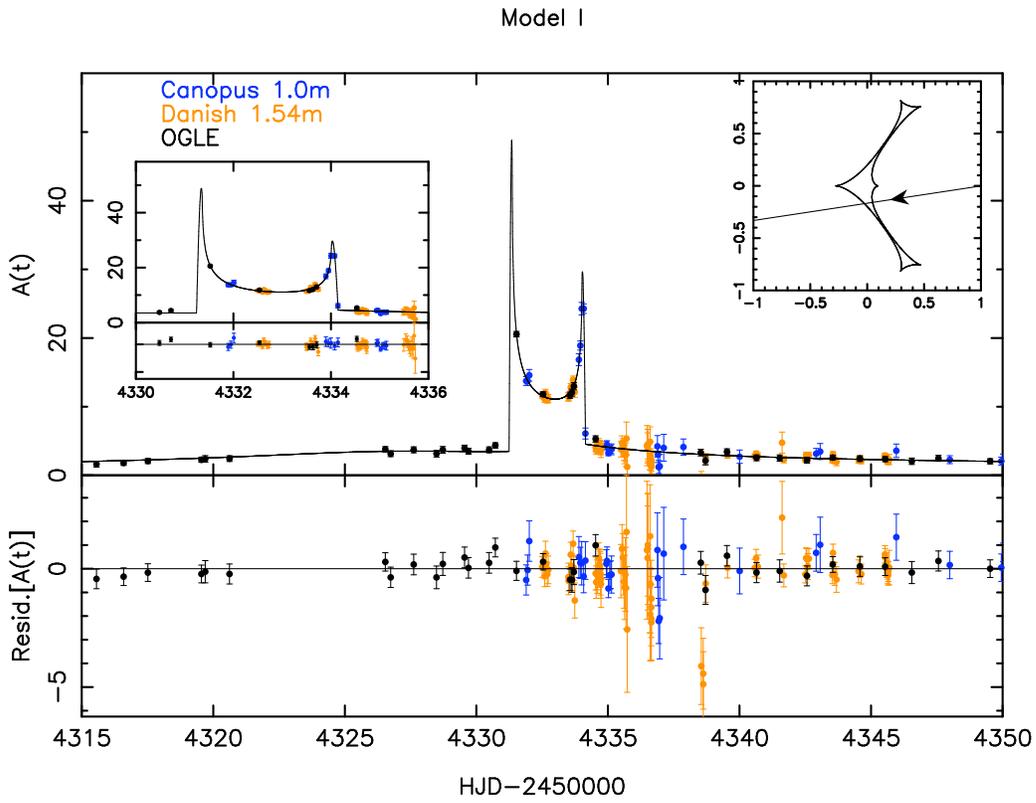


Figure 4.8: Same as Fig. 4.7 for model I.

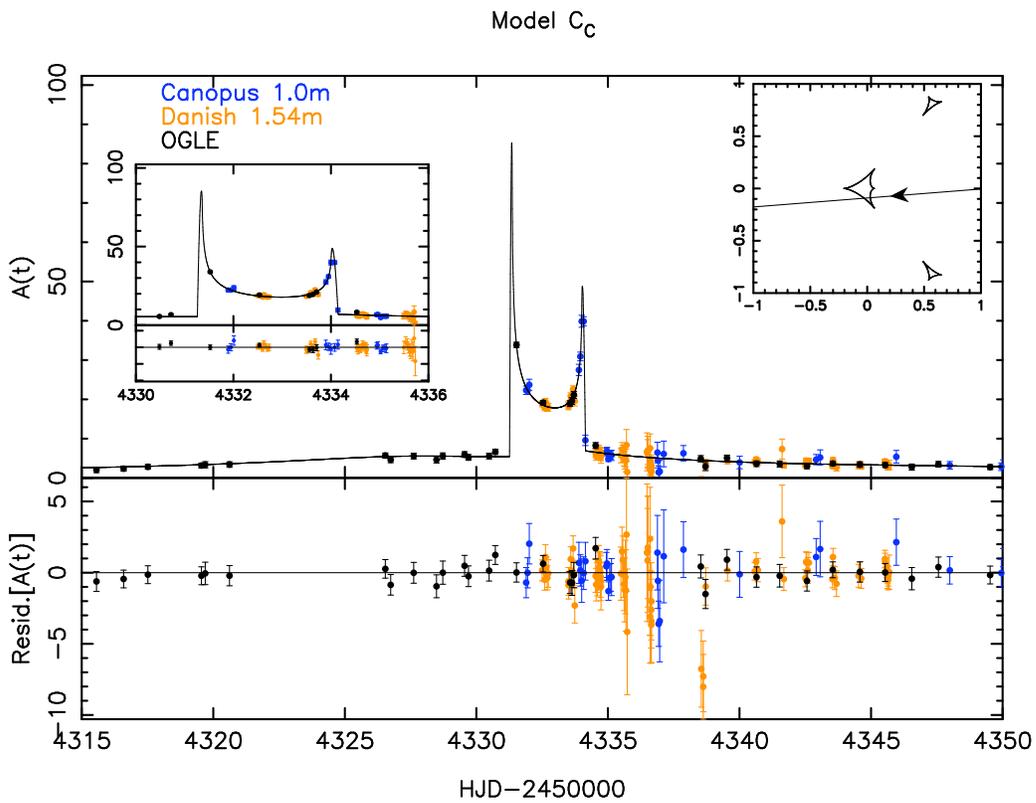


Figure 4.9: Same as Fig. 4.7 for model C_c .

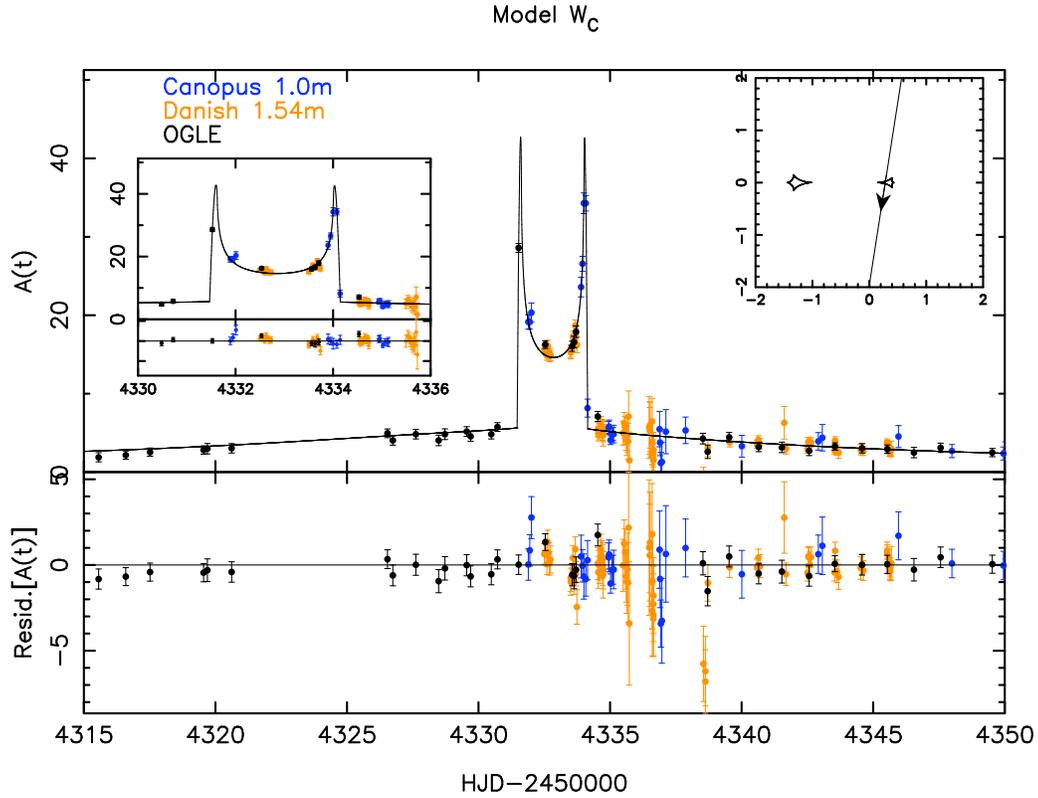


Figure 4.10: Same as Fig. 4.7 for model W_C .

4.3.6 Parameter correlations

Fig. 4.6 shows that the models with a source crossing a secondary caustic have increasingly large values of t_E as they go towards lower values of the mass ratio. This is expected since the time Δt between t_{in} and t_{out} is fixed by the data. As the size of caustics scales with $q^{1/2}$, and $t_E \sim \Delta t/q^{1/2}$, the source must therefore cross the Einstein Ring over a longer timescale as q is reduced, for Δt to be conserved. In addition to this, blending decreases for decreasing values of q , and therefore decreases with increasing t_E , contrary to what might be expected. Indeed, one would expect the blending factor $g = F_B/F_S$ (where F_B and F_S are the blend and source flux respectively) to increase with increasing t_E in order to mask long timescales and reproduce the observed timescale. However in this region of parameter space, the caustics are weak, which means that too much blending would not allow models to reproduce the observed rise in the source magnitude at the caustic entry and caustic exit. For a region of parameter space to contain satisfactory models, there must therefore be a fine balance between blending, timescale and mass ratio.

For models where the source crosses a central caustic, the impact parameter u_0 must de-

crease with decreasing mass ratio, since the size of central caustic decreases with decreasing mass ratio, and the range of allowed u_0 decreases if the source must cross the caustic. This means that for smaller mass ratios, blending will have to increase in order to mask the correspondingly higher PSPL magnification of the source that results from the smaller impact parameter.

4.3.7 Physical properties of the models

Source characteristics

A colour-magnitude diagram of the field (Fig. 4.11) was produced extracting 1497 stars from I and V images at $t = 4340.08$ (I) and $t = 4340.13$ (V) taken at the Danish 1.54m telescope. The combination of the source and the blend lies very slightly blueward of the red giant clump, at $(V - I) = 2.43$. All the models, however, are heavily blended (Table 4.2). The source magnitude and blending magnitude for each model can be found using the equations $I_s = I_{\text{base}} + 2.5 \log(1 + g)$ and $I_b = I_s - 2.5 \log(g)$.

Using this equation, we find I-band source magnitudes ranging from 17.89 (model C_s) to 20.21 (model C_c) (see Table 4.2). Our V-band data set does not allow us to determine the source's colour, but assuming that the source is a main sequence star we use the calculated I magnitude of the source for each model to estimate a colour, using the results of Holtzman et al. (1998). This then enables us to estimate the source's angular radius which we use in Section 4.3.7 to compute probability densities of the lensing system's properties.

We calibrate the baseline magnitude of our target (source and blend combined) using the location of the red clump as a reference. We find $I_{\text{base}} = 15.61 \pm 0.10$, which is in agreement with the OGLE value of $I_{\text{base}} = 16.00 \pm 0.50$. Comparing this to the location of the red clump, we can derive an estimate for the reddening coefficient A_I . From Hipparcos results, Stanek & Garnavich (1998) find an absolute magnitude for the red clump at $M_{I,RC} = -0.23 \pm 0.03$. Using a distance modulus to the galactic centre of $\mu = 14.41 \pm 0.09$ (i.e. assuming $D_S = 7.6$ kpc) (Eisenhauer et al., 2005), this translates to a dereddened magnitude for this target of $I_{\text{base}} = 14.18 \pm 0.09$. Hence using the relation $A_I = I_{\text{base}} - M_{I,RC} - \mu$, we get a value for the I-band reddening parameter of $A_I = 1.43 \pm 0.13$. Alternatively, fitting 2MASS isochrones to our CMD, we obtain a value $A_I = 1.46 \pm 0.08$ and $E(V - I) = 1.46 \pm 0.11$. We use these values of reddening to determine dereddened magnitudes and colours for the source of each model.

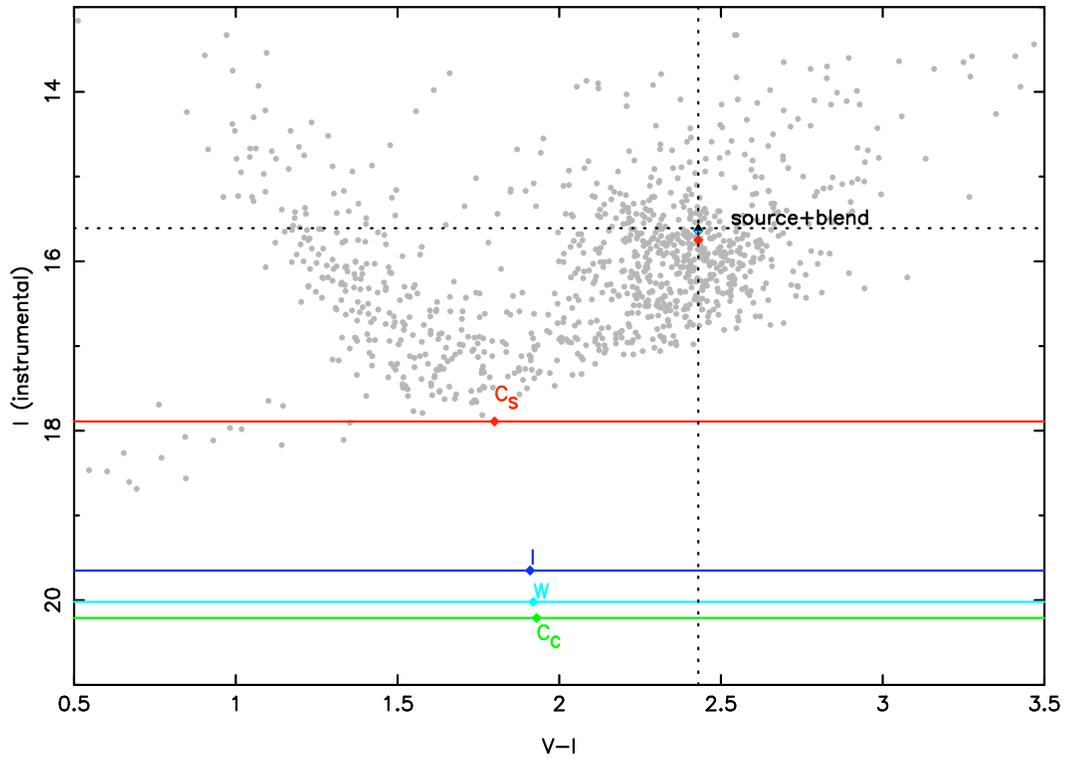


Figure 4.11: Colour - Magnitude diagram of the field. The target OGLE-2007-BLG-472 is shown as a black triangle at $(V - I, I) = (2.43, 15.61)$. The position of the deblended source for each model is labeled and indicated by a solid line and a coloured diamond, with the blend for each model also plotted as a diamond in the same colour.

These, together with the surface brightness relations from Kervella & Fouqué (2008), allow us to calculate the apparent angular radius of the source θ_* for each of the models, given in Table 4.2.

Lens characteristics

Although the characteristics of any microlensing event depend on various properties of the lensing system, including the mass of the lenses, the only measurable quantity that can be directly related to physical properties of the lens is the timescale of the event t_E . While the physical properties of the lensing system can be fully constrained when the photometry is affected by both finite source-size effects and parallax. When these are not measured, such as is the case with our analysis OGLE-2007-BLG-472, we can still use Bayesian inference to determine probability densities of physical properties of the lens, based on a chosen Galactic model. We have chosen not to include parallax in our analysis because its effect would be very small for such a low-magnification event; in addition to this, we are only seeking a first-order analysis of binary-lens events with our current method, although second-order effects such as parallax and lens rotation will be taken into account in future work.

We use our fitted value of the source size parameter ρ_* to place constraints on the mass of the lens, which can be expressed as a function of fractional distance $x = D_L/D_S$ and the source size ρ_* as (e.g. Dominik 1998a)

$$\frac{M(x)}{M_\odot} = \frac{c^2}{4GM_\odot} \frac{D_S \theta_*^2}{\rho_*^2} \frac{x}{1-x}, \quad (4.2)$$

where M is the mass of the lens, θ_* is the angular radius of the source, the value of which is given in Table 4.2, and other quantities are defined as before. The mass-distance curve showing constraints from this equation is plotted on Fig. 4.13.

Since we did not measure parallax for this event, we use a probabilistic approach following that of Dominik (2006) to derive probability densities for physical properties of lens components. The Galactic model used here is a piecewise mass spectrum (e.g. Chabrier 2003), two double exponentials for the disc mass density and a barred bulge tilted at an angle of 20° with the direction to the Galactic centre (Dwek et al. 1995), and the distribution of effective transverse velocities used in Dominik (2006).

Using this Galactic model, we infer properties for the lensing system, separating the cases

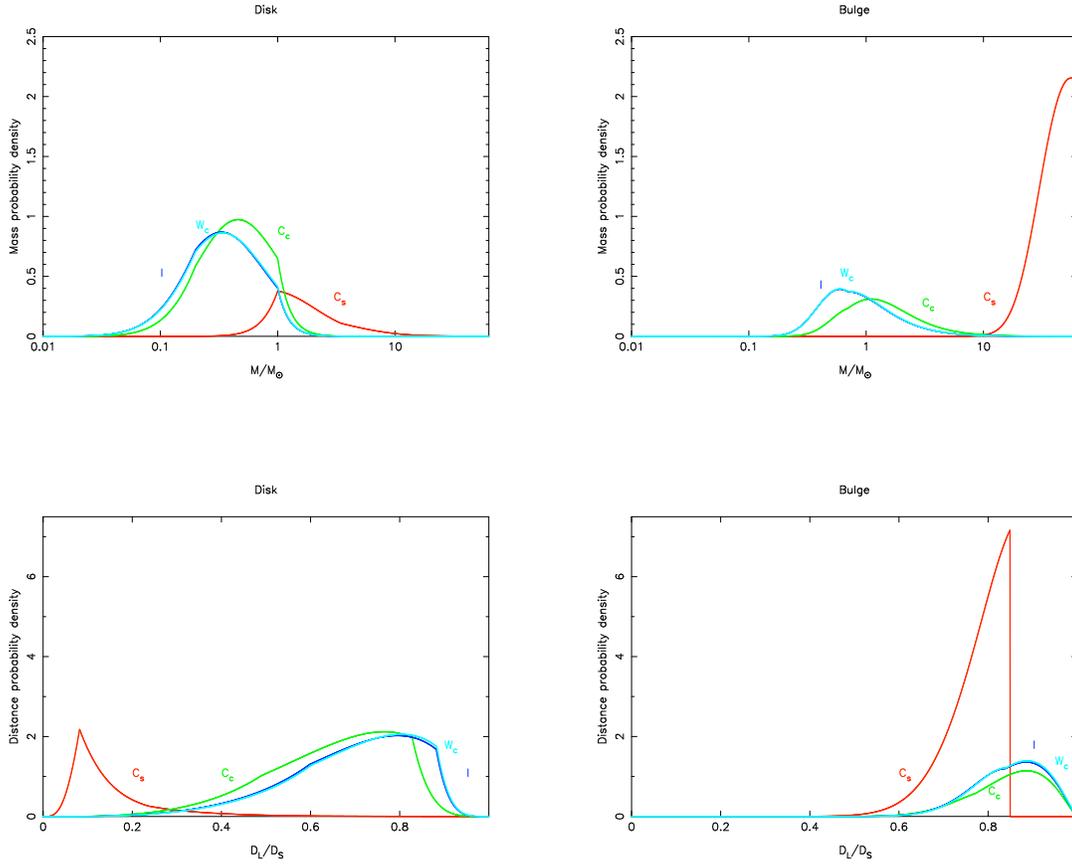


Figure 4.12: Probability densities for the mass of the primary lens star and the fractional distance D_L/D_S , for a lens in the disc (left side) and a lens in the bulge (right side). The values quoted in Tables 4.2 & 4.2 are the median value and the limits of the 68.3% confidence interval. On each plot, the probability densities are plotted for model C_s (red), model C_c (green), model I (dark blue), and model W_c (light blue).

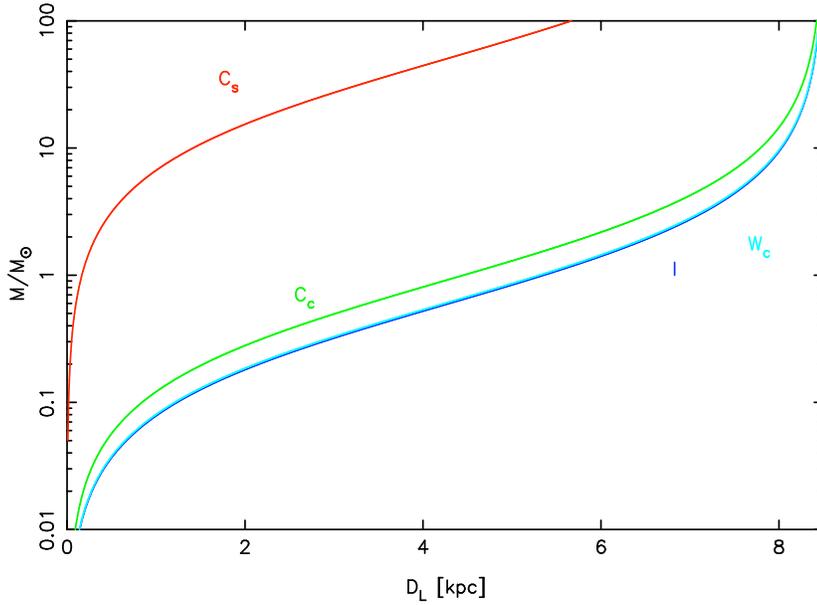


Figure 4.13: Mass-distance diagram showing the constraint on the lens mass from the source size, given by Eq. (4.2), for each model. The curves are labeled with the name of the model to which they correspond.

where the lens is in the Galactic disc and in the Galactic bulge. For a lens in the disc, we find a primary mass $1.50^{+1.85}_{-0.58}M_{\odot}$ and a secondary mass of $0.12^{+0.14}_{-0.05}M_{\odot}$, at a distance of $1.00^{+0.95}_{-0.36}$ kpc with a lens velocity of 25^{+24}_{-9} km s $^{-1}$. For a lens in the bulge, we find a primary mass $41^{+14}_{-14}M_{\odot}$ and a secondary mass of $3.2^{+1.1}_{-1.1}M_{\odot}$, at a distance of $6.7^{+0.4}_{-0.7}$ kpc with a lens velocity 167^{+10}_{-17} km s $^{-1}$. These are the physical lens properties for the lowest- χ^2 model (model C_s). The values of these physical parameters for the other models are given in Table 4.2. Probability densities of these properties for all models are plotted on Fig. 4.12.

Discussion

For our lowest- χ^2 model, the parameters we find imply very unusual properties of the lensing system. As discussed in Sec. 4.3.4, the fact that we find these types of models is a consequence of the fitting approach we are taking. Traditional fitting methods would struggle to find these minima, since most of them require a human modeller providing a starting point in parameter space. This is an issue when solely using an MCMC algorithm to fit microlensing events: although an MCMC run may be able to make its way through parameter space to find minima reasonably far away from its starting point, it is highly unlikely that a chain will be able to reach a minimum that has significantly different parameters from the starting point. As we see from Fig. 4.6, there are minima in many parts of parameter space, with values of t_E that are different by almost two orders of magnitude. These parameters are non-intuitive, since

they cannot be guessed only by looking at the lightcurve. As a result, it is improbable that this kind of parameter will be used as starting points for "classic" fitting algorithms.

We solve this problem for the static binary-lens case by using the method described in Sec. 4.2.2. Using this approach, we systematically locate minima throughout the parameter space. However, we must then be careful with interpreting the significance of the obtained model parameters. The shape of probability densities shown in Fig. 4.12 for model C_s indicates that our value of t_E pushes the lens mass towards the end of the adopted mass spectrum in the Galactic model we have adopted. This results in the abrupt transitions seen on Fig. 4.12.

Similarly, the mass-distance curve for model C_s on Fig. 4.13 shows that the mass of the lens rapidly becomes very large for lenses beyond ~ 1 kpc. These unusual curves are caused by a value of $t_E \sim 200$ days. Models with $t_E \sim 3000$ days (corresponding to the low- q minimum visible on Fig. 4.5 & 4.6) are obviously not acceptable, but how can we formally reject them? Finding these models from minima in the χ^2 surface shows the limits of using χ^2 as the sole criterion for favouring models. A solution to this would be to use prior distributions on as many of the parameters as we can. During the MCMC part of our fitting process, this would mean that we obtain posterior distributions that are different from the ones obtained without using prior distributions on the parameters, or, equivalently, assuming uniform priors for all parameters. Such priors can be obtained in various ways, such as looking at the distribution of timescales for past microlensing events or calculating these distributions from Galactic models (e.g. Dominik 2006), or by using luminosity functions of the Galactic bulge to find a prior for the blending factor g (e.g. Holtzman et al. 1998). Such work requires careful consideration of which priors are most appropriate to use, and is discussed in Chapter 5. Using these priors in combination with our method to find minima will lead to more robust determination of minima by taking into account our knowledge of physical parameter distributions.

4.4 Summary and prospects

Our analysis of OGLE-2007-BLG-472 is a good illustration of the importance and power of using parameters that are related to lightcurve features. Indeed, despite incomplete coverage of the caustic entry and high blending, a few crucial data points and an appropriate choice of non-standard parameters enable us to find several good binary-lens model fits to our data for this event by exploring the parameter space systematically with a coarse grid. Some of the

good fits that we identify have unphysical parameters, and we must then reject them. However, using this parameterisation allows us to be certain that the parameter space has been thoroughly explored. We find four models with different parameters: two close binary models, one intermediate configuration, and a wide binary model. The lowest- χ^2 model corresponds to a G dwarf star being lensed by a binary system with component masses $M_1 = 1.50_{-0.58}^{+1.85}M_\odot$ and $M_2 = 0.12_{-0.05}^{+0.14}M_\odot$, which are compatible with our blending values. However, it is obvious from physical parameter distributions that using χ^2 as a sole criterion for determining the best model is insufficient, because it does not take into account our knowledge of the distributions of physical parameters.

Since the approach presented in this chapter can form the basis for a systematic, wide ranging exploration of the parameter space to localise all possible models for a given dataset, it is particularly relevant to current efforts to automate real-time fitting of binary-lens events. This could prove useful to provide faster feedback on events being observed, and prioritise observing schedules, especially on robotic telescopes. Expanding robotic telescope networks controlled by automated intelligent algorithms are expected to play an increasingly important role in microlensing surveys in the coming years (e.g. Tsapras et al. 2009). Fitting methods such as the one described in this chapter are essential for making sure any anomalies are interpreted correctly, and that minima are located in as large a part of parameter space as possible.

5

Towards a full Bayesian approach for automatic fitting of binary-lens events

5.1 Introduction

In the previous chapter, I showed that a parameterisation such as that formalised by Cassan (2008) could be used to conduct a systematic search of the binary-lens microlensing parameter space. For this discussion, I used the special (but frequently occurring) case of caustic-crossing events, analysing such an event from the 2007 observing season, OGLE-2007-BLG-472 to illustrate the usefulness of the method (Kains et al., 2009). For that event, several minima were found in the χ^2 hypersurface, some of them corresponding to models with very unusual parameters. The reason for this was discussed in the previous chapter: since we are forcing the source to go through a caustic in order to reproduce the features seen in observational data, when caustics become very small, we are essentially forcing the source to pass through the eye of a needle. Since the timing of the caustic crossing does not change,

being well constrained by data, the only way to reproduce the observed lightcurve when the source goes through a very small caustic with the correct timing is to make it move very slowly through the caustic structure, in other words, to make the timescale t_E of the event very long, of the order of hundreds or even thousands of days.

In practice, however such long timescales are rarely observed in microlensing events, because the relative velocity between source and lens is generally too fast to lead to events lasting much longer than 100 or 200 days. Therefore we know that if an event is equally well fitted by a model with $t_E = 500$ days and a model with $t_E = 30$ days, we can conclude that the most likely model should be the latter. This information has to be quantified in some way; this can be most easily achieved in the form of Bayesian priors. The aim of using Bayesian priors on parameters is then to include as much prior information as we can without biasing our modelling procedure unfairly. One of the virtues of using an MCMC algorithm as described in chapter 3 is that it is straightforward to incorporate priors into it.

In this chapter, I discuss how priors affect the modelling process, starting from the analytical calculations performed by Cassan et al. (2009) in order to derive analytical forms for the priors on his non-standard parameters from known information on standard parameters.

5.2 Analytical priors

The modelling of OGLE-2007-BLG-472, described in Chapter 4 was performed without taking any prior into account. Equivalently, this means that priors were assumed to be uniform, or non-informative for all parameters. In this section, we recall the main steps involved in deriving priors on the Cassan parameters, following Cassan et al. (2009). Indeed, Kains et al. (2009) assumed uniform priors on both the standard and non-standard parameters. This has obvious problems: for example, since the curvature of the caustic folds which make up a caustic structure is concave (see e.g. Fig. 2.4), a source cannot enter and then exit a caustic along the same caustic fold. This needs to be reflected in suitable priors on the corresponding parameters, in this example, on the parameter s_{in} and s_{out} , which define the locus of the caustic crossings along the caustic folds.

Cassan et al. (2009) derived such priors, starting from the parameterisation of caustics derived by Witt (1990):

$$\frac{1}{1+q} \left[\frac{1}{z^2} + \frac{q}{(z+d)^2} \right] = e^{-i\phi}, \quad (5.1)$$

where z is a position on the lens plane, and ϕ is a parameter on the range $[0, 2\pi]$. For each value of ϕ , this equation can then be solved to find z , and the locations of the *critical curves*. The lens equation is then used to find the corresponding points of the *caustics* ζ in the source plane. After some algebra, one can derive a correspondance matrix between the sets of standard and non-standard parameters, i.e. the Jacobian

$$J = \left| \frac{\partial (u_0, \alpha, t_E, t_0)}{\partial (s_{\text{in}}, s_{\text{out}}, t_{\text{in}}, t_{\text{out}})} \right|. \quad (5.2)$$

Because the standard parameters u_0 and α depend only on s_{in} and s_{out} and not on t_{in} and t_{out} , whereas t_0 and t_E each depend on all four non-standard parameters, using the parameter ϕ instead of s , the Jacobian can be expressed as

$$J = \left| \frac{\partial (u_0, \alpha)}{\partial (\phi_{\text{in}}, \phi_{\text{out}})} \right| \times \left| \frac{\partial (t_E, t_0)}{\partial (t_{\text{in}}, t_{\text{out}})} \right|, \quad (5.3)$$

where ϕ_{in} and ϕ_{out} are the values of the parameter ϕ at the caustic entry and exit respectively. After more algebra, one can work out the value of each component of the cross-product in Eq. (5.3) to express it in terms of source plane positions ζ as

$$J = \frac{\left| (\zeta_{\text{out}} - \zeta_{\text{in}}) \wedge \frac{d\zeta_{\text{in}}}{d\phi_{\text{in}}} \right| \left| (\zeta_{\text{out}} - \zeta_{\text{in}}) \wedge \frac{d\zeta_{\text{out}}}{d\phi_{\text{out}}} \right|}{|\zeta_{\text{out}} - \zeta_{\text{in}}|^4}, \quad (5.4)$$

where the “wedge” product \wedge for two complex numbers w and z is defined as

$$w \wedge z = \Re(w)\Im(z) - \Im(w)\Re(z). \quad (5.5)$$

Finally, as mentioned earlier, one must take into account the fact that an entry/exit pair cannot occur on the same caustic fold, because of the concavity of caustic structures. To include this, one must then check that for a given set of parameters, the trajectory of the source at entry is indeed inward and similarly, at exit, the trajectory is pointing outward from

the caustic. To verify that this condition is satisfied, Cassan et al. (2009) compute the dot product of the source trajectory vector, N_t , with the outward vector normal to the caustic fold at the entry/exit point, N_c . If this product is negative at entry, i.e. if $N_{c,in} \cdot N_{t,in} < 0$, then the entry is indeed an entry, that is, the source is indeed moving into the caustic. Similarly, for an exit to be verified, the condition is $N_{c,out} \cdot N_{t,out} > 0$. We then have

$$P(s_{in}, s_{out}) = \begin{cases} J & \text{if } N_{c,in} \cdot N_{t,in} < 0 \text{ and } N_{c,out} \cdot N_{t,out} > 0 \\ 0 & \text{otherwise .} \end{cases}$$

These are therefore analytical priors on s_{in} and s_{out} , assuming uninformative priors on all standard parameters. Examples of such prior maps are plotted on Figs. 5.1 - 5.3. In the next section, we discuss the effect on $P(s_{in}, s_{out})$ of adding priors on standard parameters.

The structure visible is explained as such: The caustics are evenly sampled in the s parameter, meaning that the cusps are located at varying intervals along the caustic arc length parameter s . Because the caustic lines are concave, a source cannot enter the caustic between two cusps and exit the caustic between the same two cusps, meaning that the prior distribution has to be equal to 0 for these regions of the (s_{in}, s_{out}) plane. The high probability of trajectories seen near the cusps is due to the fact that a trajectory entering a caustic near a cusp is also likely to exit it near the cusp. Since the length of such trajectories inside caustics are very short, the timescale of the event must be very long in order to reproduce the correct time observed between caustic crossings. This is the origin of long-timescale models discussed in Chapter 4, and, as mentioned in that discussion, these features need to be addressed by useful informative priors.

5.3 Additional prior information

The priors shown on Figs. 5.1-5.3 can be modified to include information on parameters other than s_{in} and s_{out} . In practice, we can obtain substantial additional information by looking at Galactic models, including Galactic dynamics, and at distribution of observed parameters. When using the latter, one must also be careful not to construct a prior that would reflect an observational bias or selection effect, as this could exclude models on the grounds that the type of events to which they correspond are not being observed. Cassan et al. (2009) used two different priors on the parameter t_E to illustrate their effect on $P(s_{in}, s_{out})$. One was a

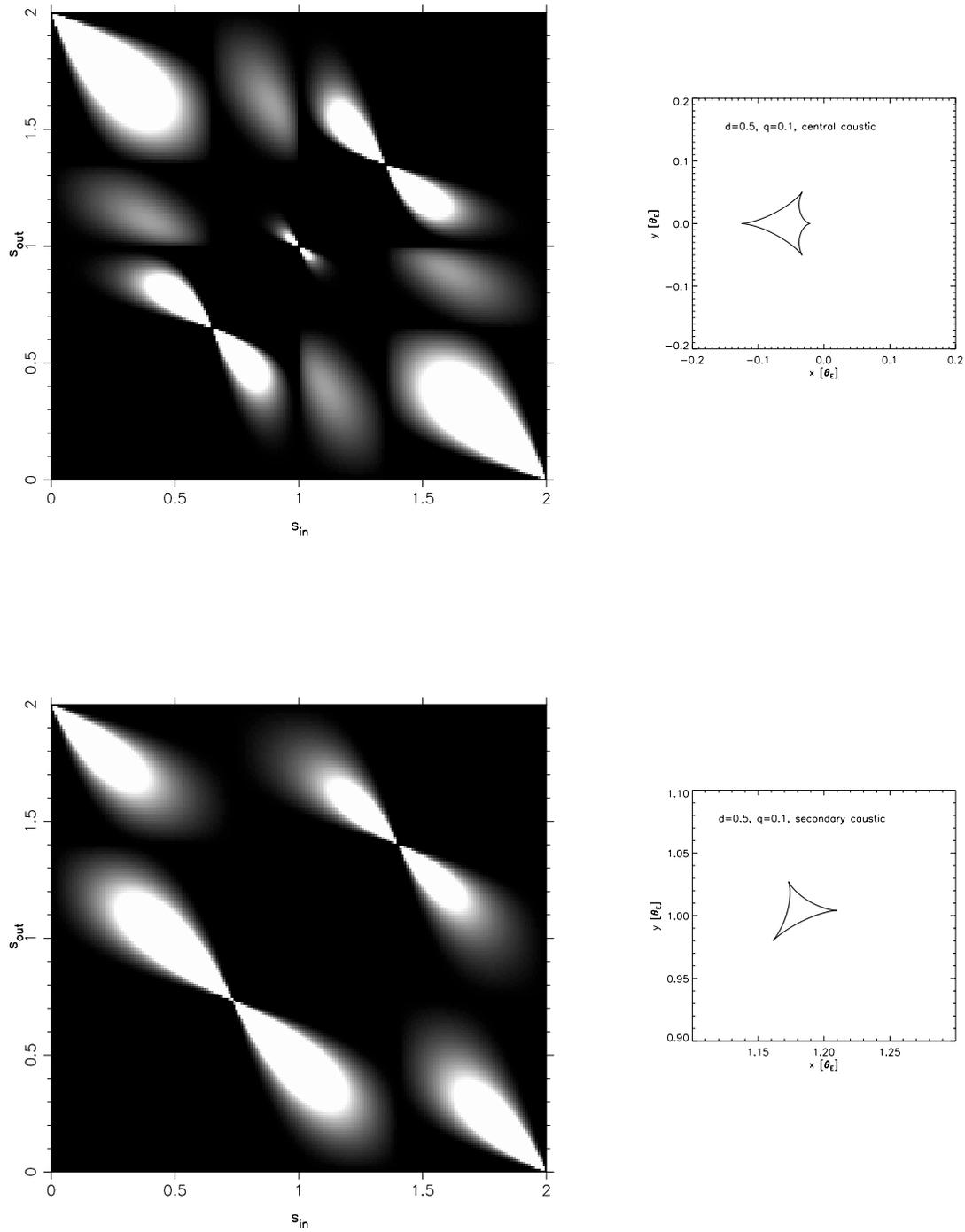


Figure 5.1: Prior maps $P(s_{\text{in}}, s_{\text{out}})$ and caustics for a close configuration (here $d = 0.5, q = 0.1$), with an uninformative prior on all other parameters. *Top:* Prior map for the central caustic. *Bottom:* Prior map for the (upper) secondary caustic (upper or lower secondary caustics have identical prior maps). Whiter shades represent a higher prior probability density.

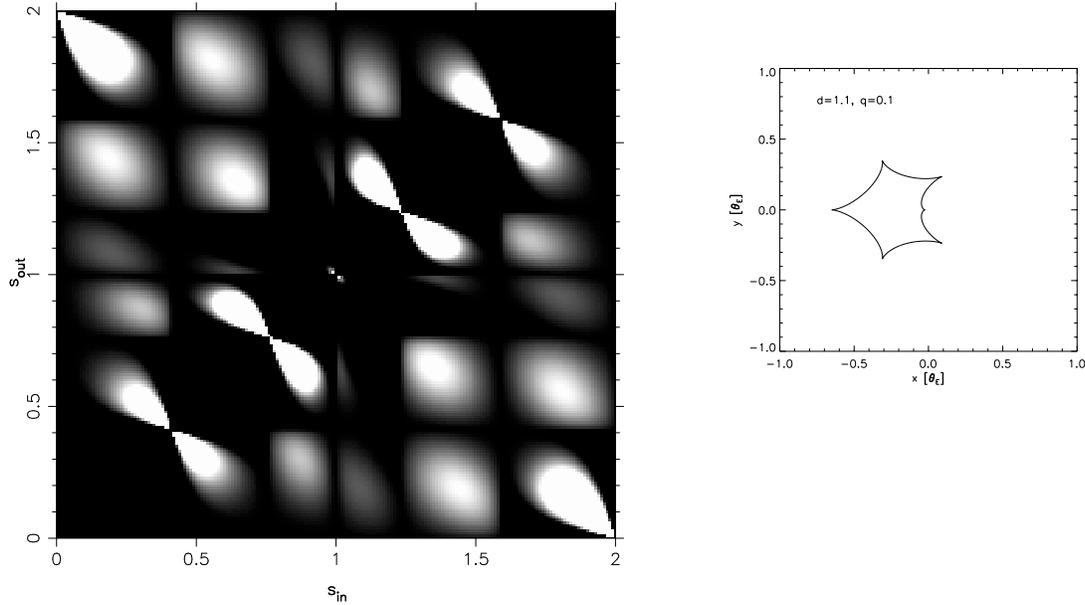


Figure 5.2: Same as Fig. 5.1 for an intermediate caustics configuration ($d = 1.1, q = 0.1$).

distribution of event timescales observed in past microlensing seasons, and another was the model distribution of Wood & Mao (2005). These two distributions are in fact in excellent agreement with each other, as shown on Fig. 5.4.

The effect of adding this prior on t_E is clear when comparing Figs. 5.1 - 5.3 to Figs. 5.5-5.7: since the events with very large values of t_E are suppressed by the priors, so are the corresponding regions of the $s_{\text{in}}, s_{\text{out}}$ plane which correspond to these long timescales. That is, regions of $s_{\text{in}}, s_{\text{out}}$ where the source enters and exit the caustic structure near cusps are suppressed by this prior.

An interesting feature to note in the plot of $P(s_{\text{in}}, s_{\text{out}})$ is that the prior distributions maps are divided into sub-boxes, only a few of which contain acceptable models. This leads us to propose the modelling strategy outlines in the next section.

5.4 An exhaustive modelling strategy

In Chapter 4, I outlined a method to fit caustic-crossing events systematically. However, because no priors were used on any of the parameters, and we used standard parameters to optimise the final best-fit models for each (d, q) grid points, this was essentially equivalent

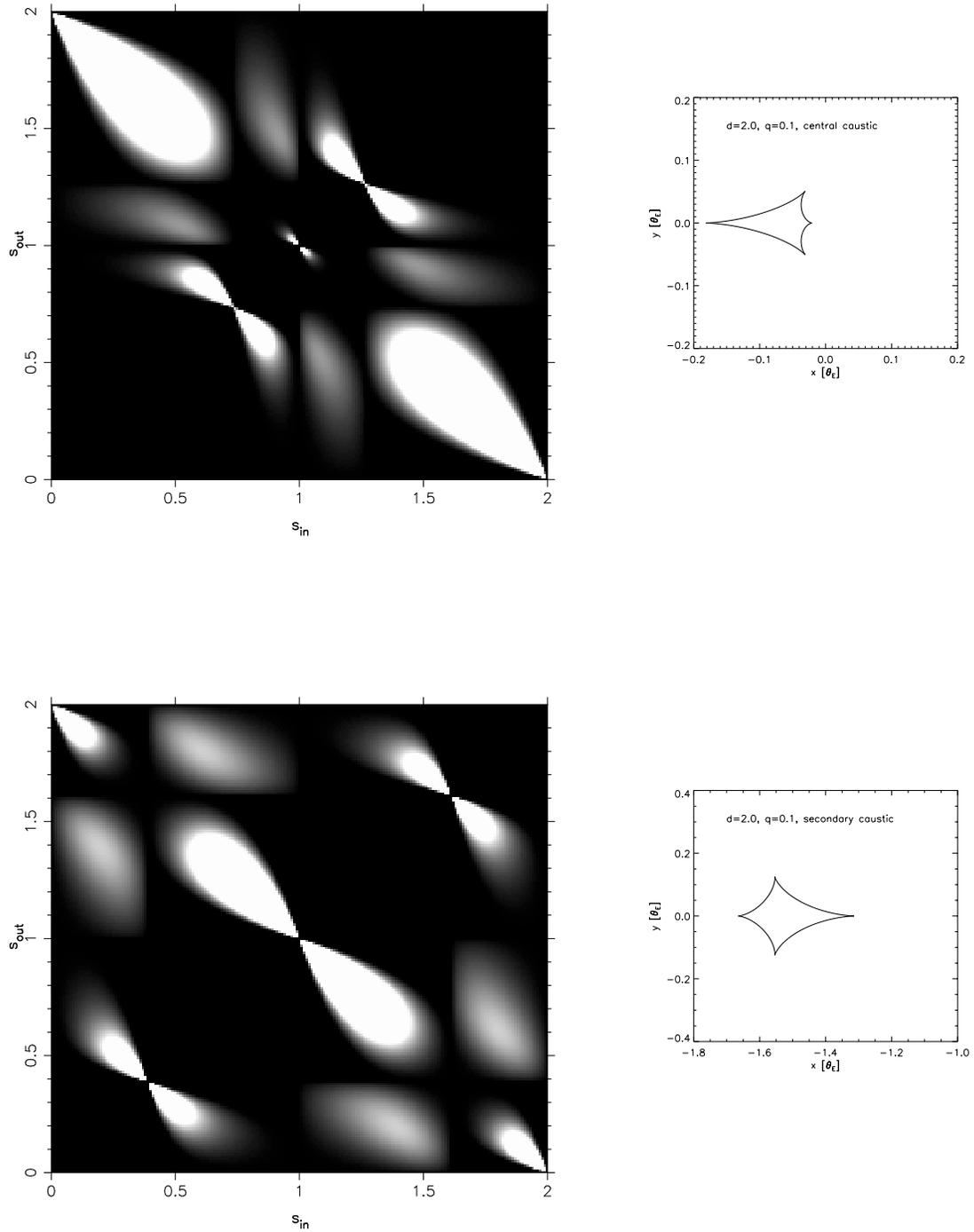


Figure 5.3: Same as Fig. 5.1 for a wide caustics configuration ($d = 2.0, q = 0.1$). *Top:* Priors for the central caustic. *Bottom:* Prior for the upper secondary caustic. The prior map for the lower secondary caustic is identical.

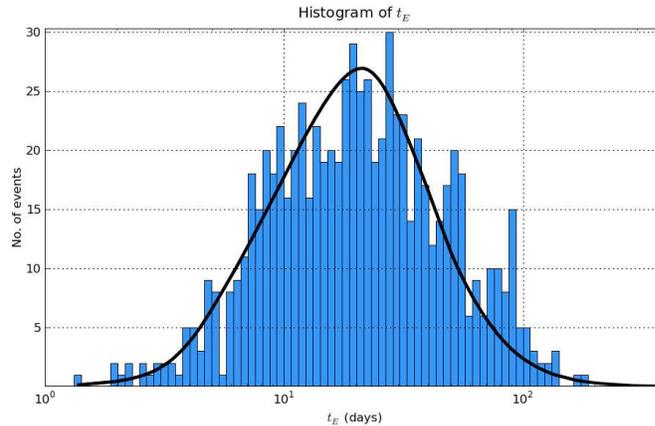


Figure 5.4: A histogram of microlensing event timescales for the OGLE 2006 and 2007 observing seasons, with the model distribution of Wood & Mao (1995) shown as a solid black line. Taken from Cassan et al. (2009).

to starting our modelling in one of the sub-boxes that can be seen in the prior maps (Figs 5.5-5.7). Therefore, it is possible that the global minimum might not have been found for each (d, q) pair in that search, since we did not carry out a thorough systematic search of all sub-boxes. In order to be completely exhaustive, I therefore suggest adopting the following modelling strategy.

We start by building a grid in (d, q) , regularly sampled in $\log(d)$ and $\log(q)$. For each of these grid points, we must then consider separately the cases where the source crosses a central caustic and the cases where it crosses a secondary caustic, and, in the case of close caustics configurations, we must also consider separately the cases of the source crossing an upper and lower secondary caustic. For each case, the prior map $P(s_{\text{in}}, s_{\text{out}})$ is then computed, and divided into sub-boxes. The limiting values for each sub-box simply correspond to the locations of the cusps on the caustic structures. For each sub-box in which $P(s_{\text{in}}, s_{\text{out}})$ is non-zero at least somewhere within the box, we then compute the maximum of $P(s_{\text{in}}, s_{\text{out}})$. We use the location of the maximum as the starting point of our MCMC algorithm, and optimise parameters within the sub-box (using its boundaries as constraints on the parameter values); we then repeat this for each of the prior map's sub-boxes. When all sub-boxes have been explored, we keep the best overall model(s). Finally, we repeat this for each central and secondary caustic configuration(s) of the (d, q) pair.

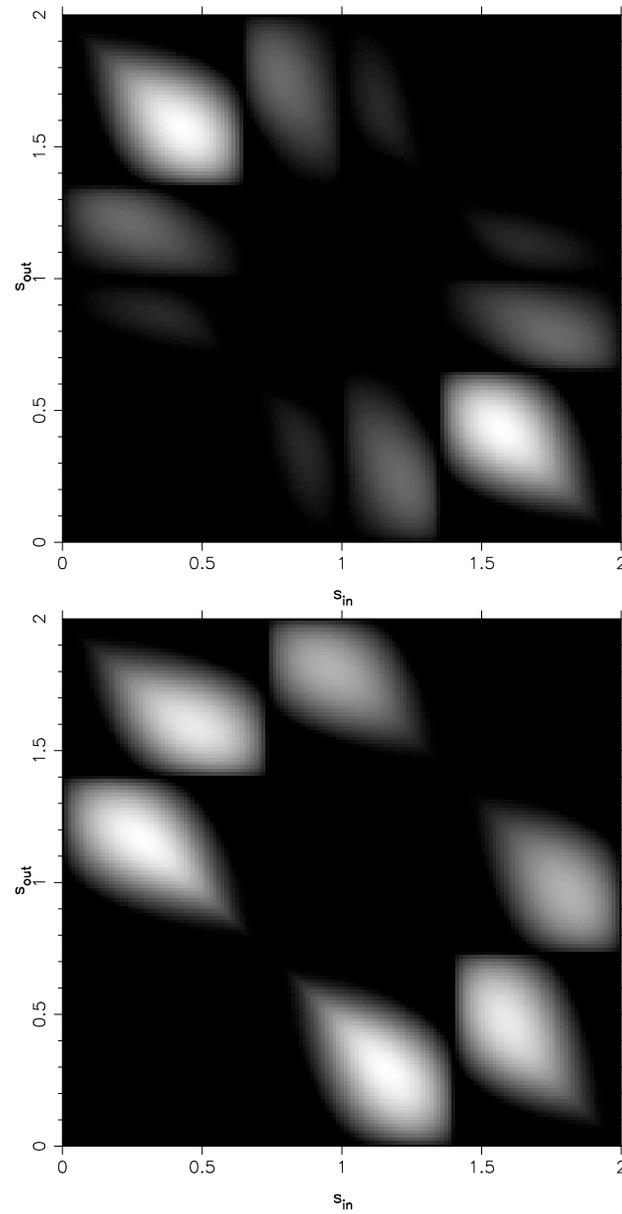


Figure 5.5: Maps of the prior $P(s_{in}, s_{out})$, for close caustics (here $d = 0.5, q = 0.1$) with the prior on t_E advocated by Wood & Mao (2005). *Top:* Priors for the central caustic. *Bottom:* Prior for the secondary caustic (upper or lower secondary caustics have identical prior maps). Whiter shades represent a higher prior probability density.

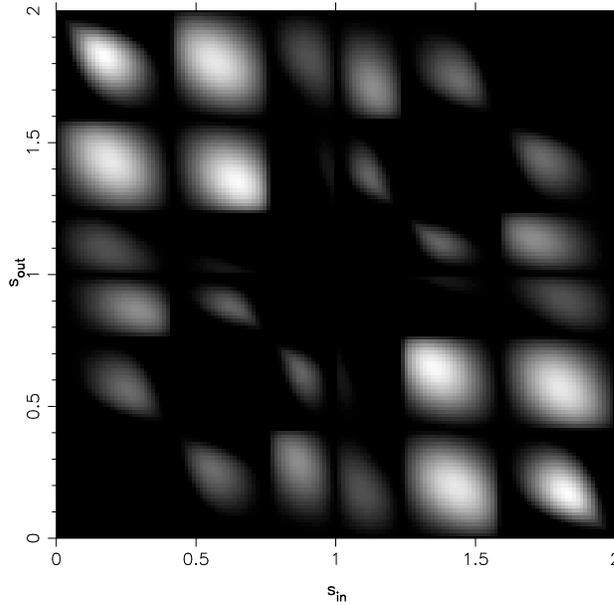


Figure 5.6: Same as Fig. 5.6 for an intermediate caustic configuration ($d = 1.1, q = 0.1$).

A flowchart of this method is shown on Fig. 5.8. Following this procedure over the whole (d, q) grid should ensure that no model can be missed by the fitting algorithm, and therefore that we can truly identify all possible models for a given microlensing event. But is such a procedure realistic, given the high demands of microlensing modelling on computational resources?

Let us go back to the grid we used to analyse the event OGLE-2007-BLG-472 (see e.g. Fig. 4.4). For this event, we used a 14×12 grid, evenly spaced in $\log(d)$ and $\log(q)$, giving a total of 168 grid points. Of these, 66 correspond to close caustic configurations, 46 to intermediate caustics, and 56 to wide caustics. Using the maps plotted on Figs 5.5-5.7 as representative of each caustic regime, we see that the regimes have non-zero prior sub-boxes distributed as follows:

- Close caustics: 10 for central caustics and 6 for each of the secondary caustics
- Intermediate caustics: 28
- Wide caustics: 10 for central caustics and 12 for secondary caustics.

This means that the total number of MCMC runs needed to explore each sub-box of each caustic configuration of each (d, q) grid point *once* is equal to $66 \times (10 + 12) + 46 \times 28 + 56 \times (10 + 12) = 3972$ chains. If on average an optimisation run takes 6 hours of computing time (a conservative estimate when finite source effects are being taken into account), this equates

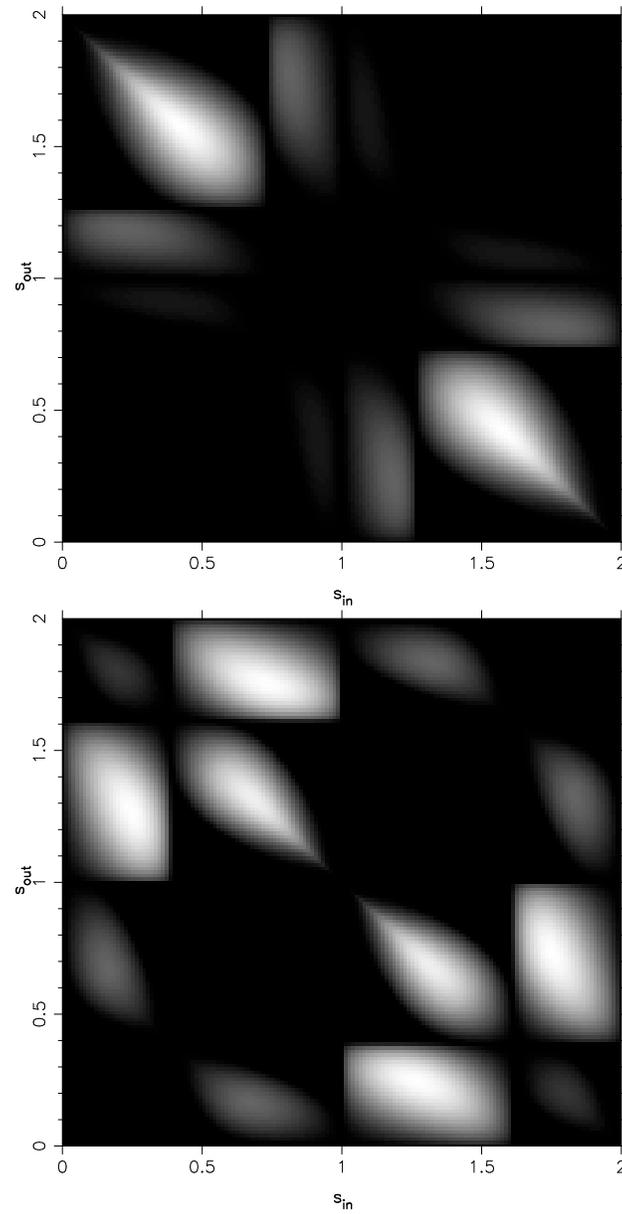


Figure 5.7: Same as Fig. 5.5 for a wide caustics configuration ($d = 2.0, q = 0.1$). *Top:* Priors for the central caustic. *Bottom:* Prior for the secondary caustic.

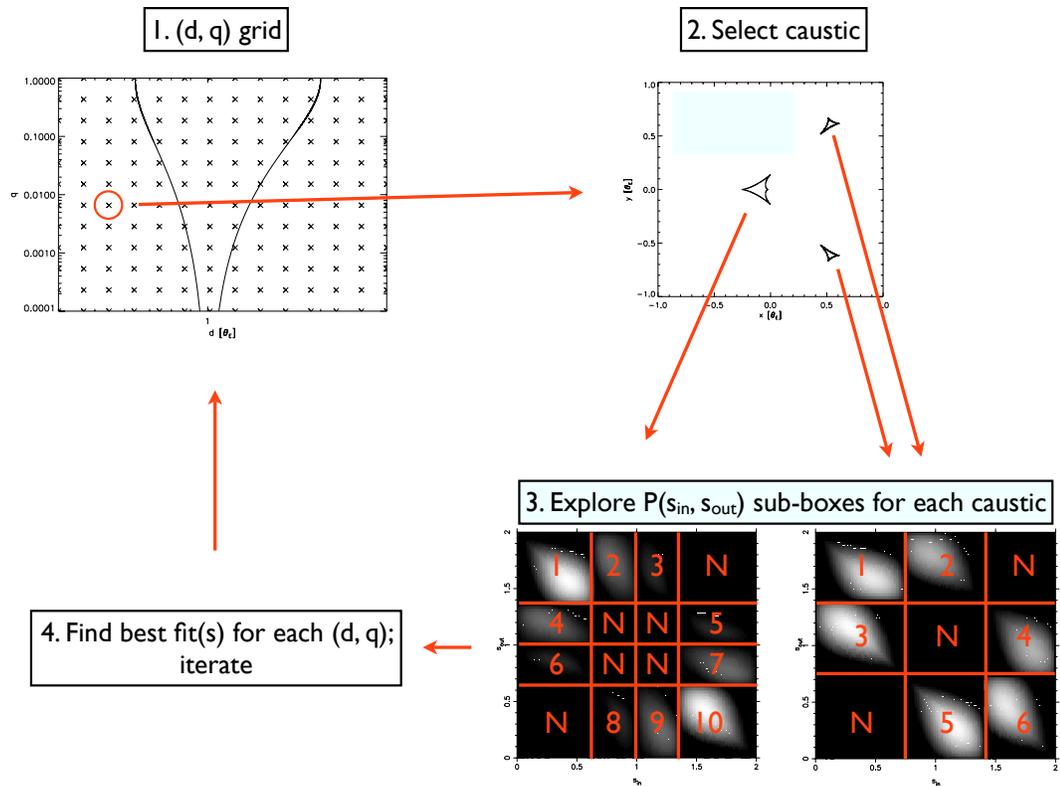


Figure 5.8: Flow chart of our proposed automatic fitting algorithm for caustic-crossing binary-lens microlensing events. The example here is the flow for a close caustics case. Each caustic is explored separately, and each sub-box of the $P(s_{in}, s_{out})$ map is labeled with a number (if the sub-box has regions of non-zero prior), or N (if the prior is null everywhere in the sub-box). In this case, the central caustic prior map has 10 non-zero prior sub-boxes, while the secondary caustic prior maps have 6.

to roughly 1000 days of total computing time. This is obviously a problem, since this is for analysing a single event. When dealing with simultaneous anomalous events, we are faced with an obvious issue with computational resources. Therefore, when fitting events in real time, we must find a balance between being thorough and efficient, so that when necessary, we can extract useful information rapidly in order to provide feedback to observing telescopes.

The next step in developing efficient real-time modelling will need to address these issues. Some possible solutions that could be explored would be to avoid fitting entire datasets when a few additional observations have been made on an event, but instead to be able to build on previous models to speed up calculations, or using pre-calculated magnification maps, although this rapidly becomes impractical when dealing with such a complex parameter space.

6

Modelling the evolution of debris discs around solar-type stars

This chapter is based on Kains, Wyatt & Greaves (2010, submitted for publication).

In the paper on which this chapter is based, I contributed to the data reduction (with Jane Greaves), modelling and discussion of the results (with Mark Wyatt and Jane Greaves).

6.1 Introduction

Since the first detection of infrared excess attributable to the presence of a circumstellar disc, around the star Vega, by Aumann et al. (1984), hundreds more discs have been found around main-sequence stars in our Galaxy. These discs sometimes bear signs of past or ongoing formation of planetary systems around their host stars. Understanding their evolution is therefore crucial to improving our understanding of planet formation, and placing our own Solar System in perspective. Indeed the Solar System has its own debris disc, of which two components

can be distinguished: the asteroid belt, located between Mars and Jupiter, and the Kuiper Belt, located beyond the orbit of Neptune and extending for around 15 AU. The observation and modelling of Solar System analogues, for which we can detect both a disc and a planetary system, therefore provide us with important insights into the history and evolution of our own debris disc.

The rapid development of observing methods and improvements in instruments over the last 15 years has seen more and more discs discovered with increasingly accurate measurements. Usually this discovery comes through measuring the spectral energy distribution (SED) of a star and searching for the signs of infrared excess emission caused by circumstellar dust; some discs have now also been imaged directly. Today, instruments such as the *Spitzer Space Telescope* allow us to make high-accuracy observations of debris discs, both in terms of signal-to-noise and spatial resolution, and to use these to test current models of disc evolution and planet formation.

Most evolutionary models of debris discs present a simplified picture of debris discs. Dust grains are usually assumed to be perfect blackbodies, which is a good first-order approximation (Hillenbrand et al., 2008). Investigating how debris discs evolve essentially means investigating how their luminosity and temperature change with time, and therefore in addition to accurate flux measurements, good determination of the ages of disc host stars is essential to using observations to constrain evolutionary models.

In this chapter, I present an analysis of debris disc data around Solar-type stars (spectral types F0-K5) using the steady-state analytical model of Wyatt et al. (2007a). Models are fitted to data from the FEPS (*Formation and Evolution of Planetary Systems*) and SIMTPF projects taken with MIPS and IRAC on the *Spitzer* telescope at $24\mu\text{m}$ and $70\mu\text{m}$, and compared to a previously published analysis of debris discs around A stars using the same evolutionary model. We find that the model reproduces most features found in the data sets, but suggest that debris discs around solar-type stars have different effective properties from their counterparts around earlier-type stars. However, we also hint that despite these differences in effective values, intrinsic properties might not be significantly different. We also identify several objects which are not fitted well by our models, including some which we classify as being transient (i.e. not in steady state), including the first transient classification of HD 101259.

6.2 Debris discs around solar-type stars

Recent surveys looking for excess emission around solar-type (F, G, K spectral types, hereafter FGK) stars have found that over 15% of these objects are surrounded by debris discs (Bryden et al., 2006; Beichman et al., 2006a; Moro-Martín et al., 2007; Trilling et al., 2008; Hillenbrand et al., 2008; Greaves et al., 2009). Although these discs are believed to be analogues of the Solar System’s asteroid and Kuiper belts, the infrared luminosity of most detected discs around solar-type stars is usually around 100 times higher than that of the Solar System’s debris structures (Greaves & Wyatt, 2010). This suggests that the Solar System may have unusually sparse debris, or that the detected discs around other stars are anomalously luminous, perhaps as a result of recent collisions in their planetesimal belts, producing transient excess dust. This may be caused by a process similar to the Late Heavy Bombardment (LHB) period which occurred ~ 700 Myr after the formation of the Sun (Booth et al., 2009).

Among the sample of FGK stars with detected discs, several are known to harbour extrasolar planets. Interestingly, searches for asteroid and Kuiper belts analogues around some of the known extrasolar planetary systems orbiting solar-type stars, such as τ Boo, π Men (HD 39091), ν And, 55 Cnc and 51 Peg did not find detectable levels of dust (Bryden et al., 2009; Beichman et al., 2006b), which could indicate that dust luminosity has fallen to near the luminosity of the Kuiper and asteroid belts of the Solar System and that these systems are past their own LHB equivalent. Booth et al. (2009), however, suggest that such events are rare, and an alternative explanation could be that these systems might also have been born with lower-mass planetesimal belts. Studying extrasolar systems therefore offers an important perspective for understanding the history of the Solar System. In particular, it is essential to gain an understanding of how planetesimal belts evolve around their host stars, and which mechanisms affect their structure and their luminosity in order to be able to compare observed extrasolar systems to the Kuiper belt.

Wyatt et al. (2007a) described a simple model of steady-state debris disc evolution, based on an earlier model by Dominik & Decin (2003) by considering the collisional grinding down of planetesimal belts. The basic features of this model are recalled in Sec. 6.3. In Sec. 6.4, we outline our modelling procedure. In Sec. 6.5, we briefly describe the 24 and $70\mu\text{m}$ data we use, taken from the studies of Trilling et al. (2008), Hillenbrand et al. (2008) and Beichman et al. (2006a). In Sec. 6.6, we fit this data with our model and discuss our results in Sec. 6.7,

comparing them to the results of the analogous study that was carried out for A stars by Wyatt et al. (2007b). We also discuss the possible reasons for differences between the evolution of debris discs around FGK stars and that of their earlier-type counterparts, and consider individual systems that are not well fitted by our models. We conclude by looking at the implications of our models for the properties of discs around FGK stars compared to those around A stars.

6.3 Debris disc model

In this section we recall the main features of the analytical model derived by Wyatt et al. (2007a) and revised by Wyatt et al. (2007b), as well as the assumptions made in applying this model to populations of debris discs. We start with a planetesimal belt characterised by a size distribution

$$n(D) \propto D^{-3.5}, \quad (6.1)$$

where D is the diameter of the planetesimals; this is the distribution expected for a planetesimal belt in collisional equilibrium (Dohnanyi, 1969). This distribution is assumed to be valid from the largest planetesimals, of diameter D_c down to the *blowout* diameter D_{bl} , below which particles are blown away by radiation pressure.

For a planetesimal belt at radius r , with a width dr , the fractional luminosity of the dust emission, $f = L_{\text{IR}}/L_*$ can be expressed in terms of the total cross-sectional area σ_{tot} as

$$f = \frac{\sigma_{\text{tot}}}{4\pi r^2}. \quad (6.2)$$

Given the size distribution in Eq. (6.1), σ_{tot} is proportional to M_{tot} through a constant that depends on D_{bl} and D_c (Wyatt et al., 2007a).

Assuming the dust particles act as blackbody emitters, the blowout diameter is given by

$$D_{\text{bl}} = 0.8 \frac{L_*}{M_*} \frac{2700}{\rho}, \quad (6.3)$$

where D_{bl} is in μm , L_* and M_* are in solar units, and ρ is the density of the dust particles in

kg m^{-3} . The temperature T of the dust at radius r is given by

$$T = 278.3 \left(\frac{L_*}{L_\odot} \right)^{0.25} \left(\frac{r}{\text{AU}} \right)^{-0.5}. \quad (6.4)$$

For blackbody emitters, we also know that emission from the disc at wavelength λ is

$$F_{\nu, \text{disc}} = 2.35 \times 10^{-11} B_\nu(\lambda, T) \sigma_{\text{tot}} d^{-2}, \quad (6.5)$$

where d is the distance to the star in pc and F_ν is in Jy if the Planck function B_ν is in Jy sr^{-1} .

Considering a collisional cascade in which the population within a given size range is being destroyed in collisions with other members of the cascade, but is being replenished by fragmentation of larger objects, the collision of the larger objects in the cascade solely determine the long-term evolution of the population. Hence the long-term evolution depends on the collisional lifetime t_c of the planetesimals in the disc with size D_c . For the eccentricity and inclination of the planetesimals' orbits having $e = I$ (which we will assume throughout this chapter¹), the expression for t_c can be reduced to

$$t_c = \frac{3.8 \rho r^{3.5} (dr/r) D_c}{M_*^{0.5} M_{\text{tot}}} \left\{ \frac{8}{9 G(X_c)} \right\}, \quad (6.6)$$

where t_c is in Myr, D_c is in km, M_{tot} is in units of M_\oplus and the factor $G(X_c)$ is defined in Wyatt et al. (2007a). X_c is defined as $X_c = D_{cc}/D_c$, where D_{cc} is the diameter of the smallest planetesimal that has sufficient energy to destroy a planetesimal of size D_c . This factor can be calculated from the value of the dispersal threshold Q_D^* , which is defined as the specific incident energy required to catastrophically destroy a particle. It follows (Wyatt & Dent, 2002) that with Q_D^* given in J kg^{-1} ,

$$X_c = 1.3 \times 10^{-3} \left(\frac{Q_D^* r M_*^{-1}}{2.25 e^2} \right)^{1/3}. \quad (6.7)$$

Ignoring non-collisional processes, the time-dependence of the disc mass can be worked out by solving the differential equation $dM_{\text{tot}}/dt = -M_{\text{tot}}/t_c$, yielding

¹This assumption stems from a desire to simplify the full analytical expressions of Wyatt et al. (2007a); $e = I$ was also assumed by Wyatt et al. (2007b).

$$M_{\text{tot}}(t) = M_{\text{tot}}(0)/[1 + t/t_c(0)]. \quad (6.8)$$

As noted by Wyatt et al. (2007a), the mass of the disc at $t \gg t_c$ does not depend on the initial disc mass, since $t_c(0)$ depends on $M_{\text{tot}}(0)$ (see Eq. (6.6)). As a result, at a given age, there is a maximum mass that can remain after collisional evolution, and therefore a maximum infrared luminosity f_{max} , given by (Wyatt et al., 2007a)

$$f_{\text{max}} = \left[\frac{10^{-6} r^{1.5} (dr/r)}{4\pi M_*^{0.5} t_{\text{age}}} \right] \left\{ \frac{4}{3 G(X_c)} \right\} \left(\frac{D_{\text{bl}}}{D_c} \right)^{-0.5}, \quad (6.9)$$

for r in AU, D_{bl} in μm and t_{age} in Myr.

For observations limited by a calibration limit expressed as a flux ratio $R_{\text{det}}(\lambda) = F_{\nu, \text{disc}}/F_{\nu, \text{phot}}$, the corresponding detection limit in terms of fractional dust luminosity, f_{det} , is given by

$$f_{\text{det}} = 6 \times 10^9 R_{\text{det}} r^{-2} L_* T_*^{-4} B_\nu(\lambda, T_*)/B_\nu(\lambda, 278.3 L_*^{0.25} r^{-0.5}). \quad (6.10)$$

This is the threshold value we use to determine which model radii are detectable at given wavelengths in the rest of this chapter. Finally, the limit at which Poynting-Robertson (PR) drag becomes important, i.e. when radiation and gravitational forces become similar, is given (Wyatt et al., 2007a) by

$$f_{\text{PR}} = 50 \times 10^{-6} (dr/r) \sqrt{M_*/r}, \quad (6.11)$$

for r in AU. That is, if a disc has an infrared flux $f < f_{\text{PR}}$, then it is likely that it will be affected by PR drag, which could mean that in such a disc the dust component of the disc becomes spatially separated from the planetesimal belt.

6.4 Modelling procedure

As done by Wyatt et al. (2007b) for A stars, we apply this model to fit the flux evolution of debris disc populations, assuming that all stars have a planetesimal belt which is undergoing collisional processes as described by the above equations. This means that the initial state of the belts is completely determined by the parameters $r, dr, \rho, D_{\text{bl}}, M_{\text{tot}}(0), L_*$ and M_* (of

which the latter two determine D_{bl} through Eq.6.3), while their evolution also depends on Q_D^* , e , D_c and I . We also set $\rho = 2700 \text{ kg m}^{-3}$, $dr = r/2$ and $e = I = 0.05$. Values of Q_D^* and D_c were then constrained by fits to observational data of debris discs.

To simulate a population of stars with debris discs, values of $M_{\text{tot}}(0)$, r , L_* and M_* were drawn from model distributions of properties of those stars and debris discs. We use a power law distribution for disc radii, $N(r) \propto r^\gamma$ with γ and the limits of the radius distribution, r_{min} and r_{max} , being treated as free parameters. The distribution of discs at both 24 and $70\mu\text{m}$ is affected by strong observational bias, especially for solar-type stars, for which discs with radii larger than ~ 50 AU can only be detected at $24\mu\text{m}$ if they have a large fractional excess (e.g. Fig. 6.6), meaning that there is an observational bias towards the detection of discs with radii lower than ~ 50 AU, and this must be accounted for by a satisfactory model. We constrain γ by comparing the fit of the observed radius distribution to a fit to the subsample of the model population which could be detected at 24 and $70\mu\text{m}$. This comparison also yielded constraints on r_{min} and r_{max} , although these parameters are also sensitive to the fraction of discs detected in different age bins.

For the distribution of initial disc masses $M_{\text{tot}}(0)$, we use the results of Andrews & Williams (2005), who derived a lognormal distribution of dust masses centred on a value M_{mid} , with a standard deviation of 0.8 dex, from submillimetre observations of protoplanetary discs in Taurus-Auriga, i.e. for \sim solar-mass stars. We use their value for the distribution's dispersion but fit M_{mid} as a free parameter. We choose this approach because their submillimetre data do not detect all sizes of dust grains.

A spectral type was drawn at random on the range of spectral type F0 - K5, chosen to correspond to the range observed by the various *Spitzer* programmes (Trilling et al., 2008; Hillenbrand et al., 2008; Beichman et al., 2006a). This allowed us to determine a value of D_{bl} for each spectral type, using corresponding values of L_* and M_* and Eq. (6.3). We also assigned values to other parameters using the relevant distributions, determining the initial conditions of each disc. It should be noted that this approach assumes that the mass and radius of the disc are independent both of each other and of the properties of the star around which they are located.

The systems were then evolved using Equations 6.6 and 6.8, and values were drawn at random for the age and distance (for a constant volume) of each system between 0 and 10

Gyr, and between 0 and 100 pc respectively (both chosen to cover the same ranges as the observational data used in this analysis), in order to determine its current "observed" properties. Although we assign a distance to each system, its value has no impact on the flux statistics on which the analysis presented in this chapter is based. This left Q_D^* , D_c , γ , M_{mid} , r_{min} , r_{max} as free parameters of the model. Finally, another constraint was that a good-fit model (based on its χ^2 statistic) was only retained if the parameter values were realistic, i.e. consistent with ranges of values predicted by models of catastrophic collisions (Benz & Asphaug, 1999) and planet formation (Kenyon & Bromley, 2002).

6.5 Data

A list of all stars with $70\mu\text{m}$ excess emission only in the sample we use in this chapter is given in Table 6.1 (p.96), while a list of sources that show excess at both 24 and $70\mu\text{m}$ is given in Table 6.2 (p.97). Stellar properties (L_* , M_* , d , t_{age}) for these are published values resulting from Kurucz model fits, or in cases for which these were not available, were computed using Schmidt-Kaler relations (Aller et al., 1982). Using the 24 – $70\mu\text{m}$ colour temperature of the dust, and assuming that the dust acts as a blackbody, we also calculated a disc radius for the discs in Table 6.2 with Eq. (6.4). We do not include the debris discs published recently by Koerner et al. (2010) because they do not report stellar ages or non-detections in their data. This sample would only add one disc detected at both 24 and $70\mu\text{m}$, and therefore would not influence our results significantly.

We made a cut in stellar age, excluding objects with an age below 30 Myr in order to avoid protoplanetary discs affecting our statistics. We also excluded discs for which $1-\sigma$ error bars were above a threshold of $\sigma = 5$ mJy. This threshold was chosen empirically to avoid having noisy flux upper limits being wrongly counted as large excesses. These cuts affected mostly the FEPS data (Hillenbrand et al., 2008), with 6 discs detected at both wavelengths being removed from their published sample. The reason for this is the difference in the observational approaches of the SIMTPF and FEPS observations: the observations of Trilling et al. (2008) and Beichman et al. (2006a) at $70\mu\text{m}$ are calibration-limited while those of Hillenbrand et al. (2008) are sensitivity-limited; at $24\mu\text{m}$, all observations are calibration-limited (see Table 6.2). The consequence of this is that most of the FEPS discs are very bright discs, since these are the only ones for which good photometry could be obtained.

The resulting sample consists of 46 discs detected at $70\mu\text{m}$ only and 17 detected at both 24 and $70\mu\text{m}$. Below is a short description of each subsample we used.

6.5.1 Trilling et al. (2008) sample

Trilling et al. (2008) collected a sample of 193 F, G and K stars observed with the Multiband Imaging Photometer for *Spitzer* (MIPS). Together with data already published by Bryden et al. (2006), the whole sample gives a view of debris discs around stars with masses and ages similar to that of the Sun, covering a range of spectral types between F0 and K5. Their “solar type” sample was selected using criteria on the spectral type (F5-K5) and luminosity class (IV or V), as well as setting a minimum photospheric $70\mu\text{m}$ flux and signal-to-noise (S/N) ratio. From the combined FGK sample, 27 show significant excess at $70\mu\text{m}$, meaning an excess with $\chi_{70} \geq 3$, where χ_{70} is defined as

$$\chi_{70} = \frac{F70 - P70}{\sigma_{70}}, \quad (6.12)$$

where $F70$ is the total $70\mu\text{m}$ flux measured, $P70$ is the predicted photospheric flux at $70\mu\text{m}$, and σ_{70} is the error bar associated with the flux measurement. On top of these 27 objects, 3 more are classified as excess sources for reasons detailed in Trilling et al. (2008), but one is removed due to the stellar age not being determined, bringing the total number of excess objects in the sample to 29. This includes 7 systems for which significant excess flux is detected at both 24 and $70\mu\text{m}$.

6.5.2 Beichman et al. (2006) sample

The sample of Beichman et al. (2006a) includes objects observed in the frame of other projects, including radial velocity search teams, coronagraphy and interferometry missions. As a result, it includes some low-mass close stars. They selected stars within 25 pc (with a few exceptions for the earlier-type stars) and excluded targets with binary companions within 100 AU on the grounds that binarity might prevent the formation or long-term stability of planetary systems. Their selected sample was then observed with MIPS at 24 and $70\mu\text{m}$ and four stars were observed further with the *Infrared Array Camera* (IRAC) in order to help determine their photospheric flux.

Based on the criterion expressed by Eq. (6.12), Beichman et al. (2006a) identify 12 stars

Table 6.1: List of sources with $70\mu\text{m}$ excess only. Flux ratios and limits are given as total flux divided by photospheric flux. $R24$ is defined as $R24 = F24/F24_*$, and similarly, $R70 = F70/F70_*$. $C24$ and $C70$ are the $3 - \sigma$ calibration limits at 24 and $70\mu\text{m}$, defined as $C_\lambda = 1 + 3\sigma_{*,\lambda}/F_{*,\lambda}$.

Star name	Sp. type	d (pc)	t_{age} (Gyr)	$R24$	$C24$ (3σ)	$R70$	$C70$ (3σ)
Trilling et al.							
HD 1581	F9 V	8.6	3.02	0.98	1.10	1.4	1.3
HD 3296	F5	47	2.5	1.02	1.10	4.8	2.8
HD 17925	K1 V	10	0.19	1.05	1.10	3.4	1.7
HD 19994	F8 V	22	3.55	1.00	1.10	1.7	1.5
HD 20807	G1 V	12	7.88	1.04	1.10	1.8	1.5
HD 22484	F8 V	17	8.32	1.02	1.10	1.9	1.3
HD 30495	G1 V	13	1.32	1.06	1.10	5.5	1.6
HD 33262	F7 V	12	3.52	1.03	1.10	1.7	1.4
HD 33636	G0	29	3.24	1.00	1.10	7.0	2.2
HD 50554	F8 V	31	4.68	1.00	1.10	8.4	3.4
HD 52265	G0	28	6.03	0.99	1.10	4.8	2.9
HD 57703	F2	44	2.3	1.03	1.10	9.3	3.3
HD 72905	G1.5 V	14	0.42	1.06	1.10	2.3	1.5
HD 75616	F5	36	4.8	1.03	1.10	10	2.5
HD 76151	G3 V	11	1.84	1.03	1.10	2.4	1.6
HD 82943	G0	27	4.07	1.02	1.10	17	3.1
HD 110897	G0 V	17	9.7	0.98	1.10	4.3	1.9
HD 115617	G5 V	8.5	6.31	1.04	1.10	4.0	1.5
HD 117176	G5 V	18	5.37	0.98	1.10	1.8	1.3
HD 128311	K0	167	0.39	0.94	1.10	3.0	2.3
HD 206860	G0 V	18	5.00	1.03	1.10	2.3	1.5
HD 212695	F5	51	2.3	1.00	1.10	9.5	3.3
Hillenbrand et al.							
HD 6963	G7 V	27	1.00	1.05	1.13	13	8.5
HD 8907	F8	34	0.32	1.05	1.13	46	12
HD 31392	K0 V	26	1.00	1.02	1.12	20	8.5
HD 35850	F7/8 V	27	0.03	1.14	1.14	4.9	3.9
HD 38529	G8 III/IV	42	3.16	0.96	1.12	4.3	3.1
HD 72905	G1.5	14	0.10	0.84	1.10	2.7	2.1
HD 122652	F8	37	3.16	1.08	1.13	24	10
HD 145229	G0	33	1.00	1.09	1.14	20	9.1
HD 150706	G3 V	27	1.00	1.05	1.13	8.7	6.4
HD 187897	G5	33	1.00	1.03	1.12	14	7.5
HD 201219	G5	36	1.00	1.07	1.13	18	11
HD 209253	F6/7 V	30	0.10	1.14	1.14	14	6.8
Beichman et al.							
HD 38858	G4 V	16	4.57	1.00	1.30	10	3.0
HD 48682	G0 V	17	3.31	1.00	1.59	12	2.0
HD 90089	F2 V	21	1.78	1.00	1.59	2.3	1.7
HD 105211	F2 V	20	2.53	1.01	1.19	11	2.4
HD 139664	F5 IV-V	18	0.15	1.10	1.24	5.9	1.6
HD 158633	K0 V	13	4.27	0.90	1.53	18	2.0
HD 219623	F7 V	20	5.06	1.04	1.12	3.0	1.7

Table 6.2: List of sources with 24 and 70 μ m excesses. Flux ratios and limits are given as total flux divided by photospheric flux. R_{24} , R_{70} , C_{24} and C_{70} as defined as in Table 6.1. Stellar properties are either published values found by fitting Kurucz models to the data when these were available, or calculated using Schmidt-Kaler relations for main sequence stars.

Star name	Sp. type	d (pc)	t_{age} (Gyr)	L_* L_{\odot}	M_* M_{\odot}	f (10^{-5})	r (AU)	f/f_{max}	f/f_{PR}	R_{24}	C_{24} (3σ)	R_{70}	C_{70} (3σ)
Trilling et al.													
HD 166	K0 V	14	5.0	0.42	0.79	5.9	9.1	1.23	8.0	1.14	1.10	6.9	1.8
HD 3126	F2	42	3.5	2.9	1.5	13	21.8	1.14	19.7	1.16	1.10	27	3.3
HD 10647	F9 V	17	6.3	1.8	1.1	34	21	3.66	75	1.21	1.10	51	2.1
HD 69830	K0 V	13	4.7	0.42	0.79	20	1.0	509	9.0	1.47	1.10	1.5	1.5
HD 101259	G6/8 V	65	11	0.74	0.89	6.0	0.98	544	2.5	1.25	1.10	1.6	1.7
HD 105912	F5	50	1.8	3.2	1.4	7.9	7.7	3.52	7.4	1.52	1.10	11	3.0
HD 207129	G0 V	16	5.8	1.5	1.1	12	15.3	2.10	18	1.17	1.10	16	2.8
Hillenbrand et al.													
HD 25457	F7 V	19	0.10	2.4	1.3	10	17	0.04	14.8	1.31	1.16	18	5.0
HD 37484	F3 V	60	3.0	1.5	0.10	32	17	0.13	43.5	1.43	1.29	46	14
HD 85301	G5	32	0.79	0.92	1.00	13	8	1.1	14.8	1.36	1.17	13	8.4
HD 202917	G7 V	46	0.79	0.92	0.03	25	9	0.05	31.0	1.63	1.20	28	16
HD 219498	G5	150	0.79	0.92	0.32	20	31	0.03	46.4	1.29	1.15	25	14
Beichman et al.													
HD 25998	F7 V	21	0.6	2.4	1.2	4.5	13	0.17	5.8	1.14	1.12	4.2	2.2
HD 40136	F1 V	15	1.3	4.3	1.6	1.9	5.9	1.4	1.5	1.13	1.12	1.6	1.5
HD 109085	F2 V	18	1.3	2.9	1.5	15	5.9	8.7	11.8	1.99	1.12	5.9	1.6
HD 199260	F7 V	21	3.2	2.4	1.2	3.3	14	0.59	4.4	1.11	1.12	3.5	2.0
HD 219482	F7 V	21	6.1	2.4	1.2	3.6	18	0.64	5.5	1.08	1.12	4.4	1.7

out of the 88 in their sample to have a $70\mu\text{m}$ excess at the 3σ level or better. Amongst these, 5 also display significant excess at $24\mu\text{m}$.

6.5.3 Hillenbrand et al. (2008) sample

Hillenbrand et al. (2008) presented data obtained using MIPS, IRS and IRAC, with most of their systems observed at $3.6, 4.5, 8.0, 13, 24, 33, 70$ and $160\mu\text{m}$, as part of the FEPS program. They report 25 systems with excess flux at $70\mu\text{m}$ and $S/N_{70\mu\text{m}} \geq 3$. Amongst this sample, 13 also have excess emission at $24\mu\text{m}$. Since photospheric sensitivity at $70\mu\text{m}$ could not be achieved without very long integration times except for the closest stars in the sample, the sensitivity of their observations was determined by a target detection threshold of dust emission, which they expressed relative to dust emission in a young Solar System model. For most targets, the survey was sensitive to 5-10 times the dust emission predicted by the model.

The sample of Hillenbrand et al. (2008) is different from those of Trilling et al. (2008) and Beichman et al. (2006a), as the former sample is age-selected, leading to an even distribution in logarithmic age bins, while the latter are volume-limited, leading to a linear age distribution. This difference is taken into account in the analysis that follows, and the data cuts we made were chosen to minimise the effect of poorly constrained photospheric fluxes in the FEPS sample on our modelling of the statistics.

6.6 Best-fit models

6.6.1 Fit to the $70\mu\text{m}$ statistics

Following the procedure described in Sec. 6.4, we found best-fit values of $D_c = 600\text{ km}$ and $Q_D^* = 6000\text{ Jkg}^{-1}$. These are consistent with models of catastrophic collisions (Benz & Asphaug, 1999). We also found best-fit parameters for the radius distribution of $r_{\text{min}} = 1\text{ AU}$, $r_{\text{max}} = 120\text{ AU}$ and $\gamma = -1.5 \pm 0.4$ compared to $\gamma \sim -0.8 \pm 0.3$ for the A stars study (Wyatt et al., 2007b); however the values of γ are within 2σ of each other. Error bars were worked out using uncertainties from blackbody fits to the 24 and $70\mu\text{m}$ data. These were obtained using error bars on measured fluxes reported in the literature. A histogram of the distribution of disc radii for the model and observed data is shown on Fig. 6.1. Our radius distribution extends to lower radii than the one found by Löhne et al. (2008), who used a value of $r_{\text{min}} = 20\text{ AU}$. However they limited their data sample to G stars, and their analysis used radius values

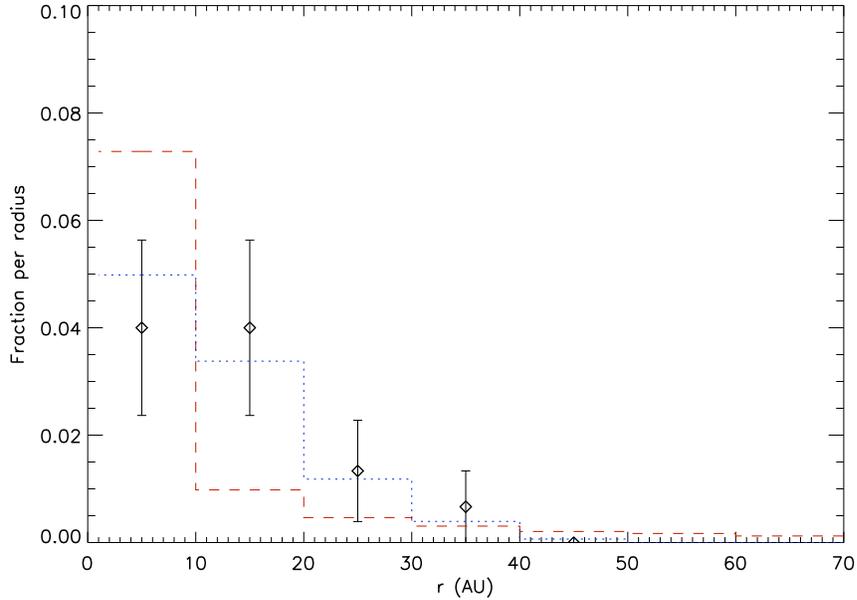


Figure 6.1: Histogram for the distribution of disc radii of the combined observational samples (diamonds) for discs detected at both 24 and 70 μm , shown with \sqrt{N} error bars, and the distributions for the entire model population (red, dashed line) and the model population that could be detected at 24 and 70 μm (blue, dotted line).

calculated by assuming realistic grain emission rather than blackbody emission. The true radius distribution should be larger than the one we derive, as is discussed in Sec. 6.7.2. Finally, we find a best-fit value for the median of the disc mass distribution of $M_{\text{mid}} = 15M_{\oplus}$, which is consistent with results for the disc mass-stellar mass relation of Natta (2004), who find an approximate range for solar-mass stars of $-2.5 < \log(M_{\text{disc}}/M_{*}) < -0.5$.

The models found with these parameters are plotted on Fig. 6.2, and the statistics are reproduced by the model convincingly, as they were as for A stars (Wyatt et al., 2007b); this is shown on the bottom panel of Fig. 6.2. The flux ratio shows very slow evolution past the earliest age bins, which is what is seen in the data as well. We use two excess categories, small ($R70 < 15$) and large ($R70 > 15$). The threshold value $R70 = 15$ was chosen empirically to avoid upper limits contaminating the large excess sample. The large excess fraction makes up $\sim 20\%$ of excesses at early ages, falling to a few percent within ~ 3 Gyr.

6.6.2 Fit to the 24 μm statistics

Fig. 6.3 shows the model fit (with the parameter values given in the previous section) to the observed 24 μm statistics. The statistics both in the model and the data samples show that 24 μm excess flux evolves on a timescale of $\sim 1.5 - 2$ Gyr, similar to what was found by

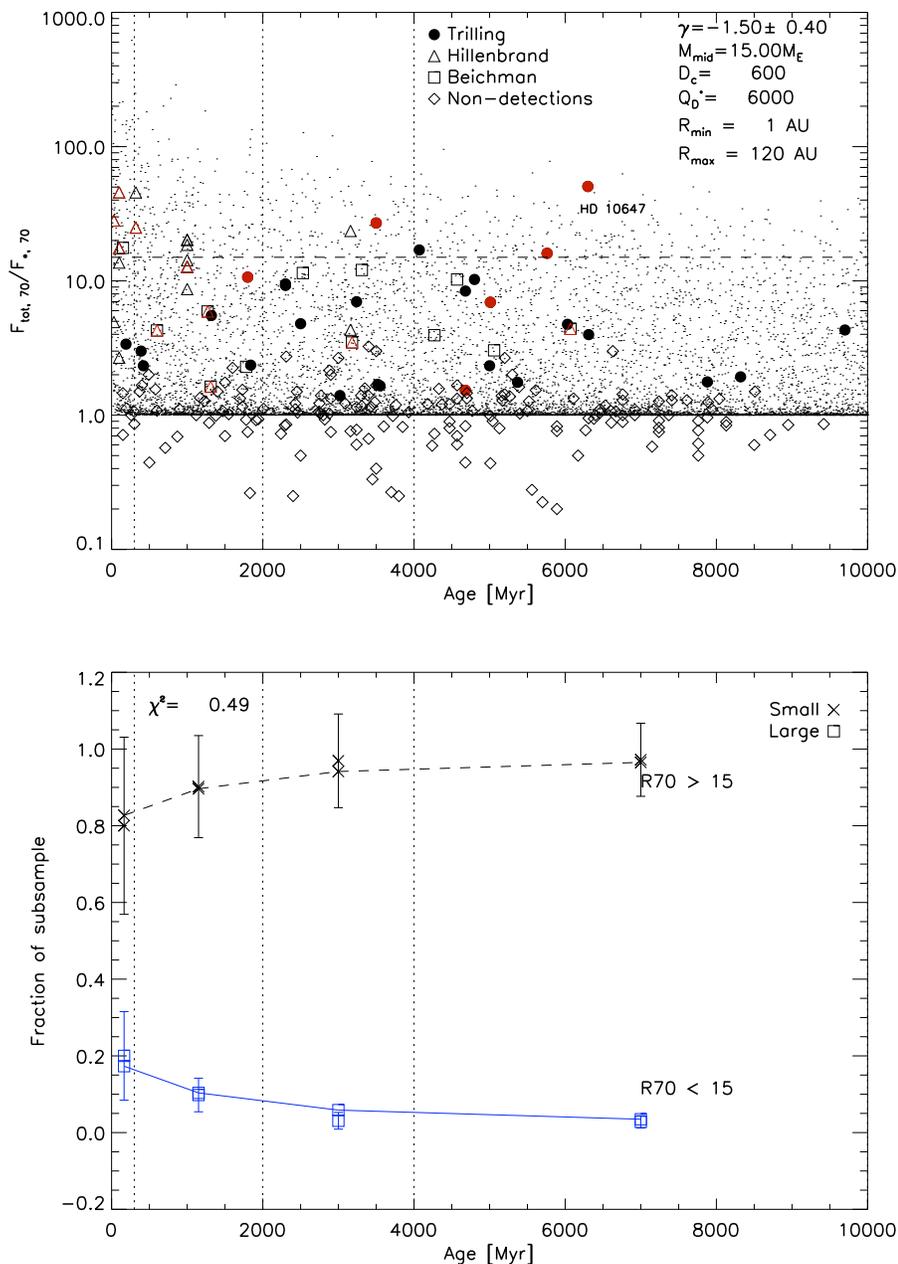


Figure 6.2: *Top:* total 70μm flux divided by the photospheric flux against age. The model population is shown with small dots, and the observations of Trilling, Hillenbrand and Beichman are shown as filled dots, triangles and squares respectively. Disks detected at 70μm with $< 3\sigma$ confidence are shown with open diamonds. Red symbols indicate disc emission detected at $> 3\sigma$ at both 24μm and 70μm. The horizontal dashed line separates the populations with small and large excess flux. *Bottom:* fractional populations of stars with different flux ratios for different age bins (0.03–0.3 Gyr, 0.3–2 Gyr, 2–4 Gyr, 4–10 Gyr). Small ($R_{70} = F_{70,\text{tot}}/F_{70,\text{phot}} < 15$) and large excesses ($R_{70} > 15$) are shown as crosses and squares respectively. Model predictions are connected with lines and observed values (Trilling et al., 2008; Hillenbrand et al., 2008; Beichman et al., 2006a) are plotted with \sqrt{N} error bars. On both panels, the dotted vertical lines indicate the limits of the age bins.

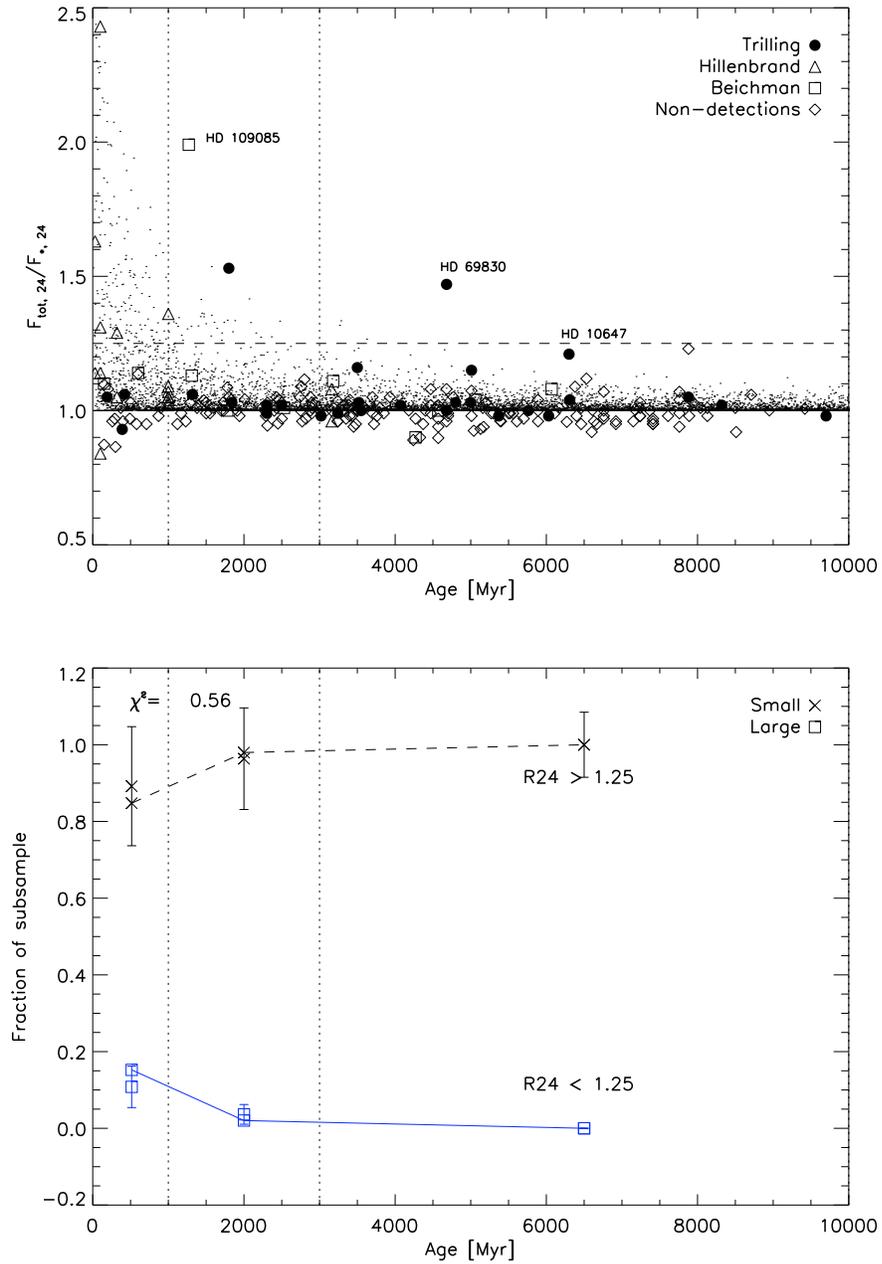


Figure 6.3: *Top:* total $24\mu\text{m}$ flux divided by the photospheric flux against age. Symbols are the same as for Fig. 6.2. Horizontal dashed lines separate the populations with small and large excess flux. *Bottom:* fractional populations of stars with different flux ratios for different age bins (0.03 – 1 Gyr, 1 – 3 Gyr, 3 – 13 Gyr). Small ($R_{24} = F_{24,\text{tot}}/F_{24,\text{phot}} < 1.25$), and large excesses ($R_{24} > 1.25$) are shown as crosses, triangles and squares respectively. Model predictions are connected with lines and observed values (Trilling et al., 2008; Hillenbrand et al., 2008; Beichman et al., 2006a) are plotted with \sqrt{N} error bars. On both panels, the dotted vertical lines indicate the limits of the age bins.

Löhne et al. (2008), but slower than what is seen in the observations of Siegler et al. (2007) and Meyer et al. (2008). $24\mu\text{m}$ statistics are well fitted by the model and are particularly crucial in constraining the parameters which determine the evolution timescale of the models (Q_D^* , e , D_c , M_{mid}), as they show clearer time evolution than the $70\mu\text{m}$ data; this is clearly visible on the top panel of Fig. 6.3. It is also clear from Fig. 6.3, however, that a number of systems are not well fitted by our models. We discuss shortly the systems labeled on the plot: HD 109085, HD 69830 and HD 10647, as well as HD 101259 (which is not shown on Figs. 6.2 and 6.3 because of its age of 10.96 Gyr).

The disc around HD 109085 was found by Sheret et al. (2004) to have a radius of ~ 180 AU, compared to the blackbody radius of 6 AU used in this analysis. This system was also reported to have 2 separate disc components: a component with $r < 3.5$ AU (recent work also shows that for this component $r > 0.5$ AU, e.g. Smith et al. 2009) and a cold component at ~ 150 AU (Smith et al., 2008), with the middle region possibly cleared by a planetary system (Wyatt et al., 2005). Our results suggest that the hotter disc component is transient, whereas the cold component is evolving in steady state. HD 69830 has 3 known planetary companions (Lovis et al., 2006), as well as an asteroid belt around 1 AU from the star (Beichman et al., 2005; Lisse et al., 2007). Since a disc with such a small radius is expected to have processed its material at 4.7 Gyr (the age of HD 69830), this system can be considered anomalous; fitting it in our models would require unrealistic parameter values.

HD 10647 also has a known Jupiter-mass planetary companion and a large cold disc at a radius of ~ 300 AU has been detected at sub-millimetre wavelengths (Liseau et al., 2008); this system has the largest $70\mu\text{m}$ excess in our whole sample. As for HD 69830, finding a model that fits this object would require extreme parameter values, as it is unusually bright compared to other discs of similar age and colour temperature. This could be caused by the disc around HD 10647 having unusually strong planetesimals, or having been recently stirred. It is also possible that a significant fraction of the $24\mu\text{m}$ excess emission comes from an additional hot component of the disc.

Finally, we classify HD 101259 as transient. This is the first identification of this object as an anomalous disc, and rests on the fact that its infrared luminosity is over 500 times the maximum theoretical value for a steady-state disc. As mentioned before, this is due to the fact that this is the oldest disc in our sample as well as that with the smallest radius, despite

the fact that the smaller discs are expected to evolve faster. However there is a degree of uncertainty as to the age of this system, as Takeda et al. (2007) report a best-estimate stellar age for HD 101259 of 1.88 Gyr, while Valenti & Fischer (2005) only limit its age to be between 2.9 and 11.6 Gyr. A better age determination is needed to make strong conclusions as to the transient nature of HD 101259, but our best-fit model implies that it is indeed anomalous if it is older than ~ 4 Gyr.

6.7 Discussion of results

6.7.1 Significance of best-fit parameters

The best-fit values we find for Q_D^* and D_c ($Q_D^* = 6000 \text{ J kg}^{-1}$, $D_c = 600 \text{ km}$) are significantly different from those found for the models of Wyatt et al. (2007b) for debris discs around A stars. Both Q_D^* and D_c are an order of magnitude larger (values for these parameters in the A stars study are 150 J kg^{-1} and 60 km respectively). This could suggest either that our models are incomplete and these parameter values compensate for the value of one of several other free parameters being poorly chosen or fitted, or that the properties of debris discs around later-type stars are actually different. Other regions of parameter space were also explored, and our best fit constitutes the best minimum in the χ^2 goodness-of-fit statistic which also corresponds to realistic parameter values, as stated in Sec. 6.4. The difference in parameters is illustrated on Figs. 6.4 and Fig. 6.5, which show the FGK observations with the best-fit model found for A stars (Wyatt et al., 2007b) at 24 and $70 \mu\text{m}$ respectively.

In comparison with the study of A stars, we find different values for r_{min} , γ , M_{mid} , D_c and Q_D^* (see Sec. 6.6). The most significant differences are the higher values of Q_D^* and D_c , although the exact values of the individual free parameters are poorly constrained and are not as informative as combinations of these parameters. In this case, the values point to a slower evolution of $F_{70,\text{tot}}/F_{70,\text{phot}}$ and of $F_{24,\text{tot}}/F_{24,\text{phot}}$ compared to models found for discs around earlier-type stars.

Q_D^* might vary because the composition of the disc may be different around later-type stars, or the bodies making up the disc may be more compacted by previous collisions, making them stronger and accounting for a higher intrinsic strength of the discs. The value of D_c indicates an initial population made of large objects, up to Pluto-size asteroids rather than smaller planetesimals, hinting at Kuiper-belt-like properties. However, the difference between

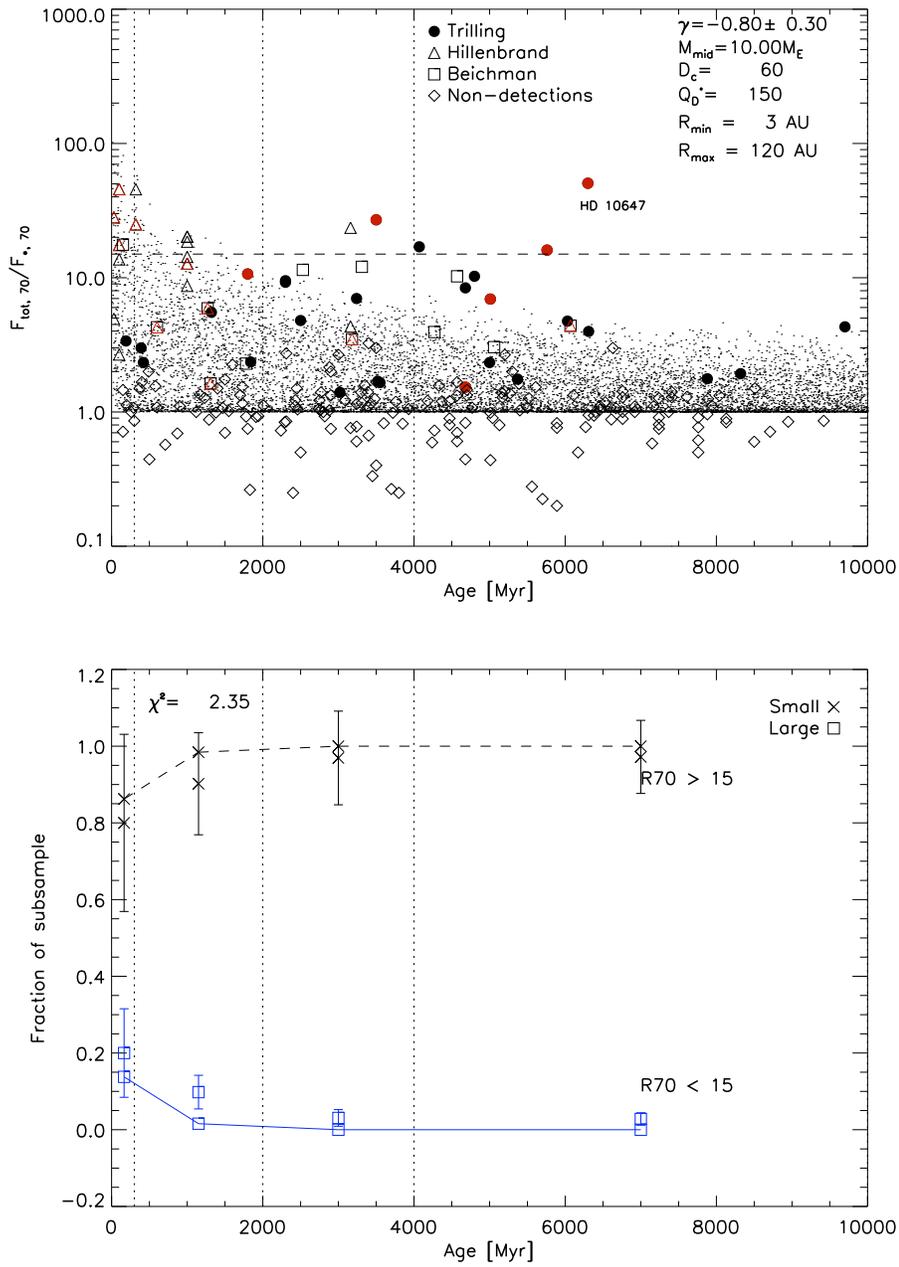


Figure 6.4: Same as Fig. 6.2, but with the parameters found for A stars (Wyatt et al. 2007). This illustrates the significant difference between the evolutionary timescales corresponding to the models found in this analysis compared to those found for A stars; this significant time evolution is not seen in observations of discs around FGK stars.

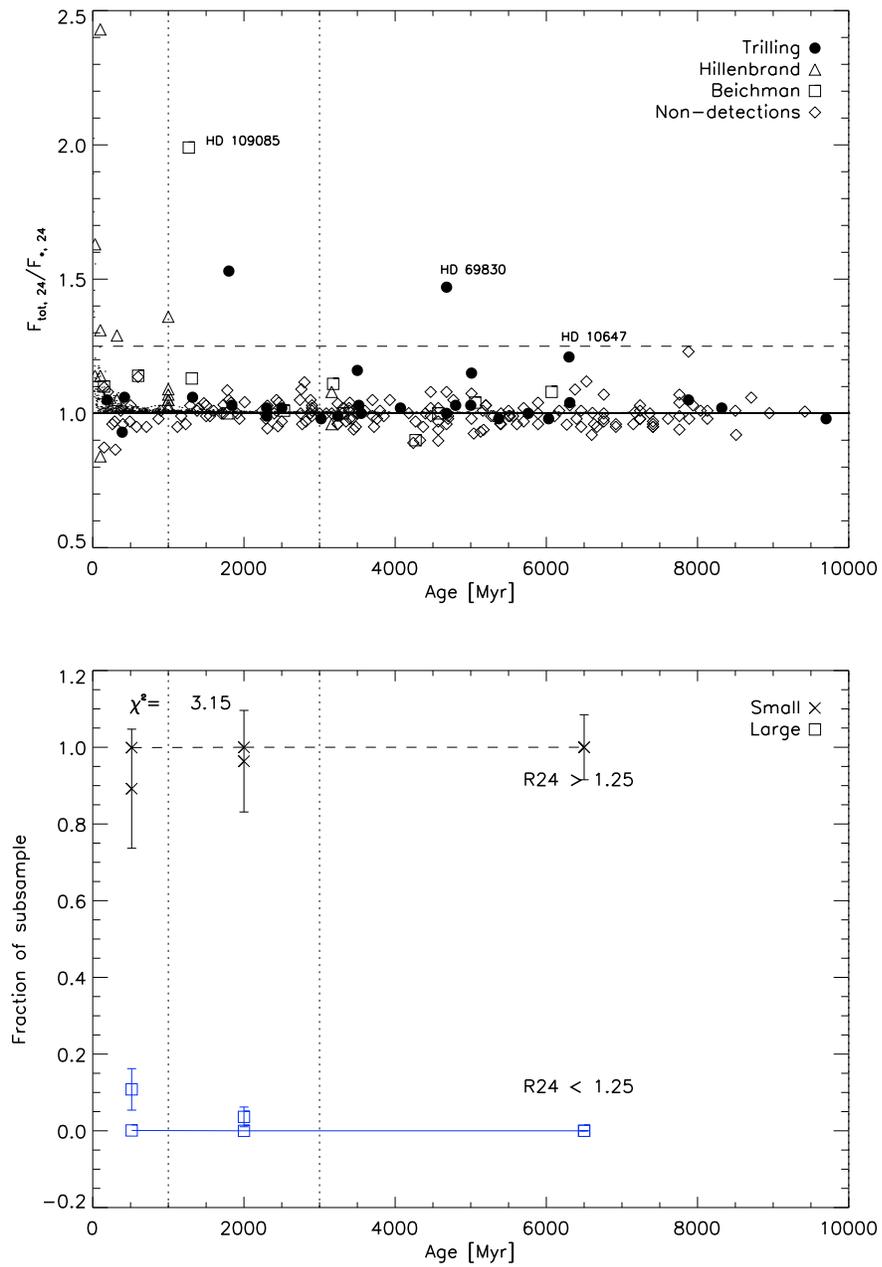


Figure 6.5: Same as Fig. 6.3, but with the parameters found for A stars (Wyatt et al. 2007). This illustrates the significant difference between the evolutionary timescales corresponding to the models found in this analysis compared to those found for A stars; the lack of a significant time evolution, not seen in observations of discs around FGK stars, is most obvious at $24\mu\text{m}$, where no significant excess is predicted beyond ~ 500 Myr.

solar-type and A stars results may also be due to inaccuracies in deriving the radius of discs with the $24 - 70\mu\text{m}$ colour temperature.

Because small grains are inefficient emitters, and we are assuming blackbody radiation from all grains in the disc, the actual radii may be larger than the derived values. This is confirmed by results of disc imaging which show that discs are generally 2-3 times larger than the blackbody value (e.g. Maness et al. 2009). We therefore emphasise that in this analysis we are finding an *effective* value for the parameters Q_D^* and D_c , as was done for the A stars study. The discrepancy in effective values between A and FGK stars could then be caused by a dearth of small grains around A stars compared to solar-type stars. This is because the blowout size of grains is larger around earlier-type stars ($\sim 10\mu\text{m}$ for A stars and $\sim 1\mu\text{m}$ for a G0 star), meaning that with larger number of inefficient emitters, the emission is *less* well described by blackbody emission for discs around FGK stars. Hence the ratio of real radius to blackbody radius could be larger for solar-type stars than for A type stars. Despite *effective* values being different, the real values of the planetesimal strength and largest planetesimal size may therefore be similar for solar-type and A stars. In the next section, we discuss and quantify this.

6.7.2 Effective and real parameters

Bonsor & Wyatt (2010) used grain modelling, taking into account the non-blackbody nature of the grains and a realistic size distribution, to show that for A stars, real disc radii are expected to be larger than radii derived by fitting blackbody curves to spectral energy distributions (SEDs) by a factor of ~ 2 (with some dependence on radius and spectral type), in agreement with observations of resolved discs. They also derived scaling laws between realistic and effective values for parameters within the context of the modelling presented here. Assuming that the real radii are larger by a factor X_r , compared to blackbody radii, then the real parameters can be calculated from parameters which were found using the blackbody assumption, by making sure that all discs have the same luminosity evolution when the radii are changed. In the following discussion we use a simplified notation so that $M \equiv M_{\text{mid}}$, $Q \equiv Q_D^*$ and $D \equiv D_c$. Furthermore, we use the subscript r to denote a real value, while its absence denotes an effective value: Q_r refers to the real planetesimal strength, while Q refers to the fitted (effective) parameter.

Bonsor & Wyatt (2010) find that discs start off with the same luminosity if

$$M_r D_r^{0.5} X_r^{-2} = M D^{0.5}, \quad (6.13)$$

from Eq. 3-4 of Wyatt et al. (2007b) (see also Eq. 16 of that paper), and discs evolve on the same timescale if

$$Q_r^{5/6} D_r^{0.5} X_r^{7/3} = Q^{5/6} D^{0.5}. \quad (6.14)$$

The real parameters for each population cannot be derived from the parameters of the blackbody fit, because there are 4 free parameters and 2 constraints. However, it is possible to compare the real parameters of the FGK and A star populations from their blackbody fits. We rearrange Eq. (6.13) and (6.14) to derive expressions for X_r and Q_r for the FGK stars population in terms of the equivalent A stars parameters. We use the additional subscripts A for A stars and F for FGK stars in the following comparison.

By rearranging Eqs. 6.13 and 6.14 for A and FGK stars, we find that the parameters for the two populations are related through

$$\frac{X_{r,F}}{X_{r,A}} = \left(\frac{M_F M_{r,A}}{M_A M_{r,F}} \right)^{-1/2} \left(\frac{D_F D_{r,A}}{D_A D_{r,F}} \right)^{1/4} \quad (6.15)$$

$$\frac{Q_{r,F}}{Q_{r,A}} = \frac{Q_F}{Q_A} \left(\frac{M_F M_{r,A}}{M_A M_{r,F}} \right)^{7/5} \left(\frac{D_F D_{r,A}}{D_A D_{r,F}} \right)^{-1/10}. \quad (6.16)$$

If we assume that the strength of the planetesimals follows $Q_r \propto D_r^{1.5}$ (i.e. that the strength of the planetesimals is mainly gravitational), and that planetesimals have the same composition around A and FGK stars, we can also rearrange these equations to find that, with the parameter values found in this analysis and in Wyatt et al. (2007b) ($Q_A = 150 \text{ Jkg}^{-1}$, $D_A = 60 \text{ km}$ and $M_A = 10 M_\oplus$) as well as the value of $X_{r,A} = 2$ found for A stars by Bonsor & Wyatt (2010),

$$\frac{M_{r,F}}{M_{r,A}} \simeq 18 \frac{D_{r,A}}{D_{r,F}}, \quad (6.17)$$

and using this to derive an expression for the scaling factor yields

$$\frac{X_{r,F}}{X_{r,A}} \simeq 6.1 \left(\frac{D_{r,A}}{D_{r,F}} \right)^{-3/4}. \quad (6.18)$$

We can expect X_r to be larger for FGK stars than for A stars for the reason mentioned earlier that there are more inefficient emitters around FGK stars. Unfortunately the overlap between the sample of discs resolved with *Spitzer* and the data sample used in this analysis is too small for this to be useful to constrain X_r for FGK stars, although imaged discs such as HD 181327 (Schneider et al., 2006) suggest that a value of ~ 3 is a good estimate for $X_{r,F}$. Using $X_{r,A} = 2$ (Bonsor & Wyatt, 2010), the value of $X_{r,F}$ is consistent with the estimate of ~ 3 that is expected from imaged discs (e.g. Schneider et al. 2006) if the planetesimals around FGK stars are larger than those around A stars by a factor of $\sim 6 - 7$. Putting this number into Eq. (6.17), this means that the median disc mass around FGK stars is larger than for A stars by a factor of ~ 3 . This is surprising, since earlier-type stars are expected to have more massive discs (Natta, 2004). Further observations will allow us to better constrain the best-fit parameters for FGK stars, test our finding that $M_{r,F} > M_{r,A}$, and possible reasons for this.

Finally, the values found for γ , r_{\min} and r_{\max} point to more numerous small discs for FGK stars compared to their earlier-type counterparts. This, however, does not take into account the possible systematic difference in actual disc radius contained in the scaling factor $X_{r,F}$, which may result in a similar size distribution to that found for the A stars.

6.7.3 Fractional luminosity vs. radius

Fig. 6.6 shows a plot of fractional luminosity f against disc radius, with the observed data plotted over the model population. Also plotted are the detection limits at both 24 and $70\mu\text{m}$, given by Eq. (6.10), and the (dashed) lines of maximum fractional luminosity f_{\max} at 2 and 10 Gyr, given by Eq. (6.9), for an F0 dwarf and the best-fit model parameters. The detection thresholds are calculated assuming calibration limits of $R_{24} = 0.11$ and $R_{70} = 0.5$. In theory, discs should lie on or below the line of maximum luminosity for their age and spectral type, although as noted by Wyatt et al. (2007b), the precise location of these lines depends on parameters which may vary between discs such as Q_D^* , e , D_c , and the spectral type of the disc's host star. The sharp increase in the $24\mu\text{m}$ threshold for discs with radii larger than ~ 50 AU means that the discs with larger radii, i.e. those discs which are expected to have the strongest emission at $70\mu\text{m}$ will be more difficult to detect at $24\mu\text{m}$ as well as they will require high fractional dust luminosity.

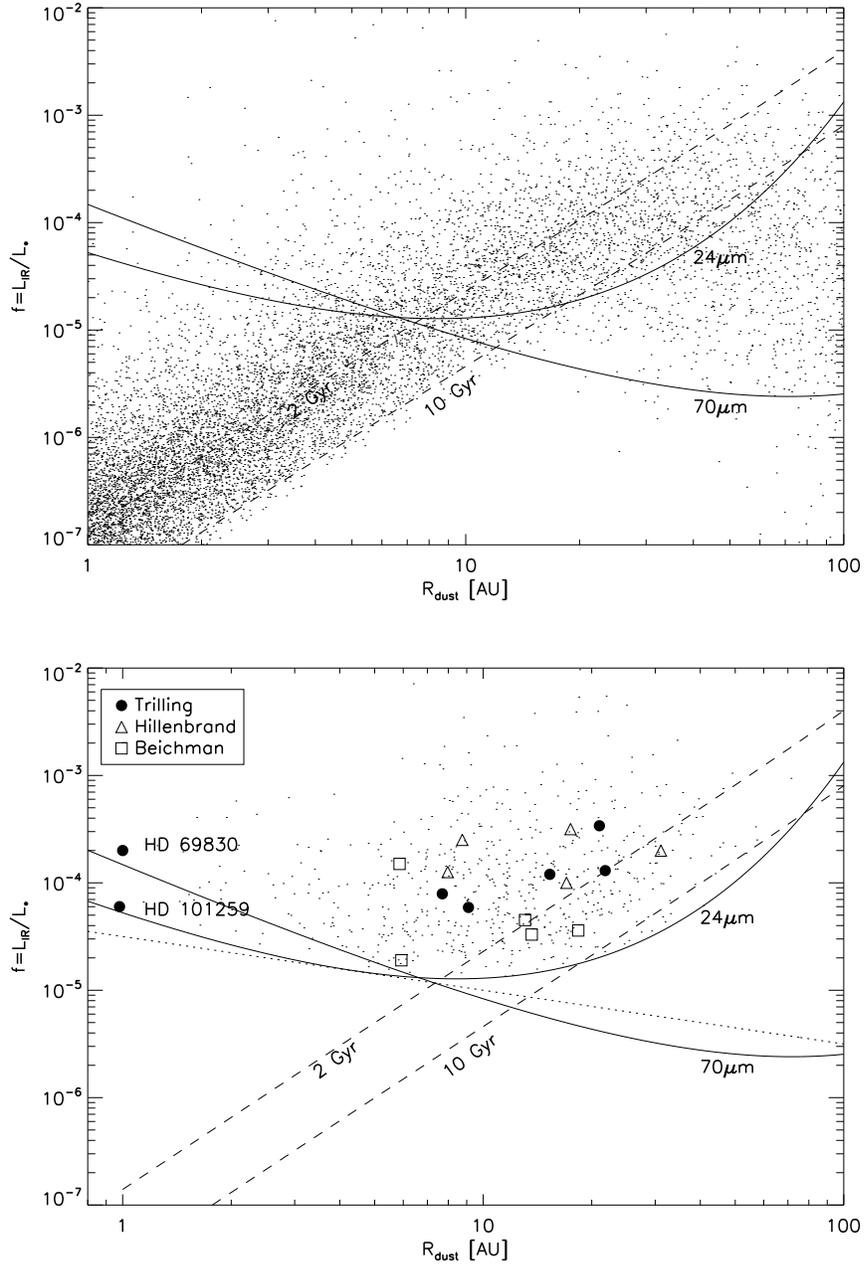


Figure 6.6: Dust luminosity $f = L_{IR}/L_*$ plotted as a function of disc radius R_{dust} . The top panel shows the whole model population, while the bottom panel only has the model population that could be detected at 24 and $70\mu\text{m}$. The model population is shown with small dots, and data symbols are the same as for Fig. 6.3. Detection thresholds at 24 and $70\mu\text{m}$ for an F0 dwarf are indicated by solid lines. Lines showing the maximum possible fractional luminosity for an F0 dwarf are shown as dashed lines for 2 and 10 Gyr. The dotted line indicates the limit, given by Eq. (6.11), where Poynting-Robertson drag becomes important for an F0 dwarf.

Values of f/f_{\max} for discs detected at 24 and $70\mu\text{m}$ are given in Table 6.2. Since f_{\max} depends on the disc radius², a good estimate can only be made for those discs detected at both wavelengths, so we do not include a value f/f_{\max} for discs detected at $70\mu\text{m}$ only. Two discs stand out in Table 6.2 and on Fig. 6.6: HD 69830 and HD 101259 (Trilling et al., 2008) have luminosities f over 500 times the maximum theoretical value f_{\max} for their age, radius and spectral type. This comes from a very small disc radius combined with an old stellar age: with an age of 10.96 Gyr, HD 101259 is the oldest star in the observed sample (although the age of this object is poorly determined, a metallicity measurement of $[M/H]=-0.6$ reported by Trilling et al. (2008) suggests that this system is old), yet the radius of its disc is the smallest, with a disc radius calculated to be 0.98 AU. A fit to the model population yields $f \propto r^{0.56}$, while a fit to the observed population of discs in the Trilling et al. (2008) and Beichman et al. (2006a) samples, without these anomalously bright systems, gives a relation $f \propto r^{0.56 \pm 0.08}$, in excellent agreement.

6.7.4 Fractional luminosity vs. age

Fig. 6.7 shows the fractional (disc/star) luminosity of discs as a function of age, as well as theoretical evolution lines for discs of initial mass 3, 30 and $300 M_{\oplus}$ of radii 10 and 100 AU. The larger discs do not reach steady state in the timescale considered here, while the smaller discs have reached their collisional equilibrium by ~ 1 Gyr.

A fit to the model population gives the relation $f \propto t^{-0.50}$ while a fit to the observed population yields $f \propto t^{-0.18 \pm 0.10}$, just outside 3σ of the best-fit model. Fitting of the observed discs is strongly affected by the low number of discs in our data set, especially by the relatively few young discs (< 1 Gyr) in the sample, but one can conclude from Fig. 6.7 that the model population is a good fit for the observed population. The validity of these fits will benefit from additional observations, which will allow more significant conclusions to be drawn with larger samples. Interesting to notice is the absence of a peak in excess flux analogous to the one seen in the observed A star samples around 10-15 Myr (Currie et al., 2008).

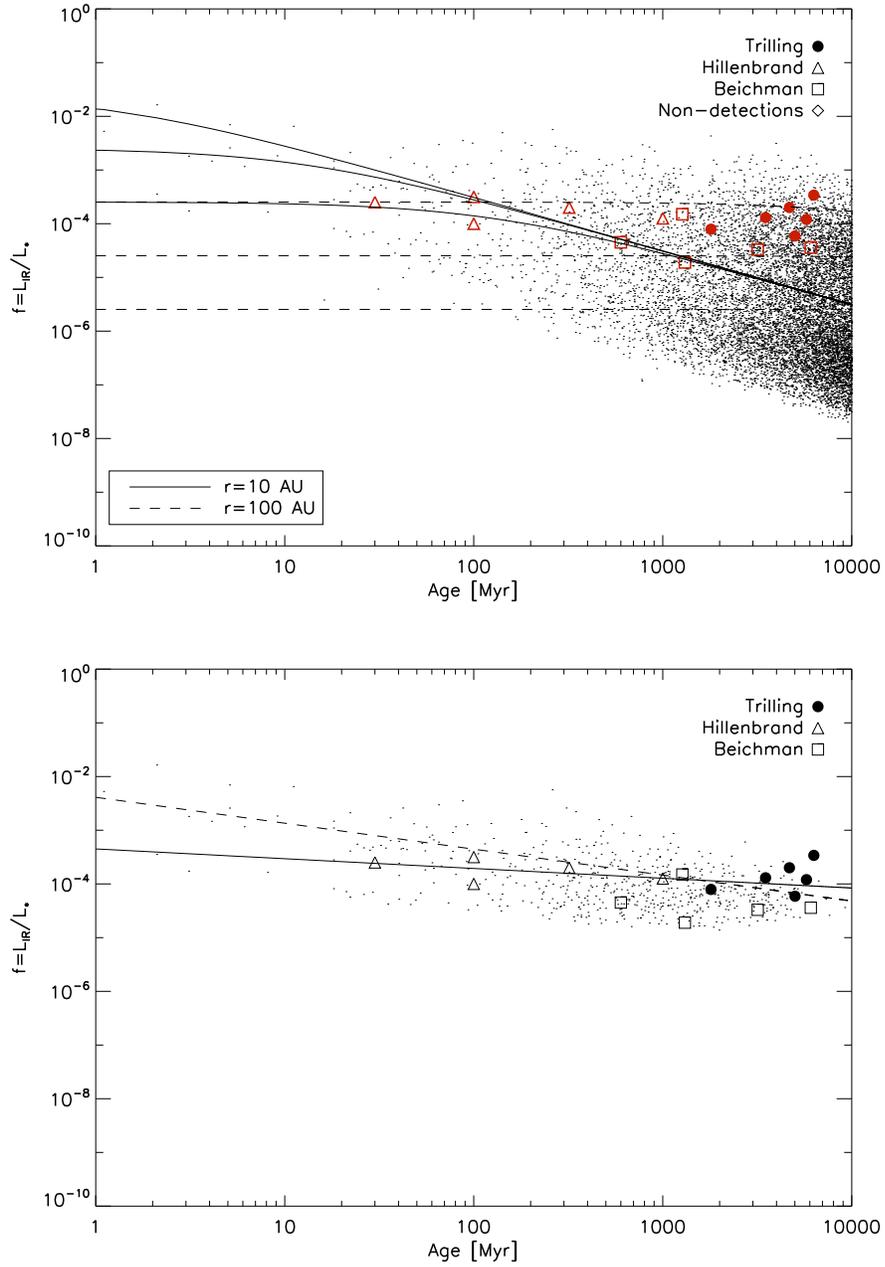


Figure 6.7: Dust luminosity $f = L_{\text{IR}}/L_*$ plotted as a function of age. The model population is shown with small dots, and data symbols are the same as for Fig. 6.3. For $< 3\sigma$ detections, the fluxes plotted are maximum values. *Top:* Entire model population, and all the $70\mu\text{m}$ excess detections, with objects detected at both 24 and $70\mu\text{m}$. Also shown are theoretical evolutionary tracks for discs of radii 10 (solid line) and 100 (dashed line) AU, and for initial disc mass of 3, 30 and $300 M_{\oplus}$ (with the higher masses corresponding to the lines with higher early-age fractional luminosities), around an F0 star. *Bottom:* Same with only the populations that could be detected at both 24 and $70\mu\text{m}$ plotted. Also plotted on the bottom panel are power-law fits to the model (dashed line) and the data (solid line).

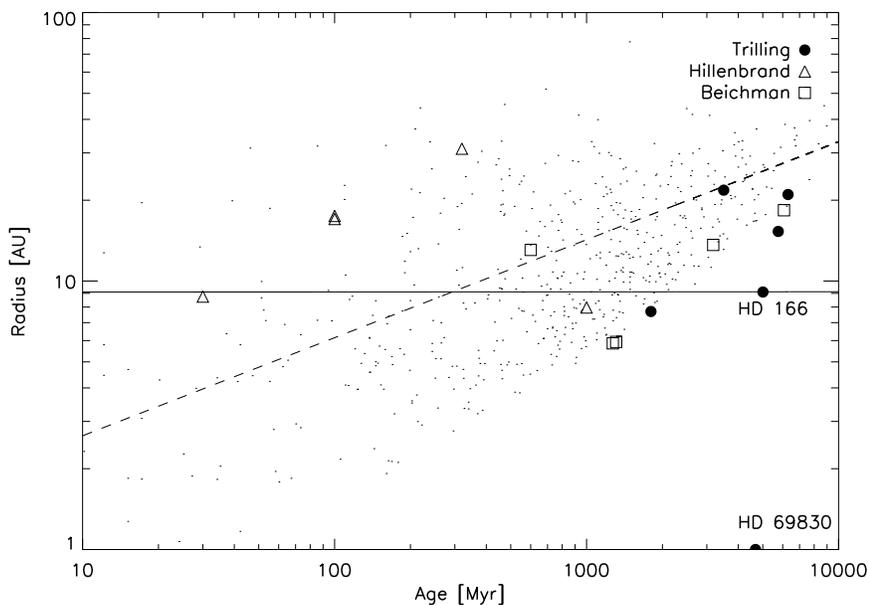


Figure 6.8: Disk radius plotted as a function of age. The model population is shown with small dots, and data symbols are the same as for Fig. 6.3. Only the populations which could be detected at both 24 and 70 μm are plotted. Also shown are power-law fits to the model (dashed line) and observed (solid line) populations.

6.7.5 Radius vs. age

Fig. 6.8 is a plot of disc radius as a function of age, for those discs that could be detected at 24 and 70 μm .

A lower limit for detectable radii is visible in the model population, with this lower limit increasing with age beyond ~ 100 Myr. Older discs must therefore have larger radii to be detected at both 24 and 70 μm . This comes from the fact that discs with larger radii are slower in processing their mass and can therefore remain above detection thresholds for a longer time than discs with small radii. This radius increase with age is therefore only due to detection sensitivity considerations. In fact, there is as yet no evidence for a correlation between radius and age (Najita & Williams, 2005).

A fit to the model population confirms the apparent increase of disc radius with age, with the relation $r \propto t^{0.34}$, while a fit to the observed population yields $r \propto t^{0.00 \pm 0.04}$. The fit to the data is strongly influenced by the discs from the Hillenbrand et al. (2008) sample, which consists of younger, and mostly larger discs. We also see a lack of observed small and young

²Note that the calculation of f_{max} is unaffected by the discussion in section 6.7.2 on the difference between real and effective radius, as long as f_{max} is calculated using both effective parameters and blackbody radius (which would give the same value that would be calculated from the real parameters and real radius).

discs, which are predicted by the model. Although this could be evidence for rapid inner clearing, the small number of discs of the observed sample makes it difficult to draw robust conclusions as to the cause of this. Thus we note that it is worth scrutinising <300 Myr stars to see if there really is an absence of <10 AU discs that would be the precursors of older systems such as HD 40136 and η Corvi, and whether the radii derived from 24-70 μm colour temperatures appear smaller than reality for a significant number of older discs, perhaps due to the presence of both hot and cold dust (as is known to be the case for η Corvi; Wyatt et al. 2005), or due to the action of P-R drag (discussed later).

The observed discs HD 69830 and HD 166 (Trilling et al., 2008) appear on the plot where the model predicts that no discs should be detected (bottom right corner of Fig. 6.8); HD 101259 does not appear on this plot due to its age of 11 Gyr, but it is also anomalous. For HD 101259 and HD 69830, we attribute this to an unusually high disc luminosity; indeed these discs are found to have over 500 times the maximum theoretical luminosity for its radius, age and the spectral type of their host, as was already mentioned in Sec. 6.7.3. HD 166 lies just outside the model population. Its disc is small for its age and its luminosity is just over its maximum theoretical luminosity; the unusual brightness for a disc of this age and colour temperature may mean that the disc around HD 166 has unusually strong or massive planetesimals compared to its peers.

Dust might also find itself closer to the star than the planetesimal belt under the action of Poynting-Robertson drag. This becomes important when P-R drag and collisional timescales are comparable, or in terms of luminosity, when $f < f_{\text{PR}}$, with f_{PR} being given by Eq. (6.11).

This limit is shown on Fig. 6.6. From this plot, it emerges that discs with small radii could be affected by PR drag, given the proximity of the limit for PR drag to the limit of detection at 24 μm for these discs. In our sample, no system has $f/f_{\text{PR}} < 1$, meaning that PR drag is not significant for the observed discs we are considering; the most likely system to be affected by this effect is HD 40136. We also note that if the disc radius has been underestimated by a factor X_r , then the limit at which P-R drag becomes important given in Eq. (6.11) should be lower by a factor $X_r^{-1/2}$. Therefore we do not expect P-R drag to have significantly influenced the blackbody radius, derived from dust temperature, which is thus a good proxy for the location of the planetesimal belt (with the caveats mentioned in Sec. 6.7.2).

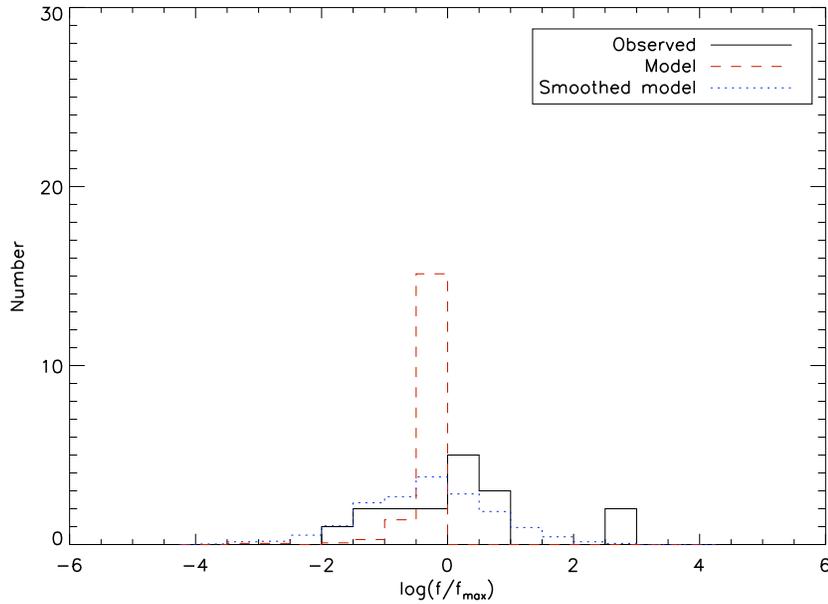


Figure 6.9: Histogram of the dust fractional luminosity for the observed population with an excess at both 24 and 70 μm (solid line), the model population (dashed, red) and the model population with a Gaussian smoothing of 1 dex (dotted, blue). Only the model population which could be detected at both 24 and 70 μm is plotted.

6.7.6 Histogram of f/f_{\max}

Fig. 6.9 shows a histogram of the quantity f/f_{\max} , plotted for our data sets and for the model. We also plot the model with an arbitrary Gaussian smoothing of 1 dex to account for variations of disc properties between individual systems, which we have not considered in our model. The range of fluxes is broader in the data, which, as stated by Wyatt et al. (2007b) can be explained by the fact that we have assumed in our model that all the planetesimal belts have the same properties, where in reality the values of parameters such as D_c , Q_D^* and e are expected to vary from one belt to another. The parameters we used would then be correct as an average over a whole population of stars, but not necessarily for individual debris discs. This is particularly relevant when the sample of observed discs is as small as it is here. Although our data set is small, main features are reproduced by the smoothed model. Disagreements are obvious too, in particular the high proportion of low-luminosity discs in our data sample compared to the model prediction. We attribute this to small statistics; future observations will help build up more significant statistics.

6.8 Conclusions

In this chapter, we showed that the main features seen in observations of debris discs around FGK stars can be attributed to collisional grinding of planetesimals. We modelled data collected by several teams with the Spitzer Space Telescope, and used 24 and $70\mu\text{m}$ colours to derive simple properties of the discs for each of these systems. Our results are consistent with models of planetesimal strengths and their size distributions, although we find parameters corresponding to these properties to be an order of magnitude different from those found in the study of A stars that was done using the same model (Wyatt et al., 2007b). We discussed whether these values might be effective values due to differences in stellar environment, and found that properties of discs around later-type stars may be somewhat different from those around earlier-type stars, because they evolve on longer timescales and therefore have had more time to form larger planetesimals, which are then also stronger if the strength is mainly due to gravitational pull. Therefore planetesimals around FGK stars may be stronger and larger than around A stars. We also find that discs around FGK stars are more massive than those around A stars, by a factor of a few. Future observations to increase the sample of debris discs detected at multiple wavelengths will allow us to test our assumption that the blackbody disc radii underestimate the true radii of discs around FGK stars by a larger factor than for those around A stars. Realistic grain modelling analogous to the work done for A stars by Bonsor & Wyatt (2010) could also help constrain this factor.

Finally, we have also identified a potential transient candidate in HD 101259, for which data suggest that its debris disc cannot be in steady state; however a better age determination of that system is needed to confirm the dust emission from the disc as transient.

7

Conclusions and outlook

My doctoral work has been concerned with two aspects of planetary systems, and with making a contribution to understanding how the Earth and the Solar System fit in our Galaxy. Firstly, by trying to detect and characterise extrasolar planets, and working on making the methods used to model anomalous microlensing lightcurves more systematic, I aim to strengthen the impact that microlensing observations has on our understanding of the make-up of planet populations, as through the improvement of modelling techniques, we can extract more meaningful information out of our observations. Secondly, I have also investigated the earliest stages of planet formation, by attempting to determine which mechanisms dominate the evolution of debris discs, the birth place of planets and how properties of discs vary around different types of stars. In particular, I have modelled discs around stars of spectral type F, G and K, "solar-type" stars, using data from the *Spitzer Space Telescope*. Working out how the early stages of planet formation take place around stars similar to our Sun is essential to gaining an insight into the origins of the structure of our own Solar System, including planets, as well as the asteroid and Kuiper belts.

Below, I include suggestions and an outline for possible future work. Much of this will form the starting point of the work I intend to carry out as an ESO postdoctoral fellow.

7.1 Automatic fitting: the future for microlensing

As I have already briefly mentioned in the introduction and in Chapter 4, the next few years should see the deployment of many robotic telescopes, increasing the amount of observing time devoted to following microlensing events, and of data to be analysed. The current way of modelling each individual microlensing event by hand will rapidly become impractical, considering that the current data flow is already stretching available human and computational resources. Clearly, then, a more efficient approach needs to be taken.

The work I have presented in Chapters 4 and 5, in particular, constitute very promising bases to develop a real automated binary-lens fitting algorithm. However, this work was carried out for a special class of anomalous microlensing events, in which caustic crossings are clearly visible in observational datasets. Although many anomalous events indeed fall into that category, some important events do not- an example of this being the event through which the first cool, rocky extrasolar planet was discovered, OGLE-2005-BLG-390 (Beaulieu et al., 2006). We will therefore need to extend the parameterisation developed by Cassan (2008) to include anomalies which are caused by sources approaching caustics but not necessarily crossing them.

Once a working algorithm has been developed including priors, I aim to test it by re-analysing OGLE-2007-BLG-472, which is the event used as an illustration in Chapter 4. Using the same event will allow us to directly compare results of our algorithm with and without priors. I then hope to develop a full algorithm with several collaborators, and to implement it to fit events in real-time during the next observing season. This will allow us to take a leading role in microlensing modelling, as well as to optimise the use of our observing time, an essential step towards the long-term objective of obtaining a more robust and coherent sample of detected planets.

7.2 Towards a census of debris discs around all types of stars

M stars are by far the most numerous types of stars in the Galaxy, with about 70% of its population, yet very few discs have been detected around these objects, due in part to the difficulty

of measuring accurate photospheric fluxes for these stars. Recent searches (e.g. Lestrade et al. 2009) have found the fraction of cold discs around M stars to be smaller than for A and FGK samples, although these results are mitigated by their marginal statistical significance. Gathering additional observational data of discs around M stars in future years will be crucial to determining whether discs are really more scarce around these types of stars, and will also allow us to conduct proper modelling of the discs' evolution. This will enable us to draw conclusions as to the correlation between evolutionary processes and timescales with spectral type. Moreover, future observations will also help verify the findings made in this thesis, especially regarding the properties of discs around later-type stars. Even for FGK stars, the sample of discs detected at multiple wavelengths is small, and additional observations, including with missions and observing facilities like the Herschel Space Observatory and ALMA, will no doubt prove crucial to strengthening or modifying the current understanding of the earliest stages of planet formation.

7.3 Conclusion

The overall underlying motivation for my doctoral work has been to put the Solar System in perspective. Twenty-five years ago, only a handful of debris discs and no extrasolar planets had been detected. Some were even doubting that such objects would ever be found, hinting that the corner of the Galaxy that we inhabit was rather lonely. Today offers a strikingly different picture: hundreds of debris discs and extrasolar planets have been detected, and it is surely only a matter of time until the first truly Earth-like planet residing in the "habitable zone" of a star is found, i.e. a planet which might have the potential to sustain life as we know it. Such a discovery will bring with it a plethora of questions, and challenge our understanding of life itself; enough to keep astronomers busy and excited for decades to come!

References

- Alard, C., Guibert, J., Bienayme, O., Valls-Gabaud, D., Robin, A. C., Terzan, A., & Bertin, E. 1995, *The Messenger*, 80, 31
- Albrow, M. D. et al. 2001, *ApJ*, 549, 759, arXiv:astro-ph/0004243
- . 1999, *ApJ*, 522, 1022, arXiv:astro-ph/9903008
- . 2000a, *ApJ*, 534, 894, arXiv:astro-ph/9910307
- . 2000b, *ApJ*, 534, 894, arXiv:astro-ph/9910307
- . 2009, *MNRAS*, 397, 2099, 0905.3003
- Alcock, C. et al. 1993, *Nature*, 365, 621, arXiv:astro-ph/9309052
- . 1995, *ApJ*, 454, L125+, arXiv:astro-ph/9506114
- . 2001, *Nature*, 414, 617
- Aller, L. H. et al. 1982, *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, vol.2 (Springer-Verlag)
- Andrews, S. M., & Williams, J. P. 2005, *ApJ*, 631, 1134, arXiv:astro-ph/0506187
- Aubourg, E. et al. 1993, *Nature*, 365, 623
- Aumann, H. H. et al. 1984, *ApJ*, 278, L23
- Beaulieu, J.-P., et al. 2006, *Nature*, 439, 437, arXiv:astro-ph/0601563
- Beichman, C. A. et al. 2005, *ApJ*, 626, 1061, arXiv:astro-ph/0504491
- . 2006a, *ApJ*, 652, 1674, arXiv:astro-ph/0611682

- . 2006b, *ApJ*, 639, 1166, arXiv:astro-ph/0601467
- Bennett, D. 1998, *Phys. Rep.*, 307, 97, arXiv:astro-ph/9808121
- Benz, W., & Asphaug, E. 1999, *Icarus*, 142, 5, arXiv:astro-ph/9907117
- Bond, I. A., et al. 2004, *ApJ*, 606, L155, arXiv:astro-ph/0404309
- Bonsor, A., & Wyatt, M. C. 2010, in preparation
- Booth, M., Wyatt, M. C., Morbidelli, A., Moro-Martín, A., & Levison, H. F. 2009, *MNRAS*, 399, 385, 0906.3755
- Bramich, D. M. 2008, *MNRAS*, 386, L77, arXiv:astro-ph/08021273
- Bryden, G. et al. 2009, *ApJ*, 705, 1226
- . 2006, *ApJ*, 636, 1098, arXiv:astro-ph/0509199
- Burke, C. J. et al. 2007, *ApJ*, 671, 2115, 0705.0003
- Byalko, A. V. 1969, *AZh*, 46, 998
- Cassan, A. 2005, PhD thesis, Institut d'Astrophysique de Paris / Université Pierre et Marie Curie (Paris VI)
- . 2008, *A&A*, 491, 587, arXiv:astro-ph/08081527
- Cassan, A., Horne, K., Kains, N., Tsapras, Y., & Browne, P. 2009, *ArXiv e-prints*, 0911.5285
- Chabrier, G. 2003, *PASP*, 115, 763
- Chang, K., & Refsdal, S. 1979, *Nature*, 282, 561
- Charbonneau, P. 1995, *ApJS*, 101, 309
- Chwolson, O. 1924, *Astronomische Nachrichten*, 221, 329
- Collinge, M. J. 2004, *ArXiv Astrophysics e-prints*, arXiv:astro-ph/0402385
- Crowe, M. 2008, *The extraterrestrial life debate, antiquity to 1915 : a source book* (University of Notre Dame, Notre Dame, Ind., M.J.Crowe (ed)), 198–200

- Currie, T., Kenyon, S. J., Balog, Z., Rieke, G., Bragg, A., & Bromley, B. 2008, *ApJ*, 672, 558, 0709.2510
- Dohnanyi, J. S. 1969, *J. Geophys. Res.*, 74, 2531
- Dominik, C., & Decin, G. 2003, *ApJ*, 598, 626, arXiv:astro-ph/0308364
- Dominik, M. 1998a, *A&A*, 330, 963
- . 1998b, *A&A*, 329, 361, arXiv:astro-ph/9702039
- . 1998c, *A&A*, 333, 893, arXiv:astro-ph/9801119
- . 1999, *A&A*, 349, 108, arXiv:astro-ph/9903014
- . 2004, *MNRAS*, 352, 1315, arXiv:astro-ph/0403212
- . 2006, *MNRAS*, 367, 669, arXiv:astro-ph/0507540
- . 2007, *MNRAS*, 377, 1679, arXiv:astro-ph/0703305
- Dwek, E., et al. 1995, *ApJ*, 445, 716
- Eddington, A. S. 1920, *The Observatory*, 43, 228
- Einstein, A. 1911, *Annalen der Physik*, 35, 898
- . 1916, *Annalen der Physik*, 354, 769
- . 1936, *Science*, 84, 506
- Eisenhauer, F., et al. 2005, *ApJ*, 628, 246, arXiv:astro-ph/0502129
- Erdl, H., & Schneider, P. 1993, *A&A*, 268, 453
- Gaudi, B. S. 1998, *ApJ*, 506, 533
- Gaudi, B. S., & Han, C. 2004, *ApJ*, 611, 528, arXiv:astro-ph/0402417
- Gaudi, B. S., et al. 2008, *Science*, 319, 927, arXiv:astro-ph:08021920
- Gelman, A., & Rubin, D. 1992, *Statistical Science*, 7, 457
- Geweke, J. 1992, *Bayesian Statistics 4* (Oxford University Press, J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (eds)), 169–193

- Ghosh, H. et al. 2004, ApJ, 615, 450, arXiv:astro-ph/0405500
- Gould, A. 2008, ApJ, 681, 1593, 0801.2578
- Gould, A., & Loeb, A. 1992, ApJ, 396, 104
- Gould, A., Miralda-Escude, J., & Bahcall, J. N. 1994, ApJ, 423, L105+
- Greaves, J. S., & Wyatt, M. C. 2010, ArXiv e-prints, 1001.5177
- Greaves, J. S., Wyatt, M. C., & Bryden, G. 2009, MNRAS, 397, 757, 0907.3677
- Griest, K., & Hu, W. 1992, ApJ, 397, 362
- Griest, K., & Safizadeh, N. 1998, ApJ, 500, 37
- Han, C., & Gould, A. 2003, ApJ, 592, 172, arXiv:astro-ph/0303309
- Han, C., & Jeong, Y. 1998, MNRAS, 301, 231, arXiv:astro-ph/9804273
- Hanson, K. 2002, talk presented at the 7th International Meeting on Bayesian Statistics, Valencia, Spain, 57, 97
- Harvey, P. M., Wilking, B. A., & Joy, M. 1984, Nature, 307, 441
- Hastings, W. K. 1970, Biometrika, 57, 97
- Hillenbrand, L. A. et al. 2008, ApJ, 677, 630, 0801.0163
- Holtzman, J. A., Watson, A. M., Baum, W. A., Grillmair, C. J., Groth, E. J., Light, R. M., Lynds, R., & O'Neil, Jr., E. J. 1998, AJ, 115, 1946, arXiv:astro-ph/9801321
- Jaroszynski, M. et al. 2004, Acta Astronomica, 54, 103, arXiv:astro-ph/0408243
- Jaroszynski, M., et al. 2006, Acta Astronomica, 56, 307, arXiv:astro-ph/0701919
- Kains, N. et al. 2009, MNRAS, 395, 787, 0901.1285
- Kenyon, S. J., & Bromley, B. C. 2002, AJ, 123, 1757, arXiv:astro-ph/0111384
- Kervella, P., & Fouqué, P. 2008, A&A, 491, 855, 0809.1732
- Koerner, D. W. et al. 2010, ApJ, 710, L26
- Kubas, D., et al. 2005, A&A, 435, 941, arXiv:astro-ph/0502018

- Lestrade, J., Wyatt, M. C., Bertoldi, F., Menten, K. M., & Labaigt, G. 2009, *A&A*, 506, 1455, 0907.4782
- Liseau, R. et al. 2008, *A&A*, 480, L47, 0803.1294
- Lisse, C. M., Beichman, C. A., Bryden, G., & Wyatt, M. C. 2007, *ApJ*, 658, 584, arXiv:astro-ph/0611452
- Löhne, T., Krivov, A. V., & Rodmann, J. 2008, *ApJ*, 673, 1123, 0710.4294
- Lovis, C. et al. 2006, *Nature*, 441, 305, arXiv:astro-ph/0703024
- Maness, H. L. et al. 2009, *ApJ*, 707, 1098, 0910.5223
- Mao, S., & Paczyński, B. 1991, *ApJ*, 374, L37
- Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355
- Metropolis, N., Rosenbluth, A., Teller, A., & Teller, E. 1953, *J. Chem. Phys.*, 21, 1087
- Meyer, M. R. et al. 2008, *ApJ*, 673, L181, 0712.1057
- Moro-Martín, A. et al. 2007, *ApJ*, 668, 1165, 0706.3368
- Najita, J., & Williams, J. P. 2005, *ApJ*, 635, 625, arXiv:astro-ph/0508165
- Natta, A. 2004, in *Astronomical Society of the Pacific Conference Series*, Vol. 324, *Debris Disks and the Formation of Planets*, ed. L. Caroff, L. J. Moon, D. Backman, & E. Praton, 20–+
- Nemiroff, R. J., & Wickramasinghe, W. A. D. T. 1994, *ApJ*, 424, L21, arXiv:astro-ph/9401005
- Paczynski, B. 1986, *ApJ*, 304, 1
- Poindexter, S., Afonso, C., Bennett, D. P., Glicenstein, J., Gould, A., Szymański, M. K., & Udalski, A. 2005, *ApJ*, 633, 914, arXiv:astro-ph/0506183
- Raftery, A., & Lewis, S. 1992, *Bayesian Statistics 4* (Oxford University Press, J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (eds)), 763–773
- Roberts, G. 1992, *Markov Chain Monte Carlo in Practice* (Chapman & Hall, London, W.R.Gilks, S.Richardson & D.J.Spiegelhalter (eds)), 45–57
- Schneider, G. et al. 2006, *ApJ*, 650, 414, arXiv:astro-ph/0606213

- Schneider, P., & Weiss, A. 1992, *A&A*, 260, 1
- Sheret, I., Dent, W. R. F., & Wyatt, M. C. 2004, *MNRAS*, 348, 1282, arXiv:astro-ph/0311593
- Siegler, N., Muzerolle, J., Young, E. T., Rieke, G. H., Mamajek, E. E., Trilling, D. E., Gorlova, N., & Su, K. Y. L. 2007, *ApJ*, 654, 580, arXiv:astro-ph/0609141
- Smith, B. A., & Terrile, R. J. 1984, *Science*, 226, 1421
- Smith, R., Wyatt, M. C., & Dent, W. R. F. 2008, *A&A*, 485, 897, 0804.4580
- Smith, R., Wyatt, M. C., & Haniff, C. A. 2009, *A&A*, 503, 265, 0906.3704
- Soldner, J. 1804, *Berliner Astronomisches Jahrbuch*, 161
- Stanek, K. Z., & Garnavich, P. M. 1998, *ApJ*, 503, L131+, arXiv:astro-ph/9802121
- Takeda, G., Ford, E. B., Sills, A., Rasio, F. A., Fischer, D. A., & Valenti, J. A. 2007, *ApJS*, 168, 297, arXiv:astro-ph/0607235
- Tegmark, M. et al. 2004, *Phys. Rev. D*, 69, 103501, arXiv:astro-ph/0310723
- Trilling, D. E. et al. 2008, *ApJ*, 674, 1086, 0710.5498
- Tsapras, Y., et al. 2009, *Astronomische Nachrichten*, 330, 4, 0808.0813
- Udalski, A. 2003, *Acta Astronomica*, 53, 291
- Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., & Mateo, M. 1992, *Acta Astronomica*, 42, 253
- Valenti, J. A., & Fischer, D. A. 2005, *ApJS*, 159, 141
- Walsh, B. 2002, *Lecture Notes for EEB 596z, University of Arizona*, 21, 1087
- Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, *Nature*, 279, 381
- Witt, H. J. 1990, *A&A*, 236, 311
- Witt, H. J., & Mao, S. 1994, *ApJ*, 430, 505
- Wolszczan, A., & Frail, D. A. 1992, *Nature*, 355, 145
- Wood, A., & Mao, S. 2005, *MNRAS*, 362, 945, arXiv:astro-ph/0507210

- Wyatt, M. C., & Dent, W. R. F. 2002, MNRAS, 334, 589, arXiv:astro-ph/0204034
- Wyatt, M. C., Greaves, J. S., Dent, W. R. F., & Coulson, I. M. 2005, ApJ, 620, 492, arXiv:astro-ph/0411061
- Wyatt, M. C., Smith, R., Greaves, J. S., Beichman, C. A., Bryden, G., & Lisse, C. M. 2007a, ApJ, 658, 569, arXiv:astro-ph/0610102
- Wyatt, M. C., Smith, R., Su, K. Y. L., Rieke, G. H., Greaves, J. S., Beichman, C. A., & Bryden, G. 2007b, ApJ, 663, 365, arXiv:astro-ph/0703608
- Zwicky, F. 1937, ApJ, 86, 217